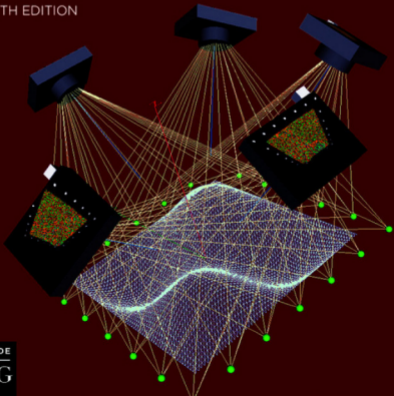*Thomas Luhmann, Stuart Robson,*
Stephen Kyle, Jan Boehm

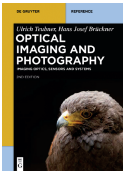# CLOSE-RANGE PHOTOGRAMMETRY AND 3D IMAGING

4TH EDITION

Thomas Luhmann, Stuart Robson, Stephen Kyle, Jan Boehm
Close-Range Photogrammetry and 3D Imaging

Also of interest

Thomas Luhmann, Stuart Robson,
Stephen Kyle, Jan Boehm

Close-Range Photogrammetry and 3D Imaging

—

4th Edition

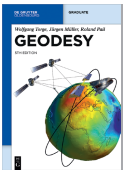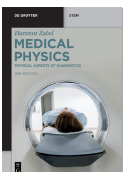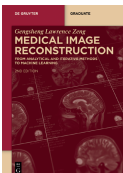**DE GRUYTER**

Authors

Prof. Dr. Ing. habil. Thomas Luhmann
Jade Hochschule
Institut für Angewandte Photogrammetrie
und Geoinformatik (IAPG)
Ofener Str. 16/19
26121 Oldenburg
luhmann@jade-hs.de

Dr. Stephen Kyle
University College London
Dept. of Geomatic Engineering
Chadwick Building
Gower Street
London WC1E 6BT
Great Britain
s.kyle@ucl.ac.uk

Prof. Dr. Stuart Robson
University College London
Dept. of Geomatic Engineering
Chadwick Building
Gower Street
London WC1E 6BT
Great Britain
s.robson@ucl.ac.uk

Prof. Dr. Jan Boehm
University College London
Dept. of Geomatic Engineering
Chadwick Building
Gower Street
London WC1E 6BT
Great Britain
j.boehm@ucl.ac.uk

Preface

The first edition of "Close Range Photogrammetry" was published in 2006 by Whittles Publishing. This was a translated and extended version of the original German book "Nahbereichsphotogrammetrie" and was well received by the large international community of photogrammetrists, metrologists and computer vision experts. This success was further recognized by the International Society of Photogrammetry and Remote Sensing (ISPRS) which awarded the authors the then newly inaugurated Karl Kraus Medal for excellence in authorship (2010).

The second and third editions, entitled "Close-Range Photogrammetry and 3D Imaging", were published by de Gruyter in 2014 and 2019. They were based on the latest German versions of "Nahbereichsphotogrammetrie" but extended to reflect new methods and systems for 3D imaging, particularly in the field of image analysis. Currently also, versions in Russian and Arabic are available and the preparation of a Spanish version is in progress.

Due to the rapid pace of development in our field, there is a constant need to review the progress made in our technologies and their application. Hence, only four years after edition 3, this current 4th edition again incorporates state-of-the-art updates on diverse topics such as colour processing. There is also new content, for example covering developments in 3D sensors, mobile scanning, metrology systems, Augmented and Virtual Reality (AR/VR) and applications.

Three-dimensional information acquired from imaging sensors is widely used and accepted. The field of photogrammetry, optical 3D metrology and 3D imaging is still growing, especially in areas which have no traditional link to photogrammetry and geodesy. However, whilst 3D imaging methods are established in many scientific communities, photogrammetry is still an engineering-driven technique where quality and accuracy play an important role.

It is the expressed objective of the authors to appeal to non-photogrammetrists and experts from many other fields in order to transfer knowledge and avoid re-invention of the wheel. The structure of the book therefore assumes different levels of pre-existing knowledge, from beginner to scientific expert. For this reason, the book also presents a number of fundamental techniques and methods in mathematics, adjustment techniques, physics, optics, image processing and others. Although this information may also be found in other textbooks, the objective here is to create a closer link between different fields and present a common notation for equations and parameters.

The authors would also like to express their gratitude to the many generous colleagues who have helped complete the work. In addition to all those who have already contributed to the previous editions, on this occasion we would particularly like to thank: Maria Chizhova, Matevz Domajnko, Francesco Fassi, Andreas Georgopoulos, Ute Greve-Luhmann, Albrecht Grimm, Heidi Hastedt, Oliver Kahmen,

Oldenburg /London/Guernsey, June 2023

Thomas Luhmann, Stuart Robson, Stephen Kyle, Jan Boehm

Contents

1 Introduction

1.1 Overview

1.1.1 Content

Chapter 1 provides an overview of the fundamentals of photogrammetry, with particular reference to close-range measurement. After a brief discussion of the principal methods and systems, typical areas of applications are presented. The chapter ends with a short historical review of close-range photogrammetry.

Chapter 2 deals with mathematical basics. These include the definition of some important coordinate systems and the derivation of geometric transformations which are needed for a deeper understanding of topics presented later. In addition, a number of geometrical elements important for object representation are discussed. The chapter concludes with a summary of least squares adjustment and statistics.

Chapter 3 is concerned with photogrammetric image acquisition for close-range applications. After an introduction to physical basics and the principles of image acquisition, geometric fundamentals and imaging models are presented. There follow discussions of digital imaging equipment as well as specialist areas of image recording. The chapter ends with a summary of illumination techniques and some general targeting principles.

Analytical methods of image orientation and object reconstruction are presented in Chapter 4. The emphasis here is on bundle triangulation. The chapter also presents methods for dealing with single, stereo and multiple image configurations based on measured image coordinates, and concludes with a review of panoramic and multi-media (underwater) photogrammetry.

Chapter 5 brings together many of the relevant methods of digital photogrammetric image analysis. Those which are most useful to dimensional analysis and three-dimensional (3D) object reconstruction are presented, in particular methods for feature extraction and image matching.

Photogrammetric systems developed for close-range measurement are discussed in Chapter 6. As targeting is an integral part of the system, some detailed targeting concepts are first presented, followed by solutions for scale and reference system definition. The systems themselves are classified into interactive systems, tactile and laser-based measuring systems, systems for the measurement of points and surfaces, systems for dynamic processes, systems on mobile platforms such as drones and, finally, 3D visualization and projection systems.

Chapter 7 discusses imaging project planning and quality criteria for practical measurement tasks. After an introduction to network planning and optimization, quality criteria and approaches to accuracy assessment are discussed. The chapter concludes with strategies for camera and system calibration.

Finally, Chapter 8 uses case studies and examples to demonstrate the potential for close-range photogrammetry in fields such as architecture, heritage conservation, construction, manufacturing industry, medicine and science.

1.1.2 References

Relevant literature is directly referenced within the text in cases where it is highly recommended for the understanding of particular sections. In general, however, further reading is presented in Chapter 9 which provides an extensive list of thematically ordered literature. Here each chapter in the book is assigned a structured list of reference texts and additional reading. Efforts have been made to suggest reference literature which is easy to access. In addition, the reader is advised to make use of conference proceedings, journals and the webpages of universities, scientific societies and commercial companies for up-to-date information.

1.2 Fundamental methods

1.2.1 The photogrammetric process

Photogrammetry encompasses methods of image measurement and interpretation in order to derive the shape and location of an object from one or more photographs of that object. In principle, photogrammetric methods can be applied in any situation where the object to be measured can be photographically recorded. The primary purpose of a photogrammetric measurement is the three-dimensional reconstruction of an object in digital form (coordinates, point clouds, 3D models and derived geometric elements) or graphical form (images, drawings, maps and 3D visualisations). The photograph or image represents a store of information which can be re-accessed at any time.

Fig. 1.1: Photogrammetric images.

Fig. 1.1 shows examples of photogrammetric images. The reduction of a three-dimensional object to a two-dimensional image implies a loss of information. In the first place, object areas which are not visible in the image cannot be reconstructed from it. This not only includes hidden parts of an object such as the rear of a building, but also regions which cannot be interpreted due to lack of contrast or size limitations. Whereas the position in space of each point on the object may be defined by three coordinates, there are only two coordinates available to define the position of its image. There are geometric changes caused by the shape of the object, the relative positioning of camera and object, perspective imaging and optical lens defects. Finally, there are also radiometric (colour) changes since the reflected electromagnetic radiation recorded in the image is affected by the light source, the transmission media (air, glass) and the light-sensitive recording medium.
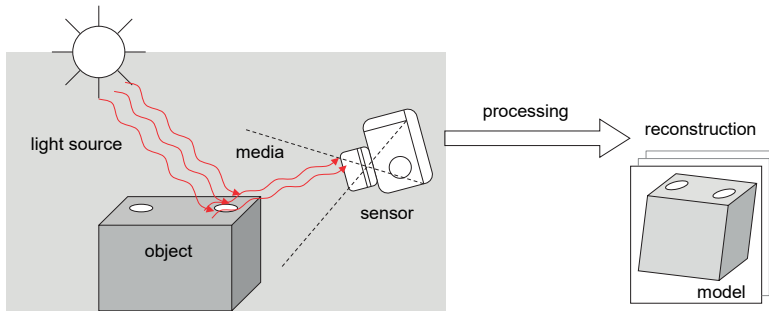
Fig. 1.2: From object to image.

For the reconstruction of an object from images it is therefore necessary to describe the optical process by which an image is created. This includes all elements which contribute to this process, such as light sources, properties of the surface of the object, the medium through which the light travels, sensor and camera technology, image processing, and further processing (Fig. 1.2).

Methods of image interpretation and measurement are then required which permit the identification of an object by its form, brightness or colour distribution in the image. For every image point, values in the form of radiometric data (intensity, grey value, colour value) and geometric data (position in image) can then be obtained. Optionally, this requires appropriate geometric and radiometric calibration procedures.

From these measurements and a mathematical transformation between image and object space, the object can finally be modelled.

Fig. 1.3 simplifies and summarizes this sequence. The left-hand side indicates the principal instrumentation used whilst the right hand side indicates the methods involved. Together with the physical and mathematical models, human knowledge,

experience and skill play a significant role. They determine the extent to which the reconstructed model corresponds to the imaged object or fulfils the task objectives.
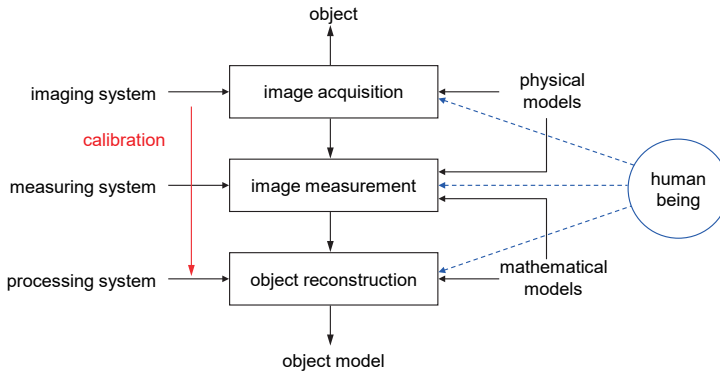
Fig. 1.3: The photogrammetric process: from object to model.

1.2.2 Aspects of photogrammetry

Because of its varied application areas, close-range photogrammetry has a strong interdisciplinary character. There are not only close connections with other measurement techniques but also with fundamental sciences such as mathematics, physics, information sciences and biology.

Close-range photogrammetry also has significant links with aspects of graphics and photographic science, for example computer graphics and computer vision, digital image processing, computer aided design (CAD), augmented and virtual reality (AR/VR) , geographic information systems (GIS) and cartography.

Traditionally, there are further strong associations of close-range photogrammetry with the techniques of surveying, particularly in the areas of adjustment methods and engineering surveying. With the increasing application of photogrammetry to industrial metrology and quality control, links have been created in other directions, too. Furthermore, there is increasing demand for optical 3D measurement systems in rapidly developing fields such as autonomous vehicles and medical technology.

Fig. 1.4 gives an indication of the relationship between size of measured object, required measurement accuracy and relevant technology. Although there is no hard-and-fast definition, it may be said that close-range photogrammetry usually applies to objects ranging from about 0.1 m to 200 m in size, with accuracies under 0.1 mm at the smaller end (manufacturing industry) and around 1 cm at the larger end (AEC - architecture, engineering and construction).
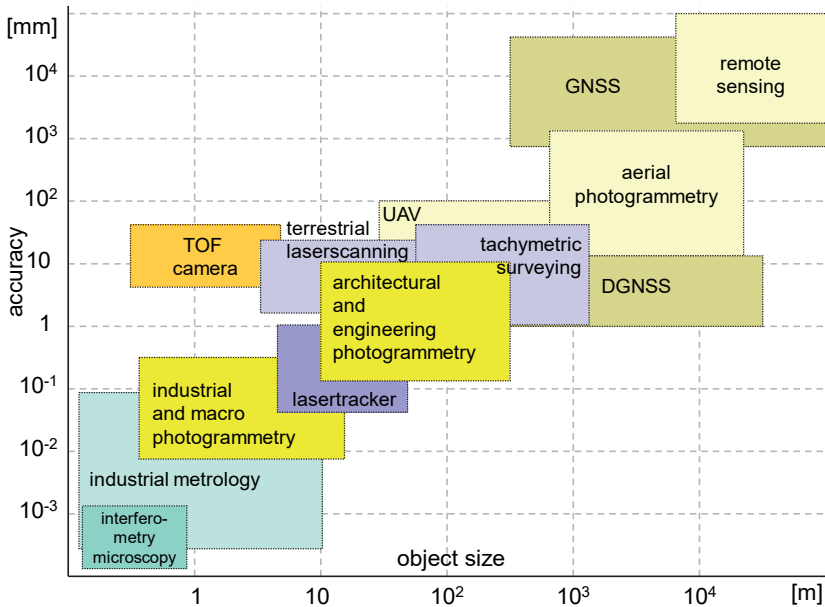
Fig. 1.4: Relationship between measurement methods and object size and accuracy (The extent of the named methods and application areas is indicative only).

Optical methods using light as the information carrier lie at the heart of non-contact 3D measurement techniques. Measurement techniques using electromagnetic waves may be subdivided in the manner illustrated in Fig. 1.5. Techniques based on light waves are as follows:

– Triangulation techniques:
  Photogrammetry (single, stereo and multiple imaging), Structure-from-Motion (SfM) and vision-based Simultaneous Localisation and Mapping (Visual SLAM), structured light (light section procedures, fringe projection, phase measurement, moiré topography), focusing methods, shadow methods, etc.
– Interferometry:
  Optically coherent time-of-flight measurement, holography, speckle interferometry, coherent radar.
– Time-of-flight measurement:
  Distance measurement by optical modulation methods, pulse modulation, etc.

The clear structure of Fig. 1.5 is blurred in practice since multi-sensor and hybrid measurement systems utilize different principles in order to combine the advantages of each.
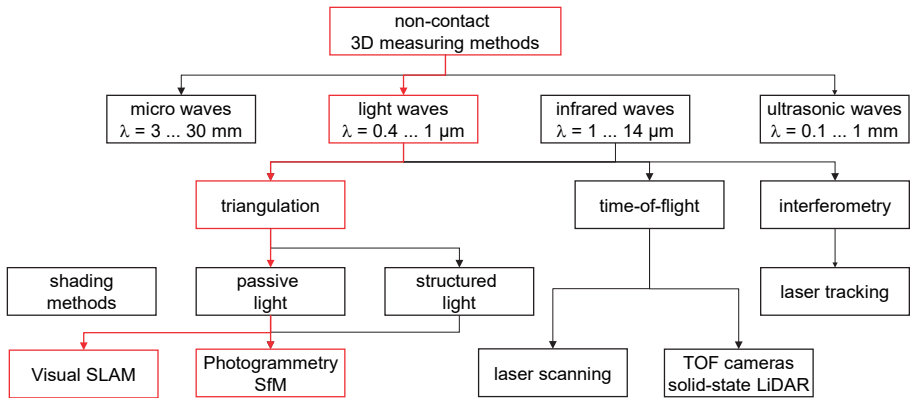
Fig. 1.5: Non-contact 3D measuring methods.

Photogrammetry can be categorized in a multiplicity of ways:

— By camera position and object distance *h*:

  – satellite photogrammetry: processing of remote sensing and
                                            satellite images, *h* > ca. 200 km
  – aerial photogrammetry: processing of aerial photographs, *h* > ca. 300 m
  – UAV photogrammetry: processing of aerial photographs from drones,
                                              h < ca. 120 m
  – terrestrial photogrammetry: measurements from a fixed terrestrial location
  – close-range photogrammetry: imaging distance *h* < ca. 300 m
  – underwater photogrammetry: object recording in or through water
  – macro photogrammetry: image scale > 1 (microscope imaging)
  – mobile mapping: data acquisition from moving vehicles,
                                              h < ca. 100 m

— By method of recording and processing:

  – plane table photogrammetry: graphical evaluation (until ca. 1930)
  – analogue photogrammetry: analogue cameras, opto-mechanical
                                            measurement systems (until ca. 1980)
  – analytical photogrammetry: analogue images, computer-controlled measurement
  – digital photogrammetry: digital images, computer-controlled measurement
  – videogrammetry: digital image acquisition and measurement
  – panoramic photogrammetry: panoramic imaging and processing
  – line photogrammetry: analytical methods based on straight lines and
                                          polynomials
  – multi-media photogrammetry: recording through media of different refraction

–   By number of measurement images:

– single-image photogrammetry:	single-image processing, mono-plotting, rectification, orthophotos
– stereo photogrammetry:	dual image processing, stereoscopic measurement
– multi-image photogrammetry:	n images where n>2, bundle triangulation

–   By availability of measurement results:

– offline photogrammetry:	sequential, digital image recording, separated in time or location from measurement
– online photogrammetry:	simultaneous, multiple, digital image recording, immediate measurement
– real-time photogrammetry:	recording and measurement completed within a specified time period particular to the application

–   By application or specialist area:

– architectural photogrammetry:	architecture, heritage conservation, archaeology
– engineering photogrammetry:	general engineering (construction) applications
– industrial photogrammetry:	industrial (manufacturing) applications
– forensic photogrammetry:	applications to diverse legal problems
– shape from stereo:	stereo image processing (computer vision)
– structure-from-motion:	multi-image processing (computer vision)

1.2.3 Image-forming model

Photogrammetry is a three-dimensional measurement technique which uses *central projection imaging* as its fundamental mathematical model (Fig. 1.6). Shape and position of an object are determined by reconstructing bundles of rays in which, for each camera, each image point P', together with the corresponding *perspective centre* O', defines the spatial direction of the ray to the corresponding object point P. Provided the imaging geometry within the camera and the location of the imaging system in object space are known, then every image ray can be defined in 3D object space. Since photogrammetry is basically a method for measurement of spatial directions (angles), i.e. it measures shape not scale, at least one absolute scale information in object space must therefore be given, e.g. by a measured distance or by control points with known coordinates.

From the intersection of at least two corresponding (homologous), spatially separated image rays, an object point can be located in three dimensions. In stereo photogrammetry two images are used to achieve this. In multi-image photogrammetry the number of images involved is, in principle, unlimited.
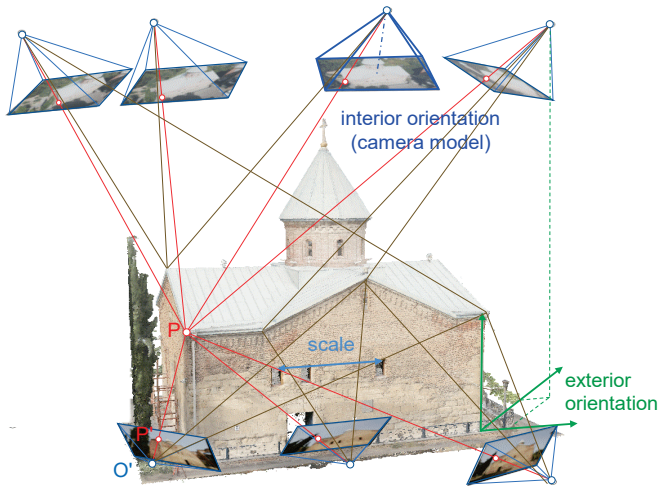
Fig. 1.6: Principle of photogrammetric measurement.

The *interior orientation* parameters describe the internal geometric model of a camera. Photogrammetric usage, deriving from German, applies the word to groups of camera parameters. Exterior (extrinsic) orientation parameters incorporate this angular meaning but extend it to include position. Interior (intrinsic) orientation parameters, which include a distance, two coordinates and a number of polynomial coefficients, involve no angular values. The use of the terminology here underlines the connection between two very important, basic groups of parameters.
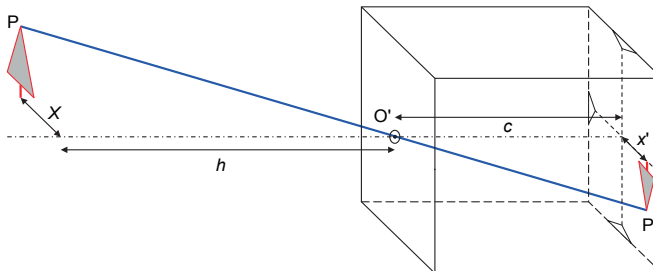
Fig. 1.7: Pinhole camera model.

With the model of the pinhole camera as its basis (Fig. 1.7), the most important reference location is the perspective centre O, through which all image rays pass. The interior orientation defines the position of the perspective centre relative to a reference system fixed in the camera (image coordinate system), as well as departures

from the ideal central projection (image distortion). The most important parameter of interior orientation is the principal distance, c, which defines the distance between image plane and perspective centre (see section 3.3.2).

A real and practical photogrammetric camera will differ from the pinhole camera model. The necessity of using a relatively complex objective lens, a camera housing which is not built for stability and an image recording surface which may be neither planar nor perpendicular to the optical axis of the lens will all give rise to departures from the ideal imaging geometry. The interior orientation, which will include parameters defining these departures, must be determined by calibration for every camera (see section 3.3).

A fundamental property of a photogrammetric image is the *image scale* or *photo scale*. The photo scale factor m defines the relationship between the object distance, h, and principal distance, c. Alternatively it is the relationship between a distance, X, parallel to the image plane in the object, and the corresponding distance in image space, x':

$$m = \frac{h}{c} = \frac{X}{x'} \tag{1.1}$$

Since every object point in 3D space is, in general, at a different distance to the camera, an image does not have a uniform image scale. Only in the special case of an orthogonal view above a planar object, is the image scale equal across the whole image. In many applications, however, a mean image scale is given for the entire image, e.g. for planning purposes or accuracy estimates.

The image scale is in every case the deciding factor in resolving image details, defined by the ground sample distance (GSD) which is derived from the pixel spacing $\Delta s'$ in the camera:

$$GSD = m \cdot \Delta s' \tag{1.2}$$

The image scale also determines the photogrammetric measurement accuracy, since any measurement error in the image is multiplied in the object space by the scale factor (see section 3.3.1).

The *exterior orientation* parameters specify the spatial position and orientation of the camera in a global coordinate system. The exterior orientation is described by the coordinates of the perspective centre in the global system and, commonly, three suitably defined angles expressing the rotation of the image coordinate system with respect to the global system (see section 4.2.1). The exterior orientation parameters are calculated indirectly, after measuring image coordinates of well identified object points.

Every measured image point corresponds to a spatial direction from projection centre to object point. The length of the direction vector is initially unknown, i.e. every object point lying on the line of this vector generates the same image point. In

other words, although every three-dimensional object point transforms to a unique image point for given orientation parameters, a unique reversal of the projection is not possible. The object point can be located on the image ray, and thereby absolutely determined in object space, only by intersecting the ray with an additional known geometric element such as a second spatial direction or an object plane.

Every image generates a spatial *bundle of rays*, defined by the imaged points and the perspective centre, in which the rays were all recorded at the same point in time. If all the bundles of rays from multiple images are intersected as described above, a dense network is created. For an appropriate imaging configuration, such a network has the potential for high geometric strength. Using the method of *bundle triangulation* any number of images (ray bundles) can be simultaneously oriented, together with the calculation of the associated three-dimensional object point locations (Fig. 1.6, Fig. 1.8, see section 4.4).

For absolute scaling of a photogrammetric network, at least one known distance in object space must be known, e.g. directly from measured distances or calibrated baselines between cameras, or derived from coordinates of known reference points.
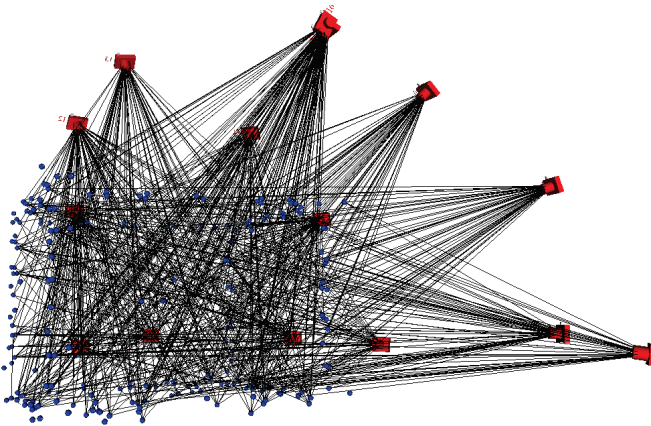
Fig. 1.8: Bundle of rays from multiple images.

1.2.4 Photogrammetric systems and procedures

1.2.4.1 Digital system

With few exceptions, photogrammetric image recording today uses digital cameras supported by image processing based on methods of visual and digital image analysis. A closed digital system is therefore possible which can completely measure an object directly on site and without any significant time loss between image acquisition and delivery of results (Fig. 1.9).

By using suitably targeted object points and automatic pattern recognition, complex photogrammetric tasks can be executed fully automatically, hence eliminating the need for manual image measurement, orientation and processing. This approach is particularly important in industrial applications where, in the first instance, 3D coordinates of discrete points are required. The measurement of free-form surfaces through the use of dense point clouds is performed by stereo or multi-image matching of textured object areas. By adopting the method of structure-from-motion (SfM), arbitrary configurations of images can be oriented fully automatically. In contrast, the measurement of linear object structures (e.g. for creating maps or drawings) largely remains a visual, interactive process.

Digital image recording and processing offer the possibility of a fast, closed data flow from taking the images to presenting the results. Two general procedures are distinguished here. *Offline photogrammetry* uses a single camera with measurement results generated after all images have first been recorded and then evaluated together. *Online photogrammetry* records simultaneously using at least two cameras, with immediate generation of results. If the result is delivered within a certain process-specific time period, the term *real-time photogrammetry* is commonly used.
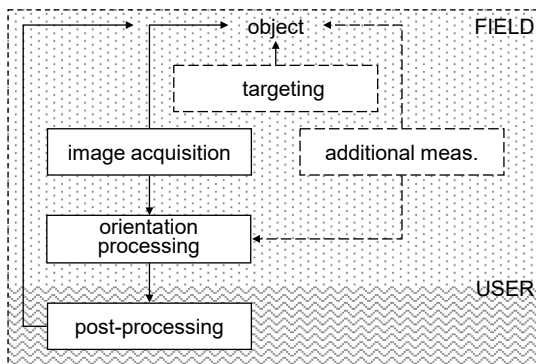
Fig. 1.9: Digital photogrammetric system.

Automation and short processing cycles enable a direct integration with other processes where decisions can be made on the basis of feedback of the photogrammetric results. Digital systems are therefore critical to the application of photogrammetry in complex real-time processes, in particular industrial manufacturing and assembly, robotics and medicine where feedback with the object or surroundings takes place.

When imaging scenes with purely natural features, without the addition of artificial targets, the potential for automation is much lower. An intelligent evaluation of object structures and component forms demands a high degree of visual

interpretation which is conditional on a corresponding knowledge of the application and further processing requirements. However, even here simple software interfaces, and robust techniques of image orientation and camera calibration, make it possible for non-expert users to carry out photogrammetric recording and analysis.

1.2.4.2 Recording and analysis procedures

Fig. 1.10 shows the principal procedures in close-range photogrammetry which are briefly summarized in the following sections.
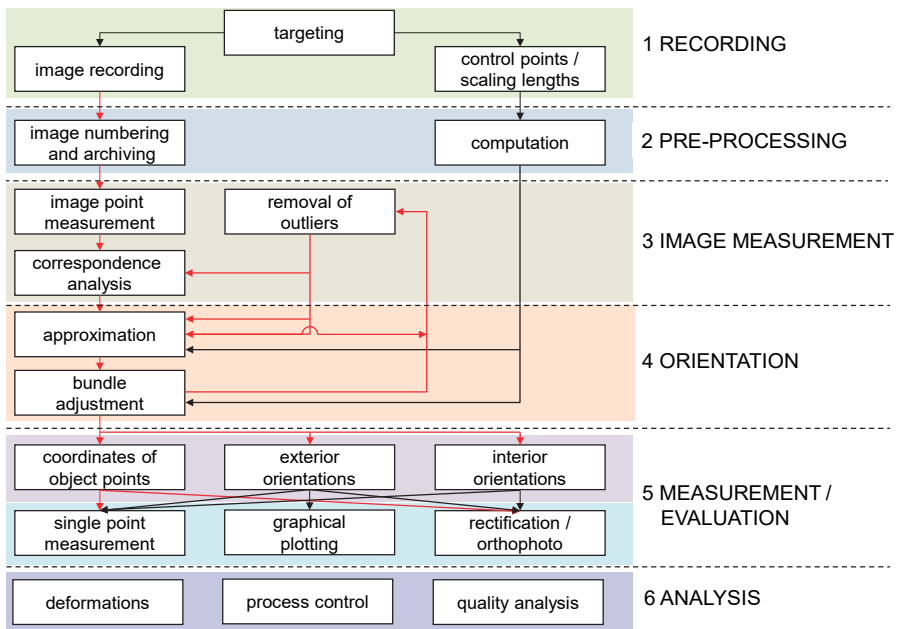
Fig. 1.10: Recording and analysis procedures (red - can be automated).

1. Recording
   a. Targeting (signalizing):
      target selection and attachment to object features to improve automation and increase the accuracy of target measurement in the image; definition of control points.
   b. Determination of control points or scaling lengths:
      creation of a global object coordinate system by definition of reference (control) points and/or reference lengths (scales).
   c. Image recording:
      digital image recording of the object with a photogrammetric system.

2. Pre-processing
   a. Numbering and archiving:
      assigning photo numbers to identify individual images and archiving or
      storing the photographs.
   b. Computation:
      calculation of reference point coordinates and/or distances from survey
      observations, e.g. using network adjustment.
3. Orientation
   a. Measurement of image points:
      identification and measurement of reference and scale points,
      identification and measurement of tie points.
   b. Correspondence analysis:
      matching of identical points (features) in all images.
   c. Approximation:
      calculation of approximate (starting) values for unknown quantities to be
      calculated by the bundle adjustment.
   d. Bundle adjustment:
      adjustment program which simultaneously calculates parameters of both
      interior and exterior orientation as well as the object point coordinates
      which are required for subsequent analysis.
   e. Removal of outliers:
      detection and removal of gross errors which mainly arise during
      measurement of image points.
4. Measurement and evaluation
   a. Single point measurement:
      creation of three-dimensional object point coordinates, e.g. point clouds,
      for further numerical processing.
   b. Graphical plotting:
      production of scaled maps or plans in analogue or digital form, e.g. hard
      copies for maps and electronic files for CAD models or GIS.
   c. Rectification/Orthophoto/3D visualization:
      generation of transformed images or image mosaics which remove the
      effects of tilt relative to a reference plane (rectification) and/or remove the
      effects of perspective (orthophoto).

5. Analysis
   a. Deformations:
      Analysis of changes with respect to previous epochs.
   b. Process control:
      Control of external systems using the measured 3D data.

c. Quality analysis:
    Comparison against nominal values in shape and size or accuracy
    verifications.

To a significant extent, this sequence can be automated (see connections in red in Fig. 1.10). This automation requires that either object features are suitably marked and identified using coded targets or, if there are sufficient textured and dense images available, processing can be done using structure-from-motion. In both cases the calculation of initial values and removal of outliers (gross errors) must be done by robust estimation methods.

1.2.5 Photogrammetric products

In general, photogrammetric systems supply three-dimensional object coordinates derived from image measurements (Fig. 1.11). From these, further elements and dimensions can be derived, for example lines, distances, areas and surface definitions, parameters of 6 degrees of freedom (6DOF) or navigation data as well as quality information such as comparisons against design and machine control data. The direct determination of geometric elements such as straight lines, planes and cylinders is also possible without explicit calculation of point coordinates. In addition, the recorded image is an objective data store which documents the state of the object at the time of recording. Hence, semantic information can be derived to create maps or qualified 3D models. Last but not least, the visual data can be provided as corrected camera images, rectifications or orthophotos.
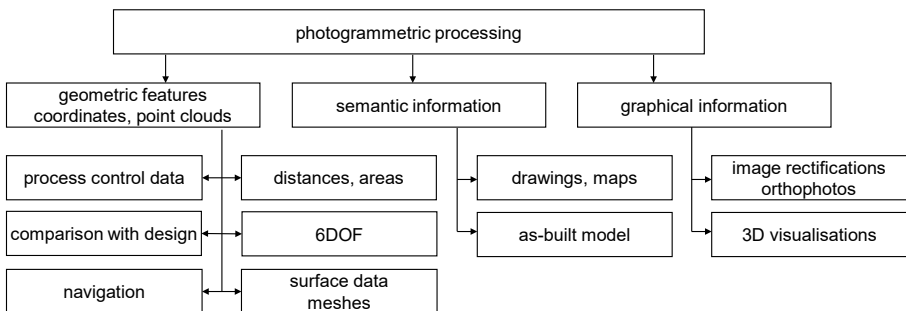
Fig. 1.11: Typical photogrammetric products.

Fig. 1.12 shows an example of two measurement images of a multi-image network for the 3D digitisation of a statue (State Museum Nature and Human, Oldenburg). Fig.

1.13 shows the visualised 3D model and the back-projection of part of the model's surface triangulation into the original images.

Fig. 1.12: Measurement images of an African statue.

Fig. 1.13: Reconstructed 3D model (left) and back-projection into original images.

1.3 Application areas

Much shorter imaging ranges, typically from a few centimetres to a few hundred metres, and alternative recording techniques, differentiate close-range photo-grammetry from its aerial and satellite equivalents.

The following comments, based on ones made by Thompson as long ago as 1963, identify applications in general terms by indicating that photogrammetry and optical 3D measurement techniques are potentially useful when:
– the object to be measured is difficult to access but images are easily available;
– the object is not rigid and its instantaneous dimensions are required;

- it is not certain that measurement will be required at all, or even what measurements are required (i.e. the data is preserved for possible later evaluation);
- the object is very small;
- the use of direct measurement would influence the measured object or disturb events around it;
- real-time results are required;
- the simultaneous recording and the measurement of a very large number of points is required.

The following specific application areas (with examples) are amongst the most important in close-range photogrammetry:

Architecture, heritage conservation, archaeology:
- façade measurement
- historic building documentation
- deformation measurement
- reconstruction of damaged buildings
- mapping of excavation sites
- modelling monuments and sculptures
- 3D models and texturing

Fig. 1.14: Building record.

Automotive, machine and ship building industries:
- inspection of tooling jigs
- reverse engineering of design models
- manufacturing control
- optical shape measurement
- recording and analysing car safety tests
- robot calibration
- driver assistance systems
- autonomous driving

Fig. 1.15: Car safety test.

Aerospace industry:
- measurement of parabolic antennae and mirrors
- control of component assembly
- inspection of tooling jigs
- space simulations

Fig. 1.16: Parabolic mirror.

Engineering:
- as-built measurement of process plants
- measurement of large civil engineering sites
- deformation measurements
- pipework and tunnel measurement
- mining
- road and railway track measurement
- wind power systems

Fig. 1.17: Engineering.

Medicine and physiology:
- dental measurements
- spinal deformation
- plastic surgery
- neuro surgery
- motion analysis and ergonomics
- microscopic analysis
- computer-assisted surgery (navigation)
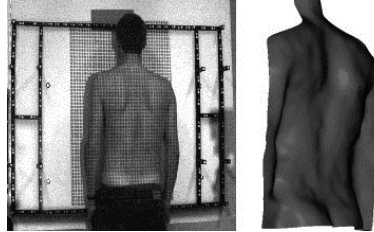
Fig. 1.18: Spinal analysis.

Animation and movie/film industries:
- body shape recording
- motion capture (of actors)
- 3D movies
- virtual reality (VR)
- augmented reality (AR)
- mixed reality (MR)
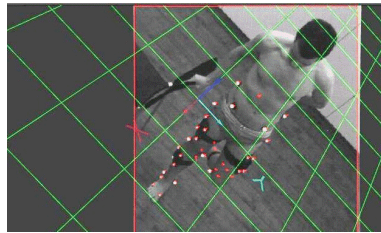
Fig. 1.19: Motion capture.

Police work and forensic analysis:
- accident recording
- scene-of-crime measurement
- legal records
- measurement of individuals

Fig. 1.20: Accident recording.

Information systems:
– building information modelling (BIM)
– facility management
– production planning
– image databases
– internet applications (digital globes)
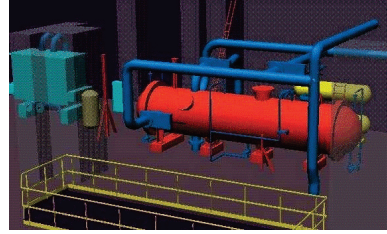
Fig. 1.21: Pipework measurement.

Natural sciences:
– liquid flow measurement
– wave topography
– crystal growth
– material testing
– glacier and soil movements
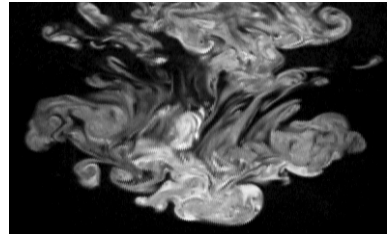
Fig. 1.22: Flow measurement.

In general, similar methods of recording and analysis are used for all application areas of close-range photogrammetry and the following features are shared:
– powerful image recording systems;
– freely chosen imaging configuration with almost unlimited numbers of pictures;
– photo orientation based on the technique of bundle triangulation;
– visual and digital analysis of the images;
– presentation of results in the form of 3D models, 3D coordinate files, CAD data, photographs or drawings.

Industrial and engineering applications make special demands of the photogrammetric technique:
– limited recording time on site (no significant interruption of industrial processes);
– delivery of results for analysis after only a brief time;
– high accuracy requirements;
– traceability of results to standard unit of dimension, the metre;
– proof of accuracy attained.

1.4 Historical development

It comes as a surprise to many that the history of photogrammetry is almost as long as that of photography itself and that, for at least the first fifty years, the predominant application of photogrammetry was to close-range, architectural measurement rather

than to topographical mapping. Only a few years after the invention of photography during the 1830s and 1840s by Fox Talbot in England, by Niepce and Daguerre in France, and by others, the French military officer Laussedat began experiments in 1849 into measuring from perspective views by working on the image of a façade of the Hotel des Invalides. Admittedly Laussedat, usually described as the first photogrammetrist, was in this instance using a camera lucida for he did not obtain photographic equipment until 1852.

Fig. 1.23 shows an early example of a photogrammetric camera, built by Brunner and by Laussedat, with a stable construction without moving components. Fig. 1.24 shows an example of Laussedat's work for military field mapping by "metrophotographie". As early as 1858 the German civil engineer Meydenbauer used photographs to draw plans of the cathedral of Wetzlar and by 1865 he had constructed his "great photogrammeter", a forerunner of the phototheodolite. In fact, it was Meydenbauer and Kersten, a geographer, who coined the word "photogrammetry", this first appearing in print in 1867.
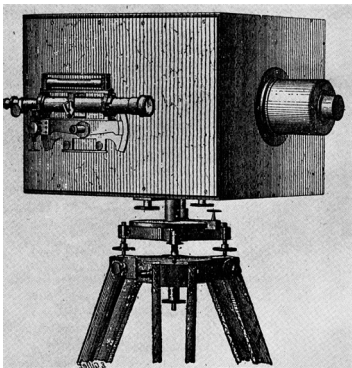
Fig. 1.23: One of the first photogrammetric cameras by Brunner, 1859 (Laussedat 1899).
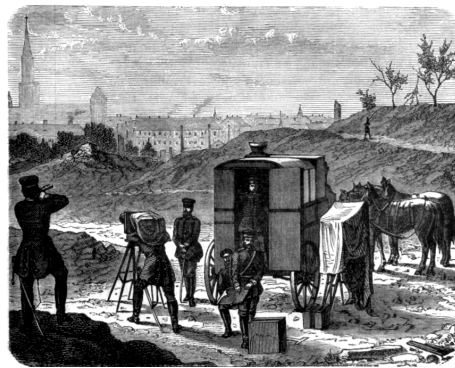
Fig. 1.24: Early example of photogrammetric field recording, about 1867 (Laussedat 1899).

Meydenbauer also made topographical surveys by means of plane table photogrammetry, for example in 1867 for the mapping of Freyburg a.d. Unstrut at a scale of 1:5000 (Fig. 1.25). He also used photography as an alternative to manual methods of measuring façades. For this he developed his own large-format, glass-plate cameras (see Fig. 1.26) and, between 1885 and 1909, compiled an archive of around 13000 metric[1] images of the most important Prussian architectural monuments. This represents a very early example of cultural heritage preservation by

---

1 A "metric" camera is defined as one with known and stable interior orientation.

photogrammetry. Fig. 1.27 shows an original metric image, Fig. 1.28 displays the façade map derived from it.
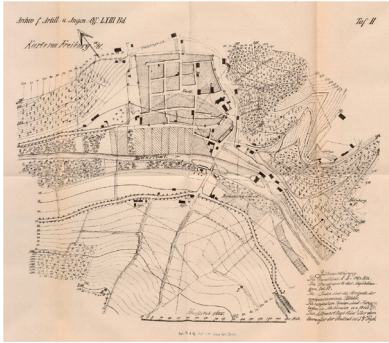
Fig. 1.25: Topographic surveying by Meydenbauer, 1867 (Burchardi & Meydenbauer 1868).
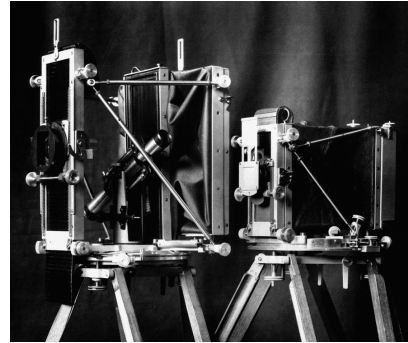
Fig. 1.26: Metric cameras by Meydenbauer (ca. 1890); left: 30x30 cm², right: 20x20 cm² (Albertz 2009).

Fig. 1.27: Metric image of the Marburg Elisabeth Church by Meydenbauer (1883).
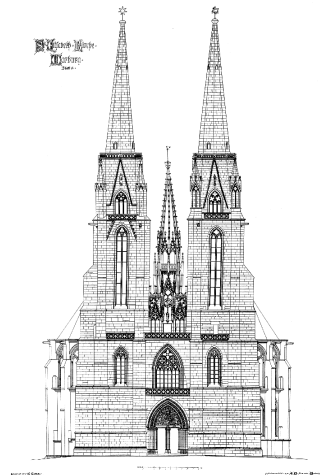
Fig. 1.28: Façade drawing of the Marburg Elisabeth Church (1883).

The phototheodolite, as its name suggests, represents a combination of camera and theodolite. The direct measurement of orientation angles leads to a simple photogrammetric orientation. A number of inventors, such as Porro and Paganini in

Italy, in 1865 and 1884 respectively, and Koppe in Germany, 1896, developed such instruments (Fig. 1.29).

Horizontal bundles of rays can be constructed from terrestrial photographs, with two or more permitting a point-by-point survey using intersecting rays. This technique, often called *plane table photogrammetry*, works well for architectural subjects which have regular and distinct features. However, for topographic mapping it can be difficult identifying the same feature in different images, particularly when they are well separated to improve accuracy. Nevertheless, despite the early predominance of architectural photogrammetry, mapping was still undertaken. For example, in the latter part of the 19th century, Paganini mapped the Alps, Stolze the Friday Mosque in Shiraz, Deville the Rockies and Jordan the Dachel oasis, whilst Finsterwalder developed analytical solutions.
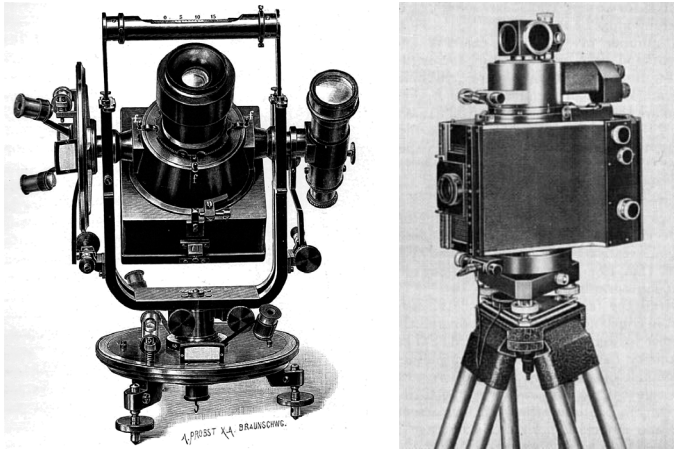
Fig. 1.29: Phototheodolite by Koppe (1889) and Zeiss Jena 19/1318 (ca. 1904).

The development of stereoscopic measurement around the turn of the century was a major breakthrough in photogrammetry. Following the invention of the stereoscope around 1830, and Stolze's principle of the floating measuring mark in 1893, Pulfrich in Germany and Fourcade in South Africa, at the same time but independently[2], developed the stereocomparator which implemented Stolze's principle. These enabled the simultaneous setting of measuring marks in the two comparator images, with calculation and recording of individual point coordinates (Fig. 1.30).

---

2 Pulfrich's lecture in Hamburg announcing his invention was given on 23rd September 1901, while Fourcade delivered his paper in Cape Town nine days later on 2nd October 1901.

Photogrammetry then entered the era of *analogue computation*, very different to the numerical methods of surveying. Digital computation was too slow at that time to compete with continuous plotting from stereo instruments, particularly of contours, and analogue computation became very successful for a large part of the 20th century.
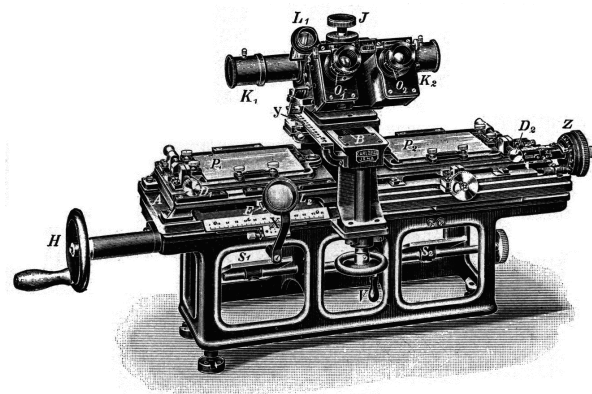
Fig. 1.30: Pulfrich's stereocomparator (Zeiss, 1901).

In fact, during the latter part of the 19th century much effort was invested in developing stereoplotting instruments for the accurate and continuous plotting of topography. In Germany, Hauck proposed a device and, in Canada, Deville claimed "the first automatic plotting instrument in the history of photogrammetry". Deville's instrument had several defects, but they inspired many developers such as Pulfrich and Santoni to overcome them.

In Germany, conceivably the most active country in the early days of photogrammetry, Pulfrich's methods were very successfully used in mapping. This inspired von Orel in Vienna to design an instrument for the "automatic" plotting of contours, which lead to the Orel-Zeiss Stereoautograph in 1909. In England, F. V. Thompson anticipated von Orel in the design of the Vivian Thompson stereoplotter and subsequently the Vivian Thompson Stereoplanigraph (1908). This was described by E. H. Thompson (1974) as "the first design for a completely automatic and thoroughly rigorous photogrammetric plotting instrument".

The rapid development of aviation, which began shortly after this, was another decisive influence on the course of photogrammetry. Not only is the Earth, photographed vertically from above, an almost ideal subject for the photogrammetric method, but also aircraft made almost all parts of the Earth accessible at high speed. In the first half, and more, of the 20th century these favourable circumstances allowed impressive development in photogrammetry, with tremendous economic benefit in air survey. On the other hand, the application of stereo photogrammetry to the

complex surfaces relevant to close-range work was impeded by far-from-ideal geometry and a lack of economic advantage.

Although there was considerable opposition from surveyors to the use of photographs and analogue instruments for mapping, the development of stereoscopic measuring instruments forged ahead in very many countries during the period between the First World War and the early 1930s. Meanwhile, non-topographic use was sporadic for the reasons that there were few suitable cameras and that analogue plotters imposed severe restrictions on principal distance, on image format and on disposition and tilts of cameras. Instrumentally complex systems were being developed using optical projection (for example Multiplex), opto-mechanical principles (Zeiss Stereoplanigraph) and mechanical projection using space rods (for example Wild A5, Santoni Stereocartograph), designed for use with aerial photography. By 1930 the Stereoplanigraph C5 was in production, a sophisticated instrument able to use oblique and convergent photography. Even if makeshift cameras had to be used at close range, experimenters at least had freedom in the orientation and placement of these cameras and this considerable advantage led to some noteworthy work.

Fig. 1.31: Wild P32 metric camera mounted on a theodolite.

Fig. 1.32: Wild C120 stereometric camera.

As early as 1933 Wild stereometric cameras were being manufactured and used by Swiss police for the mapping of accident sites, using the Wild A4 Stereoautograph, a plotter especially designed for this purpose. Such stereometric cameras comprise two identical metric cameras fixed to a rigid base of known length such that their axes are coplanar, perpendicular to the base and, usually, horizontal[3] (Fig. 3.38a, see section

---

3 This is sometimes referred to as the "normal case" of photogrammetry.

4.3.1.4). Other manufacturers have also made stereometric cameras (Fig. 1.32) and associated plotters (Fig. 1.34) and a great deal of close-range work has been carried out with this type of equipment. Initially glass plates were used in metric cameras in order to provide a flat image surface without significant mechanical effort (see example in Fig. 1.31, Fig. 1.33). From the 1950s, film was increasingly used in metric cameras which were then equipped with a mechanical film-flattening device.

The 1950s were the start of the period of *analytical photogrammetry*. The expanding use of digital, electronic computers in that decade shifted interest from prevailing analogue methods to a purely analytical or numerical approach to photogrammetry. While analogue computation is inflexible, in regard to both input parameters and output results, and its accuracy is limited by physical properties, a numerical method allows virtually unlimited accuracy of computation and its flexibility is limited only by the mathematical model on which it is based. Above all, it permits over-determination which may improve precision, lead to the detection of gross errors and provide valuable statistical information about the measurements and the results. The first analytical applications were to photogrammetric triangulation. As numerical methods in photogrammetry improved, the above advantages, but above all their flexibility, were to prove invaluable at close range.

Fig. 1.33: Jenoptik UMK 1318.
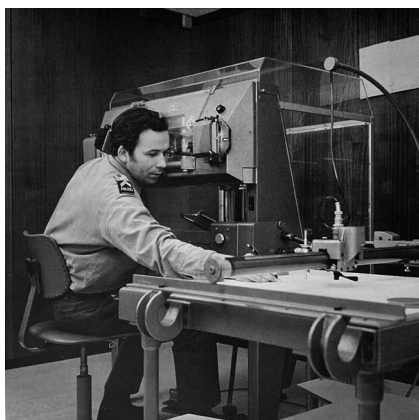
Fig. 1.34: Wild A40 stereoplotter.

Subsequently, stereoplotters were equipped with devices to record model coordinates for input to electronic computers. Arising from the pioneering ideas of Helava (1957), computers were incorporated in stereoplotters themselves, resulting in *analytical stereoplotters* with fully numerical reconstruction of the photogrammetric models. Bendix/OMI developed the first analytical plotter, the AP/C, in 1964 and, during the following two decades, analytical stereoplotters were produced by the major

instrument companies and others (example in Fig. 1.35). While the adaptability of such instruments has been of advantage in close-range photogrammetry, triangulation programs with even greater flexibility were soon to be developed, which were more suited to the requirements of close-range work.
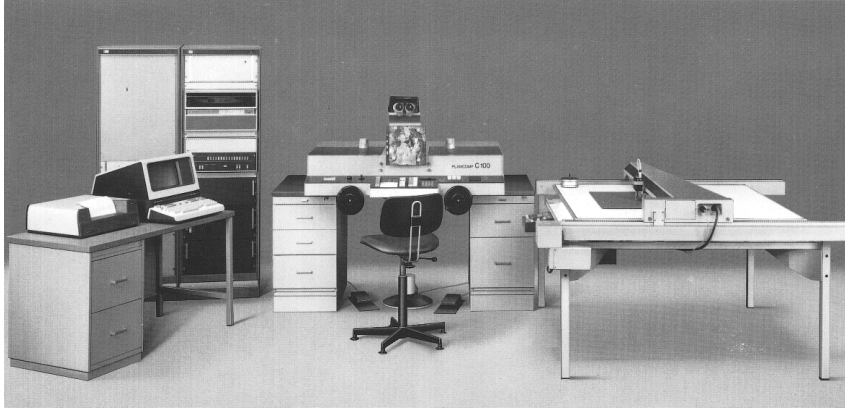
Fig. 1.35: Analytical Stereoplotter Zeiss Planicomp (ca. 1980).

Analytical photogrammetric triangulation is a method, using numerical data, of point determination involving the simultaneous orientation of all the photographs and taking all inter-relations into account. Work on this line of development, for example by the Ordnance Survey of Great Britain, had appeared before World War II, long before the development of electronic computers. Analytical triangulation required instruments to measure photo coordinates. The first stereocomparator designed specifically for use with aerial photographs was the Cambridge Stereocomparator designed in 1937 by E. H. Thompson. By 1955 there were five stereocomparators on the market and monocomparators designed for use with aerial photographs also appeared.

In the 1950s many mapping organizations were also experimenting with the new automatic computers, but it was the ballistic missile industry which gave the impetus for the development of the bundle method of photogrammetric triangulation. This is commonly known simply as the bundle adjustment and is today the dominant technique for triangulation in close-range photogrammetry. Seminal papers by Schmid (1956–57, 1958) and Brown (1958) laid the foundations for theoretically rigorous block adjustment. A number of bundle adjustment programs for air survey were developed and became commercially available, such as those by Ackermann et al. (1970) and Brown (1976). Programs designed specifically for close-range work have appeared since the 1980s, such as STARS (Fraser & Brown 1986), BINGO (Kruck 1983), MOR (Wester-Ebbinghaus 1981) or CAP (Hinsken 1989).

The importance of bundle adjustment in close-range photogrammetry can hardly be overstated. The method imposes no restrictions on the positions or the orientations of the cameras, nor is there any necessity to limit the imaging system to central projection. Of equal or greater importance, the parameters of interior orientation of all the cameras may be included as unknowns in the solution. Until the 1960s many experimenters appear to have given little attention to the calibration[4] of their cameras. This may well have been because the direct calibration of cameras focused for near objects is usually much more difficult than that of cameras focused for distant objects. At the same time, the interior orientation must usually be known more accurately than is necessary for vertical aerial photographs because the geometry of non-topographical work is frequently far from ideal. In applying the standard methods of calibration in the past, difficulties arose because of the finite distance of the targets, either real objects or virtual images. While indirect, numerical methods to overcome this difficulty were suggested by Torlegård (1967) and others, bundle adjustment now removes this concern. For high precision work, it is no longer necessary to use metric cameras which, while having the advantage of known and constant interior orientation, are usually cumbersome and expensive. Virtually any camera can now be used. Calibration via bundle adjustment is usually known as *self-calibration* (see sections 3.3 and 4.4).

Fig. 1.36: Rolleiflex SLX semi-metric camera (ca. 1980).

Fig. 1.37: Rollei MR2 multi-image restitution system (ca. 1990).

Many special cameras have been developed to extend the tools available to the photogrammetrist. One example promoted by Wester-Ebbinghaus (1981) was a modified professional photographic camera with an inbuilt réseau, an array of engraved crosses on a glass plate which appear on each image (see Fig. 1.36).

---

4 In photogrammetry, unlike computer vision, "calibration" refers only to interior orientation. Exterior orientation is not regarded as part of calibration.

The use of traditional stereo photogrammetry at close ranges has declined. As an alternative to the use of comparators, multi-photo analysis systems which use a digitizing pad as a measuring device for photo enlargements, for example the Rollei MR2 from 1986 (Fig. 1.37) have been widely used for architectural and accident recording.
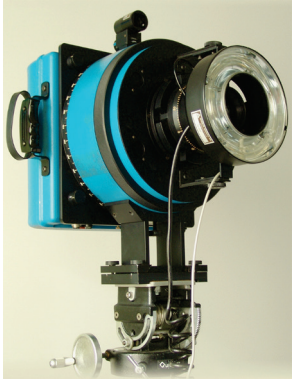
Fig. 1.38: Partial-metric camera GSI CRC-1 (ca. 1986).
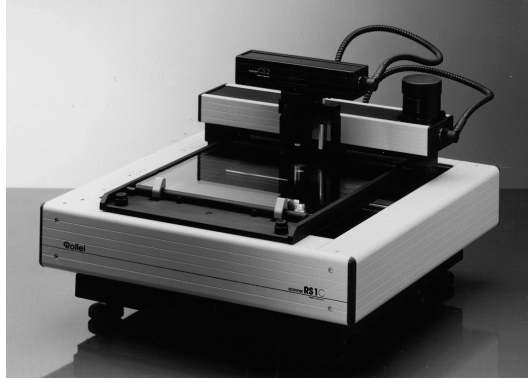
Fig. 1.39: Réseau-Scanner Rollei RS1 (ca. 1986).

Fig. 1.40: Online multi-image system Mapvision (1987).
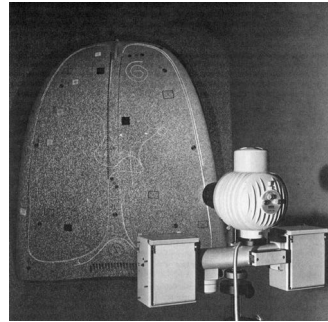
Fig. 1.41: Zeiss Indusurf (1987).

Since the middle of the 1980s, the use of opto-electronic image sensors has increased dramatically. Advanced computer technology enables the processing of digital images, particularly for automatic recognition and measurement of image features, including pattern correlation for determining object surfaces. Procedures in which both the image and its photogrammetric processing are digital are often referred to as *digital photogrammetry*. Automated precision monocomparators, in combination with large format réseau cameras, were developed for high-precision, industrial

applications, e.g. by Fraser and Brown (1986) or Luhmann and Wester-Ebbinghaus (1986), see Fig. 1.38 and Fig. 1.39.Initially, standard video cameras were employed. These generated analogue video signals which could be digitized with resolutions up to 780 x 580 picture elements (pixels) and processed in real time (*real-time photogrammetry, videogrammetry*). The first operational online multi-image systems became available in the late 1980s (example in Fig. 1.40). Analytical plotters were enhanced with video cameras to become analytical correlators, used for example in car body measurement (Zeiss Indusurf 1987, Fig. 1.41). Closed procedures for simultaneous multi-image processing of grey level values and object data based on least squares methods were developed, e.g. by Förstner (1982) and Gruen (1985).

The limitations of video cameras in respect of their small image format and low resolution led to the development of scanning cameras which enabled the high resolution recording of static objects to around 6000 x 4500 pixels. In parallel with this development, electronic theodolites were equipped with video cameras to enable the automatic recording of directions to targets (Kern SPACE). With the Leica/Rollei system POM (Programmable Optical Measuring system, Fig. 1.42) a complex online system for the measurement of automotive parts was developed which used réseau-scanning cameras (Fig. 1.43) and a rotary table for all-round measurements.
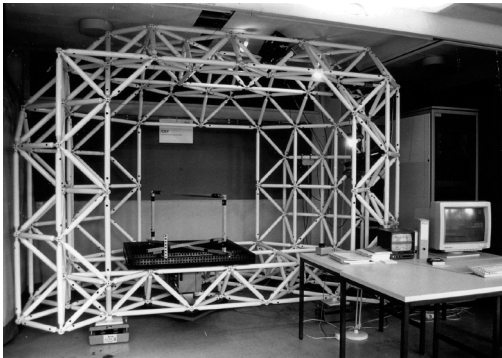
Fig. 1.42: POM online system with digital rotary table (1990).

Fig. 1.43: Réseau-scanning camera Rollei RSC (1990).

Digital cameras with high resolution, which can provide a digital image without analogue signal processing, have been available since the beginning of the 1990s. Resolutions ranged from about 1000 x 1000 pixels, e.g. the Kodak Megaplus (1986), to over 4000 x 4000 pixels. Easily portable, still-video cameras could store high resolution images directly in the camera, e.g. the Kodak DCS 460 (Fig. 1.44). They have led to a significant expansion of photogrammetric measurement technology, particularly in the industrial field. See, for example, systems from GSI, AICON (now part of Hexagon) and GOM (now part of Zeiss). *Online photogrammetric systems* (Fig.

1.45) have been brought into practical use, in addition to *offline systems*, both as mobile systems and in stationary configurations. Coded targets allowed the fully automatic identification and assignment of object features and orientation of the image sequences. Surface measurement of large objects were now possible with the development of pattern projection methods combined with photogrammetric techniques.

Fig. 1.44: Still-video camera Kodak DCS 460 (ca. 1996).
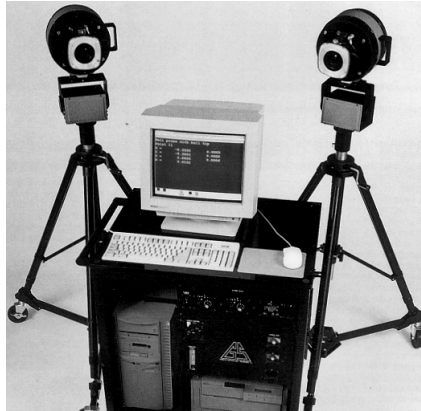
Fig. 1.45: GSI VSTARS online industrial measurement system (ca. 1991).

Interactive digital stereo systems, such as the Leica/Helava DSP and Zeiss PHODIS, have existed since around 1988 (Kern DSP-1). They have replaced analytical plotters, but they are rarely employed for close-range use. Interactive, graphical multi-image processing systems are of more importance here as they offer processing of freely chosen image configurations in a CAD environment, for example the Phocad PHIDIAS (Fig. 1.46). Easy-to-use, low-cost software packages, such as the EOS PhotoModeler (Fig. 1.47) or Photometrix iWitness, provide object reconstruction and creation of virtual 3D models from digital images without the need for a deep understanding of photogrammetry. Since around 2010 computer vision algorithms (interest operators, structure-from-motion approaches, see section 5.5.2) have become very popular and provide fully automated 3D modelling for arbitrary imagery without any pre-knowledge or on-site measurements. These systems provide dense point clouds and true orthophotos as well. See, for example, systems from AgiSoft, Pix4D, RealityCapture or MicMac, and the example output in Fig. 1.48.
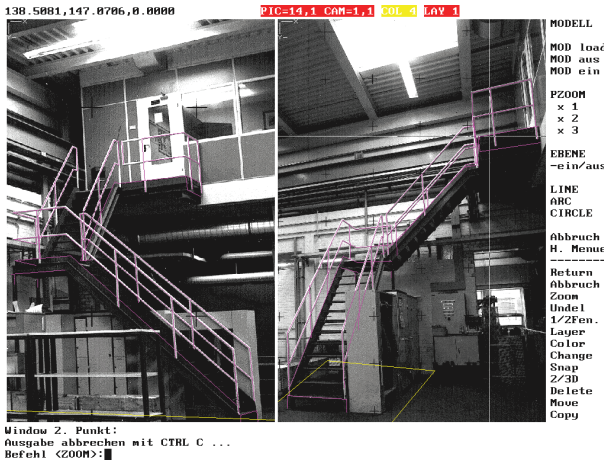
Fig. 1.46: PHIDIAS-MS multi-image analysis system (Phocad, 1994, see also Fig. 6.20).
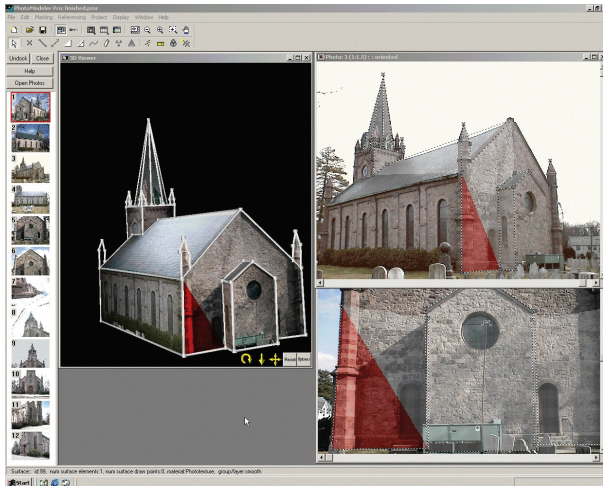
Fig. 1.47: Multi-image analysis system PhotoModeler (EOS Systems, 2008).

A trend in close-range photogrammetry is now towards the integration or embedding of photogrammetric components in application-oriented hybrid systems. This includes links to such packages as 3D CAD systems, databases and information systems, quality analysis and control systems for production, navigation systems for autonomous robots and vehicles, 3D visualization systems, internet applications, 3D animations and virtual reality. Another trend is the increasing use of methods from computer vision, such as projective geometry or pattern recognition, for rapid solutions which do not require high accuracy. Multi-sensor systems such as laser

scanners combined with cameras, GNSS-enabled cameras and cameras with integrated range finders are growing in importance. There is increased interest, too, in mobile and dynamic applications, including UAV applications. Finally, the continuing fall in the cost of digital cameras and processing software will ensure that photogrammetry is open to everyone.
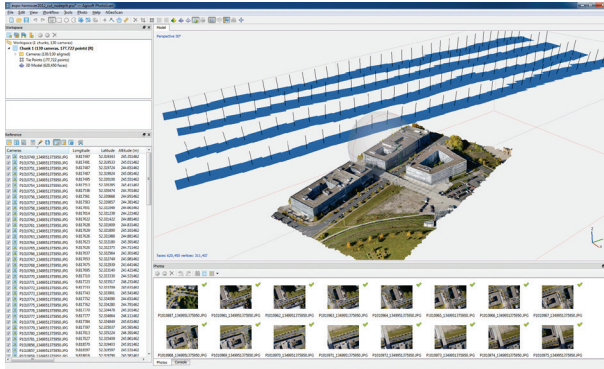
Fig. 1.48: Structure-from-Motion software PhotoScan (AgiSoft, 2016).

Close-range photogrammetry is today a well-established, universal 3D measuring technique, routinely applied in a wide range of interdisciplinary fields. There is every reason to expect its continued development long into the future.

2 Mathematical fundamentals

This chapter presents mathematical fundamentals which are essential for a deeper understanding of close-range photogrammetry. After defining some common coordinate systems, the most important plane and spatial coordinate transformations are summarized. An introduction to homogeneous coordinates and graphical projections then follows and the chapter concludes with the basic theory of least-squares adjustment.

2.1 Coordinate systems

2.1.1 Pixel and sensor coordinate system

Digital cameras incorporating electronic image sensors are routinely used for image recording. They deliver a positive digital image of an object or scene in the form of a matrix of pixels whose rows and columns are defined in a *pixel coordinate system*. It is a left-handed system, *u,v*, with its origin in the upper left element (Fig. 2.1 and section 5.1.2). The digital image can be viewed as a two-dimensional matrix with *m* columns and *n* rows which, in the case of multiple stored channels such as colour channels, can also be defined as multi-dimensional (see also section 5.1.3). A digital image only has a relationship to the physical image sensor in the camera when there is a 1:1 correspondence between individual pixels and individual sensor elements. For transformation into a metric image coordinate system, the physical pixel dimensions $\Delta s'_u$, $\Delta s'_v$ must be known. Together with a knowledge of pixel numbers *m* and *n*, this information then enables the origin to be shifted to the centre of the sensor (centre of image) and converted to a right-handed system *x',y'* (see sections 2.1.2 and 3.3.2.1 and eqn. 2.2).
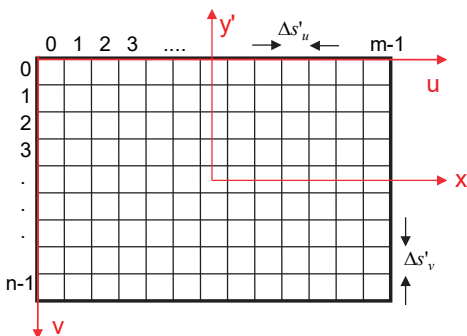
Fig. 2.1: Pixel coordinate system.

2.1.2 Image and camera coordinate systems

The *image coordinate system* defines a two-dimensional, image-based reference system of right-handed rectangular Cartesian coordinates, x',y'. In a film camera its physical relationship to the camera is defined by reference points, either fiducial marks or a réseau, which are projected into the acquired image (see section 3.3.2.1). For a digital imaging system, the sensor matrix in the camera's imaging plane is the source of the matrix of imaging pixels. These define an image coordinate system where the origin of metric image coordinates is located at the image centre (see section 2.1.1).
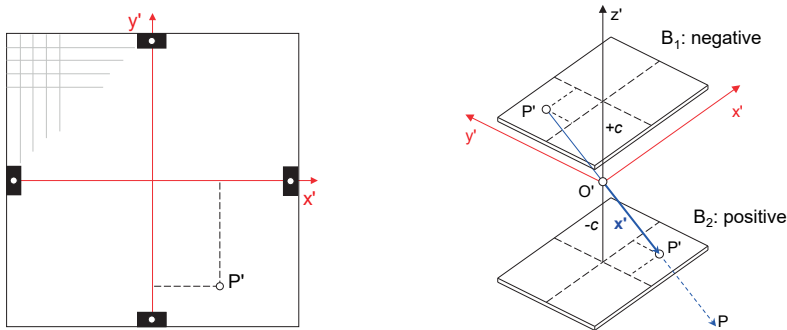
Fig. 2.2: Image and camera coordinate system.
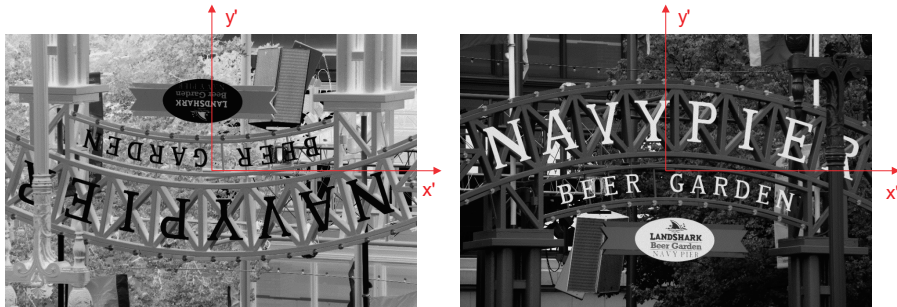
Fig. 2.3: Image coordinate system in negative (left) and positive image (right).

The relationship between the positive plane image and the camera, regarded as a spatial object, can be established when the image coordinate system is extended by a z' axis normal to the image plane, with a positive direction as indicated in order to preserve a right-handed system (see Fig. 2.2). This 3D coordinate system will be called the *camera coordinate system* and its origin is located at the perspective centre O'. The

z' axis coincides approximately with the optical axis. An image position B_1 corresponds to a location in the physically acquired image, which is the image negative. With respect to the positive, this is upside down (Fig. 2.3 left). For a number of mathematical calculations, it is easier to use the corresponding image position B_2, in the equivalent positive image (upright, see Fig. 2.3 right).

For points in the positive image, a 3D vector of image coordinates **x'** points from the origin at the perspective centre to the object point P. Note that this requires a negative z' value which has the same magnitude as the principal distance, c. Hence the 3D image vector **x'** as follows:

$$\mathbf{x'} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ -c \end{bmatrix} \tag{2.1}$$

Thus, the image vector **x'** describes the projection ray, with respect to the image coordinate system, from the positive image point to the object point. The spatial position of the perspective centre in the image coordinate system is given by the parameters of *interior orientation* under consideration of the principal point and imaging errors (see section 3.3.2).

The transformation between pixel and metric image using the physical pixel sizes $\Delta s'_u$, $\Delta s'_v$ and pixel numbers m and n, is given as follows:

$$
\begin{aligned}
s'_x &= m \cdot \Delta s'_u & s'_y &= n \cdot \Delta s'_v \\
x' &= -\frac{s'_x}{2} + u \cdot \Delta s'_u & y' &= \frac{s'_y}{2} - v \cdot \Delta s'_v
\end{aligned}
\tag{2.2}
$$

Here, s'_x and s'_y normally define the sensor format. Where the photogrammetric calculation is required in a right-handed pixel coordinate system, with the origin located at the image centre, the transformation is defined by $\Delta s'_u = \Delta s'_v = 1$.

2.1.3 Model coordinate system

The spatial Cartesian *model coordinate system* xyz is used to describe the relative position and orientation of two or more images (image coordinate systems). Normally its origin is at the perspective centre of one of the images (Fig. 2.4). In addition, the model coordinate system may be parallel to the related image coordinate system (see section 4.3.3).
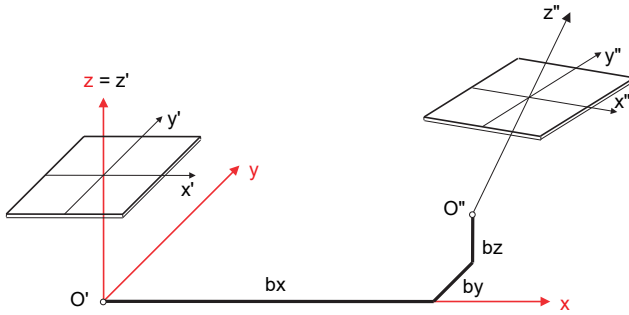
Fig. 2.4: Model coordinate system.

2.1.4 Object coordinate system

The term *object coordinate system*, also known as the world coordinate system, is here used for every spatial Cartesian coordinate system XYZ that is defined by reference points of the object (examples in Fig. 2.5). For example, national geodetic coordinate systems (X = easting, Y = northing, Z = altitude, origin at the equator) are defined by geodetically measured reference points[1]. Another example is the local object or workpiece coordinate system of a car body that is defined by the constructional axes (X = longitudinal car axis, Y = front axle, Z = height, origin at centre of front axle).
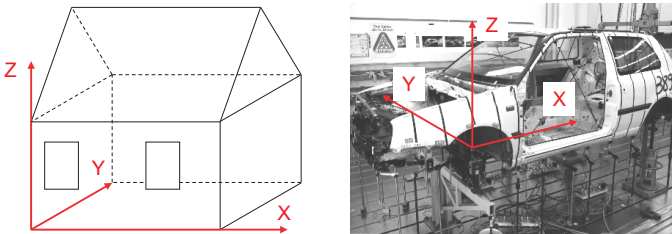
Fig. 2.5: Object coordinate systems.

A special case of three-dimensional coordinate system is an arbitrarily oriented one used by a 3D measuring system such as a camera or a scanner. This is not directly related to any superior system or particular object but if, for instance, just one

---

1 National systems of geodetic coordinates which use the geoid as a reference surface are equivalent to a Cartesian coordinate system only over small areas.

reference scale is given (Fig. 2.6), then it is still possible to measure spatial object coordinates.

The definition of origin, axes and scale of a coordinate system is also known as the *datum*.
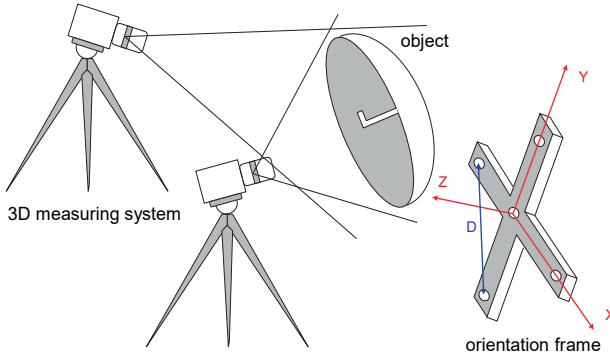
Fig. 2.6: 3D instrument coordinate system.

2.2 Coordinate transformations

2.2.1 Plane transformations

2.2.1.1 Homogenous coordinates

Homogenous coordinates can be derived from Cartesian coordinates by adding one dimension and scaling by any non-zero factor λ. In two dimensions this leads to:

$$\boldsymbol{x} = \lambda \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad \text{where } x = u/w, y = v/w, \lambda \neq 0 \qquad (2.3)$$

Three-dimensional Cartesian coordinates are converted to homogenous coordinates in an analogous way[2].

The homogenous coordinate transformation

$$\boldsymbol{x}' = \lambda \mathbf{T} \boldsymbol{x} \qquad (2.4)$$

maintains its projection properties independently of λ. Consequently, all major coordinate transformations (translation, rotation, similarity, central projection) can be formed in a consistent way and can be combined in an arbitrary order to a total

---

2 Homogenous vectors are denoted in bold and italic text.

transformation **T** (see section 2.2.3). The photogrammetric projection equations can also be elegantly expressed in homogenous coordinates (see section 4.2.4.2).

2.2.1.2 Similarity transformation

The *plane similarity transformation* is used for the mapping of two plane Cartesian coordinate systems (Fig. 2.7). Generally a 4-parameter transformation is employed which defines two translations, one rotation and a scaling factor between the two systems. Angles and distance proportions are maintained.
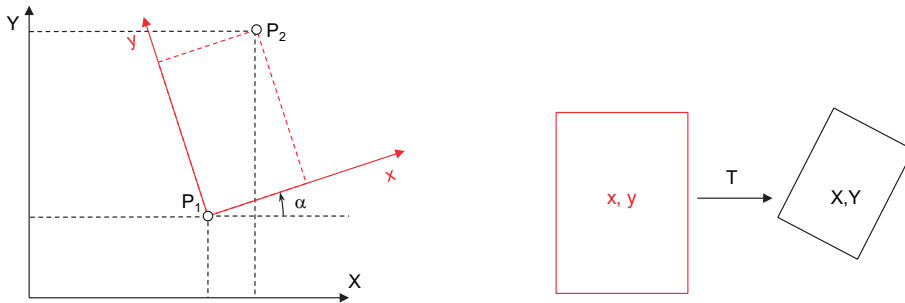
Fig. 2.7: Plane similarity transformation.

Given a point P in the xy source system, the XY coordinates in the target system are

$$X = a_0 + a_1 \cdot x - b_1 \cdot y \qquad\qquad Y = b_0 + b_1 \cdot x + a_1 \cdot y \qquad (2.5)$$

or

$$X = a_0 + m \cdot (x \cdot \cos\alpha - y \cdot \sin\alpha) \qquad Y = b_0 + m \cdot (x \cdot \sin\alpha + y \cdot \cos\alpha) \qquad (2.6)$$

Here a_0 and b_0 define the translation of the origin, α is the rotation angle and m is the global scaling factor. In order to determine the four coefficients, a minimum of two identical points is required in both systems. With more than two identical points, the transformation parameters can be calculated by an over-determined, least-squares adjustment.

In matrix notation (2.5) is expressed as

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{x} + \mathbf{a}$$

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} \qquad (2.7)$$

or in non-linear form with $a_0 = X_0$ und $b_0 = Y_0$:

$$\mathbf{X} = m \cdot \mathbf{R} \cdot \mathbf{x} + \mathbf{X}_0$$

$$\begin{bmatrix} X \\ Y \end{bmatrix} = m \cdot \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} \tag{2.8}$$

$\mathbf{R}$ is the rotation matrix corresponding to rotation angle α. This is an orthogonal matrix having orthonormal column (or row) vectors and it has the properties:

$$\mathbf{R}^{-1} = \mathbf{R}^T \qquad \text{and} \qquad \mathbf{R}^T \cdot \mathbf{R} = \mathbf{I} \tag{2.9}$$

For the reverse transformation of coordinates from the target system into the source system, the transformation equations (2.8) are re-arranged as follows:

$$\mathbf{x} = \frac{1}{m} \cdot \mathbf{R}^{-1} \cdot (\mathbf{X} - \mathbf{X}_0)$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{m} \cdot \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} X - X_0 \\ Y - Y_0 \end{bmatrix} \tag{2.10}$$

or explicitly with the coefficients of the forward transformation:

$$x = \frac{a_1(X - a_0) + b_1(Y - b_0)}{a_1^2 + b_1^2} \qquad\qquad y = \frac{a_1(Y - b_0) - b_1(X - a_0)}{a_1^2 + b_1^2} \tag{2.11}$$

2.2.1.3 Affine transformation

The *plane affine transformation* is also used for the mapping of two plane coordinate systems (Fig. 2.8). This 6-parameter transformation defines two displacements, one rotation, one shearing angle between the axes and two separate scaling factors.
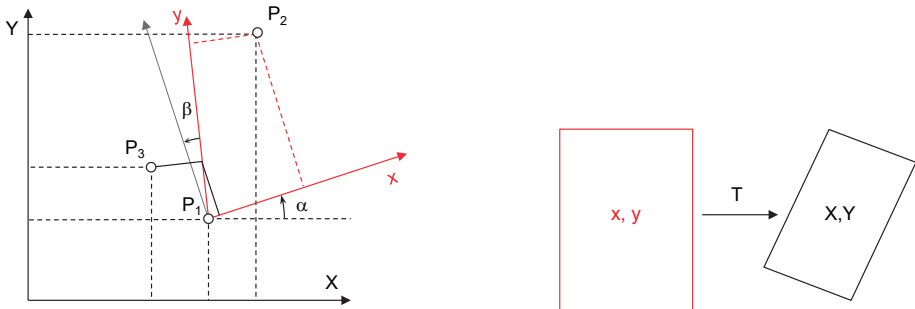
Fig. 2.8: Plane affine transformation.

For a point P in the source system, the XY coordinates in the target system are given by

$$X = a_0 + a_1 \cdot x + a_2 \cdot y \qquad\qquad Y = b_0 + b_1 \cdot x + b_2 \cdot y \qquad\qquad (2.12)$$

or in non-linear form with $a_0 = X_0$ und $b_0 = Y_0$:

$$X = X_0 + m_X \cdot x \cdot \cos\alpha - m_Y \cdot y \cdot \sin(\alpha + \beta)$$
$$Y = Y_0 + m_X \cdot x \cdot \sin\alpha + m_Y \cdot y \cdot \cos(\alpha + \beta) \qquad\qquad (2.13)$$

The parameters a_0 and b_0 (X_0 and Y_0) define the displacement of the origin, α is the rotation angle, β is the shearing angle between the axes and m_X, m_Y are the scaling factors for x and y. In order to determine the six coefficients, a minimum of three identical points is required in both systems. With more than three identical points, the transformation parameters can be calculated by over-determined least-squares adjustment.

In matrix notation the affine transformation can be written as:

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{x} + \mathbf{a}$$
$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_0 \\ b_0 \end{bmatrix}$$

or

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} m_X \cdot \cos\alpha & -m_Y \cdot \sin(\alpha + \beta) \\ m_X \cdot \sin\alpha & m_Y \cdot \cos(\alpha + \beta) \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} \qquad (2.14)$$

A is the affine transformation matrix. For transformations with small values of rotation and shear, the parameter a_1 corresponds to the scaling factor m_X and the parameter b_2 to the scaling factor m_Y.

For the reverse transformation from coordinates in the target system to coordinates in the source system, eqn. (2.14) is re-arranged as follows

$$\mathbf{x} = \mathbf{A}^{-1} \cdot (\mathbf{X} - \mathbf{a}) \qquad\qquad (2.15)$$

or explicitly with the coefficients with the original, forward transformation:

$$x = \frac{a_2(Y - b_0) - b_2(X - a_0)}{a_2 b_1 - a_1 b_2} \qquad\qquad y = \frac{b_1(X - a_0) - a_1(Y - b_0)}{a_2 b_1 - a_1 b_2} \qquad (2.16)$$

2.2.1.4 Polynomial transformation

Non-linear deformations (Fig. 2.9) can be described by polynomials of degree n: In general, the transformation model can be written as:

$$X = \sum_{j=0}^{n} \sum_{i=0}^{j} a_{ji} x^{j-i} y^i \qquad\qquad Y = \sum_{j=0}^{n} \sum_{i=0}^{j} b_{ji} x^{j-i} y^i \qquad (2.17)$$

where n = degree of polynomial

A polynomial with $n = 2$ is given by:

$$X = a_{00} + a_{10} \cdot x + a_{11} \cdot y + a_{20} \cdot x^2 + a_{21} \cdot x \cdot y + a_{22} \cdot y^2$$
$$Y = b_{00} + b_{10} \cdot x + b_{11} \cdot y + b_{20} \cdot x^2 + b_{21} \cdot x \cdot y + b_{22} \cdot y^2$$

(2.18)

The polynomial with $n = 1$ is identical to the affine transformation (2.12). In general, the number of coefficients required to define a polynomial transformation of degree n is $u = (n+1) \cdot (n+2)$. In order to determine the u coefficients, a minimum of $u/2$ identical points is required in both systems.
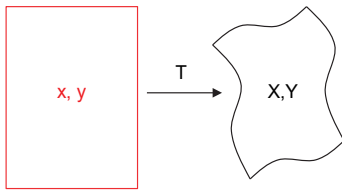
Fig. 2.9: Plane polynomial transformation.

2.2.1.5 Bilinear transformation

The *bilinear transformation* is similar to the affine transformation but extended by a mixed term:

$$X = a_0 + a_1 x + a_2 y + a_3 xy \qquad\qquad Y = b_0 + b_1 x + b_2 y + b_3 xy$$

(2.19)

In order to determine the eight coefficients, a minimum of four identical points is required.

The bilinear transformation can be used in the unconstrained transformation and interpolation of quadrilaterals, for example in réseau grids or digital surface models.
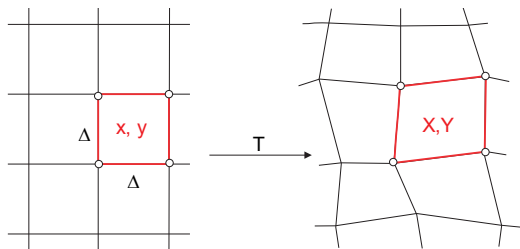
Fig. 2.10: Bilinear transformation.

For the transformation of a square with side length Δ (Fig. 2.10), the coefficients can be calculated as follows:

$$
\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}
\tag{2.20}
$$

$$
\text{where} \qquad \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1/\Delta & 1/\Delta & 0 & 0 \\ -1/\Delta & 0 & 1/\Delta & 0 \\ 1/\Delta^2 & -1/\Delta^2 & -1/\Delta^2 & 1/\Delta^2 \end{bmatrix}
$$

2.2.1.6 Projective transformation

The *plane projective transformation* maps two plane coordinate systems using a central projection. All projection rays are straight lines through the perspective centre (Fig. 2.11).

The transformation model is:

$$
X = \frac{a_0 + a_1 \cdot x + a_2 \cdot y}{1 + c_1 \cdot x + c_2 \cdot y} \qquad\qquad Y = \frac{b_0 + b_1 \cdot x + b_2 \cdot y}{1 + c_1 \cdot x + c_2 \cdot y}
\tag{2.21}
$$

The system of equations (2.21) is not linear. By multiplying by the denominator and rearranging, the following linear form can be derived. This is suitable as an observation equation in an adjustment procedure:

$$
\begin{aligned}
a_0 + a_1 x + a_2 y - X - c_1 xX - c_2 yX = 0 \\
b_0 + b_1 x + b_2 y - Y - c_1 xY - c_2 yY = 0
\end{aligned}
\tag{2.22}
$$

In order to determine the eight coefficients, four identical points are required where no three may lay on a common straight line. With more than four points, the system of equations can be solved by adjustment (see calculation scheme in section 4.2.6). For the derivation of (2.21) the spatial similarity transformation can be used (see section 2.2.3).

The reverse transformation can be calculated by re-arrangement of equations (2.21):

$$
\begin{aligned}
x = \frac{a_2 b_0 - a_0 b_2 + (b_2 - b_0 c_2)X + (a_0 c_2 - a_2)Y}{a_1 b_2 - a_2 b_1 + (b_1 c_2 - b_2 c_1)X + (a_2 c_1 - a_1 c_2)Y} \\
y = \frac{a_0 b_1 - a_1 b_0 + (b_0 c_1 - b_1)X + (a_1 - a_0 c_1)Y}{a_1 b_2 - a_2 b_1 + (b_1 c_2 - b_2 c_1)X + (a_2 c_1 - a_1 c_2)Y}
\end{aligned}
\tag{2.23}
$$

In this form the equations again express a projective transformation. By substitution of terms the following form is derived:

$$x = \frac{a'_0 + a'_1 X + a'_2 Y}{1 + c'_1 X + c'_2 Y} \qquad\qquad y = \frac{b'_0 + b'_1 X + b'_2 Y}{1 + c'_1 X + c'_2 Y} \qquad\qquad (2.24)$$

where

$$a'_0 = \frac{a_2 b_0 - a_0 b_2}{N} \qquad b'_0 = \frac{a_0 b_1 - a_1 b_0}{N} \qquad c'_1 = \frac{b_1 c_2 - b_2 c_1}{N}$$

$$a'_1 = \frac{b_2 - b_0 c_2}{N} \qquad b'_1 = \frac{b_0 c_1 - b_1}{N} \qquad c'_2 = \frac{a_2 c_1 - a_1 c_2}{N}$$

$$a'_2 = \frac{a_0 c_2 - a_2}{N} \qquad b'_2 = \frac{a_1 - a_0 c_1}{N} \qquad N = a_1 b_2 - a_2 b_1$$
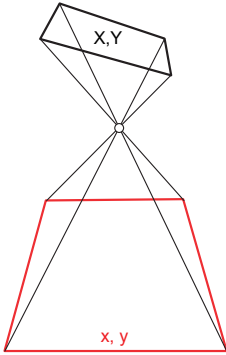
Fig. 2.11: Plane projective transformation.
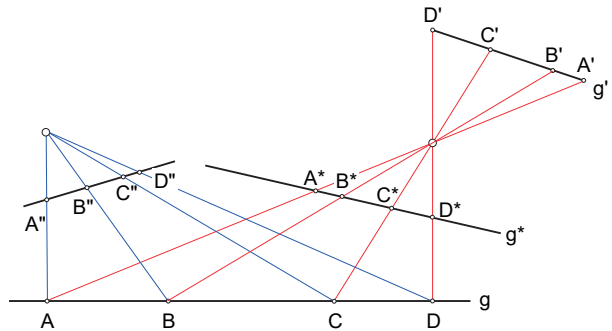
Fig. 2.12: Cross ratios.

The plane projective transformation preserves rectilinear properties and intersection points of straight lines. In contrast, angles, length and area proportions are not invariant. An additional invariant property of the central projection are the *cross ratios* of distances between points on a straight line. They are defined as follows:

$$\lambda = \frac{\overline{AB}}{\overline{BC}} \div \frac{\overline{AD}}{\overline{CD}} = \frac{\overline{A^{*}B^{*}}}{\overline{B^{*}C^{*}}} \div \frac{\overline{A^{*}D^{*}}}{\overline{C^{*}D^{*}}} = \frac{\overline{A'B'}}{\overline{B'C'}} \div \frac{\overline{A'D'}}{\overline{C'D'}} = \frac{\overline{A''B''}}{\overline{B''C''}} \div \frac{\overline{A''D''}}{\overline{C''D''}} \qquad (2.25)$$

The cross ratios apply to all straight lines that intersect a bundle of perspective rays in an arbitrary position (Fig. 2.11).

The plane projective transformation is applied to single image analysis, e.g. for rectification or coordinate measurement in single images (see section 4.2.6).

Example 2.1:

Given 8 points in the source and target coordinate systems with the following plane coordinates:

No.	x	y	X	Y
1	−12.3705	−10.5075	0	0
2	−10.7865	15.4305	0	5800
3	8.6985	10.8675	4900	5800
4	11.4975	−9.5715	4900	0
5	7.8435	7.4835	4479	4580
6	−5.3325	6.5025	1176	3660
7	6.7905	−6.3765	3754	790
8	−6.1695	−0.8235	1024	1931

These correspond to the image and control point coordinates in Fig. 5.56.
The plane transformations described in section 2.2.1.1 to section 2.2.1.6 then give rise to the following transformation parameters:

Coeff.	4-param transf.	6-param transf.	Bilinear transf.	Projective transf.	Polynomial 2nd order
a0	2524.3404	2509.3317	2522.4233	2275.9445	2287.8878
a1	237.2887	226.9203	228.0485	195.1373	230.9799
a2		7.3472	11.5751	−11.5864	16.4830
a3			2.1778		2.9171
a4					2.2887
a5					−0.0654
b0	2536.0460	2519.9142	2537.9164	2321.9622	2348.9782
b1	5.6218	21.5689	23.1202	−9.0076	28.0384
b2		250.1298	255.9436	222.6108	250.7228
b3			2.9947		−0.2463
b4					3.5332
b5					2.5667
c1				−0.0131	
c2				−0.0097	
s0 [mm]	369.7427	345.3880	178.1125	3.1888	38.3827

The standard deviation s_0 indicates the spread of the transformed points in the XY system. It can be seen that the projective transformation has the best fit, with the 2nd order polynomial as second best. The other transformations are not suitable for this particular distribution of points.

Using homogenous coordinates the plane projective transformation can be expressed as:

$$U = H \cdot x$$
$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{aligned} X &= 1/W \cdot U \\ \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} &= \begin{bmatrix} U/W \\ V/W \\ W/W \end{bmatrix} \end{aligned} \tag{2.26}$$

This formulation is known as *homography*. Since the matrix H can be scaled without altering its projective properties (see section 2.2.1.1), there are eight degrees of

freedom as there are in the plane projective transformation of eqn. (2.21). When seeking a direct solution to $\boldsymbol{h} = [h_{11}, ..., h_{33}]$ the elements should be normalized to deal with this rank deficiency and to avoid the trivial solution $\boldsymbol{h}=\boldsymbol{0}$. For example, this can be done by setting $h_{33}=1$ which gives the same result as in eqn. (2.21). It can be numerically advantageous to seek a normalisation via the norm of the vector e.g. $|\boldsymbol{h}|=1$, which is implicitly the case if a solution is sought via an eigenvalue or singular value decomposition.

2.2.2 Spatial rotations

2.2.2.1 Rotation matrix using trigonometric functions

For plane transformations, rotations take effect about a single point. In contrast, spatial rotations are performed successively about the three axes of a spatial coordinate system. Consider a point P in the source system xyz which is rotated with respect to the target system XYZ. Using trigonometric functions, individual rotations about the three axes of the target system are defined as follows (see Fig. 2.13):
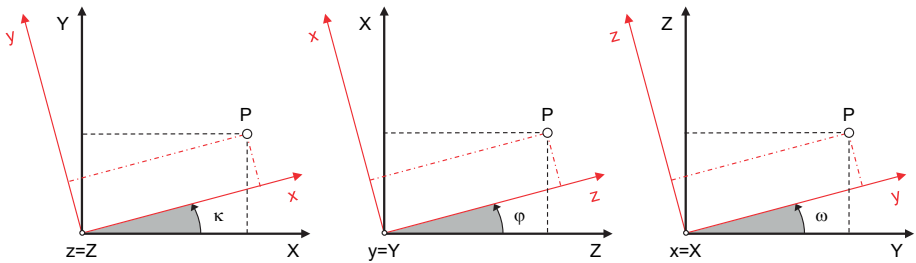
Fig. 2.13: Definition of spatial rotation angles.

1. Rotation about Z-axis:

A Z-axis rotation is conventionally designated by angle κ. This is positive in an anticlockwise direction when viewed down the positive Z axis towards the origin. From eqn. (2.8), this results in the following point coordinates in the target system XYZ:

$$
\begin{aligned}
X &= x\cdot\cos\kappa - y\cdot\sin\kappa \\
Y &= x\cdot\sin\kappa + y\cdot\cos\kappa \\
Z &= z
\end{aligned}
\qquad \text{or} \qquad
\begin{aligned}
\mathbf{X} &= \mathbf{R}_{\kappa}\cdot\mathbf{x} \\[4pt]
\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} &=
\begin{bmatrix} \cos\kappa & -\sin\kappa & 0 \\ \sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{bmatrix}
\cdot
\begin{bmatrix} x \\ y \\ z \end{bmatrix}
\end{aligned}
\qquad (2.27)
$$

2. Rotation about Y-axis:

The corresponding rotation about the Y-axis is designated by rotation angle φ. This results in the following XYZ target point coordinates:

$$X = x \cdot \cos\varphi + z \cdot \sin\varphi \qquad \text{or} \qquad \mathbf{X} = \mathbf{R}_\varphi \cdot \mathbf{x}$$
$$Y = y$$
$$Z = -x \cdot \sin\varphi + z \cdot \cos\varphi \qquad \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad (2.28)$$

3. Rotation about X-axis:

Finally, the X axis rotation is designated by angle ω, which results in XYZ values:

$$X = x \qquad \text{or} \qquad \mathbf{X} = \mathbf{R}_\omega \cdot \mathbf{x}$$
$$Y = y \cdot \cos\omega - z \cdot \sin\omega$$
$$Z = y \cdot \sin\omega + z \cdot \cos\omega \qquad \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & -\sin\omega \\ 0 & \sin\omega & \cos\omega \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad (2.29)$$

The given rotation matrices are orthonormal, i.e.

$$\mathbf{R} \cdot \mathbf{R}^T = \mathbf{R}^T \cdot \mathbf{R} = \mathbf{I} \qquad\qquad \mathbf{R}^{-1} = \mathbf{R}^T \quad \text{and} \quad \det(\mathbf{R}) = 1 \qquad (2.30)$$

The complete rotation **R** of a spatial coordinate transformation can be defined by the successive application of 3 individual rotations, as defined above. Only certain combinations of these 3 rotations are possible and these may be applied about either the fixed axial directions of the target system or the moving axes of the source system. If a general rotation is defined about *moving axes* in the order ω φ κ, then the complete rotation is given by:

$$\mathbf{X} = \mathbf{R} \cdot \mathbf{x} \qquad (2.31)$$

where

$$\mathbf{R} = \mathbf{R}_\omega \cdot \mathbf{R}_\varphi \cdot \mathbf{R}_\kappa \qquad (2.32)$$

and

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$
$$= \begin{bmatrix} \cos\varphi\cos\kappa & -\cos\varphi\sin\kappa & \sin\varphi \\ \cos\omega\sin\kappa + \sin\omega\sin\varphi\cos\kappa & \cos\omega\cos\kappa - \sin\omega\sin\varphi\sin\kappa & -\sin\omega\cos\varphi \\ \sin\omega\sin\kappa - \cos\omega\sin\varphi\cos\kappa & \sin\omega\cos\kappa + \cos\omega\sin\varphi\sin\kappa & \cos\omega\cos\varphi \end{bmatrix}$$

If the rotation is alternatively defined about *fixed* axes in the order ω φ κ, then the rotation matrix is given by:

$$\mathbf{R}^{*} = \mathbf{R}_{\kappa} \cdot \mathbf{R}_{\varphi} \cdot \mathbf{R}_{\omega} \tag{2.33}$$

This is mathematically equivalent to applying the same rotations about moving axes but in the reverse order.

From eqn. (2.31) the inverse transformation which generates the coordinates of a point P in the *rotated* system xyz from its XYZ values is therefore given by:

$$\mathbf{x} = \mathbf{R}^{T} \cdot \mathbf{X} \tag{2.34}$$

where

$$\mathbf{R}^{T} = \mathbf{R}_{\kappa}^{T} \cdot \mathbf{R}_{\varphi}^{T} \cdot \mathbf{R}_{\omega}^{T} \tag{2.35}$$

Note that in this inverse transformation, the individually inverted rotation matrices are multiplied in the reverse order.

From the matrix coefficients $r_{11}...r_{33}$ in eqn. (2.32), the individual rotation angles can be calculated as follows:

$$
\begin{array}{ccc}
\sin\varphi = r_{13} & & \sin\varphi = r_{13} \\[2mm]
\tan\omega = -\dfrac{r_{23}}{r_{33}} & \text{or} & \cos\omega = \dfrac{r_{33}}{\cos\varphi} \\[3mm]
\tan\kappa = -\dfrac{r_{12}}{r_{11}} & & \cos\kappa = \dfrac{r_{11}}{\cos\varphi}
\end{array}
\tag{2.36}
$$

Eqn. (2.36) shows that the determination of φ is ambiguous due to solutions for $\sin\varphi$ in two quadrants. In addition, there is no unique solution for the rotation angles if the second rotation (φ in this case) is equal to 90° or 270° (cosine φ in r_{11} and r_{33} then causes division by zero). This effect also exists in gimbal systems (gyroscopes) where it is known as *gimbal lock*.
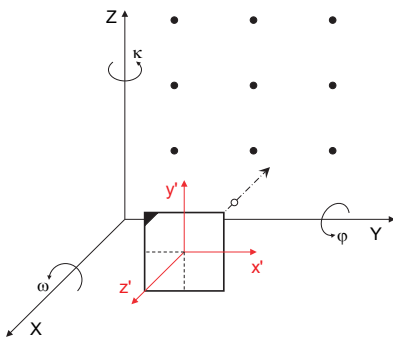
Fig. 2.14: Image configuration where $\omega = 0°$, $\varphi = 90°$ and $\kappa = 90°$.

A simple solution to this ambiguity problem is to alter the order of rotation. In the case that the secondary rotation is close to 90°, the primary and secondary rotations can be exchanged, leading to the new order φ ω κ. This procedure is used in close-range photogrammetry when the viewing direction of the camera is approximately horizontal (see Fig. 2.14 and also section 4.2.1.2). The resulting rotation matrix is then given by:

$$\mathbf{R}_{\varphi\omega\kappa} = \mathbf{R}_{\varphi} \cdot \mathbf{R}_{\omega} \cdot \mathbf{R}_{\kappa} \tag{2.37}$$

where

$$\mathbf{R}_{\varphi\omega\kappa} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

$$= \begin{bmatrix} \cos\varphi\cos\kappa + \sin\varphi\sin\omega\sin\kappa & -\cos\varphi\sin\kappa + \sin\varphi\sin\omega\cos\kappa & \sin\varphi\cos\omega \\ \cos\omega\sin\kappa & \cos\omega\cos\kappa & -\sin\omega \\ -\sin\varphi\cos\kappa + \cos\varphi\sin\omega\sin\kappa & \sin\varphi\sin\kappa + \cos\varphi\sin\omega\cos\kappa & \cos\varphi\cos\omega \end{bmatrix}$$

Example 2.2:

Referring to Fig. 2.14, an image configuration is shown where the primary rotation $\omega = 0°$, the secondary rotation $\varphi = 90°$ and the tertiary rotation $\kappa = 90°$. In this case the $\mathbf{R}_{\varphi\omega\kappa}$ reduces to

$$\mathbf{R}_{\varphi\omega\kappa} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

This rotation matrix represents an exchange of coordinate axes. The first row describes the transformation of the X axis. Its x, y and z elements are respectively 0, 0 and 1, indicating a transformation of X to z. Correspondingly, the second row shows Y transforming to x and the third row transforms Z to y.

However, the exchange of rotation orders is not a suitable solution for arbitrarily oriented images (see Fig. 3.39 and Fig. 4.57). Firstly, the rotation angles of images freely located in 3D space are not easy to visualize. Secondly, ambiguities cannot be avoided, which leads to singularities when calculating orientations. The effects can be avoided by rotation matrices based on algebraic functions (see next sections).

2.2.2.2 Rotation matrix using quaternions

The ambiguities for trigonometric functions (above) can be avoided when a rotation matrix with algebraic functions is used and where the rotation itself is a single rotation about a specific axis in space. The direction of the rotation axis is defined by three vector components a, b, c and the rotation defined by angular value d. This four-dimensional vector is known as a *quaternion* and it gives rise to the following rotation matrix:

$$\mathbf{R}^T = \begin{bmatrix} d^2 + a^2 - b^2 - c^2 & 2(ab - cd) & 2(ac + bd) \\ 2(ab + cd) & d^2 - a^2 + b^2 - c^2 & 2(bc - ad) \\ 2(ac - bd) & 2(bc + ad) & d^2 - a^2 - b^2 + c^2 \end{bmatrix} \tag{2.38}$$

Implicitly, this rotation matrix contains a common scaling factor:

$$m = a^2 + b^2 + c^2 + d^2 \tag{2.39}$$

Using the constraint $m = 1$, an orthogonal rotation matrix with three independent parameters is obtained. Normalization with m gives the unit quaternion $\mathbf{q}$:

$$\mathbf{q} = \begin{bmatrix} a/m \\ b/m \\ c/m \\ d/m \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_0 \end{bmatrix} = \begin{bmatrix} n_x \sin(\alpha/2) \\ n_y \sin(\alpha/2) \\ n_z \sin(\alpha/2) \\ \cos(\alpha/2) \end{bmatrix} \tag{2.40}$$

with the orthonormal rotation matrix:

$$\mathbf{R} = \begin{bmatrix} 1 - 2(q_2^2 + q_3^2) & 2(q_1 q_2 - q_0 q_3) & 2(q_0 q_2 + q_1 q_3) \\ 2(q_1 q_2 + q_0 q_3) & 1 - 2(q_1^2 + q_3^2) & 2(q_2 q_3 - q_0 q_1) \\ 2(q_1 q_3 - q_0 q_2) & 2(q_0 q_1 + q_2 q_3) & 1 - 2(q_1^2 + q_2^2) \end{bmatrix} \tag{2.41}$$

The parameters $a...c$, or $q_1...q_3$, are called the vector components of the quaternion and the parameter d, or q_0, is called the scalar component. The rotation matrix becomes a unity matrix when either $q_0 = 1$ (corresponding to $\alpha = 0$) or $q_1 = q_2 = q_3 = 0$.
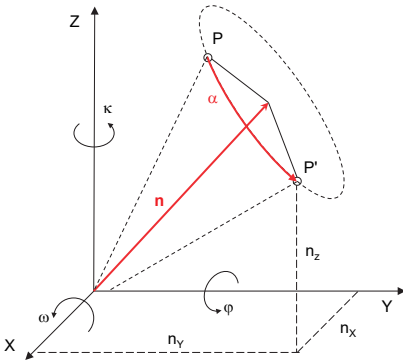
Fig. 2.15: Rotation around an axis in space.

The formulation with quaternions corresponds to a rotation angle α about an axis in space defined by the normalized direction vector $\mathbf{n} = [n_x, n_y, n_z]^T$ (see Fig. 2.15). Since the

axis only specifies direction, and its length has no importance, only two of its parameters are independent. Together with the rotation angle, three independent parameters therefore still remain to describe a rotation in space. This form of rotation is often used in computer graphics, e.g. OpenGL or VRML. The only ambiguity associated with quaternions is the fact that a rotation defined by **q** is identical to a rotation defined by $\mathbf{q}^{-1}$, i.e. a rotation can be formed equally in the reversed viewing direction using the inverted quaternion.

The quaternion can be calculated from a given orthonormal rotation matrix **R** as follows:

$$q_0 = \pm\frac{1}{2}\sqrt{r_{11} + r_{22} + r_{33}} = \cos\frac{\alpha}{2} \tag{2.42}$$

$$q_1 = \frac{r_{32} - r_{23}}{4q_0} \qquad\qquad q_2 = \frac{r_{13} - r_{31}}{4q_0} \qquad\qquad q_3 = \frac{r_{21} - r_{12}}{4q_0}$$

The sign of q_0, or equivalently the value of angle α, cannot be uniquely defined (see above). The transformation of the coefficients q into Euler angles of the rotation matrix (2.41) is done analogously to (2.36) or directly by

$$\omega = -\arctan\left(\frac{2(q_2 q_3 - q_0 q_1)}{q_0^2 - q_1^2 - q_2^2 + q_3^2}\right)$$

$$\varphi = \arcsin\left(2(q_0 q_2 + q_1 q_3)\right) \tag{2.43}$$

$$\kappa = -\arctan\left(\frac{2(q_1 q_2 - q_0 q_3)}{q_0^2 + q_1^2 - q_2^2 - q_3^2}\right)$$

whereby the ambiguities described in section 2.2.2.1 still exist.

Example 2.3:
Given the rotation matrix

$$\mathbf{R} = \begin{bmatrix} 0.996911 & -0.013541 & -0.077361 \\ 0.030706 & 0.973820 & 0.225238 \\ 0.072285 & -0.226918 & 0.971228 \end{bmatrix}$$

Application of eqn. (2.36) results in the following rotation angles:
 ω=−13.0567°, φ=−4.4369°, κ=0.7782°
Application of eqns. (2.42) and (2.38) results in the following quaternion:
 q_1 = −0.113868, q_2 = −0.037686, q_3 = 0.011143, q_0 = 0.927183 and $α$ = 13.834°
See also example 4.2 in section 4.2.3.1

In summary, a rotation matrix with algebraic functions offers the following benefits in contrast to trigonometric functions:
– no singularities, (i.e. no gimbal lock);
– no dependency on the sequence of rotations;

– no dependency on the definition of coordinate axes;
– simplified computation of the design matrix (the first derivatives of a, b, c, d are linear);
– faster convergence in adjustment systems;
– faster computation by avoiding power series for internal trigonometric calculations.

However, the geometric interpretation of quaternions is more complex, e.g. in error analysis of rotation parameters around particular rotation axes.

2.2.2.3 Rodrigues rotation matrix

The rotation matrix according to Rodrigues is also based on a rotation around an axis in space. Using the quaternion in (2.42) and the parameters

$$a' = \frac{2q_1 \cdot \tan(\alpha/2)}{\sqrt{q_1^2 + q_2^2 + q_3^2}} \qquad b' = \frac{2q_2 \cdot \tan(\alpha/2)}{\sqrt{q_1^2 + q_2^2 + q_3^2}} \qquad c' = \frac{2q_3 \cdot \tan(\alpha/2)}{\sqrt{q_1^2 + q_2^2 + q_3^2}} \qquad (2.44)$$

the Rodrigues matrix is derived:

$$\mathbf{R} = \frac{1}{4 + a'^2 + b'^2 + c'^2} \begin{bmatrix} 4 + a'^2 - b'^2 - c'^2 & 2a'b' - 4c' & 2a'c' + 4b' \\ 2a'b' + 4c' & 4 - a'^2 + b'^2 - c'^2 & 2b'c' - 4a' \\ 2a'c' - 4b' & 2b'c' + 4a' & 4 - a'^2 - b'^2 + c'^2 \end{bmatrix} \qquad (2.45)$$

The Rodrigues matrix consists of three independent parameters but cannot describe rotations where $\alpha = 180°$ as the tangent function is undefined at 90° ($\tan(\alpha/2)$).

2.2.2.4 Rotation matrix with direction cosines

The spatial rotation matrix can also be regarded as a matrix of *direction cosines* of the angles δ between the original and the rotated coordinate axes. The unit vectors **i,j,k** are defined in the direction of the rotated axes (Fig. 2.16).

$$\mathbf{R} = \begin{bmatrix} \cos\delta_{xX} & \cos\delta_{yX} & \cos\delta_{zX} \\ \cos\delta_{xY} & \cos\delta_{yY} & \cos\delta_{zY} \\ \cos\delta_{xZ} & \cos\delta_{yZ} & \cos\delta_{zZ} \end{bmatrix} = \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \end{bmatrix} \qquad (2.46)$$
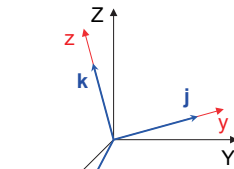
Fig. 2.16: Direction cosines.

2.2.2.5 Normalization of rotation matrices

If the coefficients of a rotation matrix are not explicitly derived from three rotational values, but instead are the result of a calculation process such as the determination

of exterior orientation or a spatial similarity transformation, then the matrix can show departures from orthogonality and orthonormality. Possible causes are systematic errors in the input data or limits to computational precision. In this case, the matrix can be orthonormalized by methods such as the Gram-Schmidt procedure or the following similar method:

With the initial rotation matrix (to be orthonormalized)

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{u} & \mathbf{v} & \mathbf{w} \end{bmatrix} \tag{2.47}$$

create direction vectors which have unit length (unit vectors), and are mutually orthogonal, and which form the new (orthonormal) matrix as follows:

$$\mathbf{u'} = \frac{\mathbf{u}}{|\mathbf{u}|} \qquad\qquad \mathbf{s} = \mathbf{v} - \frac{\mathbf{v} \cdot \mathbf{u'}}{\mathbf{u'}} \qquad \mathbf{v'} = \frac{\mathbf{s}}{|\mathbf{s}|} \qquad\qquad \mathbf{w'} = \mathbf{u'} \times \mathbf{v'} \tag{2.48}$$

$$\mathbf{R'} = \begin{bmatrix} \mathbf{u'} & \mathbf{v'} & \mathbf{w'} \end{bmatrix} \qquad \text{: orthonormalized rotation matrix}$$

Example 2.4:
A rotation matrix **R** is defined by angles $\omega = 35°$, $\varphi = 60°$, $\kappa = 30°$ according to eqn. (2.32). In this example, the values of the coefficients after the third decimal place are subject to computational error (see also example 2.4):

$$\mathbf{R} = \begin{bmatrix} 0.433273 & 0.844569 & -0.324209 \\ -0.248825 & 0.468893 & 0.855810 \\ 0.876000 & -0.284795 & 0.409708 \end{bmatrix} \qquad \text{and} \qquad \det(\mathbf{R}) = 1.018296$$

which, when multiplied by its transpose, does not result in a unit matrix:

$$\mathbf{R}^T\mathbf{R} = \begin{bmatrix} 1.017015 & -0.000224 & 0.005486 \\ -0.000224 & 1.014265 & 0.010784 \\ 0.005486 & 0.010784 & 1.005383 \end{bmatrix} \qquad \text{and} \qquad \det(\mathbf{R}^T\mathbf{R}) = 1.036927$$

The matrix orthonormalized according to (2.48) is given by:

$$\mathbf{R'} = \begin{bmatrix} 0.429633 & 0.838703 & -0.334652 \\ -0.246735 & 0.465529 & 0.849944 \\ 0.868641 & -0.282594 & 0.406944 \end{bmatrix} \qquad \text{and} \qquad \det(\mathbf{R'}) = 1.000000$$

The three column vectors are now orthogonal to one another in pairs and all have unit length.

2.2.2.6 Comparison of coefficients

The spatial rotation defined in

$$\mathbf{X} = \mathbf{R} \cdot \mathbf{x} \tag{2.49}$$

depends on the nine coefficients $r_{11}...r_{33}$ of **R**. See, for example, the rotation order ω φ κ about rotated axes which defines **R** in eqn. (2.32). If the identical transformation result is to be achieved by a rotation matrix **R'** using a different rotation order, the coefficients of **R'** must be equal to those of **R**:

$$\mathbf{R} = \mathbf{R'} \tag{2.50}$$

If the rotation angles ω',φ',κ' of rotation matrix **R'** are to be calculated from the explicitly given angles ω,φ,κ of **R**, this can be achieved by a comparison of matrix coefficients and a subsequent reverse calculation of the trigonometric functions.

Example 2.5:

Given the rotation matrix of eqn. (2.32) defined by angles ω = 35°, φ = 60°, κ = 30°, determine the rotation angles ω',φ',κ' belonging to the equivalent rotation matrix **R'** defined by eqn. (2.37):

1. Evaluate the coefficients $r_{11}...r_{33}$ of **R** by multiplying out the individual rotation matrices in the order $\mathbf{R} = \mathbf{R}_\omega \cdot \mathbf{R}_\varphi \cdot \mathbf{R}_\kappa$, substituting the given values of ω φ κ:

$$\mathbf{R} = \begin{bmatrix} 0.433013 & -0.250000 & 0.866025 \\ 0.839758 & 0.461041 & -0.286788 \\ -0.327576 & 0.851435 & 0.409576 \end{bmatrix}$$

2. Write the coefficients $r'_{11}...r'_{33}$ of **R'** in trigonometric form by multiplying the individual rotation matrices in the order $\mathbf{R'} = \mathbf{R}_\varphi \cdot \mathbf{R}_\omega \cdot \mathbf{R}_\kappa$. Assign to each coefficient the values from **R**, i.e. $r'_{11} = r_{11}$, $r'_{12} = r_{12}$, and so on.

3. Calculate the rotation angles ω',φ',κ' of **R'** by solution of trigonometric equations:
   ω' = 16.666°   φ' = 64.689°   κ' = 61.232°

2.2.3 Spatial transformations

2.2.3.1 General transformations

The general linear transformation of homogeneous coordinates is given by:

$$\boldsymbol{X} = \lambda \cdot \mathbf{T} \cdot \boldsymbol{x} \tag{2.51}$$

where λ is an arbitrary scale factor not equal to zero and **T** is the transformation or projection matrix[3].

$$\mathbf{T} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ {}_{3,3} & {}_{1,3} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \\ {}_{3,1} & {}_{1,1} \end{bmatrix} \tag{2.52}$$

---

3 Note that **T** is a homogeneous matrix whilst the four sub-matrices are not.

Due to the arbitrary scale factor, there are 15 independent parameters remaining. By selecting parameter subsets from this general transformation, special transformations such as the affine transformation, projective transformation or similarity transformation can be derived. The result of this transformation always results in a new homogeneous coordinate vector. The four sub-matrices contain information as follows:

$\mathbf{T}_{11}$: scaling, reflection in a line, rotation

$\mathbf{T}_{12}$: translation

$\mathbf{T}_{21}$: perspective

$\mathbf{T}_{22}$: homogeneous scaling

Scaling or reflection about a line is performed by the factors s_X, s_Y, s_Z:

$$\mathbf{T}_S = \begin{bmatrix} s_X & 0 & 0 & 0 \\ 0 & s_Y & 0 & 0 \\ 0 & 0 & s_Z & 0 \\ \hline 0 & 0 & 0 & 1 \end{bmatrix} \quad : \text{scaling, reflection in a line} \qquad (2.53)$$

A spatial rotation results if $\mathbf{T}_{11}$ is replaced by the rotation matrix derived in section 2.2.2:

$$\mathbf{T}_R = \begin{bmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ \hline 0 & 0 & 0 & 1 \end{bmatrix} \quad : \text{spatial rotation} \qquad (2.54)$$

Translation by a vector x_T, y_T, z_T is performed by the matrix:

$$\mathbf{T}_T = \begin{bmatrix} 1 & 0 & 0 & x_T \\ 0 & 1 & 0 & y_T \\ 0 & 0 & 1 & z_T \\ \hline 0 & 0 & 0 & 1 \end{bmatrix} \quad : \text{translation} \qquad (2.55)$$

Combined transformations $\mathbf{T}_1$, $\mathbf{T}_2$ etc. can be created by sequential multiplication of single projection matrices as follows:

$$\boldsymbol{X} = \mathbf{T} \cdot \boldsymbol{x} = \mathbf{T}_n \cdot \ldots \cdot \mathbf{T}_2 \cdot \mathbf{T}_1 \cdot \boldsymbol{x} \qquad (2.56)$$

In general, the multiplication order may not be changed because the projections are not necessarily commutative.

The reverse transformation is given by:

$$\boldsymbol{x} = \mathbf{T}^{-1} \cdot \boldsymbol{X} = \mathbf{T}_1^{-1} \cdot \mathbf{T}_2^{-1} \cdot \ldots \cdot \mathbf{T}_n^{-1} \cdot \boldsymbol{X} \tag{2.57}$$

This inversion is only possible if the projection matrix is not singular, as is the normal case for the transformation of one 3D system into another. However, if the vector $\boldsymbol{x}$ is projected onto a plane, the projection matrix does become singular. The original coordinates cannot then be calculated from the transformed plane coordinates $\boldsymbol{X}$.

2.2.3.2 Central projection

The central projection is of fundamental importance in photogrammetry and it can also be expressed by a homogenous transformation.
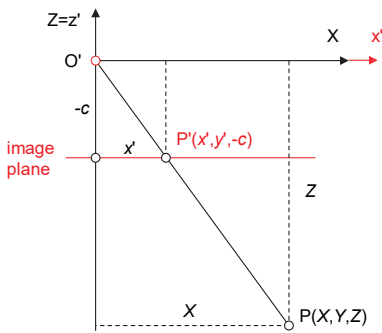
Fig. 2.17: Central projection.

The central projection is modelled firstly for the following special case. The projection plane is oriented normal to the viewing direction Z with the distance $-c$ to the perspective centre at O'. Here the origins of the image coordinate system x'y'z' and the object coordinate system XYZ are both located in the perspective centre. Referring to Fig. 2.17, the following ratios can be derived.

$$\frac{x'}{-c} = \frac{X}{Z} \qquad\qquad \frac{y'}{-c} = \frac{Y}{Z} \qquad\qquad \frac{z'}{-c} = \frac{Z}{Z} = 1 \tag{2.58}$$

and further rearranged to give x', y' and z':

$$x' = -c\frac{X}{Z} = m \cdot X \qquad\quad y' = -c\frac{Y}{Z} = m \cdot Y \qquad\quad z' = -c \tag{2.59}$$

If the perspective centre moves to infinity, c becomes infinite and the term c/Z approaches the value 1. The central projection then changes to a parallel projection.

In homogeneous matrix notation, the perspective transformation is firstly written as

$$\begin{bmatrix} U' \\ V' \\ W' \\ \hline T' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & | & 0 \\ \hline 0 & 0 & 1/-c & | & 0 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \\ Z/-c \end{bmatrix} \qquad (2.60)$$

$$U' = \mathbf{T} \cdot X$$

and for the resulting Cartesian coordinates after division by $Z/-c$:

$$x' = \frac{-c}{-c/Z} U' = \mathbf{T}_P \cdot X$$

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} -c/Z & 0 & 0 & | & 0 \\ 0 & -c/Z & 0 & | & 0 \\ 0 & 0 & -c/Z & | & 0 \\ \hline 0 & 0 & 1/Z & | & 0 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} -cX/Z \\ -cY/Z \\ -c \\ 1 \end{bmatrix} \qquad (2.61)$$

The fourth row of $\mathbf{T}_P$ (sub-matrices $\mathbf{T}_{21}$ and $\mathbf{T}_{22}$ in eqn. 2.52) implies the central projective effect. Eqn. (2.61) also leads to the projection equations (2.59) that are already known for the pinhole camera model (compare section 1.2.3).

If the above-mentioned special case is extended to an arbitrary exterior orientation of the image plane (position and orientation in space), the transformation of object coordinates into image coordinates can be performed by the following matrix operation, which is the inverse of (2.66):

$$x' = \mathbf{T}_P \cdot \mathbf{T}_R^{-1} \cdot \mathbf{T}_T^{-1} \cdot X \qquad (2.62)$$

2.2.3.3 General affine transformation

The general affine transformation describes a spatial transformation of points by 12 parameters which for each axis comprise a translation, scale factor, shear parameter and rotation.

$$\mathbf{X} = \mathbf{X}_0 + \mathbf{A} \cdot \mathbf{x} \qquad (2.63)$$

or

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} a_0 \\ b_0 \\ c_0 \end{bmatrix} + \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad \begin{aligned} X &= a_0 + a_1 x + a_2 y + a_3 z \\ Y &= b_0 + b_1 x + b_2 y + b_3 z \\ Z &= c_0 + c_1 x + c_2 y + c_3 z \end{aligned}$$

In homogeneous notation:

$$X = \mathbf{T}_A \cdot x$$

$$
\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & a_0 \\ 0 & 1 & 0 & b_0 \\ 0 & 0 & 1 & c_0 \\ \hdashline 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} a_1 & a_2 & a_3 & 0 \\ b_1 & b_2 & b_3 & 0 \\ c_1 & c_2 & c_3 & 0 \\ \hdashline 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 & a_0 \\ b_1 & b_2 & b_3 & b_0 \\ c_1 & c_2 & c_3 & c_0 \\ \hdashline 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{2.64}
$$

This system of equations is linear and can be solved without approximate values. The transformation is therefore also suitable for obtaining approximate values for non-linear spatial transformations.

2.2.4 Spatial similarity transformation

2.2.4.1 Mathematical model

The *spatial similarity transformation* is used for the shape-invariant mapping of a three-dimensional Cartesian coordinate system xyz into a corresponding target system XYZ. Both systems can be arbitrarily rotated, shifted and scaled with respect to each other. It is important to note that the rectangularity of the coordinate axes is preserved. This transformation is therefore a special case of the general affine transformation (section 2.2.3.3).
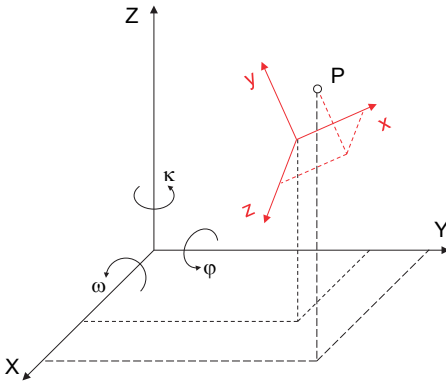
Fig. 2.18: Spatial similarity transformation.

The spatial similarity transformation, also known as a 3D Helmert transformation, is defined by 7 parameters, namely 3 translations to the origin of the xyz system (vector $\mathbf{X_0}$ defined by X_0, Y_0, Z_0), 3 rotation angles ω, φ, κ about the axes XYZ (implied by orthogonal rotation matrix $\mathbf{R}$) and one scaling factor m (Fig. 2.18). The 6 parameters for translation and rotation correspond to the parameters of exterior orientation (see

section 4.2.1). Parameters are applied in the order rotate - scale - shift and the transformation function for a point P(x,y,z), defined by vector **x**, is given by:

$$\mathbf{X} = \mathbf{X_0} + m \cdot \mathbf{R} \cdot \mathbf{x} \tag{2.65}$$

or

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + m \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Using homogenous coordinates, the spatial similarity transformation of eqn. (2.65) is given by ($m = s_x = s_y = s_z$, see eqn. 2.51):

$$\mathbf{X} = \boldsymbol{T}_T \cdot \boldsymbol{T}_S \cdot \boldsymbol{T}_R \cdot \boldsymbol{x}$$

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} m & 0 & 0 & 0 \\ 0 & m & 0 & 0 \\ 0 & 0 & m & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{2.66}$$

$$= \begin{bmatrix} mr_{11} & mr_{12} & mr_{13} & X_0 \\ mr_{21} & mr_{22} & mr_{23} & Y_0 \\ mr_{31} & mr_{32} & mr_{33} & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

In order to determine the seven parameters, a minimum of seven observations is required. These observations can be derived from the coordinate components of at least three spatially distributed reference points (control points). They must contain at least 2 X, 2 Y and 3 Z components[4] and they must not lie on a common straight line in object space.

The spatial similarity transformation is of fundamental importance to photogrammetry for two reasons. Firstly, it is a key element in the derivation of the collinearity equations, which are the fundamental equations of analytical photogrammetry (see section 4.2.2). Secondly, it is used for the transformation of local 3D coordinates such as model coordinates or 3D measuring machine coordinates, into an arbitrary superior system, for example an object or world coordinate system, as required, say, for absolute orientation (see section 4.3.5) or bundle adjustment (see section 4.4). It can also be used to detect deviations or deformations between two groups of points.

There are simplified solutions for a transformation between two systems that are approximately parallel. In the general case both source and target system have an

---

4 It is assumed that the viewing direction is approximately parallel to the Z axis. For other image orientations appropriately positioned minimum control information is required.

arbitrary relative orientation, i.e. any possible translation and rotation may occur. The calculation of transformation parameters then requires linearization of the system of equations defined by the similarity transformation (2.65). Sufficiently accurate initial values are then required in order to determine the unknown parameters (see below). An alternative solution is presented in section 2.2.4.3.

The system of equations is normally over-determined and the solution is performed by least-squares adjustment (see section 2.4). This derives an optimal fit between both coordinate systems. According to eqn. (2.65) every reference point defined in both systems generates up to three equations (compare with eqn. 2.63):

$$
\begin{aligned}
X &= X_0 + m \cdot (r_{11} \cdot x + r_{12} \cdot y + r_{13} \cdot z) \\
Y &= Y_0 + m \cdot (r_{21} \cdot x + r_{22} \cdot y + r_{23} \cdot z) \\
Z &= Z_0 + m \cdot (r_{31} \cdot x + r_{32} \cdot y + r_{33} \cdot z)
\end{aligned}
\tag{2.67}
$$

By linearizing the equations at approximate parameter values, corresponding correction equations are built up. Any reference point with defined X, Y and Z coordinates (full reference point) provides three observation equations. Correspondingly, reference points with fewer coordinate components generate fewer observation equations but they can still be used for parameter estimation. Thus, a transformation involving 3 full reference points already provides 2 redundant observations. The 3-2-1 method (see section 4.4.3), used in industrial metrology, is based on 6 observations, does not derive a scale change, and therefore results in zero redundancy.

Each reference point or each observation can be weighted individually (see section 2.4.1.2). For example, this can be based on an a priori known accuracy of the reference point measurement. If there is no reliable information to indicate that reference coordinates have different accuracies, all observations should be weighted equally. Otherwise transformation parameters may be biased and, as a result, transformed points may be subject to deformation.

There is a special case of the 3D similarity transformation when the scale factor is fixed, i.e. 6 unknown parameters remain. This transformation is then often known as a *rigid-body transformation*.

2.2.4.2 Approximate values

In order to calculate approximate values of the translation and rotation parameters of the similarity transformation, an intermediate coordinate system is formed. This is derived from 3 reference points $P_1, P_2, P_3,$ defined in an intermediate system uvw and known in both the target system XYZ and the source system xyz (Fig. 2.19). The purpose at this stage is to calculate the parameters which transform the reference points from intermediate system uvw to coordinate systems XYZ and xyz.
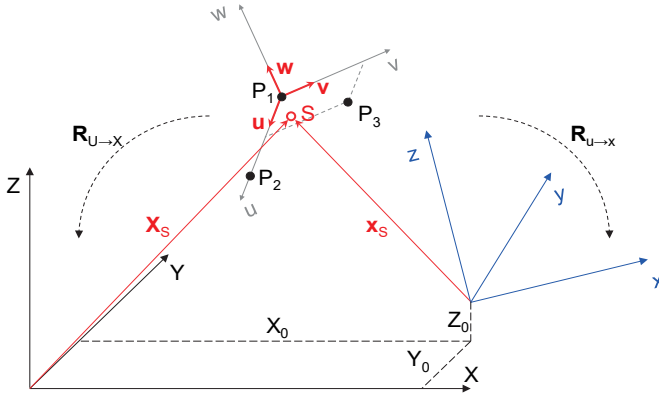
Fig. 2.19: Calculation of approximate values for 3D similarity transformation.

$$\mathbf{P}_{XYZ} = \mathbf{R}_{u \to X} \cdot \mathbf{P}_{uvw} + \mathbf{T}_{u \to X} \qquad\qquad \mathbf{P}_{xyz} = \mathbf{R}_{u \to x} \cdot \mathbf{P}_{uvw} + \mathbf{T}_{u \to x} \qquad (2.68)$$

Solving both equations for $\mathbf{P}_{uvw}$ and re-arranging:

$$\mathbf{R}_{u \to X}^{T} \cdot (\mathbf{P}_{XYZ} - \mathbf{T}_{u \to X}) = \mathbf{R}_{u \to x}^{T} \cdot (\mathbf{P}_{xyz} - \mathbf{T}_{u \to x}) \qquad (2.69)$$

and finally, for the coordinates of a point in system XYZ:

$$\begin{aligned}
\mathbf{P}_{XYZ} &= \mathbf{R}_{u \to X} \cdot \mathbf{R}_{u \to x}^{T} \cdot \mathbf{P}_{xyz} + \mathbf{T}_{u \to X} - \mathbf{R}_{u \to X} \cdot \mathbf{R}_{u \to x}^{T} \cdot \mathbf{T}_{u \to x} \\
&= \mathbf{R}_{x \to X}^{0} \cdot \mathbf{P}_{xyz} + (\mathbf{T}_{u \to X} - \mathbf{R}_{x \to X}^{0} \cdot \mathbf{T}_{u \to x})
\end{aligned} \qquad (2.70)$$

Here matrices $\mathbf{R}_{u \to X}$ and $\mathbf{R}_{u \to x}$ describe the rotation of each system under analysis with respect to the intermediate system. The vectors $\mathbf{T}_{u \to X}$ and $\mathbf{T}_{u \to x}$ describe the corresponding translations. The expression in brackets describes the translation between systems XYZ and xyz:

$$\mathbf{X}_{x \to X}^{0} = \mathbf{T}_{u \to X} - \mathbf{R}_{x \to X}^{0} \cdot \mathbf{T}_{u \to x} \qquad (2.71)$$

To calculate the required parameters, the u axis of the intermediate system is constructed through P_{13} and P_2 and the uv plane through P_3 (corresponds to the 3-2-1 method). From the local vectors defined by the reference points $\mathbf{P}_i(X_i, Y_i, Z_i)$, $i = 1...3$, normalized direction vectors are calculated. Here vectors **u,v,w** are derived from the coordinates of $\mathbf{P}_i$ in the source system xyz, while **U,V,W** are calculated from the target system coordinates XYZ:

$$\mathbf{U} = \frac{\mathbf{P}_2 - \mathbf{P}_1}{\left|\mathbf{P}_2 - \mathbf{P}_1\right|} \qquad \mathbf{W} = \frac{\mathbf{U} \times (\mathbf{P}_3 - \mathbf{P}_1)}{\left|\mathbf{U} \times (\mathbf{P}_3 - \mathbf{P}_1)\right|} \qquad \mathbf{V} = \mathbf{W} \times \mathbf{U} \tag{2.72}$$

$$\mathbf{u} = \frac{\mathbf{p}_2 - \mathbf{p}_1}{\left|\mathbf{p}_2 - \mathbf{p}_1\right|} \qquad \mathbf{w} = \frac{\mathbf{u} \times (\mathbf{p}_3 - \mathbf{p}_1)}{\left|\mathbf{u} \times (\mathbf{p}_3 - \mathbf{p}_1)\right|} \qquad \mathbf{v} = \mathbf{w} \times \mathbf{u}$$

Vector $\mathbf{u}$ is a unit vector on the u axis, $\mathbf{w}$ is perpendicular to the uv plane and $\mathbf{v}$ is perpendicular to $\mathbf{u}$ and $\mathbf{w}$. These 3 vectors directly define the rotation matrix from uvw to XYZ (see eqn. 2.46):

$$\mathbf{R}_{U \to X} = \begin{bmatrix} \mathbf{U} & \mathbf{V} & \mathbf{W} \end{bmatrix} \qquad \mathbf{R}_{u \to x} = \begin{bmatrix} \mathbf{u} & \mathbf{v} & \mathbf{w} \end{bmatrix} \tag{2.73}$$

The approximate rotation matrix from the xyz to the XYZ system is obtained from successive application of the above two matrices as follows:

$$\mathbf{R}^0_{x \to X} = \mathbf{R}_{U \to X} \cdot \mathbf{R}^T_{u \to x} \tag{2.74}$$

The approximate scale factor can be calculated from the point separations:

$$m^0 = \frac{\left|\mathbf{P}_2 - \mathbf{P}_1\right|}{\left|\mathbf{p}_2 - \mathbf{p}_1\right|} = \frac{\sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2 + (Z_2 - Z_1)^2}}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}} \tag{2.75}$$

Using the centroid of the reference points in both coordinate systems, approximate values for the translation parameters of the similarity transformation can be calculated:

$$\mathbf{X}_S = \begin{bmatrix} X_S \\ Y_S \\ Z_S \end{bmatrix} = \mathbf{T}_{u \to X} \qquad \text{: centroid in XYZ system} \tag{2.76}$$

$$\mathbf{x}_S = \begin{bmatrix} x_S \\ y_S \\ z_S \end{bmatrix} = \mathbf{T}_{u \to x} \qquad \text{: centroid in xyz system} \tag{2.77}$$

According to (2.71) the translation can then be calculated:

$$\mathbf{X}^0_{x \to X} = \mathbf{X}_S - m^0 \cdot \mathbf{R}^0_{x \to X} \cdot \mathbf{x}_S \tag{2.78}$$

Example 2.6:

5 points are known in the source and target systems and have the following 3D coordinates:

No.	x	y	z	X	Y	Z
1	110.0	100.0	110.0	153.559	170.747	150.768
2	150.0	280.0	100.0	99.026	350.313	354.912
3	300.0	300.0	120.0	215.054	544.420	319.003
4	170.0	100.0	100.0	179.413	251.030	115.601
5	200.0	200.0	140.0	213.431	340.349	253.036

Approximate values, calculated using points 1, 2 and 3 as above, are:

Rotation: $\quad \mathbf{R}^0_{x \to X} = \begin{bmatrix} 0.433558 & -0.250339 & 0.865654 \\ 0.839451 & 0.461481 & -0.286979 \\ -0.327641 & 0.851097 & 0.410226 \end{bmatrix}$

Scale factor: $\quad m^0 = 1.501637$

Translation: $\quad \mathbf{X}^0_0 = \begin{bmatrix} -23.430 \\ 10.185 \\ 9.284 \end{bmatrix}$

The adjusted parameters are given in example 2.7.

2.2.4.3 Calculation with eigenvalues and quaternions

The rotation matrix of the spatial similarity transformation can also be derived directly from the two sets of points as the related quaternion can be determined by eigenvalue analysis. Firstly, the 3D coordinates of points P_i are reduced to their centroid:

$$\overline{\mathbf{X}}_i = \mathbf{X}_i - \mathbf{X}_S \qquad\qquad \overline{\mathbf{x}}_i = \mathbf{x}_i - \mathbf{x}_S \qquad\qquad (2.79)$$

Using the matrices $\mathbf{S}_X$ and $\mathbf{S}_x$ formed by the coordinate components of all n points P_i

$$\mathbf{S}_X = \begin{bmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_3 \\ \vdots & \vdots & \vdots \\ X_n & Y_n & Z_n \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_n \end{bmatrix} \qquad \mathbf{S}_x = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_3 \\ \vdots & \vdots & \vdots \\ x_n & y_n & z_n \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} \qquad (2.80)$$

the 3x3 matrix $\mathbf{M}$ is calculated

$$\mathbf{M} = \mathbf{S}_X^T \cdot \mathbf{S}_x \qquad\qquad (2.81)$$

which then is used to form the following symmetrical matrix $\mathbf{N}$:

$$N = \begin{bmatrix} m_{11}+m_{22}+m_{33} & m_{23}-m_{32} & m_{31}-m_{13} & m_{12}-m_{21} \\ & m_{11}-m_{22}-m_{33} & m_{12}+m_{21} & m_{31}+m_{13} \\ & & -m_{11}+m_{22}-m_{33} & m_{23}+m_{32} \\ & & & -m_{11}-m_{22}+m_{33} \end{bmatrix} \quad (2.82)$$

The eigenvector of N with the largest eigenvalue λ_{max} gives the required quaternion of the rotation between both systems.

Translation and scale are calculated according to section 2.2.4.1 or by:

$$m = \lambda_{max} / \sum_{i=1}^{n} \mathbf{x}_i^T \mathbf{x}_i \qquad\qquad \mathbf{X}_{x \to X}^0 = \mathbf{X}_S - m^0 \cdot \mathbf{R}_{x \to X}^0 \cdot \mathbf{x}_S \qquad (2.83)$$

Example 2.7:

Using the five points from example 2.6 the following transformation parameters are calculated:

Least-squares adjustment	Eigenvalues and quaternions
$m = 1.500050$	$m = 1.500050$
$X_0 = -23.4154$	$X_0 = -23.4155$
$Y_0 = 10.6115$	$Y_0 = 10.6115$
$Z_0 = 9.7112$	$Z_0 = 9.7122$
$R = \begin{bmatrix} 0.433878 & -0.250183 & 0.865539 \\ 0.839270 & 0.461625 & -0.287278 \\ -0.327682 & 0.851065 & 0.410260 \end{bmatrix}$	$R = \begin{bmatrix} 0.433878 & -0.250183 & 0.865540 \\ 0.839270 & 0.461625 & -0.287278 \\ -0.327682 & 0.851065 & 0.410260 \end{bmatrix}$
$\omega = 35.00102°$	$\omega = 35.00102°$
$\varphi = 59.94435°$	$\varphi = 59.94435°$
$\kappa = 29.96861°$	$\kappa = 29.96861°$
$S_0 = 0.204$	$S_0 = 0.204$
RMS X;Y;Z = 0.075; 0.177; 0.173	RMS X;Y;Z = 0.075; 0.177; 0.173

Example 2.7 demonstrates that the calculation using quaternions generates the same result as least-squares adjustment based on the observation equations (2.67). However, the possible need for individual weighting of observations is much more complex if eigenvalues are used. Where applicable, the eigenvalue computation should be followed by a least-squares adjustment with a suitable stochastic model.

2.2.5 Additional coordinate transformations

2.2.5.1 Spherical coordinates

Spherical coordinates use two angles α, β and the distance r to the origin to define a point in 3D space. (Fig. 2.20). For r = const. all points lie on a sphere. For a right-handed convention, α is positive from X towards Y. β increases from zero on the Z axis.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = r \begin{bmatrix} \cos\alpha\sin\beta \\ \sin\alpha\sin\beta \\ \cos\beta \end{bmatrix} \tag{2.84}$$
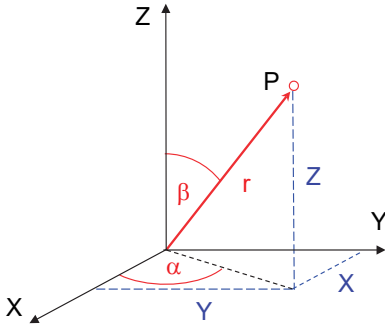
Fig. 2.20: Spherical coordinates.   **Fig. 2.21:** Cylindrical coordinates.
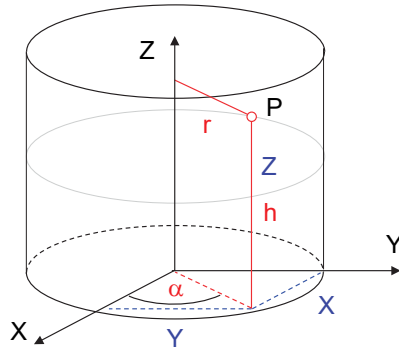
The reverse transformation from Cartesian coordinates to the angles α, β and the distance *r* depends on quadrants, hence is not unique. Spherical coordinates are used, for instance, for the calculation of 3D coordinates from total stations or laser scanners (polar measuring systems).

2.2.5.2 Cylindrical coordinates

Cylindrical coordinates define a 3D point using one angle α, the cylinder radius *r* and the height *h* along the cylinder axis (Fig. 2.21). Again, for a right-handed convention, α is positive from X towards Y.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} r\cos\alpha \\ r\sin\alpha \\ h \end{bmatrix} \tag{2.85}$$

As for spherical coordinates, the calculation of α from Cartesian coordinates is not unique. Cylindrical coordinates can be used, for example, in colour transformations (section 5.2.2.2) or in panoramic photogrammetry (section 4.5.1).

2.3 Geometric elements

The geometric reconstruction of a measured object is the major goal of a photogrammetric process. This section therefore gives a short summary of geometric elements and their mathematical definition. It distinguishes between planar

elements, spatial elements and surface descriptions that are the basic result of a photogrammetric measurement. For a detailed description of the methods of analytical geometry, the reader should refer to specialist literature on geometry and 3D computer graphics.

Process	Example 1	Example 2
points XYZ	points on cylinders	points on a free-form surface
adjustment		
calculated geometric elements	cylinder, axis	triangulation mesh planes
combination		
derived geometric elements	intersection point of cylinder axes	intersection with rays of sight
derived quantities	distance between two intersetcion points	supression of hidden lines

Fig. 2.22: Calculation progress for geometric elements.

Except in very few cases, photogrammetric methods are based on measurement of discrete object points. Geometric elements such as straight lines, planes, cylinders etc. are normally calculated in a post-processing step using the measured 3D points. For over-determined solutions, least-squares fitting methods are used. Computed geometric elements can then either be combined or intersected in order to create additional geometric elements such as the intersection line between two planes. Alternatively, specific dimensions can be derived from them, such as the distance between two points (Fig. 2.22).

In addition to the determination of regular geometric shapes, the determination and visualization of arbitrary three-dimensional surfaces (free-form surfaces) is of increasing importance. This requires a basic knowledge of different ways to represent 3D surfaces, involving point grids, triangle meshing, analytical curves, voxels etc.

Many of these calculations are embedded in state-of-the-art 3D CAD systems or programs for geometric quality analysis. CAD and photogrammetric systems are therefore often combined. However, geometric elements may also be directly employed in photogrammetric calculations, e.g. as conditions for the location of object points (see section 4.4.2.3). In addition, some evaluation techniques enable the direct calculation of geometric 3D elements without the use of discrete points (e.g. contour method, section 4.4.7.2).

2.3.1 Analytical geometry in the plane

2.3.1.1 Straight line

2.3.1.1.1 Parametric form

The straight line g between two points P_1 and P_2 (Fig. 2.23) is to be determined. For all points $P(x,y)$ belonging to g, the proportional relationship

$$\frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1} \tag{2.86}$$

leads to the parametric form of the straight line:

$$\mathbf{x} = \mathbf{x}_1 + t \cdot (\mathbf{x}_2 - \mathbf{x}_1)$$
$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + t \cdot \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \end{bmatrix} \tag{2.87}$$

Point $P_1(x_1,y_1)$ is defined at $t = 0$ and point $P_2(x_2,y_2)$ at $t = 1$.

The distance d of a point $Q(x_Q,y_Q)$ from the straight line is defined by:

$$d = \frac{(y_2 - y_1)(x_Q - x_1) - (x_2 - x_1)(y_Q - y_1)}{l} \tag{2.88}$$

with

$$l = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

The foot $F(x,y)$ of the perpendicular from Q to the line is given by:

$$\begin{bmatrix} x_F \\ y_F \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + s \cdot \begin{bmatrix} y_1 - y_2 \\ x_2 - x_1 \end{bmatrix} \tag{2.89}$$

where

$$s = \frac{(y_1 - y_Q)(x_2 - x_1) - (x_1 - x_Q)(y_2 - y_1)}{l^2}$$

2.3.1.1.2 Analytical form

The analytical form of a straight line

$$A \cdot x + B \cdot y + C = 0 \tag{2.90}$$

leads to the following relations (Fig. 2.23):

$$y = m \cdot x + c$$

where

$$m = \tan \alpha = -\frac{A}{B} = \frac{y_2 - y_1}{x_2 - x_1} \qquad \text{: slope} \qquad (2.91)$$

$$c = -\frac{C}{B} \qquad \text{: intersection point on y axis}$$

The (perpendicular) distance d to point $Q(x_Q, y_Q)$ is given by:

$$d = \frac{A \cdot x_Q + B \cdot y_Q + C}{\sqrt{A^2 + B^2}} \qquad (2.92)$$
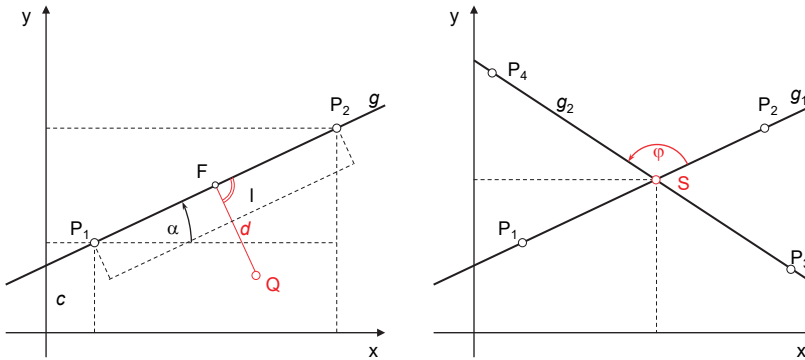
Fig. 2.23: Definition of straight lines.

2.3.1.1.3 Intersection of two straight lines

Given two straight lines g_1 and g_2, their point of intersection S is derived from (2.90) as:

$$x_S = \frac{B_1 \cdot C_2 - B_2 \cdot C_1}{A_1 \cdot B_2 - A_2 \cdot B_1} \qquad\qquad y_S = \frac{C_1 \cdot A_2 - C_2 \cdot A_1}{A_1 \cdot B_2 - A_2 \cdot B_1} \qquad (2.93)$$

Alternatively, from (2.87) two equations for parameters t_1 und t_2 are obtained:

$$\begin{aligned}
(x_2 - x_1) \cdot t_1 + (x_3 - x_4) \cdot t_2 = x_3 - x_1 \\
(y_2 - y_1) \cdot t_1 + (y_3 - y_4) \cdot t_2 = y_3 - y_1
\end{aligned} \qquad (2.94)$$

The point of intersection is obtained by substituting t_1 or t_2 into the original straight line equations:

$$\mathbf{x}_S = \mathbf{x}_1 + t_1 \cdot (\mathbf{x}_2 - \mathbf{x}_1) = \mathbf{x}_3 + t_2 \cdot (\mathbf{x}_4 - \mathbf{x}_3)$$

$$\begin{bmatrix} x_S \\ y_S \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + t_1 \cdot \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \end{bmatrix} = \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} + t_2 \cdot \begin{bmatrix} x_4 - x_3 \\ y_4 - y_3 \end{bmatrix} \tag{2.95}$$

The angle between both lines is given by:

$$\tan\varphi = \frac{A_1 \cdot B_2 - A_2 \cdot B_1}{A_1 \cdot A_2 - B_1 \cdot B_2} = \frac{m_2 - m_1}{m_1 \cdot m_2 + 1} \tag{2.96}$$

Alternatively, if both lines are defined by their direction vectors

$$\mathbf{a} = \mathbf{x}_2 - \mathbf{x}_1 \qquad \text{and} \qquad \mathbf{b} = \mathbf{x}_4 - \mathbf{x}_3$$

then the angle between them can be found from the scalar product of both vectors:

$$\cos\varphi = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| \cdot |\mathbf{b}|} = \frac{\mathbf{a}^T \mathbf{b}}{|\mathbf{a}| \cdot |\mathbf{b}|} \tag{2.97}$$

The scalar product is zero if the lines are mutually perpendicular.

2.3.1.1.4 Regression line

The generalized regression line, which is a best-fit to a set of points, is the straight line which minimizes the sum of squared distances d_i of all points P_i from the line (Fig. 2.24). For n point coordinates with equal accuracy in the x and y directions, the criterion is expressed as:

$$d_1^2 + d_2^2 + \ldots + d_n^2 = \sum_{i=1}^{n} d_i^2 \to \min \tag{2.98}$$

The regression line passes through the centroid of the points:

$$x_0 = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad \text{and} \qquad y_0 = \frac{1}{n} \sum_{i=1}^{n} y_i \tag{2.99}$$

One point on the straight line is therefore directly given. The direction of the line is defined by:

$$\tan 2\varphi = \frac{2\sum (x_i - x_0)(y_i - y_0)}{\sum (y_i - y_0)^2 - (x_i - x_0)^2} \tag{2.100}$$
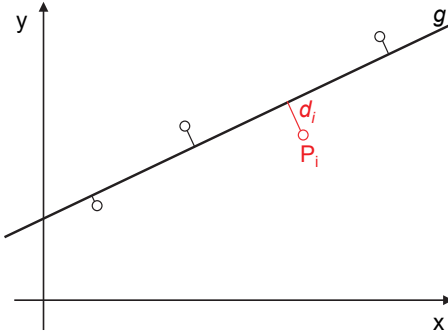
Fig. 2.24: Regression line.

Alternatively, the direction of the line can be expressed by the direction vector (a,b) which is equal to the eigenvector of the maximum eigenvalue of matrix $\mathbf{B}$:

$$\mathbf{B} = \mathbf{A}^T \cdot \mathbf{A}$$

$$\text{where} \quad \mathbf{A} = \begin{bmatrix} x_1 - x_0 & y_1 - y_0 \\ \vdots & \vdots \\ x_n - x_0 & y_n - y_0 \end{bmatrix} \tag{2.101}$$

Without restriction, the optimization principle based on minimum quadratic distances according to (2.98) can be applied to regression lines in space as well as other best-fit elements.

2.3.1.2 Circle

From the generalized equation for second order curves (conic sections)

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0 \tag{2.102}$$

the special cases of circle and ellipse are of major interest in close-range photogrammetry. For a circle with centre (x_M, y_M) and radius r, the equation is typically written as:

$$(x - x_M)^2 + (y - y_M)^2 = r^2 \tag{2.103}$$

This can be re-arranged in the form:

$$x^2 + y^2 - 2x_M x - 2y_M y + x_M^2 + y_M^2 - r^2 = 0 \tag{2.104}$$

This can be further re-arranged as:

$$x^2 + y^2 + 2D'x + 2E'y + F' = 0 \tag{2.105}$$

This is equivalent to eqn. (2.102) with $A = C = 1$ and $B = 0$. There are effectively only three independent, unknown parameters, D', E' and F' and the circle can therefore be defined with a minimum of three points. The linear form of (2.105) can be used directly to solve for D', E' and F' in a least-squares solution where there are more than three points.

By comparing eqn. (2.105) with eqn. (2.104), the radius and centre of the circle can be further derived as follows:

$$r = \sqrt{D'^2 + E'^2 - F'} \qquad\qquad x_M = -D' \qquad\qquad y_M = -E' \qquad (2.106)$$

Alternatively, the distance of any point $P_i(x_i, y_i)$ from the circumference is given by:

$$d_i = r_i - r = \sqrt{(x_i - x_M)^2 + (y_i - y_M)^2} - r \qquad (2.107)$$

The non-linear eqn. (2.107) can also be used as an observation equation after linearization. The best-fit circle is obtained by least-squares minimization of all point distances d_i. With initial approximate values for the centre coordinates and radius, the design matrix **A** consists of the derivatives

$$\frac{\partial d_i}{\partial x_M} = -\frac{x_i - x_M}{r_i} \qquad\qquad \frac{\partial d_i}{\partial y_M} = -\frac{y_i - y_M}{r_i} \qquad\qquad \frac{\partial d_i}{\partial r} = -1 \qquad (2.108)$$

Although the linear approach offers a direct solution without the requirement for initial parameter values, the non-linear approach directly generates the geometrically meaningful parameters of circle centre coordinates and circle radius. For over-determined data, the two solutions will generate slightly different circles because different parameters are used in their respective optimizations. In this case it may be advantageous to use the linear solution to find initial estimates for circle centre and radius and then apply the non-linear solution to optimize these estimates.

2.3.1.3 Ellipse

The determination of ellipse parameters, in particular the centre coordinates, is an important part of the measurement of circular targets which are projected as ellipses in the central perspective image (see sections 6.2.1.1 and 5.4.2.5). As a good approximation, the calculated ellipse centre corresponds to the required centre of the circular target (see section 6.2.1.1 for restrictions).

A simple method for the determination of the ellipse centre is based on the geometry of ellipse diameters. Ellipse diameters are chords that are bisected by the ellipse centre. A conjugate diameter is defined by the straight line through the mid-point of all chords which are parallel to a given diameter. A given diameter and its conjugate intersect at the ellipse centre (see Fig. 2.25). A possible implementation of this technique is presented in section 5.4.2.5.
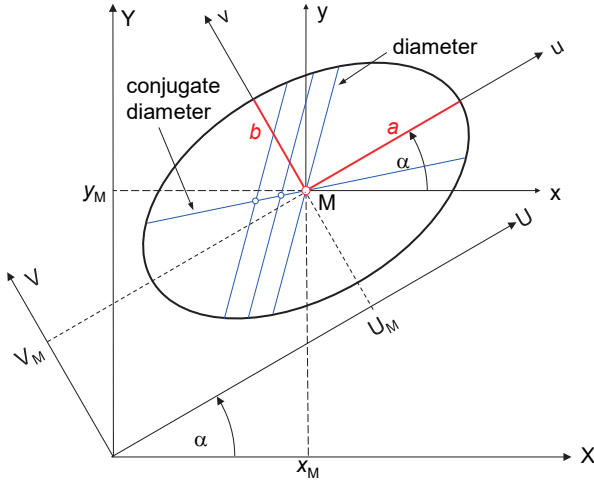
Fig. 2.25: Geometry of an ellipse.

For a full determination of ellipse parameters, a similar approach to the calculation of circle parameters can be applied. The approach is again based on fitting measured points to the generalized equation for second order curves (conic section), which is repeated here:

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0 \qquad (2.109)$$

This equation provides a direct linear solution for the unknown parameters A, B, C, D, E, F by substituting measured values for x,y in order to create an observation equation. However, as in the case of a circle where only 3 of these parameters are required, for an ellipse only 5 are required, as explained below. This requires a minimum of 5 measured points in the image which generate 5 observation equations.

The following analysis indicates a suitable modification of the generalized conic, as well as a derivation of the ellipse parameters (major and minor axes, centre position and rotation angle) from the generalized equation parameters.

The simple form of a non-rotated ellipse, with semi-major axis a, semi-minor axis b and centre at the origin of the coordinate axes, is illustrated by the uv system in Fig. 2.25 and given by:

$$\frac{u^2}{a^2} + \frac{v^2}{b^2} = 1 \qquad (2.110)$$

The axial rotation α must be applied in order to transform the equation from the uv and UV systems to the xy and XY systems respectively. In this process, it is convenient to use terms c and s where $c = \cos α$ und $s = \sin α$ as follows:

$$u = cx + sy \qquad\qquad x = cu - sv \qquad\qquad U_M = cX_M + sY_M$$
$$v = -sx + cy \qquad\qquad y = su + cv \qquad\qquad V_M = -sX_M + cY_M \qquad (2.111)$$

Substituting for u and v in (2.110), the transformed values of u and v in the xy system are:

$$\frac{(cx + sy)^2}{a^2} + \frac{(-sx + cy)^2}{b^2} = 1 \qquad (2.112)$$

Multiplying out and collecting terms gives:

$$\left(\frac{c^2}{a^2} + \frac{s^2}{b^2} \right) x^2 + 2 \left(\frac{cs}{a^2} - \frac{cs}{b^2} \right) xy + \left(\frac{s^2}{a^2} + \frac{c^2}{b^2} \right) y^2 = 1 \qquad (2.113)$$

which may be written as:

$$Ax^2 + 2Bxy + Cy^2 = 1 \qquad (2.114)$$

Applying shifts from the xy to XY system, $x = (X - X_M)$ und $y = (Y - Y_M)$,

$$AX^2 + 2BXY + CY^2 - 2(AX_M + BY_M)X - 2(BX_M + CY_M)Y$$
$$+ (AX_M^2 + 2BX_M Y_M + CY_M^2 - 1) = 0 \qquad (2.115)$$

which may be written as:

$$AX^2 + 2BXY + CY^2 + 2DX + 2EY + F = 0 \qquad (2.116)$$

Eqn. (2.116) is identical to the original generalized eqn. (2.109). Comparing (2.113) and (2.114) it can be seen that:

$$A = \left(\frac{c^2}{a^2} + \frac{s^2}{b^2} \right) \qquad\qquad C = \left(\frac{s^2}{a^2} + \frac{c^2}{b^2} \right) \qquad\qquad 2B = 2cs \left(\frac{1}{a^2} - \frac{1}{b^2} \right) \qquad (2.117)$$

Using the standard trigonometrical identities

$$2cs = 2\cos\alpha\sin\alpha = \sin 2\alpha$$
$$c^2 - s^2 = \cos^2\alpha - \sin^2\alpha = \cos 2\alpha \qquad (2.118)$$

it can further be seen that:

$$\frac{2B}{A - C} = 2\frac{cs}{c^2 - s^2} = \frac{\sin 2\alpha}{\cos 2\alpha} = \tan 2\alpha \qquad (2.119)$$

From a comparison of (2.116) and (2.115)

$$AX_M + BY_M = -D$$
$$BX_M + CY_M = -E \qquad (2.120)$$

which by standard algebraic manipulation gives:

$$X_M = \frac{-DC - EB}{AC - B^2} \qquad\qquad Y_M = \frac{-BD + AE}{B^2 - AC} \tag{2.121}$$

As can be seen from the analysis, the 6 parameters A, B, C, D, E, F of the generalized eqn. (2.116) are themselves based on only 5 parameters a, b, α, X_M, Y_M. In general, the parameters A and C are positive and one may be set to the value 1 by dividing through to obtain, for example, the following linear solution equation for measured point (X_i, Y_i):

$$X_i^2 + 2B'X_iY_i + C'Y_i^2 + 2D'X_i + 2E'Y_i + F' = 0 \tag{2.122}$$

As explained, 5 measured points on the ellipse will generate 5 linear observation equations which can be solved directly by standard matrix algebra. Expressing (2.119) and (2.121) in terms of the actual solution parameters B', C', D', E', F':

$$X_M = \frac{-D'C' - E'B'}{C' - B'^2} \qquad Y_M = \frac{-B'D' + E'}{B'^2 - C'} \qquad \tan 2\alpha = \frac{2B'}{1 - C'} \tag{2.123}$$

If required, the axial parameters of the ellipse, a and b, can be determined as follows. The ellipse equation in the UV system is found by applying the following rotational transformation from the XY system:

$$X = cU - sV \qquad\qquad Y = sU + cV \tag{2.124}$$

Substituting for X and Y in eqn. (2.122) and collecting terms results in:

$$(c^2 + 2B'cs + C's^2)U^2 + [2B'(c^2 - s^2) - 2cs(1 - C')]UV \\ + (C'c^2 - 2B'cs + s^2)V^2 + (2D'c + 2E's)U + (2E'c - 2D's)V + F' = 0 \tag{2.125}$$

This can be written as follows as follows:

$$\bar{A}U^2 + 2\bar{B}UV + \bar{C}V^2 + 2\bar{D}U + 2\bar{E}V + \bar{F} = 0 \tag{2.126}$$

In the UV system it is simple to show $\bar{B} = 0$, which leads to the same result for $\tan 2\alpha$ as expressed in eqn. (2.123). It is also possible to show that the semi axes are then given by:

$$a = \sqrt{\frac{\overline{CD}^2 + \bar{A}\bar{C}^2 - \bar{A}\bar{C}\bar{F}}{\bar{A}^2\bar{C}}} \qquad\qquad b = \sqrt{\frac{\overline{CD}^2 + \bar{A}\bar{C}^2 - \bar{A}\bar{C}\bar{F}}{\bar{A}\bar{C}^2}} \tag{2.127}$$

2.3.1.4 Curves

Consider the requirement that a polynomial with $k+1$ points $P_i(x_i, y_i)$, $i = 0 \dots k$, be described by a closed curve. If the curve should pass through the vertices of a polygon, the process is referred to as *interpolation*. If the curve should be an optimal

fit to the polygon, it is referred to as *approximation*. Curves in general are usually defined by polynomials whose order and curvature properties can be varied with respect to the application.

2.3.1.4.1 Polynomials

A polynomial of degree n is a function of the form:

$$Q(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots a_1 x^1 + a_0 \tag{2.128}$$

that is defined by $n+1$ coefficients.

All points in a data set are used to determine the polynomial coefficients, if necessary by least-squares adjustment. For over-determined solutions, the polynomial does not normally pass through the vertices of the polygon defined by the points. In particular, it does not intersect the end points. Polynomials of higher degree quickly tend to oscillate between the points (Fig. 2.26).
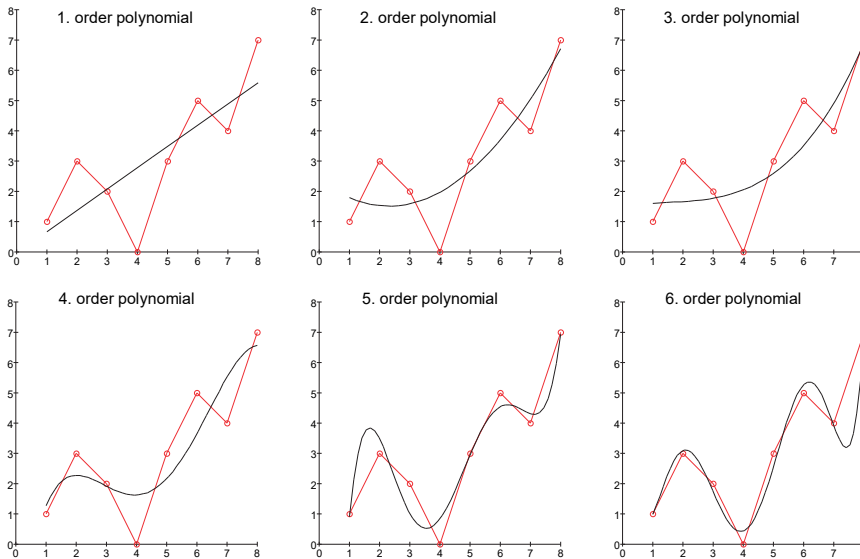
Fig. 2.26: Polygon with 8 data points and polynomial approximations of different order.

A more natural curve shape is obtained if the polygon is approximated by piecewise polynomials. A piecewise polynomial $Q(x)$ is a set of k polynomials $q_i(t)$, each of order n, and $k+1$ nodes[5] $x_0, \ldots, x_k$, with:

---

5 Nodes are the given points defining a curve, i.e. the vertices of a polygon.

$$Q(x) = \{q_i(t)\} \qquad \text{for } x_i \le t \le x_{i+1} \text{ and } i = 0, ..., k-1 \tag{2.129}$$

Using additional constraints, it is possible to generate an approximation curve that is both continuous and smooth. All methods which follow generate a curve that passes through the end points of the polygon and which can be differentiated $n-1$ times at all points. Approximations based on cubic splines ($n = 3$) are of major importance for they provide a practical level of smoothness with a minimum polynomial degree.

2.3.1.4.2 Splines

Splines are used to interpolate between the points of a polygon, i.e. the curve passes through all the points. For this purpose, basic B-spline functions of degree n and order $m = n+1$ are suitable. They are recursively defined for a set of nodes $x_0, x_1, ..., x_{k-1}$:

$$B_{i,0}(t) = \begin{cases} 1 & \text{for } x_i \le t \le x_{i+1} \\ 0 & \text{otherwise} \end{cases} \tag{2.130}$$

$$B_{i,n}(t) = \frac{t - x_i}{x_{i+n} - x_i} B_{i,n-1}(t) + \frac{x_{i+n+1} - t}{x_{i+n+1} - x_{i+1}} B_{i+1,n-1}(t) \tag{2.131}$$

for $x_i \le t \le x_{i+n+1}$.

Optimal smoothness at the data point points is required for spline interpolation, i.e. continuous derivatives up to order $n-1$ should exist. This criterion is fulfilled by the following linear combination of $k+1$ nodes:

$$S_n(t) = \sum_{i=0}^{n+k-1} a_i B_{i,n}(t) \tag{2.132}$$

For the frequently used cubic spline function ($n = 3$)

$$S_3(t) = \sum_{i=0}^{k+2} a_i B_{i,3}(t) \tag{2.133}$$

a number $k+3$ of coefficients a_i have to be determined by a corresponding number of equations. Here $k+1$ equations are provided by the data points and the remaining two equations defined by additional constraints. For example, for *natural splines* these are:

$$\begin{aligned} S_3''(x_0) &= 0 \\ S_3''(x_n) &= 0 \end{aligned} \tag{2.134}$$

Fig. 2.27a shows a polygon approximated by a cubic spline. The resulting curve continuously passes through the vertices (nodes). Splines are therefore most effective

when the vertices are free of position errors, i.e. no smoothing is desired. However, the shape of the entire curve is affected if only one point changes.

2.3.1.4.3 B-Splines

For many technical applications it is more feasible to approximate a given polygon by a curve with the following properties:

– analytical function is simple to formulate;
– can be easily extended to higher dimensions, especially for the surface approximations;
– smoothness at vertices is easy to control;
– variation of a node has only a local effect on the shape of the curve.

The requirements are met by B-spline approximations which are a combination of base functions (2.131) for each point to be interpolated P(t):

$$P(t) = \begin{cases} x(t) = \sum_{i=0}^{k} x_i B_{i,n}(t) \\ y(t) = \sum_{i=0}^{k} y_i B_{i,n}(t) \end{cases} \qquad 0 \le t \le k-n+1 \qquad (2.135)$$

It is obvious that the spline base functions are directly "weighted" by the coordinates of the vertices instead of the computed coefficients. The smoothness of the curve is controlled by the order $m = n+1$, whereby the curve becomes smoother with increasing order. Fig. 2.27c and d show B-spline approximations of order $m = 3$ and $m = 5$. In addition, the computed curve always lies inside the envelope of the polygon, in contrast to normal spline or polynomial interpolation and approximation. Moreover, the approach can be extended directly to three-dimensional polygons (surface elements).

2.3.1.4.4 Bézier approximation

The Bézier approximation has been developed by the car industry. Here a given polygon is approximated by a curve that has optimal smoothness but does not pass through the vertices. The approximation

$$P(t) = \begin{cases} x(t) = \sum_{i=0}^{k} x_i BE_{i,k}(t) \\ y(t) = \sum_{i=0}^{k} y_i BE_{i,k}(t) \end{cases} \qquad 0 \le t \le 1 \qquad (2.136)$$

is similar to the B-spline approximation but is based on the Bernstein polynomials:

$$BE_{i,k}(t) = \frac{k!}{i!(k-i)!} t^i (1-t)^{k-i} \qquad 0 \le t \le 1 \qquad (2.137)$$

All points in the polygon data set are used for the computation of the curve. The approach can be extended directly to three-dimensional polygons (section 2.3.3.2).
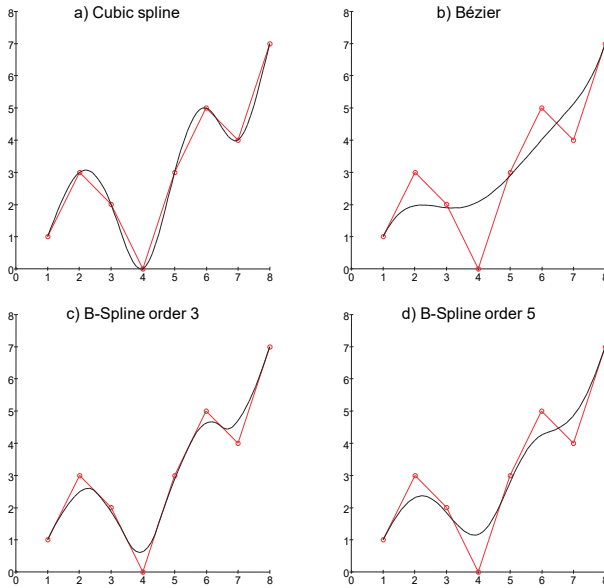
Fig. 2.27: Spline interpolation, and Bézier and B-spline approximation.

Fig. 2.27b shows the curve which results from Bézier approximation of a polygon. The continuous curve does not pass through the vertices, but shows an "averaged" shape. Bézier curves are therefore very suitable for applications where the data points are not free of error and smoothing is required.

2.3.2 Analytical geometry in 3D space

2.3.2.1 Straight line
The form of a straight line in 3D space can be derived directly from the straight line in 2D space. Thus a straight line between two points $P_1(x_1,y_1,z_1)$ and $P_2(x_2,y_2,z_2)$ is given by the proportional relationships

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1} \tag{2.138}$$

and in parametric form:

$$\mathbf{x} = \mathbf{x}_1 + t \cdot (\mathbf{x}_2 - \mathbf{x}_1)$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + t \cdot \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + t \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} \tag{2.139}$$

Here $P_0(x_0, y_0, z_0)$ is any point on the line. The direction cosines are defined by:

$$\cos\alpha = \frac{x_2 - x_1}{d} = a$$

$$\cos\beta = \frac{y_2 - y_1}{d} = b \qquad \text{where } d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \tag{2.140}$$

$$\cos\gamma = \frac{z_2 - z_1}{d} = c$$

At first glance it looks as though there are 6 independent parameters for a straight line in 3D space. However, taking into account the condition

$$a^2 + b^2 + c^2 = 1$$

there are only two direction parameters that are linearly independent. In addition, the coordinate z_0 of a point on the straight line can be derived from the corresponding x_0 and y_0 coordinates. Hence, 4 independent parameters remain in order to describe a straight line in space:

1. direction vector: $(a', b', 1)$
2. point on the line: $z_0 = -a' \cdot x_0 - b' \cdot y_0$ $\qquad$ (2.141)

For numerical reasons these two criteria are only valid for straight lines which are approximately vertical (parallel to the z axis). Arbitrarily oriented straight lines must therefore first be transformed into a vertical direction.

2.3.2.1.1 Intersection of two straight lines

The intersection point of two straight lines in space only exists if both lines lie in a common plane, otherwise the lines are skew. In this case the shortest distance e between them is defined along a direction which is perpendicular to both. For two lines g_i, $i = 1 \ldots 2$, each defined by a point $P_i(x_i, y_i, z_i)$ and direction cosine a_i, b_i, c_i the shortest distance e between them is given by:

$$e = \frac{\pm \begin{vmatrix} x_1 - x_2 & y_1 - y_2 & z_1 - z_2 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix}}{\sqrt{a^2 + b^2 + c^2}} \tag{2.142}$$

where

$$a = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \qquad\qquad b = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \qquad\qquad c = \begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix}$$
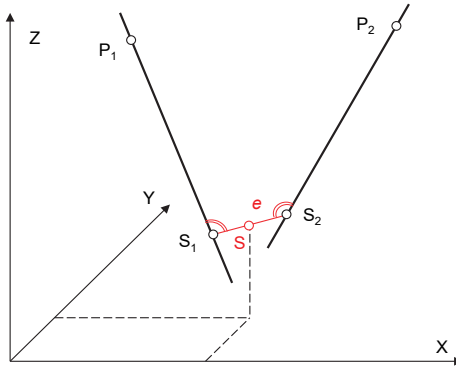
Fig. 2.28: Intersection of two straight lines in 3D space.

For consistency, the point of intersection S is then defined at half this distance, $e/2$, between both lines (Fig. 2.28). Using the factors

$$\lambda = -\frac{\begin{vmatrix} x_1 - x_2 & y_1 - y_2 & z_1 - z_2 \\ a & b & c \\ a_2 & b_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a & b & c \end{vmatrix}} \qquad\qquad \mu = -\frac{\begin{vmatrix} x_1 - x_2 & y_1 - y_2 & z_1 - z_2 \\ a & b & c \\ a_1 & b_1 & c_1 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a & b & c \end{vmatrix}}$$

the spatial coordinates of points S_1 and S_2 at the ends of the perpendicular reduce to

$$
\begin{aligned}
x_{S1} &= x_1 + \lambda \cdot a_1 \\
y_{S1} &= y_1 + \lambda \cdot b_1 \\
z_{S1} &= z_1 + \lambda \cdot c_1
\end{aligned}
\qquad\qquad
\begin{aligned}
x_{S2} &= x_2 + \mu \cdot a_2 \\
y_{S2} &= y_2 + \mu \cdot b_2 \\
z_{S2} &= z_2 + \mu \cdot c_2
\end{aligned}
$$

and hence the point of intersection S:

$$x_S = \frac{x_{S1} + x_{S2}}{2} \qquad\qquad y_S = \frac{y_{S1} + y_{S2}}{2} \qquad\qquad z_S = \frac{z_{S1} + z_{S2}}{2} \qquad (2.143)$$

The intersection angle φ between both lines is given by:

$$\cos\varphi = a_1 \cdot a_2 + b_1 \cdot b_2 + c_1 \cdot c_2 \qquad (2.144)$$

The intersection of two straight lines in space is used for spatial intersection in stereo photogrammetry (see section 4.3.6.2). Here the distance e provides a quality measure for the intersection.

2.3.2.1.2 Regression line in space

The calculation of a best-fit straight line in space can be derived directly from the algorithm presented in section 2.3.1.1. The distance of a point $P_i(x_i, y_i, z_i)$ from the straight line defined by the point $P_0(x_0, y_0, z_0)$ and the direction cosine a, b, c is given by:

$$d_i = \sqrt{u_i^2 + v_i^2 + w_i^2} \qquad (2.145)$$

where

$$u_i = c(y_i - y_0) - b(z_i - z_0)$$
$$v_i = a(z_i - z_0) - c(x_i - x_0)$$
$$w_i = b(x_i - x_0) - a(y_i - y_0)$$

The fitted line passes through P_0, the centroid of all points on the line. As in the two-dimensional case, the spatial direction of the line is defined by the eigenvector which corresponds to the largest eigenvalue of the matrix **B**:

$$\mathbf{B} = \mathbf{A}^T \cdot \mathbf{A}$$

$$\text{where} \qquad \mathbf{A} = \begin{bmatrix} x_1 - x_0 & y_1 - y_0 & z_1 - z_0 \\ \vdots & \vdots & \vdots \\ x_n - x_0 & y_n - y_0 & z_n - z_0 \end{bmatrix} \qquad (2.146)$$

2.3.2.2 Plane

2.3.2.2.1 Parameters

A plane in space is defined by $n \geq 3$ points which must not lie on a common straight line. The analytical form of a plane is given by:

$$A \cdot x + B \cdot y + C \cdot z + D = 0 \qquad (2.147)$$

A plane in 3D space is therefore analogous to a straight line in 2D (see eqn. 2.90 for comparison). The vector $\mathbf{n}(a,b,c)$ is defined as the unit vector normal to the plane with direction cosines:

$$\cos\alpha = \frac{A}{\sqrt{A^2 + B^2 + C^2}} = a$$
$$\cos\beta = \frac{B}{\sqrt{A^2 + B^2 + C^2}} = b \tag{2.148}$$
$$\cos\gamma = \frac{C}{\sqrt{A^2 + B^2 + C^2}} = c$$

Given a point $P_0(x_0,y_0,z_0)$ on the plane with normal unit vector having direction cosines (a,b,c), then all points $P(x,y,z)$ on the plane are defined by the following equation:

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \tag{2.149}$$
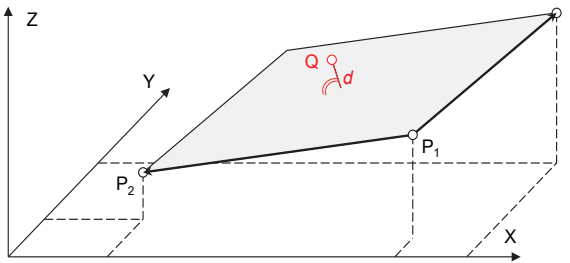
Fig. 2.29: Definition of a plane in space.

Given a plane that is formed by 3 points P_1,P_2,P_3, any other point P on the plane meets the condition (Fig. 2.29):

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0 \tag{2.150}$$

This determinant corresponds to the volume of a parallelepiped defined by its three vectors. It can be taken as a definition of the coplanarity condition used in relative orientation (see section 4.3.3.1).

The distance of a point $Q(x,y,z)$ from the plane is given by (see eqn. 2.149 for comparison):

$$d = a(x - x_0) + b(y - y_0) + c(z - z_0) \tag{2.151}$$

2.3.2.2.2 Intersection of line and plane

Given a straight line defined by point $P(x_G, y_G, z_G)$ and direction cosines (a_G, b_G, c_G)

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_G \\ y_G \\ z_G \end{bmatrix} + t \cdot \begin{bmatrix} a_G \\ b_G \\ c_G \end{bmatrix} \tag{2.152}$$

and a plane defined by point $P(x_E, y_E, z_E)$ and direction cosines (a_E, b_E, c_E):

$$a_E(x - x_E) + b_E(y - y_E) + c_E(z - z_E) = 0 \tag{2.153}$$

Substituting in (2.153) for the variable point from (2.152), the solution for line parameter t is:

$$t = \frac{a_E(x_E - x_G) + b_E(y_E - y_G) + c_E(z_E - z_G)}{a_E a_G + b_E b_G + c_E c_G} \tag{2.154}$$

The denominator becomes zero if the line is parallel to the plane. The coordinates of the point of intersection are obtained if the solution for t is substituted in (2.152).

As an example, the intersection of line and plane is used in photogrammetry for single image analysis in conjunction with object planes (monoplotting, see section 4.2.7.1).

2.3.2.2.3 Intersection of two planes

The intersection line of two non-parallel planes has a direction vector $\mathbf{a}(a,b,c)$ which is perpendicular to the unit vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ normal to the planes and can be calculated directly as the vector product of $\mathbf{n}_1$ and $\mathbf{n}_2$:

$$\mathbf{a} = \mathbf{n}_1 \times \mathbf{n}_2 \tag{2.155}$$

The magnitude of the vector product of two unit vectors is the sine of the angle θ between them. If the planes are identical, or parallel, then $\sin\theta = 0$ and the intersection line does not exist. A small value of $\sin\theta$ indicates a potentially poorly defined intersection line.

A point $\mathbf{x_0}$ on the line of intersection is defined by where it intersects a principal coordinate plane, e.g. the xy plane. The intersection line is then given in parametric form:

$$\mathbf{x} = \mathbf{x}_0 + t \cdot \mathbf{a} \tag{2.156}$$

2.3.2.2.4 Best-fit plane

In analogy with best-fitting lines, the best-fit plane is calculated by minimising the distances d_i in eqn. (2.151). The adjusted plane that fits n points is defined by the

centroid P_0 of the points and the direction cosines of the normal vector. The matrix **A**, used for the computation of eigenvalues, is identical to the matrix given in (2.146). However, here the direction cosines correspond to the eigenvector with the minimum eigenvalue.

Alternatively, a single observation equation can be formed for each point according to (2.151)

$$a(x_i - x_0) + b(y_i - y_0) + c(z_i - z_0) - d = 0 \tag{2.157}$$

which is solved within a least-squares adjustment using an additional constraint. The plane's normal vector is defined as a unit vector by setting

$$a^2 + b^2 + c^2 = 1 \tag{2.158}$$

which further defines a plane in the Hessian normal form.

2.3.2.3 Rotationally symmetric shapes

The measurement of rotationally symmetric shapes is of major importance, especially in industrial metrology. They have the common property that they can be described by a single reference axis (straight line in space) and one or more shape parameters (see Table 2.1).

Table 2.1: 3D rotationally symmetric shapes (selection).

Shape	Parameters	Degrees of freedom	Number of points
Sphere	centre point x_0,y_0,z_0 radius r	4	≥ 4
3D circle	centre point x_0,y_0,z_0 normal vector l,m,n radius r	6	≥ 3
Cylinder	axis point x_0,y_0,z_0 direction vector l,m,n radius r	5	≥ 5

These shapes (3D circle, sphere and cylinder) are often used in practical applications of close-range photogrammetry and are explained below in more detail. For the analysis of other rotationally symmetric shapes (paraboloid, ellipsoid, cone etc.) the reader is directed to further references.

2.3.2.3.1 Sphere

A sphere (Fig. 2.30) is defined by:

- the centre (x_0,y_0,z_0),
- the radius r.

Here there are 4 independent parameters which require 4 observation equations for a solution. Therefore, a minimum of 4 points must be measured on the surface of the sphere to generate these. The 4 points must not all lie on the same circle. (Any 3 will lie on a circle and the fourth point must lie off the plane of this circle.)

The equation of a sphere with centre (x_0,y_0,z_0) and radius r is given by:

$$(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2 = r^2 \tag{2.159}$$

Alternatively, a general equation for the circle is as follows:

$$x^2 + y^2 + z^2 + 2ux + 2vy + 2wz + d = 0 \tag{2.160}$$

which can be re-arranged as:

$$(x+u)^2 + (y+v)^2 + (z+w)^2 = u^2 + v^2 + w^2 - d \tag{2.161}$$

(2.161) has the same general form as (2.159) with the centre at $(-u,-v,-w)$ and radius r given by:

$$r = \sqrt{u^2 + v^2 + w^2 - d} \tag{2.162}$$
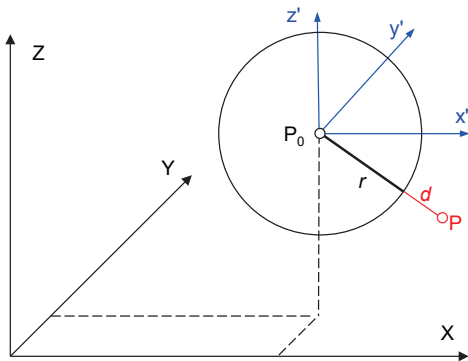
Fig. 2.30: Definition of a sphere.

Expressed in the form of (2.160), the equation is linear in the parameters u, v, w, d and can be used to compute their values directly by substituting coordinates of 4 well-chosen points to create four independent equations. From these, initial values of sphere centre and radius can be derived as indicated above.

For the over-determined case, the distance of a point from the surface of the sphere is given by:

$$d_i = r_i - r \tag{2.163}$$

where $r_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2}$

The derivatives for use in the design matrix of a least-squares analysis are then given by:

$$\frac{\partial d_i}{\partial x_0} = \frac{x_0 - x_i}{r_i} \qquad \frac{\partial d_i}{\partial y_0} = \frac{y_0 - y_i}{r_i} \qquad \frac{\partial d_i}{\partial z_0} = \frac{z_0 - z_i}{r_i} \qquad \frac{\partial d_i}{\partial r} = -1 \tag{2.164}$$

2.3.2.3.2 Circle in 3D space

A 3D circle is a circle located in an arbitrarily oriented plane in space (Fig. 2.31). It is defined by

- the centre (x_0, y_0, z_0),
- the direction cosines (a, b, c) of the normal vector to the plane,
- the radius r.

Noting that a direction has only two independent parameters, as expressed by (2.141), a 3D circle is defined by 6 independent parameters. A minimum number of 6 observations is therefore required to compute the parameters. These can be provided by 3 points on the circumference.
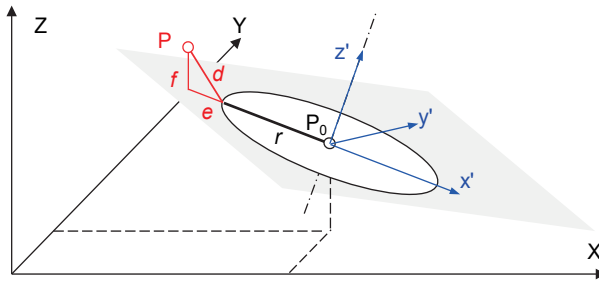
Fig. 2.31: Definition of a 3D circle.

Analogously to the best-fit sphere, a distance can be defined from a fitted point in space to the circle circumference. Here the point not only has a radial distance to the circle but also a perpendicular distance to the plane. The spatial distance is given by:

$$d_i^2 = e_i^2 + f_i^2 \tag{2.165}$$

Here e_i is the radial distance analogous to the definition in (2.107) and f_i is the distance from the plane according to (2.151). In order to calculate a best-fit circle, both components must be minimized.

$$\sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} e_i^2 + \sum_{i=1}^{n} f_i^2 \qquad (2.166)$$

Defining $c = 1$ as in (2.141), there are 6 remaining parameters x_0, y_0, z_0, a, b and r. For the special case $x_0 = y_0 = z_0 = a = b = 0$ the derivatives forming the elements of the design matrix are:

$$
\begin{aligned}
&\frac{\partial e_i}{\partial x_0} = -\frac{x_i}{r_i} & &\frac{\partial e_i}{\partial a} = -\frac{x_i z_i}{r_i} & &\frac{\partial f_i}{\partial x_0} = 0 & &\frac{\partial f_i}{\partial a} = x_i \\
&\frac{\partial e_i}{\partial y_0} = -\frac{y_i}{r_i} & &\frac{\partial e_i}{\partial b} = -\frac{y_i z_i}{r_i} & &\frac{\partial f_i}{\partial y_0} = 0 & &\frac{\partial f_i}{\partial b} = y_i \\
&\frac{\partial e_i}{\partial z_0} = -\frac{z_i}{r_i} & &\frac{\partial e_i}{\partial r} = -1 & &\frac{\partial f_i}{\partial z_0} = -1 & &\frac{\partial f_i}{\partial r} = 0
\end{aligned}
\qquad (2.167)
$$

The following procedure provides one possible algorithm for the computation of a best-fit circle with given initial values of (x_0, y_0, z_0), (a, b, c) and r:

1. Translate the data points P_i onto a local origin close to the circle centre.
2. Rotate the normal vector into an approximately vertical (z) direction.

    One method to do this requires the azimuth (α) and zenith (v) angles of the normal:

    $$\alpha = \arctan(a / b) \qquad\qquad v = \arccos(c)$$

    The rotation matrix to rotate the normal vector to the vertical is then given by:

    $$
    \mathbf{R} = \begin{bmatrix} \cos(-\alpha) & -\sin(-\alpha) & 0 \\ \sin(-\alpha) & \cos(-\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos(-v) & \sin(-v) & 0 \\ 0 & 1 & 0 \\ -\sin(-v) & 0 & \cos(-v) \end{bmatrix}^T \qquad (2.168)
    $$

    Steps 1 and 2 temporarily transform an arbitrarily oriented circle into a local system x'y'z' where the circle centre is close to the origin and the normal to the plane is vertical (see Fig. 2.31).
3. Set up and solve the normal system of equations using (2.167).
4. Correct the unknowns and reverse the transformation back into the original coordinate system.

Steps 1–4 are repeated until the unknowns do not change appreciably.

Note that discrete 3D points on the circle are not necessarily required and a circle can also be determined by photogrammetric measurement of its edges in an image (see section 4.4.7.2).

2.3.2.3.3 Cylinder

A cylinder (Fig. 2.32) is defined by
- one point (x_0, y_0, z_0) on the axis of the cylinder,
- the direction cosines (a, b, c) of the axis,
- the radius r.

Noting that a line requires only 4 independent parameters to define its location and direction, as expressed by (2.141), the position and direction of a cylinder axis is therefore defined by 4 parameters. Together with the radius, a cylinder therefore requires 5 independent parameters. A minimum of 5 observations are therefore also required to compute these parameters, i.e. a minimum of 5 points on the cylinder surface are necessary.
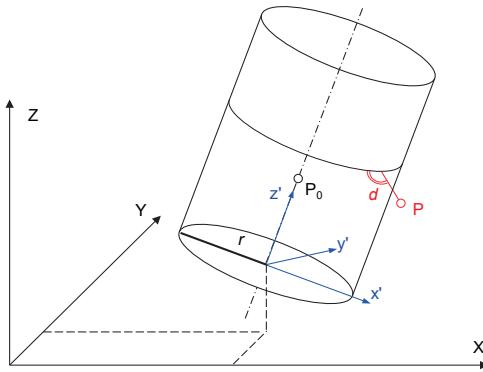
Fig. 2.32: Definition of a cylinder.

The distance of a point from the cylinder surface is given by:

$$d_i = r_i - r \tag{2.169}$$

where

$$r_i = \frac{\sqrt{u_i^2 + v_i^2 + w_i^2}}{\sqrt{a^2 + b^2 + c^2}} \quad \text{and} \quad \begin{aligned} u_i &= c(y_i - y_0) - b(z_i - z_0) \\ v_i &= a(z_i - z_0) - c(x_i - x_0) \\ w_i &= b(x_i - x_0) - a(y_i - y_0) \end{aligned}$$

In the special case where $x_0 = y_0 = a = b = 0$ the above relations simplify to:

$$r_i = \sqrt{x_i^2 + y_i^2} \tag{2.170}$$

The derivatives required to set up the design matrix **A** are given by:

$$\frac{\partial d_i}{\partial x_0} = -\frac{x_i}{r_i} \qquad\qquad \frac{\partial d_i}{\partial y_0} = -\frac{y_i}{r_i} \qquad\qquad \frac{\partial d_i}{\partial r} = -1 \qquad\qquad (2.171)$$

$$\frac{\partial d_i}{\partial a} = -\frac{x_i \cdot z_i}{r_i} \qquad\qquad \frac{\partial d_i}{\partial b} = -\frac{y_i \cdot z_i}{r_i}$$

The iterative procedure for determining the cylinder parameters is identical to the procedure given for the 3D circle (see above). This requires shifting the data points to an origin close to the cylinder axis, and then rotating this axis to be close to the vertical.

An example of the use of cylinders in close-range photogrammetry is in process plant (pipeline) modelling. Note that discrete 3D points on the cylinder surface are not necessarily required and a cylinder can also be determined by photogrammetric measurement of its edges in an image (see section 4.4.7.2).

2.3.3 Surfaces

Objects with surfaces which cannot be described by the above geometric elements are, in the first instance, usually represented by a dense distribution of 3D surface points. From these 3D point clouds, triangular mesh generation can create digital surface models of suitable detail. Analytical functions can also be used in a similar way to polynomials (see section 2.3.1.4) in order to approximate the shape of the surface.
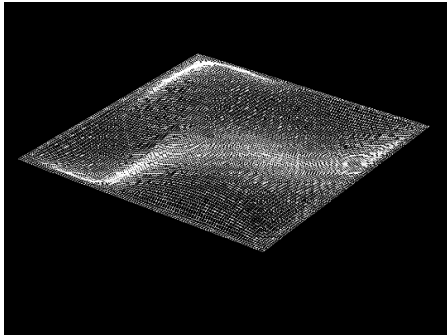
Fig. 2.33: Example of a 2½D surface.
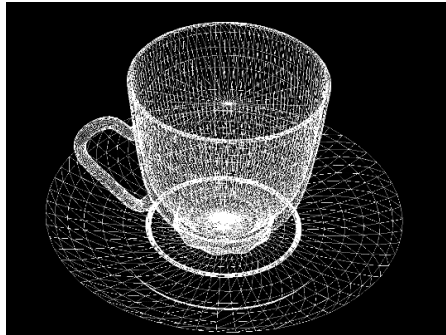
Fig. 2.34: Example of a 3D surface.

Surfaces which can be defined as a function $Z = f(X,Y)$ are known as *2½D surfaces*. Here every point on a horizontal XY plane is related to exactly one unique height value Z. Terrain models and simple component surfaces are examples of 2½D surfaces (Fig. 2.33). In contrast, objects with holes and occlusions have true *3D*

surfaces where a point on the surface is defined by a function $f(X,Y,Z) = 0$. A sculpture (Fig. 8.12) or cup with a handle (Fig. 2.34) are examples for such 3D objects.

2.3.3.1 Digital surface model

A 3D point cloud represents a *digital surface model* (DSM) if its point density (grid spacing) is sufficient for describing changes in surface shape. The point distribution can have a regular structure, e.g. $\Delta X = \Delta Y = $ const., or an irregular spacing. Object edges (breaklines) can be represented by special point codes or by additional vector-based data such as polygons.

2.3.3.1.1 Triangle meshing

The simplest way to generate a closed surface from the point cloud is by triangle meshing (Fig. 2.35), where every three adjacent 3D points combine to form a triangular surface element. Delaunay triangle meshing offers an appropriate method of creating such a triangular mesh. This identifies groups of three neighbouring points whose maximum inscribed circle does not include any other surface point.
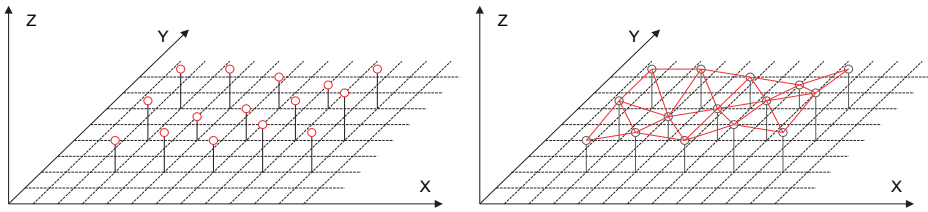
Fig. 2.35: Triangle mesh from a 3D point cloud.

Each triangle can be defined as a plane in space using eqn. (2.149) and the result is a polyhedron representation or wire-frame model of the surface.

There is normally no topological relation between the 3D points. Differential area elements must be established between adjacent points in order to generate a topologically closed surface which enables further processing as a surface description. The approximation of a surface by small planar surface elements has the advantage that it is easy to perform further calculations of, say, normal vectors or intersections with straight lines. These are required, for example, in visualization using *ray tracing* techniques or monoplotting (see section 4.2.7). If triangular elements rather than polygons are used for surface descriptions, then the planarity of surface elements is guaranteed.

Most commonly a triangular mesh is stored as a collection of triplets, where each triplet represents the three corner points of the triangle, each represented again as a triplet of its X, Y and Z coordinates. One example of a popular file format based on

this representation is the STL (stereo lithography) format used for rapid prototyping and 3D printing. It creates a block for each triangle consisting of the normal vector (to discern the inside from the outside of the object) and the vertex triplet. The content of a STL file is shown in Fig. 2.36 on the left. One problem with this format can be seen when comparing the first vertex of the two triangles stored. Obviously the format duplicates vertices which are shared by neighbouring triangles. This is not an efficient use of memory. An alternative way to represent a triangle mesh is to keep a separate list of unique vertices. The list of triangles then stores the corner points as a triplet of indices to the list of vertices. A popular file format using this scheme is the PLY format (polygon file format). Directly following the header is a list of coordinate triplets for the vertices. This is followed by a list of polygons. Each polygon starts with the number of vertices and then contains one index for each vertex. An example is provided in Fig. 2.36 on the right. From the list of vertex indices (2, 0, 1 and 2, 3, 0) it can be seen that the two triangles share two vertices (indexed 0 and 2). Thus they actually share an edge. Such formats are commonly referred to as indexed triangle meshes. A complementary form of representation centred on edges is the half-edge structure, which will not be detailed here.

```
STL ASCII file:                            PLY ASCII file:
solid example                              ply
  facet normal -0.282 0.312 0.991          format ascii 1.0
    outer loop                             element vertex 4
      vertex  -70.313 347.656 -736.759     property float x
      vertex  -70.313 345.269 -735.938     property float y
      vertex  -67.665 347.656 -735.938     property float z
    endloop                                element face 2
  endfacet                                 property list uchar int vertex_indices
  facet normal -0.849 0.172 0.500          end_header
    outer loop                             -70.313 345.269 -735.938
      vertex  -70.313 347.656 -736.759     -67.665 347.656 -735.938
      vertex  -72.130 347.656 -739.844     -70.313 347.656 -736.759
      vertex  -70.313 345.269 -735.938     -72.130 347.656 -739.844
    endloop                                3 2 0 1
  endfacet                                 3 2 3 0
```

Fig. 2.36: Two file formats storing the same triangle mesh (four vertices and two triangles).

2.3.3.1.2 Interpolation

Additional points can easily be interpolated within a given triangular element. For a tilted plane defined in a local coordinate system x'y'z', with origin located in one of the vertices of the triangle (Fig. 2.37), then the equation for the plane is given by:

$$z = a_0 + a_1 x' + a_2 y' \tag{2.172}$$

The coefficients can be calculated as follows:

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \frac{1}{x'_2\,y'_3 - x'_3\,y'_2} \begin{bmatrix} x'_2\,y'_3 - x'_3\,y'_2 & 0 & 0 \\ y'_2 - y'_3 & y'_3 & -y'_2 \\ x'_3 - x'_2 & -x'_3 & x'_2 \end{bmatrix} \cdot \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \tag{2.173}$$

For meshes defined by four points, additional points can be calculated by bilinear interpolation according to eqn. (2.19).
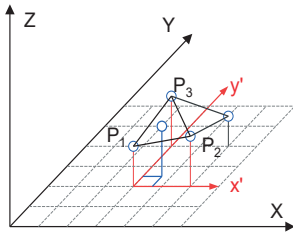
Fig. 2.37: Interpolation within a triangular mesh.
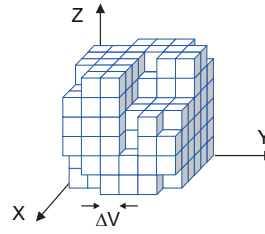
Fig. 2.38: Voxel representation.

2.3.3.2 Digital volume model (voxel)

Complex 3D objects can also be represented by sufficiently small volume elements (voxels) which are cubes of side length ΔV (Fig. 2.38). A three-dimensional matrix with column width ΔV is created and in which an attribute value is stored for every physically present object element, e.g. colour or material properties. This grid structure requires a large amount of memory but it can easily represent holes and hollow or non-connected object parts.

2.3.3.3 Range images

Point clouds acquired in raster format, e.g. by 3D cameras (see section 6.7.6), or as a result of a computation such as image matching, can be stored as range images where the intensity values correspond to the distance to the object (example in Fig. 6.67d). A *range image* in its most basic form is, analogously to eqn. (5.1):

$$R = r(x,y) \tag{2.174}$$

where the function r defines the distance from the camera to the object. While an intensity image normally contains quantized integer values, floating point numbers are usually used to represent range. In the photogrammetric context presented here, the range is defined as the distance from the projection centre of the camera to the object's surface. It should be noted that other definitions are possible, e.g. where distance is defined relative to a location on the housing of the camera.

Each pixel in the range image uniquely identifies a point in three-dimensional space. The two integer coordinate values, and the range measured at this pixel, form a triplet $(x, y, r(x,y))$. In order to obtain three-dimensional Cartesian coordinates (X,Y,Z), it is necessary to apply a projection model to the triplet, typically the pin-hole camera model (Fig. 1.7). This is very similar to transforming polar coordinates to Cartesian coordinates. To account for any geometric errors in the projection (see section 3.1.3), a full lens correction model can be applied. This is consistent with the normal workflow in photogrammetry.

Since a projection model with additional parameters can be quite complex, it can be desirable to store a representation of a range images with the projection model already applied. In this case a three-channel image is stored where the three channels contain the X, Y and Z coordinates. This representation is, for example, used in the popular ASCII exchange format PTX.

As every pixel of the image in this case holds a Cartesian coordinate triplet (X,Y,Z), this is one form of representation for a point cloud. The grid still contains the information about pixel connectivity and thus provides the neighbourhood relationship for each pixel, such as a N8 or N4 neighbourhood relationship (see section 5.1.2). This representation is therefore referred to as an *organized point cloud*. If all coordinate triplets are stored only as a list, without regard to grid structure, it is known as an *unorganized point cloud*. In this case, establishing neighbourhood relationships between points is computationally expensive and typically requires the use of a spatial search structure such as a kd-tree.

Range images can store multiple channels, e.g. with information about reflected intensity, colour or measurement uncertainty. A range image with additional true colour information is also known as an RGBD image. In principle, range images can be processed with all the usual methods of image processing (Chapter 5). Filters for image smoothing or noise reduction (section 5.2.3.2) or feature extraction methods (section 5.4.3) are commonly used.

Local curvature can be used for surface segmentation, edge extraction and feature-point detection. Either the two principal curvatures k_1 and k_2 are used directly or mean curvature H and Gaussian curvature K are derived from them:

$$H = \frac{k_1 + k_2}{2} \qquad\qquad K = (k_1 \cdot k_2) \qquad\qquad (2.175)$$

Using H and K, local surface types can be classified into eight categories as shown in Table 2.2.

Cylindrical surface types indicate the location of an edge. An alternative description of local surface shape is the shape parameter. Again using the principal curvatures k_1 and k_2, the parameter S to describe shape, and the parameter C to describe strength, are derived.

$$S = \arctan\left(\frac{k_2}{k_1}\right) \qquad\qquad C = \sqrt{k_1^2 + k_2^2} \qquad\qquad (2.176)$$

This separation of shape and strength can be seen as an analogy to the separation of colour and intensity used in alternative colour spaces (see section 5.2.2.1). The parameter C allows for the easy detection of planes (where C is smaller than a certain threshold). The parameter S can be used for the simple detection of umbilical points ($k_1 = k_2$) and minimal points ($k_1 = -k_2$).

Table 2.2: Surface types classified by mean curvature H and Gaussian curvature K.

	$K < 0$	$K = 0$	$K > 0$
$H < 0$	saddle surface (negative)	cylinder (negative)	ellipsoid (negative)
$H = 0$	minimal surface	plane	–
$H > 0$	saddle surface (positive)	cylinder (positive)	ellipsoid (positive)

2.3.3.4 B-spline and Bézier surfaces

Three-dimensional surfaces can be represented directly by a general form of the B-spline used for curves in a plane (see section 2.3.1.4). Given a three-dimensional network of $m+1$ x $n+1$ nodes (surface model, see Fig. 2.35), a B-spline surface approximation gives:

$$Q(s,t) = \begin{cases} x(s,t) = \sum\limits_{i=0}^{m}\sum\limits_{j=0}^{n} x_{ij} B_{i,\alpha}(s) B_{j,\beta}(t) \\[2mm] y(s,t) = \sum\limits_{i=0}^{m}\sum\limits_{j=0}^{n} y_{ij} B_{i,\alpha}(s) B_{j,\beta}(t) \\[2mm] z(s,t) = \sum\limits_{i=0}^{m}\sum\limits_{j=0}^{n} z_{ij} B_{i,\alpha}(s) B_{j,\beta}(t) \end{cases} \qquad (2.177)$$

where

$$0 \le s \le m - \alpha + 1 \qquad \text{and} \qquad 0 \le t \le n - \beta + 1$$

The result is a quadratic approximation when $\alpha = \beta = 2$ and a cubic spline approximation when $\alpha = \beta = 3$. The determination of the basic functions B is equivalent to the two-dimensional case.

In an analogous way, Bézier approximations can be generated for 3D elements. They are mainly used for the construction of industrial free-form surfaces, for example in the automotive industry for the representation of car body surfaces.

2.4 Adjustment techniques

2.4.1 The problem

This section provides a summary of some important techniques for the computation of over-determined, non-linear systems of equations by adjustment methods. These are essential for the understanding of numerous photogrammetric calculations. In general the task is to determine a number of unknown parameters from a number of observed (measured) values which have a functional relationship to each other. If more observations are available than required for the determination of the unknowns, there is normally no unique solution and the unknown parameters are estimated according to functional and stochastic models. See specialist literature for a more detailed explanation of adjustment methods and applications.

2.4.1.1 Functional model

A number of observations n (measured values) form an *observation vector* **L**:

$$\mathbf{L} = (L_1, L_2, ..., L_n)^\mathrm{T} \qquad\qquad \text{: observation vector} \qquad\qquad (2.178)$$

Since the elements of the observation vector are measured data they are regarded as having small random error effects but are free of systematic defects.

A number u of unknown parameters must be determined. These form the *vector of unknowns* **X**, also called the parameter vector.

$$\mathbf{X} = (X_1, X_2, ..., X_u)^\mathrm{T} \qquad\qquad \text{: vector of unknowns} \qquad\qquad (2.179)$$

The number of observations is assumed to be greater than the number of unknowns.

$$n > u$$

The *functional model* describes the relation between the "true" observation values $\tilde{\mathbf{L}}$ and the "true" values of the unknowns $\tilde{\mathbf{X}}$. This relationship is expressed by the vector of functions $\boldsymbol{\varphi}$ of the unknowns:

$$\tilde{\mathbf{L}} = \boldsymbol{\varphi}\,(\tilde{\mathbf{X}}) = \begin{bmatrix} \varphi_1(\tilde{\mathbf{X}}) \\ \varphi_2(\tilde{\mathbf{X}}) \\ \vdots \\ \varphi_n(\tilde{\mathbf{X}}) \end{bmatrix} \qquad\qquad \text{: functional model} \qquad\qquad (2.180)$$

Since the true values are normally not known, the observation vector $\tilde{\mathbf{L}}$ is replaced by the sum of the measured observations **L** and corresponding vector of residuals **v**. Similarly, the vector of unknowns is replaced by the estimated (adjusted) unknowns $\hat{\mathbf{X}}$. As a result, the following non-linear correction equations are obtained:

$$\hat{\mathbf{L}} = \mathbf{L} + \mathbf{v} = \boldsymbol{\varphi}\,(\hat{\mathbf{X}}) \tag{2.181}$$

If approximate values $\mathbf{X}^0$ of the unknowns are available, the vector of unknowns can be expressed as the following sum

$$\hat{\mathbf{X}} = \mathbf{X}^0 + \hat{\mathbf{x}} \tag{2.182}$$

i.e. only the small unknown values $\hat{\mathbf{x}}$ must be determined.

From the values in $\mathbf{X}^0$, approximate values of the observations can then be calculated using the functional model:

$$\mathbf{L}^0 = \boldsymbol{\varphi}\,(\mathbf{X}^0) \tag{2.183}$$

In this way *reduced observations* (observed minus computed) are obtained:

$$\mathbf{l} = \mathbf{L} - \mathbf{L}^0 \tag{2.184}$$

For sufficiently small values of $\hat{\mathbf{x}}$, the correction equations can be expanded into a Taylor series around the approximate values $\mathbf{X}^0$, ignoring terms after the first:

$$\begin{aligned}
\mathbf{L} + \mathbf{v} &= \boldsymbol{\varphi}\,(\mathbf{X}^0) + \left(\frac{\partial \boldsymbol{\varphi}\,(\mathbf{X})}{\partial \mathbf{X}}\right)_0 \cdot (\hat{\mathbf{X}} - \mathbf{X}^0) \\
&= \mathbf{L}^0 + \left(\frac{\partial \boldsymbol{\varphi}\,(\mathbf{X})}{\partial \mathbf{X}}\right)_0 \cdot \hat{\mathbf{x}}
\end{aligned} \tag{2.185}$$

After introduction of the *Jacobian matrix* $\mathbf{A}$, also known as the design, model or coefficient matrix

$$\mathbf{A}_{n,u} = \left(\frac{\partial \boldsymbol{\varphi}\,(\mathbf{X})}{\partial \mathbf{X}}\right)_0 = \begin{bmatrix}
\left(\dfrac{\partial \varphi_1(\mathbf{X})}{\partial X_1}\right)_0 & \left(\dfrac{\partial \varphi_1(\mathbf{X})}{\partial X_2}\right)_0 & \cdots & \left(\dfrac{\partial \varphi_1(\mathbf{X})}{\partial X_u}\right)_0 \\
\left(\dfrac{\partial \varphi_2(\mathbf{X})}{\partial X_1}\right)_0 & \left(\dfrac{\partial \varphi_2(\mathbf{X})}{\partial X_2}\right)_0 & \cdots & \left(\dfrac{\partial \varphi_2(\mathbf{X})}{\partial X_u}\right)_0 \\
\vdots & \vdots & \ddots & \vdots \\
\left(\dfrac{\partial \varphi_n(\mathbf{X})}{\partial X_1}\right)_0 & \left(\dfrac{\partial \varphi_n(\mathbf{X})}{\partial X_2}\right)_0 & \cdots & \left(\dfrac{\partial \varphi_n(\mathbf{X})}{\partial X_u}\right)_0
\end{bmatrix} \tag{2.186}$$

the linearized correction equations are obtained:

$$\hat{\mathbf{l}}_{n,1} = \mathbf{l}_{n,1} + \mathbf{v}_{n,1} = \mathbf{A}_{n,u} \cdot \hat{\mathbf{x}}_{u,1} \tag{2.187}$$

The Jacobian matrix $\mathbf{A}$ consists of derivatives which describe the functional or geometrical relation between the parameters and which are calculated from approximate values. The vector of unknowns $\hat{\mathbf{x}}$ contains the estimated parameters

and **l** is the vector of reduced observations. A computation scheme is given in section 2.4.2.2.

2.4.1.2 Stochastic model

The stochastic properties of the unknowns **L** are defined by the c*ovariance matrix* **C**$_{ll}$:

$$\mathbf{C}_{ll} \atop {}_{n,n} = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \rho_{21}\sigma_2\sigma_1 & \sigma_2^2 & \cdots & \rho_{2n}\sigma_2\sigma_n \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1}\sigma_n\sigma_1 & \cdots & \cdots & \sigma_n^2 \end{bmatrix} \tag{2.188}$$

where σ_i :   standard deviation of observation L_i, $i = 1..n$

             ρ_{ij}:   correlation coefficient between L_i and L_j, $i \neq j$

Introducing the multiplication factor σ_0^2, the c*ofactor matrix* **Q**$_{ll}$ of observations is obtained:

$$\mathbf{Q}_{ll} = \frac{1}{\sigma_0^2}\mathbf{C}_{ll} = \mathbf{P}^{-1} \tag{2.189}$$

where **P**$_{ll}$ is the weight matrix

The covariance matrix is the only component containing information about the accuracy of the functional model in the adjustment process. It is therefore called the *stochastic model*. In the case of independent observations, the correlation coefficients become zero and the covariance matrix is reduced to a diagonal matrix. This is the standard case for many adjustment problems where either independent observations are given, or no significant knowledge about correlations between observations is available.

The weight matrix **P** then becomes:

$$\mathbf{P} \atop {}_{n,n} = \begin{bmatrix} \dfrac{\sigma_0^2}{\sigma_1^2} & & & \\ & \dfrac{\sigma_0^2}{\sigma_2^2} & & \\ & & \ddots & \\ & & & \dfrac{\sigma_0^2}{\sigma_n^2} \end{bmatrix} = \begin{bmatrix} p_1 & & & \\ & p_2 & & \\ & & \ddots & \\ & & & p_n \end{bmatrix} \tag{2.190}$$

In this case an observation L_i with standard deviation $\sigma_i = \sigma_0$ has weight

$$p_i = \frac{\sigma_0^2}{\sigma_i^2} = 1 \tag{2.191}$$

and **P** becomes the identity matrix **I**. σ_0 is the true value of the *standard deviation of unit weight* (standard deviation of an observation with weight = 1). It can be regarded as a multiplication constant. Refer to sections 2.4.2.1 and 2.4.3.1 for a definition of this parameter.

Usually the true standard deviation σ is not known in practical applications and the empirical standard deviation s is used instead. Here s denotes the a priori standard deviation, while $\hat{s}$ represents the a posteriori standard deviation (adjusted standard deviation). The empirical standard deviation is only meaningful in cases of significant redundancy.

2.4.2 Least-squares method (Gauss-Markov linear model)

The Gauss-Markov adjustment model is based on the idea that the unknown parameters are estimated with maximum probability. Assuming a data set with an infinite number of measured values and normally distributed errors (non-centrality parameter $\Delta = 0$, i.e. no systematic errors), the following condition for the residuals results:

$$\mathbf{v}^T \cdot \mathbf{P} \cdot \mathbf{v} \rightarrow \min \tag{2.192}$$

For independent observations it reduces to

$$\sum_{i=1}^{n} p_i \cdot v_i^2 \rightarrow \min \tag{2.193}$$

It is known as a *least-squares adjustment* or *minimization using the L2 norm*. The Gauss-Markov model ensures that estimations of the unknown parameters are unbiased and have minimum variance.

2.4.2.1 Adjustment of direct observations

Consider a number of direct measurements of a single unknown value, e.g. from repeated measurements of the distance between two points by laser range measurement. The functional model is then reduced to the extent that the required quantity is simply the mean of the observations.

In measurements where observations are considered to be equally accurate, the weights p_i are simplified to $p_i = 1$.

For observations of varying accuracy, the corresponding weights are estimated from the a priori standard deviations of the original observations and the observation of unit weight (s_i and s_0 respectively):

$$p_i = \frac{s_0^2}{s_i^2} \qquad\qquad \text{: weight of observation } i \tag{2.194}$$

Alternatively, where measurements are considered to be equally accurate, and an improved value for a particular quantity is obtained by averaging a number of repeated measurements, then this improved average can be given a weight which corresponds to the number of measurements in the set. (A single measurement has weight 1, an average based on 6 repetitions has weight 6, etc.).

The estimated unknown is obtained by the geometric (weighted) average:

$$\hat{x} = \frac{p_1 l_1 + p_2 l_2 + \ldots + p_n l_n}{p_1 + p_2 + \ldots + p_n} = \frac{\sum_{i=1}^{n} p_i l_i}{\sum_{i=1}^{n} p_i} \tag{2.195}$$

The residual of an observation i gives:

$$v_i = \hat{x} - l_i \tag{2.196}$$

After adjustment the a posteriori standard deviation of unit weight is given by:

$$\hat{s}_0 = \sqrt{\frac{\sum p \cdot v^2}{n-1}} \tag{2.197}$$

The a posteriori standard deviation of the original observation i is given by:

$$\hat{s}_i = \frac{\hat{s}_0}{\sqrt{p_i}} \tag{2.198}$$

The standard deviation of the average value is, in this case, equal to the standard deviation of the adjusted observations:

$$\hat{s}_{\hat{x}} = \frac{\hat{s}_0}{\sqrt{\sum p_i}} \tag{2.199}$$

2.4.2.2 General least squares adjustment

Normally, values of interest must be measured indirectly. For example, photogrammetric triangulation by the intersection of measured directions produces, indirectly, the 3D coordinates of required target points. This section describes a generally applicable adjustment process.

Let the following linearized functions define an adjustment problem:

$$\underset{n,1}{\hat{\mathbf{l}}} = \underset{n,1}{\mathbf{l}} + \underset{n,1}{\mathbf{v}} = \underset{n,u}{\mathbf{A}} \cdot \underset{u,1}{\hat{\mathbf{x}}} \qquad \text{: functional model} \tag{2.200}$$

$$\mathbf{Q}_{ll} = \frac{1}{s_0^2} \mathbf{C}_{ll} = \mathbf{P}^{-1} \qquad \text{: stochastic model} \tag{2.201}$$

with n observations and u unknowns, $n>u$. To set up the weight matrix **P,** the a priori standard deviations of observations s_i, and the a priori standard deviation of unit weight s_0, are required. They could, for example, be derived from the empirically known accuracy of a measuring device:

$$p_i = \frac{s_0^2}{s_i^2}$$
: weight of observation i (2.202)

After generation of initial values, setting up of the Jacobian matrix **A,** and calculation of reduced observations **l,** the following computation scheme may be used in order to calculate the vector of unknowns $\hat{\mathbf{x}}$:

1) $\underset{n,n}{\mathbf{P}} = \underset{n,n}{\mathbf{Q}_{ll}^{-1}}$
: weight matrix (2.203)

2) $\underset{u,u}{\mathbf{N}} \cdot \underset{u,1}{\hat{\mathbf{x}}} - \underset{u,1}{\mathbf{n}} = \underset{u,1}{\mathbf{0}}$
: normal equations (2.204)

where

$\underset{u,u}{\mathbf{N}} = \underset{u,n}{\mathbf{A}^T} \cdot \underset{n,n}{\mathbf{P}} \cdot \underset{n,u}{\mathbf{A}}$
: matrix of normal equations

$\underset{u,1}{\mathbf{n}} = \underset{u,n}{\mathbf{A}^T} \cdot \underset{n,n}{\mathbf{P}} \cdot \underset{n,1}{\mathbf{l}}$
: absolute term

3) $\underset{u,u}{\mathbf{Q}} = \underset{u,u}{\mathbf{N}^{-1}}$
: solving the normal equations (2.205)

$\underset{u,1}{\hat{\mathbf{x}}} = \underset{u,u}{\mathbf{Q}} \cdot \underset{u,1}{\mathbf{n}}$

$= (\underset{u,n}{\mathbf{A}^T} \cdot \underset{n,n}{\mathbf{P}} \cdot \underset{n,u}{\mathbf{A}})^{-1} \cdot \underset{u,n}{\mathbf{A}^T} \cdot \underset{n,n}{\mathbf{P}} \cdot \underset{n,1}{\mathbf{l}}$
: where **Q**: cofactor matrix of unknowns

4) $\underset{n,1}{\mathbf{v}} = \underset{n,u}{\mathbf{A}} \cdot \underset{u,1}{\hat{\mathbf{x}}} - \underset{n,1}{\mathbf{l}}$
: residuals (2.206)

5) $\underset{n,1}{\hat{\mathbf{l}}} = \underset{n,1}{\mathbf{l}} + \underset{n,1}{\mathbf{v}}$
: adjusted observations (2.207)

$\underset{n,1}{\hat{\mathbf{L}}} = \underset{n,1}{\mathbf{L}} + \underset{n,1}{\mathbf{v}}$

6) $\underset{u,1}{\hat{\mathbf{X}}} = \underset{u,1}{\mathbf{X}^0} + \underset{u,1}{\hat{\mathbf{x}}}$
: vector of unknowns (2.208)

7) $\hat{s}_0 = \sqrt{\dfrac{\mathbf{v}^T \cdot \mathbf{P} \cdot \mathbf{v}}{n-u}}$
: standard deviation a posteriori (2.209)

8) $\underset{u,u}{\mathbf{C}} = \hat{s}_0^2 \cdot \underset{u,u}{\mathbf{Q}}$
: variance-covariance matrix (2.210)

9) $\underset{n,1}{\hat{\mathbf{L}}} \overset{!}{=} \boldsymbol{\varphi} (\underset{n,1}{\hat{\mathbf{X}}})$
: final computing test (2.211)

For most non-linear problems, initial values are only approximate and multiple iterations are required to reach an accurate solution (e.g. bundle adjustment, see section 4.4). In this case the corrected approximate values in iteration k of step (6) are used as new starting values for the linearized functional model of the next iteration $k+1$, until the sum of added corrections for the unknowns is less than a given threshold.

$$\mathbf{X}_{k+1}^0 = \mathbf{X}_k^0 + \hat{\mathbf{x}}_k \tag{2.212}$$

In order to solve the normal system of equations (2) in step (3), the Jacobian matrix $\mathbf{A}$ has to be of full column rank.

$$r = rank(\underset{n,u}{\mathbf{A}}) = u \tag{2.213}$$

This requirement means that the included observations allow a unique solution for the vector of unknowns and that the inverse of the normal equation matrix $\mathbf{N}$ exists. For adjustment problems where some observations are missing for a unique solution, a rank defect d is detected:

$$d = u - r \qquad\qquad : \text{rank defect} \tag{2.214}$$

This problem occurs, for example, in the adjustment of points in coordinate systems which are not uniquely defined by known reference points, or other suitable observations (*datum defect*).

The resulting singular system of normal equations can be solved with the help of the Moore-Penrose inverse (see section 4.4.3.4) or by including suitable constraints.

2.4.2.3 Levenberg-Marquardt algorithm

In computer vision, a bundle adjustment is often solved with the Levenberg-Marquardt algorithm (LMA). Both procedures are least-squares adjustments, but LMA offers a refinement in the form of a damping or regularization term which essentially prevents a subsequent iteration from having worse starting values than its preceding iteration.

In the Gauss-Markov model, a correction (solution) vector is calculated in the following form (compare with eqn. 2.205):

$$\hat{\mathbf{x}} = (\mathbf{A}^T \cdot \mathbf{P} \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^T \cdot \mathbf{P} \cdot \mathbf{l} \tag{2.215}$$

Using the Levenberg-Marquardt algorithm, the formulation is as follows:

$$\hat{\mathbf{x}} = (\mathbf{A}^T \cdot \mathbf{P} \cdot \mathbf{A} + \lambda \cdot diag(\mathbf{A}^T \cdot \mathbf{P} \cdot \mathbf{A}))^{-1} \cdot \mathbf{A}^T \cdot \mathbf{P} \cdot \mathbf{l} \tag{2.216}$$

Here the parameter λ regularizes the iterations. If the solution improves from one iteration to the next, then λ is reduced (often by a factor of 10) and the LM formulation becomes closely similar to the conventional formulation because the term in λ

gradually disappears. On the other hand, if the solution degrades between iterations then λ is increased, again typically by a factor of 10.

A damped iterative solution can result which may be more robust than the conventional approach.

2.4.2.4 Conditional least squares adjustment

The above method of general least squares adjustment is based on a set of observation equations that model the measured observations as a function of the unknowns. An extended adjustment model results when additional constraints are incorporated between the unknowns. This method may be called the *conditional least squares adjustment*. The following cases are examples of such constraints between unknowns (see section 4.4.2.3):

– Coordinates of a number of adjusted object points must be located on a common geometric element, e.g. a straight line, plane or cylinder.

– Two adjusted object points must have a fixed separation resulting, for example, from a high accuracy distance measurement between them.

The correction equations derived earlier are then extended by a number, r', of non-linear constraints:

$$\boldsymbol{\psi}(\tilde{\mathbf{X}}) = \begin{bmatrix} \psi_1(\tilde{\mathbf{X}}) \\ \psi_2(\tilde{\mathbf{X}}) \\ \vdots \\ \psi_{r'}(\tilde{\mathbf{X}}) \end{bmatrix} = \mathbf{0} \qquad \text{: constraints} \qquad (2.217)$$

Using approximate values, these constraint equations are linearized in an analogous way to the observation equations:

$$\underset{r',u}{\mathbf{B}} = \left(\frac{\partial \boldsymbol{\psi}(\mathbf{X})}{\partial \mathbf{X}} \right)_0 \qquad \text{: linearized constraint equations} \qquad (2.218)$$

Inconsistencies $\mathbf{w}$ result from the use of approximate values instead of expected values for the unknowns:

$$\mathbf{B} \cdot \hat{\mathbf{x}} = -\mathbf{w} \qquad \text{: vector of inconsistencies} \qquad (2.219)$$

The linearized functional model reduces to:

$$\begin{aligned} \mathbf{A} \cdot \hat{\mathbf{x}} - \mathbf{l} &= \mathbf{v} \\ \mathbf{B} \cdot \hat{\mathbf{x}} + \mathbf{w} &= \mathbf{0} \end{aligned} \qquad (2.220)$$

The Gauss-Markov model (2.192) must be extended as follows:

$$\mathbf{v}^T \cdot \mathbf{P} \cdot \mathbf{v} + 2\mathbf{k} \cdot (\mathbf{B} \cdot \hat{\mathbf{x}} + \mathbf{w}) \to \min \qquad (2.221)$$

which leads further to the following extended normal system of equations:

$$\left[\begin{array}{c|c} \mathbf{A}^T \cdot \mathbf{P} \cdot \mathbf{A} & \mathbf{B}^T \\ \hline \mathbf{B} & \mathbf{0} \end{array}\right] \cdot \left[\begin{array}{c} \hat{\mathbf{x}} \\ \hline \mathbf{k} \end{array}\right] + \left[\begin{array}{c} -\mathbf{A}^T \cdot \mathbf{P} \cdot \mathbf{l} \\ \hline \mathbf{w} \end{array}\right] = \mathbf{0} \qquad \text{: normal equations} \qquad (2.222)$$

$$\overline{\mathbf{N}} \qquad \cdot \; \overline{\mathbf{x}} \; + \qquad \overline{\mathbf{n}} \qquad = \mathbf{0}$$

Here $\mathbf{k}$ is the vector of Lagrangian multipliers. The numerical values of $\mathbf{k}$ are not normally of interest, although the condition that $\mathbf{A}^T\mathbf{Pv} + \mathbf{B}^T\mathbf{k} \equiv \mathbf{0}$ can be tested for validity. Only the first u elements of the solution vector $\overline{\mathbf{x}}$ are therefore important.

The a posteriori standard deviation is then given by:

$$\hat{s}_0 = \sqrt{\frac{\sum p \cdot v^2}{n - u + r'}} \qquad (2.223)$$

The redundancy f (degrees of freedom defined by the number of excess observations) changes to $f = n - u + r'$. Additional constraints can therefore increase redundancy or they can effectively compensate for missing observations which lead to a rank defect (see also free net adjustment, section 4.4.3.4).

2.4.3 Quality measures

Fig. 2.39 illustrates the relationship between the true value $\tilde{X}$, the expected value μ_x, the mean or adjusted value $\hat{x}$ and the single observation x_i. True value and expected value can differ due to systematic errors Δ_x. The true deviation η_i is the sum of a systematic component Δ_x and a random component ε_i.
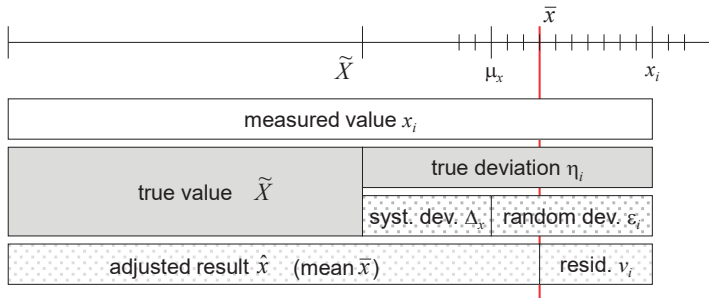
Fig. 2.39: True, stochastic and systematic deviation and residual (after Möser et al. 2000).

Since true value and expected value are unknown with a finite number of measurements, quality assessment of measured values is based on their residuals v_i. The quality values discussed below are based on statistical measures. Depending on application, the quality of a measured value, such as the fit between cylinder and

bore, must potentially be assessed by taking into account relevant associated conditions (see section 7.2).

2.4.3.1 Accuracy measures

2.4.3.1.1 Precision and accuracy

The accuracy of observations and adjusted unknowns are of prime interest when analysing quality in an adjustment procedure. The calculated stochastic values provide information about the quality of the functional model with respect to the input data. This criterion is referred to as *precision* since it describes an internal quality of the adjustment process. In contrast, the term *accuracy* should only be used if a comparison to reference data of higher accuracy is performed. However, in practice accuracy is widely used as a general term for quality.

Assuming that all observations are only affected by normally distributed random noise, hence no systematic effects exist, precision and accuracy are theoretically equal. However, in practical photogrammetry systematic effects are normally always present, for example due to illumination effects, temperature changes or imaging instabilities which have not been modelled. Since these effects cannot always be detected in an adjustment result, calculated standard deviations such as s_0 can provide an indicator of the achieved quality level but do not replace a comparison against an independent reference system or nominal values of higher accuracy (more details in sections 4.4.5 and 7.2).

2.4.3.1.2 Standard deviation

The precision of the observations and adjusted unknowns is a major contributor to the quality analysis of an adjustment result. Using the cofactor matrix **Q** or the covariance matrix **C** (see section 2.4.1.2), the standard deviations of unknowns can be obtained:

$$\mathbf{Q}_{\hat{x}\hat{x}} = \mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1u} \\ q_{21} & q_{22} & \cdots & q_{2u} \\ \vdots & \vdots & \ddots & \vdots \\ q_{u1} & q_{u2} & \cdots & q_{uu} \end{bmatrix} \qquad \text{: cofactor matrix of unknowns} \qquad (2.224)$$

The cofactor matrix of adjusted observations is derived from **Q** and the design matrix **A** as follows:

$$\mathbf{Q}_{\tilde{l}\tilde{l}} = \mathbf{A} \cdot \mathbf{Q} \cdot \mathbf{A}^T \qquad \qquad \text{: cofactor matrix of} \qquad (2.225)$$
$$\text{adjusted observations}$$

The a posteriori (empirical) *standard deviation of unit* weight is given by:

$$\hat{s}_0 = \sqrt{\frac{\mathbf{v}^T \cdot \mathbf{P} \cdot \mathbf{v}}{n-u}} \qquad (2.226)$$

with redundancy $r = n-u$

If the a posteriori standard deviation $\hat{s}_0$ diverges from the a priori standard deviation s_0, two possible sources of error are indicated. Firstly, the stochastic model may be set up incorrectly, although it should be noted that s_0 does not affect the numerical values of the adjusted unknowns. Secondly, the functional model may be incomplete. For example, unmodelled systematic errors will affect the values of the unknowns.

According to (2.189) and (2.224) the standard deviation of a single unknown x_j is given by

$$\hat{s}_j = \hat{s}_0 \cdot \sqrt{q_{jj}} \qquad (2.227)$$

where q_{jj} are the elements of the principal diagonal of matrix $\mathbf{Q}$.

2.4.3.1.3 Root mean square

In many cases, adjustment results are reported as root mean square errors instead of the above defined standard deviation. An RMS value (*root mean square*) is simply the square root of the arithmetic mean of the squares of a set of numbers X_i, $i=1...n$.

$$RMS = \sqrt{\frac{\sum X_i^2}{n}} \qquad (2.228)$$

Typically, a root mean square error is used (RMS error or RMSE) in which the numbers represent a set of differences or changes which are of some particular interest. This could perhaps indicate the RMS error of adjusted observations with respect to the mean of those adjusted observations or perhaps the difference between measured values of a set of points which have been optimally fitted to corresponding reference values. Examples are:

$$RMSE = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}} \qquad\qquad RMSE = \sqrt{\frac{\sum (X_{ref_i} - X_{meas_i})^2}{n}} \qquad (2.229)$$

For large n, the RMSE is equal to the empirical standard deviation. This is because the standard deviation of a simple set of error values would have the same form but use $(n-1)$ in place of n. As n becomes large, the difference between n and $(n-1)$ becomes negligible.

2.4.3.1.4 Span

The span R denotes the maximum separation between two observations of a set of measurements.

$$R = X_{max} - X_{min} \tag{2.230}$$

The span is not unbiased as the observations may contain blunders. However, it is important in metrology since, for manufacturing purposes, it may necessary that all measured values lie within particular limits (tolerance, see section 7.2.2.6). Hence, the span implicitly describes a confidence interval of 100 % probability (see section 2.4.3.2). The span can also be defined as the difference between the minimum and maximum residuals in a data set.

2.4.3.2 Confidence interval

It is generally assumed that the observations in an adjustment process have a normal (Gaussian) random error distribution. Given a normally distributed random variable l with expected value μ and standard deviation σ, the *probability density* is given by:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) \tag{2.231}$$

The error of the random variable is defined by:

$$\varepsilon = l - \mu \qquad\qquad : \text{random error} \tag{2.232}$$

This is valid for a normally distributed sample with an infinite number of sample points and an expected value defined as:

$$\mu = E\{x\} = \tilde{x} \qquad\qquad : \text{expected value (true value)} \tag{2.233}$$
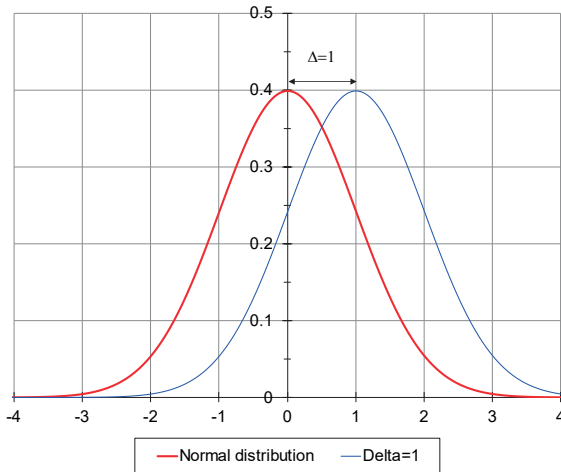
Fig. 2.40: Standardized Gaussian distribution.

Fig. 2.40 shows the probability density function of the normalized Gaussian distribution ($\mu=0$, $\sigma=1$) and, for comparison, the systematically shifted distribution corresponding to the non-centrality parameter $\Delta = 1$[6]. The area underneath the curve, between specified limits on the horizontal axis, corresponds to the probability P that the error of a random variable lies between these limits. The total area under the curve = 1 and the probability limits are usually defined as a symmetrical factor of the standard deviation.

$$P\{-k\cdot\sigma<\varepsilon<k\cdot\sigma\} \tag{2.234}$$

Table 2.3: Probability of error $|\varepsilon|$ < k·σ at different degrees of freedom.

Gaussian distribution			Student distribution			
k	P		P	P	P	P
	$f=\infty$		$f = 2$	$f = 5$	$f = 10$	$f = 20$
1	68.3%		57.7%	63.7%	65.9%	67.1%
2	95.4%		81.6%	89.8%	92.7%	94.1%
3	99.7%		90.5%	97.0%	98.7%	99.3%

Table 2.3 shows that, in the case of an infinitely large data set (degrees of freedom $f = \infty$), the probability is 68.3% that all deviations are within a single standard deviation of the true value ($k = 1$). The probability rises to 95.4% for 2 standard deviations ($k = 1.96$ for $P = 95\%$). Lastly, only 0.3% of all errors lie outside limits defined by 3 standard deviations.

In the case of large but finite data sets, the Gaussian distribution is replaced by the t-distribution (Student distribution). The probability P that a deviation is within a factor k of the standard deviation, increases with increasing degrees of freedom. For very large degrees of freedom, the t-distribution becomes equivalent to the Gaussian distribution.
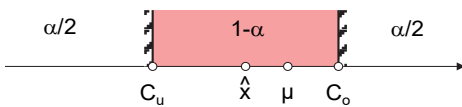
Fig. 2.41: Confidence interval.

For real (finite) data sets, only estimates $\hat{x}$ and $\hat{s}$ of the true values μ and σ can be computed. However, an interval between two limiting values C_u and C_o can be

---

6 In the following it is assumed that no systematic deviations exist, hence $\Delta = 0$.

defined, within which $\hat{x}$ is determined with probability P (Fig. 2.41). This confidence interval is given by

$$P\{C_u \leq \hat{x} \leq C_o\} = 1 - \alpha$$
$$P\{\hat{x} < C_u\} = \frac{\alpha}{2} \qquad (2.235)$$

with confidence level $1 - \alpha$.

The confidence limits for empirical estimate $\hat{x}$ with a given empirical standard deviation are defined as:

$$C_u = \hat{x} - t_{f,1-\alpha/2} \cdot s_{\hat{x}}$$
$$C_o = \hat{x} + t_{f,1-\alpha/2} \cdot s_{\hat{x}} \qquad (2.236)$$
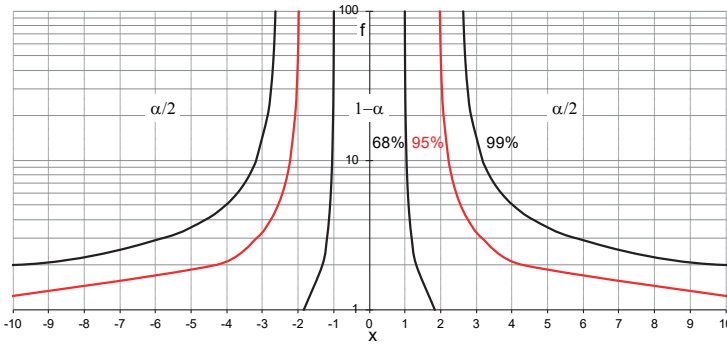
Fig. 2.42: Confidence intervals with different t-distributions.

Here t is a quantile of the t-distribution. For example, $t_{5,\,0.975} = 2.57$ corresponds to a confidence level of 95% ($\alpha = 0.05$) and $f = 5$ degrees of freedom. The confidence interval therefore increases with finite number of excess measurements, i.e. the confidence that estimate $\hat{x}$ lies between defined limits is reduced. Fig. 2.42 shows the limiting curves of confidence intervals for different degrees of freedom and different confidence levels.

2.4.3.3 Correlations
In addition to standard deviations, dependencies between adjusted parameters can also be investigated in order to assess the quality of an adjustment result. They indicate the extent to which an unknown can be calculated and hence the adequacy of the functional model and geometric configuration of the observations.

According to (2.188) the covariance matrix provides the correlations between single parameters:

$$\mathbf{C}_{\hat{x}\hat{x}} = \hat{s}_0^2 \cdot \mathbf{Q}_{\hat{x}\hat{x}} = \begin{bmatrix} \hat{s}_1^2 & \rho_{12}\hat{s}_1\hat{s}_2 & \cdots & \rho_{1u}\hat{s}_1\hat{s}_u \\ \rho_{21}\hat{s}_2\hat{s}_1 & \hat{s}_2^2 & \cdots & \rho_{2u}\hat{s}_2\hat{s}_u \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{u1}\hat{s}_i\hat{s}_1 & \rho_{u2}\hat{s}_i\hat{s}_2 & \cdots & \hat{s}_u^2 \end{bmatrix} \tag{2.237}$$

The correlation coefficient ρ_{ij} between two unknowns i and j is defined by[7]:

$$\rho_{ij} = \frac{\hat{s}_{ij}}{\hat{s}_i \cdot \hat{s}_j} \qquad\qquad -1 \le \rho_{ij} \le +1 \tag{2.238}$$

Higher correlation coefficients indicate linear dependencies between parameters. Possible reasons for this are over-parametrisation of the functional model, physical correlation within the measuring device or a weak geometric configuration. They should be avoided particularly because the inversion of the normal equation matrix, and hence the adjustment solution, can then become numerically unstable. In highly correlated solutions, adjusted parameters cannot be interpreted independently and accuracy values are often too optimistic.

2.4.3.4 Reliability
The *reliability* of an adjustment process indicates the potential to control the consistency of the observations and the adjustment model. It depends on the number of excess observations (total redundancy) and the geometric configuration (configuration of images). Reliability gives a measure of how well gross errors (outliers) can be detected in the set of observations.

Essential information about reliability can be derived from the cofactor matrix of residuals:

$$\mathbf{Q}_{vv} = \mathbf{Q}_{ll} - \mathbf{A} \cdot \mathbf{Q} \cdot \mathbf{A}^T \qquad\qquad \text{: cofactor matrix of residuals} \tag{2.239}$$

The total redundancy in an adjustment is given by

$$r = n - u = \text{trace}(\mathbf{Q}_{vv} \cdot \mathbf{P}) = \sum r_i \tag{2.240}$$

where r_i are the elements of the principal diagonal of the redundancy matrix $\mathbf{R} = \mathbf{Q}_{vv}\,\mathbf{P}$:

---

[7] Here the notation ρ is used for the empirical correlation coefficient in order to avoid confusion with the redundancy number.

$$\mathbf{R} = \mathbf{Q}_{vv} \cdot \mathbf{P} = \begin{bmatrix} r_{11} & & & \\ & r_{22} & & \\ & & \ddots & \\ & & & r_{nn} \end{bmatrix} \qquad : \text{redundancy matrix} \qquad (2.241)$$

r_i is denoted as the *redundancy number* of an observation l_i with respect to the total redundancy r where

$$0 \le r_i \le 1$$

The redundancy number of an observation indicates the relative part of an observation which is significantly used for the estimation of the unknowns $(1-r_i)$, or which is not used (r_i). Small redundancy numbers correspond to weak configurations which are hard to control, whilst high redundancy numbers enable a significant control of observations. If an observation has a redundancy number $r_i = 0$, it cannot be controlled by other observations. Hence, a gross error in this observation cannot be detected but it has a direct influence on the estimation of unknowns. If an observation has a very high redundancy number (0.8 to 1), it is very well controlled by other observations. When optimising an adjustment, such observations can initially be eliminated without a significant effect on the adjustment result.

The relation between residuals and observations is defined by:

$$\mathbf{v} = \mathbf{A} \cdot \hat{\mathbf{x}} - \mathbf{l} = -\mathbf{R} \cdot \mathbf{l} \qquad (2.242)$$

Hence, for gross (systematic) observation errors $\Delta\mathbf{l}$:

$$\Delta\mathbf{v} = -\mathbf{R} \cdot \Delta\mathbf{l} \qquad (2.243)$$

Eqn. (2.243) does permit the detection of gross errors to be quantified because gross errors do not have correspondingly large residuals when redundancy numbers are small. According to Baarda, a normalized residual is therefore used:

$$w_i = \frac{v_i}{\hat{s}_{v_i}} \qquad (2.244)$$

The standard deviation of a residual is derived either from the cofactor matrix or redundancy numbers as follows:

$$\hat{s}_{v_i} = \hat{s}_0 \sqrt{(\mathbf{Q}_{vv})_{ii}} = \hat{s}_{l_i} \sqrt{(\mathbf{Q}_{vv} \cdot \mathbf{P})_{ii}} = \hat{s}_{l_i} \sqrt{r_i} \qquad (2.245)$$

Here it is obvious that a redundancy number of $r_i = 0$ leads to an indeterminate value of w_i and no error detection is then possible. The normalized residuals are normally distributed with expectation 0 and standard deviation 1. To detect a gross error, they are compared with a threshold k:

$$|w_i| \quad \begin{cases} > k : \text{gross error} \\ \leq k : \text{no gross error} \end{cases} \tag{2.246}$$

In order to compute the threshold value k, a statistical test is used where the value δ_0 (non-centrality parameter) is defined:

$$\delta_0 = \delta_0 \ (\alpha, \beta)$$

where α : probability of identifying an *error-free* value as a gross error
        (significance level)

   β : probability of identifying a *defective* value as a gross error
        (statistical power)

This test establishes a null hypothesis which states that only randomly distributed errors may occur and the expected value of the normalized residuals must therefore be zero:

$$E\{w_0\} = 0 \tag{2.247}$$

The probability of a false decision is equal to α (type 1 error). This is the decision that residuals w_i lie outside the range $\pm k$ and are therefore excluded. Here k denotes a quantile of the t-distribution.

If a gross error occurs, the expected value of the corresponding normalized residual is not equal to zero and has standard deviation 1:

$$E\{w_a\} \neq 0 \tag{2.248}$$

Given the alternative hypothesis that only observations where $|w_i| > k$ are identified as gross errors, a possible number of outliers still remain in the data set. The probability of this false decision (type 2 error) is $1-\beta$ (Fig. 2.43).

Using (2.243) and (2.244), a lower expected value can be defined for a gross error that can be detected significantly with statistical power, β.

$$E\{\Delta l_a\} = \frac{\delta_0}{\sqrt{r_i}} \cdot \hat{s}_{l_i} = \delta'_0 \cdot \hat{s}_{l_i} \tag{2.249}$$

The following term is normalized with respect to $\hat{s}_{l_i}$:

$$\delta'_{0,i} = \frac{E\{\Delta l_a\}}{\hat{s}_{l_i}} = \frac{\delta_0}{\sqrt{r_i}} \tag{2.250}$$

It serves as a measure of *internal reliability* of an adjustment system. It defines the factor by which a gross observation error Δl_a must be larger than $\hat{s}_{l_i}$ in order to be detected with probability β.
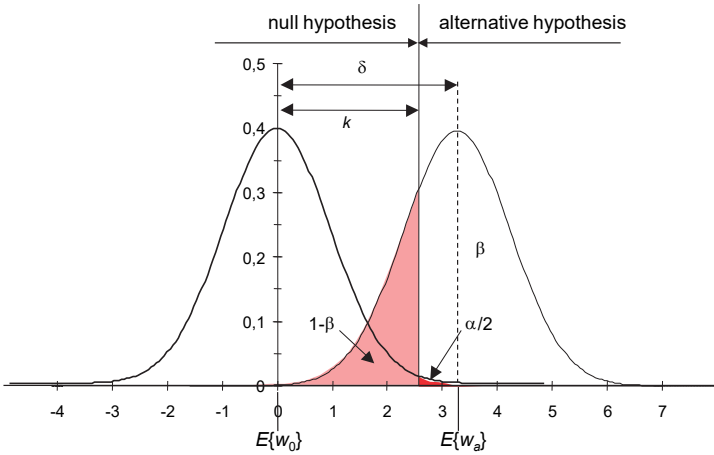
Fig. 2.43: Probability densities for null and alternative hypothesis.

Table 2.4: Test statistics and internal reliability for different significance numbers and statistical power.

	$\alpha=5\%$	$\alpha=1\%$	$\alpha=0.1\%$
	$\beta=75\%$	$\beta=93\%$	$\beta=80\%$
k	1.96	2.56	3.29
δ'_0	3.9	4.0	4.1

Table 2.4 shows some test statistics resulting from a chosen significance level α and statistical power β. The selected probabilities lead to similar measures of the *internal reliability*. It becomes clear that with increasing significance level (i.e. decreasing confidence level) the statistical power is reduced and is therefore less effective.

In general, a value of $\delta'_0 = 4$ is appropriate for photogrammetric bundle adjustment (see section 4.4). It is also an appropriate value for the decision threshold k in inequality (2.246). In order to use this standard value for bundle adjustments, a high redundancy is required. For other photogrammetric adjustment tasks, such as spatial intersection and space resection, the value of k should be modified, typically to the range $2.5 \le k \le 4$. Eqn. (2.250) clearly shows that the test value depends on the a posteriori standard deviation of a single observation and hence is a function of the standard deviation of unit weight $\hat{s}_0$. Smaller gross errors can therefore only be detected when $\hat{s}_0$ becomes sufficiently small after iteration that its value is of the order of the precision of the observations (measurement precision).

When planning a measuring project, the internal reliability can be calculated prior to knowledge of actual observations (measurement values) because the necessary information can be obtained from the Jacobian matrix **A** and the (assumed) a priori accuracy values for the observations (see section 7.1.5).

During or after an iterative adjustment, the internal reliability is used as criterion for the automated elimination of gross data errors.

The *external reliability* indicates the influence of defective observations on the estimated unknowns. For this purpose, the vector of internal reliability values, defined as in (2.249), is used in the system of equations, defined in (2.205). For each unknown it is possible to compute a total number of n values of external reliability, each dependent on an observation.

2.4.3.5 Precision of calculated coordinates

The variances and covariances of a 3D point P(X,Y,Z) are calculated by least-squares adjustment and can be found in the cofactor matrix where the values correspond to the following symbolic form:

$$\mathbf{\Sigma}_{XYZ} = \hat{s}_0^2 \cdot \mathbf{Q}_{XYZ} = \begin{bmatrix} \hat{s}_X^2 & \rho_{XY}\hat{s}_X\hat{s}_Y & \rho_{XZ}\hat{s}_X\hat{s}_Z \\ \rho_{XY}\hat{s}_X\hat{s}_Y & \hat{s}_Y^2 & \rho_{YZ}\hat{s}_Y\hat{s}_Z \\ \rho_{XZ}\hat{s}_X\hat{s}_Z & \rho_{YZ}\hat{s}_Y\hat{s}_Z & \hat{s}_Z^2 \end{bmatrix} \tag{2.251}$$

The mean point error (after Helmert) is given by:

$$\hat{s}_{XYZ} = \sqrt{\hat{s}_X^2 + \hat{s}_Y^2 + \hat{s}_Z^2} \tag{2.252}$$

This figure represents the standard error (1-sigma level) s covering all coordinate directions. The individual standard deviations in X, Y and Z depend on the definition and origin of the coordinate system. They represent a quality measure which takes no account of covariances or correlations between them. To take those also into account, a confidence ellipsoid must be calculated with the lengths and directions of its semi-axes given by an eigenvalue analysis of $\mathbf{Q}_{XYZ}$.

A spectral analysis of $\mathbf{Q}_{XYZ}$ into its eigenvalues λ_i and eigenvectors $\mathbf{s}_i$, generates an error or confidence ellipsoid within which the "true" point lies with a probability $1-\alpha$.

$$\mathbf{C}_{\hat{x}\hat{x}} = \begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \mathbf{s}_3 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} \mathbf{s}_1^T \\ \mathbf{s}_2^T \\ \mathbf{s}_3^T \end{bmatrix} \qquad : \text{spectral analysis} \tag{2.253}$$

From the eigenvalues sorted in descending order λ_i, $i=1...3$, the semi-axes a_i of the Helmert error ellipsoids become

$$a_i^2 = \hat{s}_0^2 \cdot \lambda_i \tag{2.254}$$

and the semi-axes A_i of the confidence ellipsoids:

$$A_i^2 = \hat{s}_0^2 \cdot \lambda_i \cdot F_{u,r,1-\alpha} \tag{2.255}$$

The quantile of the F-distribution defines the confidence interval for a probability $p = 1-\alpha$ with u degrees of freedom and redundancy r (here $u = 3$ for a 3D point). The directions of the semi-axes are given by the corresponding eigenvectors. The lengths of the semi-axes are independent of the selected coordinate system (compare with section 4.4.7.1).

Taking the example of a 2D point, Fig. 2.44 illustrates the relationship between single standard deviations and resulting confidence ellipses for different probabilities and redundancies. The higher the probability that the point position corresponds to the true value, the larger are the semi-axes and area of the ellipse. It is also clear that neither the mean point error (blue circle) nor the Helmert confidence ellipse (red ellipse) indicate realistic accuracy measures for coordinate components that have different accuracies and correlations. With increasing redundancy (black: $r = 5$; green: $r = 15$) the ellipses become smaller. See Fig. 4.82 for an example of a three-dimensional confidence ellipsoid.
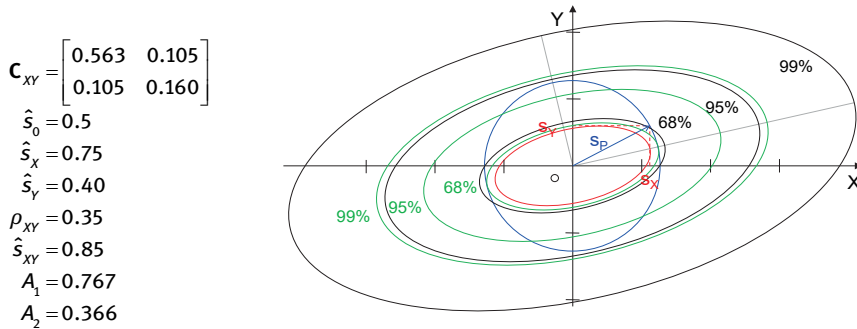
Fig. 2.44: Confidence ellipses for a 2D point ($u = 2$);
red: Helmert error ellipse; blue: mean point error; black: $r = 5$; green: $r = 15$.

2.4.4 Error detection in practice

It is difficult to avoid gross data errors (outliers) in real projects. In photogrammetric applications they typically occur as a result of faulty measurements, errors in point identification, or mistakes in image numbering. Gross errors must be eliminated from the data set because they affect all estimated unknowns and standard deviations, leading to a significantly distorted adjustment result. In summary, the following

errors in observations may occur[8], in addition to random errors and unknown systematic deviations:

– Assignment errors:
   Typical errors in this group are false point identifications, mismatches of image numbers or image coordinates measured in only one image. They must be detected and eliminated by plausibility checks (sanity checks) and data checks prior to actual calculation.

– Totally wrong measurement data:
   These are erroneous measurement values which can result from incorrect readings, operational errors or system failures and are typically in the order of (20–50) σ or more. They can affect statistical outlier tests, so that genuine outliers can no longer be detected. They must also be identified using plausibility checks.

– Gross errors (outliers):
   These deviations fall in the range from (3–5) σ up to (20–50) σ, i.e. they may occur with a reasonable probability. In photogrammetry they can caused, for instance, by defective matching of features or by false measurements of partially occluded targets.

– Rounding errors:
   Deviations caused by numerical rounding errors, a limited number of decimal places in a data display or imported data file, as well as numerical computational limits, all have a negative effect on calculated precision values, rather than adjusted mean values, if the redundancy is poor.

The residuals calculated in the adjustment should not be used directly for the detection of outliers. Residuals not only result from errors in the set of observations but also from errors in the geometric model, i.e. a functional model of the adjustment which is incomplete. Model and data errors can both be present and their effects may overlap.

Most approaches for detection and elimination of gross errors are based on the assumption that only very few outliers exist, and in extreme cases only one. The method of least-squares adjustment described above optimally disperses the observation errors over all observations in the data set, with larger deviations affecting the unknowns more than smaller. Where gross errors are present this results in a smearing effect. The ability to recognize a gross error is therefore limited, especially if several such outliers occur at the same time. Where there is an unfavourable geometric configuration of unknowns and number of outliers, even error-free observations may be identified as having gross errors. It is therefore critically important to eliminate only those observations that can be identified

---

8 After Niemeier (2008).

without doubt as errors. The elimination of outliers should always be associated with an analysis of the entire measurement task.

2.4.4.1 Error detection without adjustment

If an adjustment process does not converge it may be reasonable to check the consistency of the original observations with respect to their initial values. Here the "residuals", as defined by:

$$\mathbf{v} = \mathbf{L} - \mathbf{L}^0 \tag{2.256}$$

may indicated large discrepancies between measurement data and initial values, for example due to mistakes in point or image identification. However, a discrepancy here may simply be due to bad initial values and not necessarily to faulty measurements.

2.4.4.2 Data snooping

Baarda's *data snooping* is a method of error detection based on the value of internal reliability derived in section 2.4.3.4. It is based on the assumption that only *one* gross error can be identified in the data set at any time. The process iteratively searches for and eliminates gross errors. After each iteration of the adjustment that observation is eliminated which, on the basis of the decision function (2.246), corresponds to the largest normalized residual w_i. The complete adjustment procedure is set up again and the computation repeated until no gross errors remain in the set of observations.

In cases where several large residuals w_i exist, and where their geometric configuration ensures they are independent of each other, it is possible to detect more than one outlier simultaneously. However, one should still carefully check those observations which are suspected as gross errors.

2.4.4.3 Variance component estimation

The internal reliability value used in data snooping is a function of the standard deviation of the adjusted observations. These are derived from the covariance matrix by multiplication with the standard deviation of unit weight $\hat{s}_0$. Since $\hat{s}_0$ is a global value influenced by all residuals, it is really only useful for observations of equal accuracy. Data sets with different types of observations, or different levels of accuracy, should therefore be divided into separate groups with homogeneous accuracy.

In order to set up the weight matrix $\mathbf{P}$, each separate observation group g is assigned its own a priori variance:

$$s_{0,g}^2 \qquad\qquad\qquad \text{: a priori variance of unit weight}$$

This variance, for example, can be derived from the existing known accuracy of a measuring device used for that specific observation group.

After computing the adjustment, the a posteriori variance can be determined:

$$\hat{s}_{0,g}^2 = \frac{\mathbf{v}_g^T \mathbf{P}_g \mathbf{v}_g}{r_g} \qquad\qquad \text{: a posteriori variance of unit weight}$$

where $r_g = \sum (r_i)_g$ (2.257)

Using the a posteriori variance of unit weight, it is possible to adjust the a priori weights in succeeding adjustments until the following condition is achieved:

$$Q_g = \frac{\hat{s}_{0,g}}{s_{0,g}} = 1 \qquad\qquad\qquad (2.258)$$

Subsequently, the normal data snooping method can be used.

Taking eqn. (2.189) into account, the variance of unit weight can be used to calculate the variance of a complete observation group.

2.4.4.4 Robust estimation with weighting functions

The comments above indicate that the residuals resulting from an adjustment process are not directly suitable for the detection of one or more gross errors. Different approaches have therefore been developed for defining the weights p_i as a function of the residuals in successive iterations. If the weighting function is designed such that the influence of a gross error is reduced as the error becomes larger, then it is referred to as *robust estimation* (robust adjustment). One possible approach is given by the following function:

$$p'_i = p_i \cdot \frac{1}{1 + \left(a \cdot |v_i|\right)^b} \qquad\qquad (2.259)$$

For $v_i = 0$ it reduces to $p'_i = 1$ and for $v_i = \infty$ it reduces to $p'_i = 0$. The parameters a and b form the curve of a bell-shaped weighting function. With

$$a_i = \frac{\sqrt{p_i}}{\sqrt{r_i} \cdot \hat{s}_0 \cdot k} = \frac{1}{\hat{s}_{v_i} \cdot k} \qquad\qquad (2.260)$$

the parameter a is controlled by the redundancy number of an observation. Definition of parameters b and k is done empirically. With a correct choice of parameters, the quotient $\hat{s}_0 / s_0$ converges to 1.

Kruck proposes a weighting function that is also based on redundancy numbers:

$$p'_i = p_i \cdot \tan\left(\frac{1-r_m}{c}\right) \Big/ \tan\left(\frac{1-r_i}{c}\right) \qquad \text{where } r_m = \frac{n-u}{n} \tag{2.261}$$

The constant c is defined empirically, r_m is referred to as the *average redundancy* or *constraint density*. The weights become constant values when the stability $r_i = r_m$ is reached (*balanced observations*).

Procedures for robust estimation are primarily designed to reduce the effect of *leverage points*. Leverage points, in the sense of adjustment, are those observations which have a significant geometric meaning but only small redundancy numbers. Gross errors at leverage points affect the complete result but are difficult to detect. Using balanced weights, leverage points and observations with gross errors are temporarily assigned the same redundancy numbers as every other observation. As a result, they can be detected more reliably. After elimination of all defective observations, a final least-squares adjustment is calculated using the original weights.

2.4.4.5 Robust estimation according to L1 norm

In recent times more attention has been paid to the principle of adjustment according to the L1 norm, especially for gross error detection in weak geometric configurations. For example, it is used to calculate approximate orientation values using data sets containing gross errors.

The L1 approach is based on the minimization of the absolute values of the residuals, whilst the L2 approach (least squares, see section 2.4.2) minimizes the sum of squares of the residuals:

L1 norm: $\qquad \sum |v| \qquad \to \min$ (2.262)

L2 norm: $\qquad \sum v^2 \qquad \to \min$

The solution of the system of equations using the L1 approach is a task in linear optimization. It is much more difficult to handle than the L2 approach in terms of mathematics, error theory and computation algorithms. One solution is given by the Simplex algorithm known from linear programming.

The L1 approach is also suitable for balancing weights according to section 2.4.4.4. In theory it is possible with the L1 norm to process data sets with up to 50 % gross errors. The reason is that the L1 solution uses the median value whereas the L2 solution is based on the arithmetic mean which has a smearing effect.

After error elimination based on the L1 norm, the final adjustment should be calculated according to the least-squares approach.

2.4.4.6 RANSAC

RANSAC (random sample consensus) describes an adjustment algorithm for any functional model which is based on a voting scheme and is particularly robust in the presence of outliers, e.g. even as high as 80 %. The idea is based on the repeated calculation of a target function by using the minimum number of observations $n=u$. These are randomly selected from the full set of observations. Subsequently, all other observations are tested against this particular solution. All observations which are consistent with the calculated solution within a certain tolerance are regarded as valid measurements and form a consensus set. The solution with the maximum number of valid observations (largest consensus set) is taken as the best result. All observations are marked as outliers if they do not belong to this consensus group. Outliers are rejected and an optional final least-squares adjustment is calculated.

Fig. 2.45 shows the RANSAC principle applied to the calculation of a best-fit circle with a set of valid observations and outliers. For each of the two samples A and B, a circle is calculated from three randomly selected points. The other observations are tested against the circle tolerance d. Out of all samples, the solution with the maximum number of valid observations is finally selected (sample A in this case). Remaining outliers are rejected from the data set.
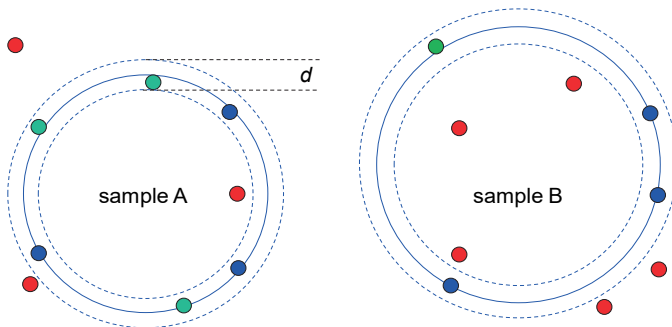
Fig. 2.45: RANSAC approach for a best-fit circle;
blue: selected RANSAC points; red: detected outliers; green: valid points.

The success of the RANSAC algorithm depends mainly on the selection of tolerance parameter d and the termination criteria, e.g. the number of iterations or the minimum size of the consensus set.

RANSAC is widely used in photogrammetry and computer vision for solving tasks such as relative orientation (section 4.3.3), feature detection in point clouds or general shape fitting of geometric primitives.

2.4.5 Computational aspects

2.4.5.1 Linearization
In order to linearize the functional model at initial values, two methods are available:
– exact calculation of the first derivative;
– numerical differentiation.

Exact calculation of the derivatives

$$\left(\frac{\partial \varphi(X)}{\partial X} \right)_0 \tag{2.263}$$

may require considerably more programming effort for complex functions such as those expressed by rotation matrices.

In contrast, numerical differentiation is based on small changes to the initial values of unknowns in order to calculate their effects on the observations:

$$\begin{aligned} L^0_{+\Delta x} &= \varphi(X^0 + \Delta X) \\ L^0_{-\Delta x} &= \varphi(X^0 - \Delta X) \end{aligned} \tag{2.264}$$

The difference quotients are then:

$$\frac{\Delta \varphi(X)}{\Delta X} = \frac{L^0_{+\Delta x} - L^0_{-\Delta x}}{2\Delta X} \approx \frac{\partial \varphi(X)}{\partial X} \tag{2.265}$$

In a computer program, the function φ can be entered directly, for example in a separate routine. The set-up of the Jacobian matrix **A** and subsequent adjustment procedure can then be programmed independently of the functional model. Only the increment ΔX need be adjusted if necessary.

Compared with numerical differentiation, the exact calculation of derivatives leads to faster convergence. If suitable initial values are available, then after a number of iterations both adjustment results are, for practical purposes, identical. Modern programming libraries for solving adjustment and optimisation tasks, such as Ceres-Solver, SciPy.Optimize or Matlab Optimization Toolbox, offer interfaces to perform these tasks automatically and efficiently.

2.4.5.2 Normal systems of equations
In order to solve solely for the solution vector $\hat{\mathbf{x}}$, efficient decomposition algorithms can be used which do not require the inverse of the normal equation matrix **N**, for example the Gaussian algorithm. However, for many photogrammetric and geodetic calculations, a quality analysis based on the covariance matrix is required and so the inverse of **N** must be computed. The dimension of the normal system of equations based on (2.204) is u x u elements. For photogrammetric bundle adjustments, the number of unknowns u can easily range from a few hundred up to a few tens of

thousand. Often a direct inversion of the normal equation matrix is not possible or else consumes too much computation time.

If the (symmetric) matrix **N** is positive definite (*trace*(**N**)>0), it can be factorized, for example, according to Cholesky . The triangular rearrangement

$$\mathbf{L} \cdot \mathbf{L}^T = \mathbf{N} \tag{2.266}$$

inserted into (2.204) yields

$$\mathbf{L} \cdot \mathbf{L}^T \cdot \hat{\mathbf{x}} = \mathbf{n} \tag{2.267}$$

With the forward substitution

$$\mathbf{L} \cdot \mathbf{g} = \mathbf{n} \tag{2.268}$$

and the subsequent backward substitution

$$\mathbf{L}^T \cdot \hat{\mathbf{x}} = \mathbf{g} \tag{2.269}$$

both **g** and $\hat{\mathbf{x}}$ can be computed by suitable solution methods, some without the need for inversion.

2.4.5.3 Sparse matrix techniques and optimization

The computational effort to solve the normal system of equations is mainly a function of the dimension of matrix **C**. Since matrix **N** can consist of numerous zero elements, relating to unknowns which are not connected by an observation, then these elements are also present in **C**.

Sparse techniques provide efficient use of RAM (Random-Access Memory). Instead of a full matrix, a profile is stored which, for each column (or row) of a matrix, only records elements from the first non-zero value up to the principal diagonal, together with a corresponding index value.

Fig. 2.46 shows an example of a network of observations and the corresponding structure of the normal equation matrix. The crosses indicate connections between unknowns while the blank fields have zero values. For example, point 2 is connected to points 1, 3, 6, 7 and 8.

In this example, the size of the profile to be stored, i.e. the number of stored matrix elements, amounts to P = 43. In order to reduce the profile size without modifying the functional relationships, the point order can be sorted (Banker's algorithm), leading to the result of Fig. 2.46 right. The profile size in this case has been reduced to P = 31.
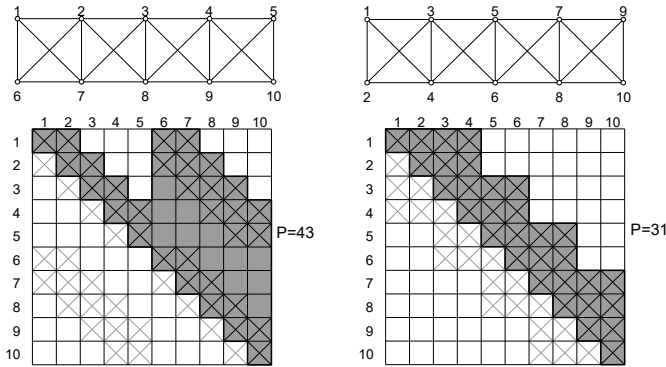
Fig. 2.46: Order of point numbers without (left) and with (right) optimization (after Kruck 1983).

Since the resulting computational effort for solving equations is a quadratic function of the profile size, optimization is of major importance for solving large systems of equations. Further information on the solution of large systems of equations, and on bundle adjustment, can be found in the relevant literature (section 9.2.2).

3 Imaging technology

Photogrammetric imaging technologies for close-range measurement purposes impact upon all elements in the measurement process, from the preparation of the measuring object prior to imaging, through image acquisition, to subsequent analysis of the image content. Following an introduction to the physics behind optical imaging, issues including distortion, resolution and sampling theory are discussed. Common photogrammetric imaging concepts are briefly presented such as online and offline approaches and imaging configurations. The key part of this chapter deals with the geometric analysis defining the camera as a measuring tool, i.e. photogrammetric camera modelling, parameters of interior orientation and correction of imaging errors. Current components and sensor technologies for 2D image acquisition are then reviewed in the sections which follow. From there, discussion moves to signalization (targeting), light projection and illumination, which are critical in achieving photogrammetric accuracy.

3.1 Physics of image formation

3.1.1 Wave optics

3.1.1.1 Electro-magnetic spectrum

In photogrammetry, the usable part of the electromagnetic spectrum (Fig. 3.1) is principally restricted to wavelengths in the visible and infra-red regions. This is due to the spectral sensitivity of the imaging sensors normally employed, such as photographic emulsions, and silicon-based CCD and CMOS sensors which respond to wavelengths in the range 380 nm to 1100 nm (visible to near infrared). In special applications, X-rays (X-ray photogrammetry), ultraviolet and longer wavelength thermal radiation (thermography) are also used. However, microwaves (radar) are generally confined to remote sensing from aircraft and satellite platforms. Fig. 3.1 summarizes the principal spectral regions, with associated sensors and applications, which are relevant to photogrammetry.

The relationship between wavelength λ, frequency υ and speed of propagation c is given by:

$$\lambda = \frac{c}{v} \tag{3.1}$$

The propagation of electromagnetic radiation is described using either a wave propagation model or a photon stream. Both models have relevance in photogrammetry. The wave properties of light are employed in the description of optical imaging and its aberrations, as well as refraction and diffraction. The particle

properties of light are useful in understanding the transformation of light energy into electrical energy in image sensors (CCD, CMOS), see also section 3.4.1.

Radiation	Wavelength	Sensor	Application area
X-ray	10 pm – 1 nm	X-ray detector	medical diagnosis, non-destructive testing
ultraviolet	300 nm – 380 nm	CCD, CMOS, silicon (Si) detector	remote sensing, UV reflectography
visible ligth	380 nm – 720 nm	CCD, CMOS, film	photography, photogrammetry, remote sensing
near infrared 1	720 nm – 1 µm	CCD, CMOS, film	IR reflectography, photogrammetry, remote sensing
near infrared 2	1 µm – 3 µm	InGaAs detectors	NIR reflectography, material classif., remote sensing
mid infrared	3 µm – 5 µm	Ge-, PbSe-, PbS detectors	remote sensing
thermal infrared	8 µm – 14 µm	micro-bolometer, quantum detector	thermography, material testing, building energy efficiency
terahertz	30 µm – 3 mm	Golay cell	spectroscopy, body scanning
microwave	1 cm – 30 cm	radar antenna	radar remote sensing
radio wave	30 cm – 10 m	coil	magnetic resonance imaging (MRI)

Fig. 3.1: Electromagnetic spectrum with example application areas.

3.1.1.2 Radiometry

According to quantum theory, all radiation is composed of quanta of energy (photons in the case of light). The radiant energy is a whole multiple of the energy in a single quantum of radiation which is related to the reciprocal of the photon's associated wavelength according to the following equation:

$$E = h \cdot v = h \cdot \frac{c}{\lambda} \tag{3.2}$$

where

h:　Planck's constant $6.62 \cdot 10^{-34}$ Js

The spectral emission of a black body at absolute temperature T is defined by Planck's law:

$$M_\lambda = \frac{c_1}{\lambda^5}\left[\exp\left(\frac{c_2}{\lambda \cdot T}\right) - 1\right]^{-1} \qquad : \text{Planck's law} \tag{3.3}$$

where

$c_1 = 3.7418 \cdot 10^{-16}$ W m^2

$c_2 = 1.4388 \cdot 10^{-2}$ K m

This states that the radiant power is dependent only on wavelength and temperature. Fig. 3.2 shows this relationship for typical black-body temperatures. Radiant power per unit area of the emitting source and per unit solid angle in the direction of

emission is defined as radiance. The area under the curves represents the total energy in Watts per m². The example of the sun at a temperature of 5800 K clearly shows the maximum radiant power at a wavelength around 580 nm, which is in the yellow part of the visible spectrum. In contrast, a body at room temperature (20° C) radiates with a maximum at a wavelength around 10 μm and a very much smaller power.

The radiant power maxima are shifted towards shorter wavelengths at higher temperatures according to Wien's displacement law (see straight line in Fig. 3.2).

$$\lambda_{max} = 2897.8 \cdot T^{-1} \qquad \text{: Wien's displacement law} \qquad (3.4)$$
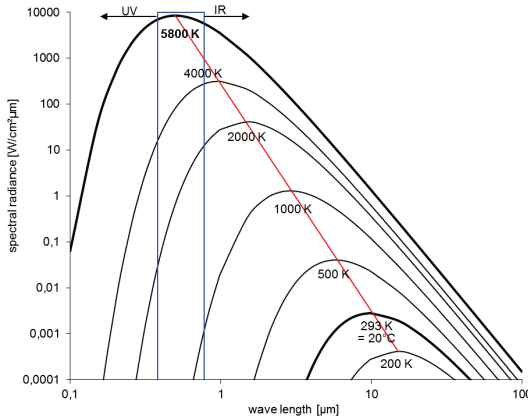
Fig. 3.2: Spectrally dependent radiant power at different black-body temperatures.

3.1.1.3 Refraction and reflection

The *refractive index n* is defined as the ratio of the velocities of light propagation through two different media (frequency is invariant):

$$n = \frac{c_1}{c_2} \qquad \text{: refractive index} \qquad (3.5)$$

In order to define refractive indices for different materials, c_1 is assigned to the velocity of propagation in a vacuum c_0 and c_{med} is the velocity of light in the medium of interest. Note that c_{med} is dependent on wavelength (colour of light):

$$n = \frac{c_0}{c_{med}} \qquad (3.6)$$

The refractive index of pure water has been determined to be $n = 1.33$ whilst for glass the value varies between 1.45 and 1.95 depending on the material constituents of the glass. In general, homogeneous and isotropic media are assumed. A ray of light

passing from a low density media to a more dense media is refracted towards the normal at the media interface *MI*, as denoted by Snell's law (Fig. 3.3, Fig. 3.4).

$$n_1 \sin \varepsilon_1 = n_2 \sin \varepsilon_2 \qquad \text{: Snell's law of refraction} \qquad (3.7)$$

The law can also be expressed as a function of the tangent:

$$\frac{\tan \varepsilon_1}{\tan \varepsilon_2} = \sqrt{n^2 + (n^2 - 1)\tan^2 \varepsilon_1} \qquad \text{where } n = \frac{n_2}{n_1} \text{ and } n_2 > n_1 \qquad (3.8)$$

Fig. 3.3: Refraction and reflection.

a) Still water surface   b) Disturbed water surface
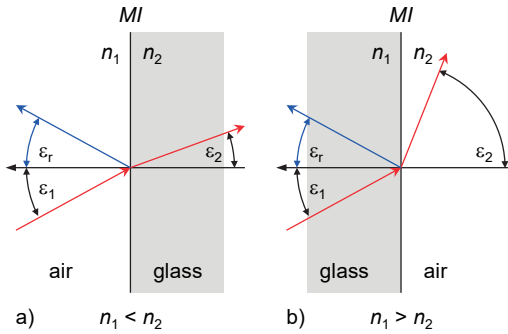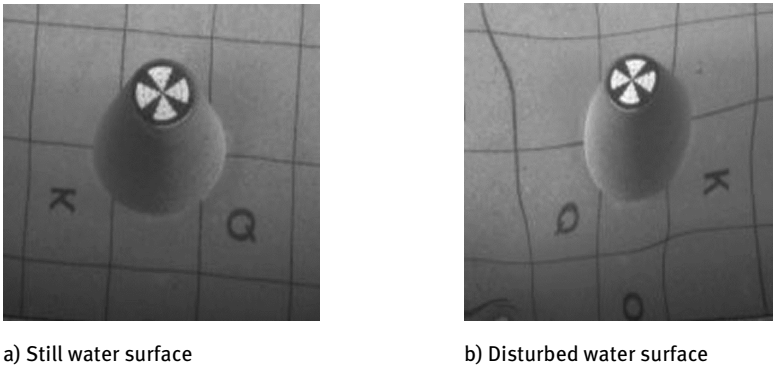
Fig. 3.4: Refraction effects caused by taking an image through a water surface.

For the case of *reflection* it holds that $n_2 = -n_1$ and the law of reflection follows:

$$\varepsilon_r = -\varepsilon_1 \qquad \text{: law of reflection} \qquad (3.9)$$

As mentioned, the velocity of propagation of light depends on the wave length. The resulting change of refraction is denoted as *dispersion* (Fig. 3.5). In an optical imaging

system this means that different wavelengths from the object are refracted at slightly different angles which lead to chromatic aberration in the image (see section 3.1.3.2).
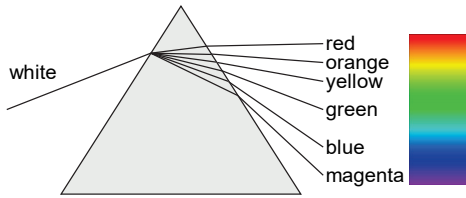
Fig. 3.5: Dispersion by a glass prism.

3.1.1.4 Diffraction

Diffraction occurs if the straight-line propagation of spherical wave fronts of light is disturbed, e.g. by passing through a slit (linear aperture) or circular aperture. The edges of the aperture can be considered as multiple point sources of light that also propagate spherically and interfere with one another to create maxima and minima (Fig. 3.6, Fig. 3.7).

The intensity *I* observed for a phase angle φ is given by:

$$I(\varphi) = \frac{\sin x}{x} = \text{sinc}(x) \tag{3.10}$$

where $\quad x = \dfrac{\pi d' \sin \varphi}{\lambda}$

$\quad\quad$ *d'*: $\quad$ slit width
$\quad\quad$ λ: $\quad$ wavelength
$\quad\quad$ φ: $\quad$ phase angle
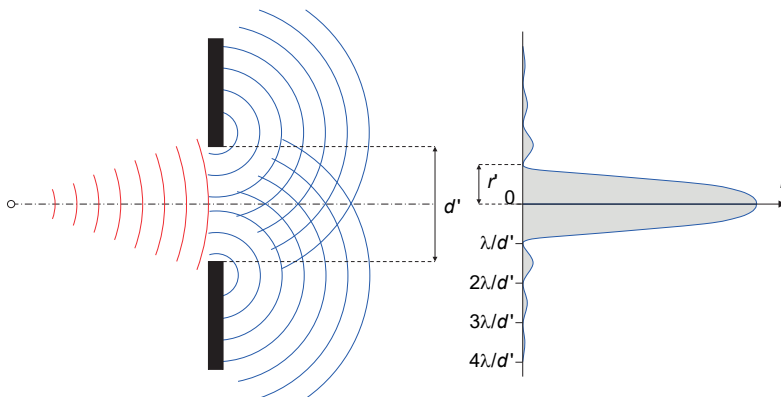
Fig. 3.6: Diffraction caused by a slit.
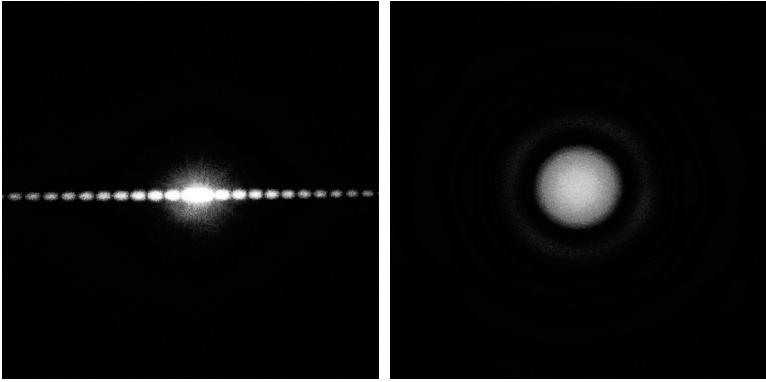
Fig. 3.7: Diffraction at a slit (left) and circular aperture (right).

The intensity becomes a minimum for the values:

$$\sin\varphi_n = n\frac{\lambda}{d'} \qquad\qquad n = 1, 2, \dots \qquad\qquad (3.11)$$

Diffraction at a circular aperture results in a diffraction pattern, known as an Airy disc, with concentric lines of interference. Bessel functions are used to calculate the phase angle of the maximum intensity of the diffraction disc which leads to:

$$\sin\varphi_n = b_n\frac{2\lambda}{d'} \qquad\qquad b_1 = 0.61;\ b_2 = 1.12;\ b_3 = 1.62 \dots \qquad (3.12)$$

The radius r of a diffraction disc at distance f' and $n = 1$ is given by:

$$r = 1.22\frac{\lambda}{d'}f' = 1.22 \cdot \lambda \cdot k \qquad\qquad\qquad (3.13)$$

where
f':   focal length
$k = f'/d'$: f/number
d':   diameter of aperture

Diffraction not only occurs at limiting circular edges such as those defined by apertures or lens mountings but also at straight edges and (regular) grid structures such as the arrangement of sensor elements on imaging sensors (see section 3.4.1). Diffraction-limited resolution in optical systems is discussed further in section 3.1.5.1.

In conjunction with deviations of the lens equation (3.14), (defocusing) diffraction yields the *point spread function* (PSF). This effect is dependent on wavelength and described by the contrast or modulation transfer function (see sections 3.1.5.3 and 3.1.6).

3.1.1.5 Polarization

Natural light has the property of propagating in the form of a transverse wave, i.e. the direction of oscillation is orthogonal to the direction of propagation. An infinite number of oscillation directions with arbitrary (unordered) wavelengths, amplitudes and phases occur circularly around the propagation vector, i.e. the light is not polarized. If the oscillation only occurs in one direction, the light has a linear polarization (Fig. 3.8a). This can be achieved, for example, when light is transmitted through a polarizer, whose transparent layer consists of parallel strings of macromolecules. The projection of this oscillation in two mutually perpendicular planes generates a sinusoidal oscillation with the same phase position.

In contrast, circularly polarized light has two perpendicular components with equal amplitudes but a phase difference of 90° (Fig. 3.8b). The direction of oscillation therefore rotates as the wave propagates.

a) Linear polarization　　　　　　b) Circular polarization
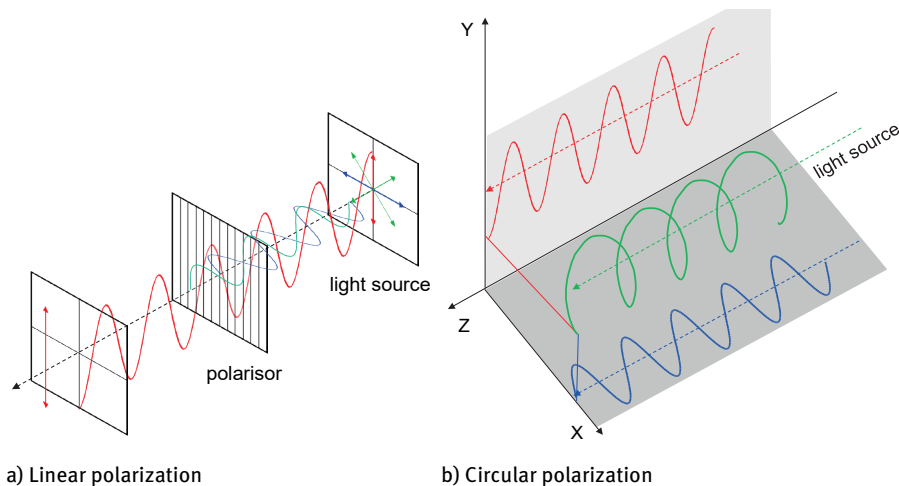
Fig. 3.8: Polarization.

Light is partly polarized by refraction, reflection or scattering. At a smooth surface, the reflected light is polarized, hence disruptive reflections in an image (hot spots) can be reduced by suitable use of polarizing filters (see section 3.4.4). Polarizing filters can be mounted both in front of the imaging lens, as well as in front of the light source. The polarization effect is also applied to the design of stereo monitors (see section 6.12.1).

Fig. 3.9 shows an image with specular reflection of a light source, as well as the result with the cross-polarization technique, in which the camera and light source are each equipped with a polarizing filter that are rotated 90° to each other. The cross-polarised image is free of specular reflections and has better colour saturation.
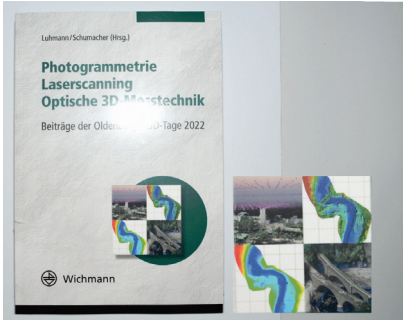
a) Bright spot due to reflective surface

b) Recording with cross-polarization filter

Fig. 3.9: Image capture with and without polarization filter.

3.1.2 Optical imaging

3.1.2.1 Geometric optics
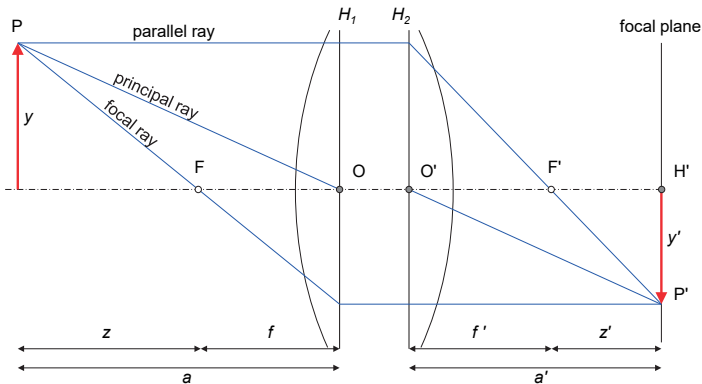
Fig. 3.10: Geometrical construction for a typical thin lens system.

The optical imaging model for thin spherical lenses is illustrated in Fig. 3.10. The well-known *thin lens equations* can be stated as follows:

$$\frac{1}{a'}+\frac{1}{a}=\frac{1}{f'} \tag{3.14}$$

$$z \cdot z' = -f'^2 \tag{3.15}$$

where

a:	object distance
a':	image distance ≈ principal distance c
f, f':	external and internal focal length
z, z':	object and image distances relative to principle foci, F and F'
H_1, H_2:	object-side and image-side principal planes
O, O':	object-side and image-side principal points

If the transmission media are different in object and image space, eqn. (3.14) is extended as follows:

$$\frac{n}{a} + \frac{n'}{a'} = \frac{n'}{f'} \qquad (3.16)$$

where

n:	refractive index in object space
n':	refractive index in image space

The *optical axis* is defined by the line joining the principal points O and O' which are the centres of the principal planes H_1 and H_2. The object and image side *nodal points* (see Fig. 3.10) are those points on the axis where an imaging ray makes the same angle with the axis in object and image space. If the refractive index is the same on the object and image sides, then the nodal points are identical with the principal points. For a centred lens system, the principal planes are parallel and the axis is perpendicular to them. The optical axis intersects the image plane at the *autocollimation point,* H'. With centred lenses and principal planes orthogonal to the axis, the autocollimation point is the point of symmetry for lens distortion and corresponds to the *principal point* in photogrammetry (see section 3.3.2).

In addition, the imaging scale or magnification is given in analogy to (3.45)[1]:

$$\beta' = \frac{y'}{y} = \frac{a'}{a} = -\frac{z'}{f'} = 1 : m \qquad (3.17)$$

or in the case of different transmission media:

$$\beta' = \frac{n}{n'} \frac{a'}{a} \qquad (3.18)$$

According to eqn. (3.14), an object point P is focused at distance a' from the image side principal plane H_2. Points at other object distances are not sharply focused (see section 3.1.2.3). For an object point at infinity, $a' = f'$. To a good approximation, image distance a' corresponds to the principal distance c in photogrammetry. m denotes image scale (magnification), as commonly used in photogrammetry.

---

1 In optics β' is used instead of $M = 1/m$.

3.1.2.2 Apertures and stops

The elements of an optical system which limit the angular sizes of the transmitted bundles of light rays (historically denoted as light pencils) can be described as *stops*. These are primarily the lens rims or mountings and the aperture and iris diaphragms themselves. Stops and diaphragms limit the extent of the incident bundles and, amongst other things, contribute to amount of transmitted light and the depth of field.

The most limiting stop is the *aperture stop,* which defines the *aperture* of a lens or imaging system. The object-side image of the aperture is termed the *entrance pupil EP* and the corresponding image-side image is known as the *exit pupil E'P.* The f/number is defined as the ratio of the image-side focal length *f'* to the diameter of the entrance pupil *d* ':

$$\text{f/number} = \frac{f'}{d'} \tag{3.19}$$

In symmetrically constructed *compound lenses*, the diameters of the pupils *E'P* and *EP* are equal. In this case *E'P* and *EP* are located in the principal planes (Fig. 3.11). In that case only, the incident and emerging rays make the same angle $\tau = \tau'$ with the optical axis. Asymmetrical lenses are produced when the component lens do not provide a symmetrical structure. For example, in wide-angle or telephoto lens designs where the asymmetrical design can place one or more of the nodes outside the physical glass boundaries or the aperture is asymmetrically positioned between the principal planes (Fig. 3.12). Such lenses are often deployed in camera systems with mirror viewing systems where the separation between the lens mount and the sensor needed to accommodate the mirror is greater than the focal length of the lens. Different angles of incidence and emergence give rise to radially symmetric lens distortion Δ*r'* (see also sections 3.1.3.1 and 3.3.2.2).
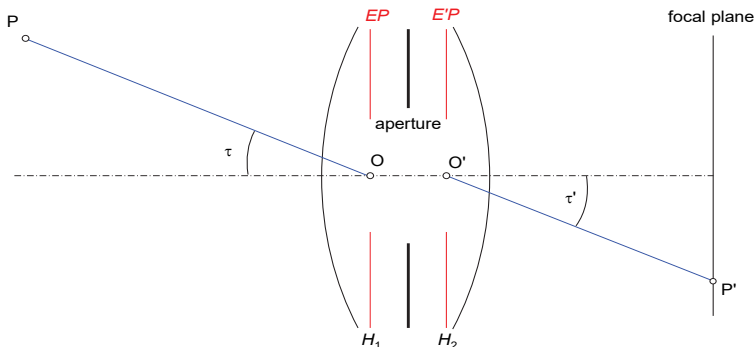
Fig. 3.11: Symmetrical compound lens.
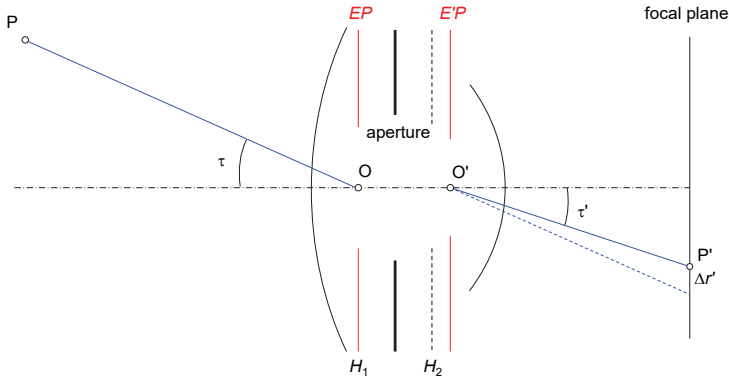
Fig. 3.12: Asymmetrical compound lens.

3.1.2.3 Focussing

In practical optical imaging, a point source of light is not focused to a point image but to a spot known as the *circle of confusion* or *blur circle*. An object point is observed by the human eye as sharply imaged if the diameter of the circle of confusion u' is under a resolution limit. For film cameras u' is normally taken as around 20–30 μm and for digital cameras as around 1–3 pixels (2–10 μm). It is also common to make the definition on the basis of the eye's visual resolution limit of $\Delta\alpha = 0.03°$. Transferring this idea to a normal-angle camera lens of focal length f' ($\approx$ image diagonal d'), a blur circle of the following diameter u' will be perceived as a sharp image:

$$u' = f' \cdot \Delta\alpha \approx \frac{1}{2000} d' \tag{3.20}$$

Example 3.1:

It is required to find the permissible diameter of the blur circle for a small-format film camera, as well as for two digital cameras with different image formats.

	Film camera	Digital camera 1	Digital camera 2
Pixels:	n/a - analogue	2560 x 1920	4368 x 2912
Pixel size:		3.5 μm	8.2 μm
Image diagonal:	43.3 mm	11.2 mm	43.0 mm
Blur circle:	22 μm	5.6 μm = 1.6 pixel	21 μm = 2.6 pixel

The blur which can be tolerated therefore becomes smaller with smaller image format. At the same time the demands on lens quality increase.

Based on the blur which can be tolerated, not only does an object point P at distance a appear to be in sharp focus but also all points between P_v and P_h (Fig. 3.13).
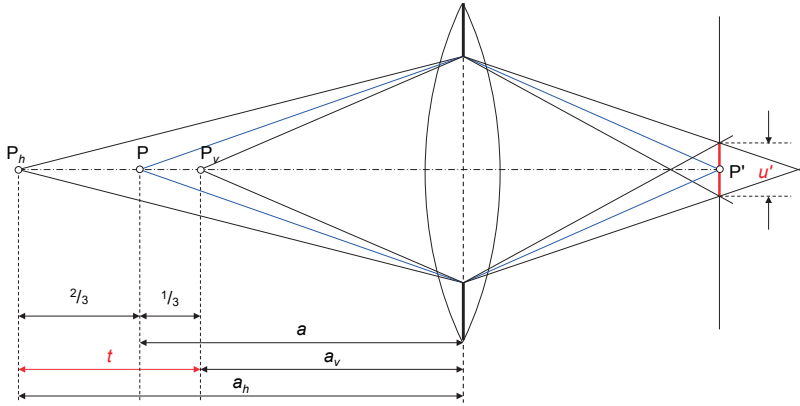
Fig. 3.13: Focusing and depth of field.

The distance to the nearest and the furthest sharply defined object point can be calculated as follows:

$$a_v = \frac{a}{1+K} \qquad\qquad a_h = \frac{a}{1-K} \qquad\qquad K = \frac{k(a-f)u'}{f^2} \qquad\qquad (3.21)$$

where

k: f/number
f: focal length
a: focused object distance

By re-arrangement of (3.21) the diameter of the blur circle u' can be calculated:

$$u' = \frac{a_h - a_v}{a_h + a_v} \cdot \frac{f^2}{k(a-f)} \qquad\qquad (3.22)$$

The *depth of field* is defined by:

$$t = a_h - a_v = \frac{2u'k(1+\beta')}{\beta'^2 - (u'k/f')^2} \qquad\qquad (3.23)$$

Hence, for a given circle of confusion diameter, depth of field depends on the f/number of the lens k and the imaging scale β'. The depth of field will increase if the aperture is reduced, the object distance is increased, or if the focal length is reduced through the use of a wider-angle lens. Fig. 3.14 shows the non-linear curve of the resulting depth of field at different scales and apertures. At very small scale numbers, the depth of field becomes extremely small, as is typical for macro-photogrammetry. Fig. 3.16a,b shows an example with different f-stop settings.
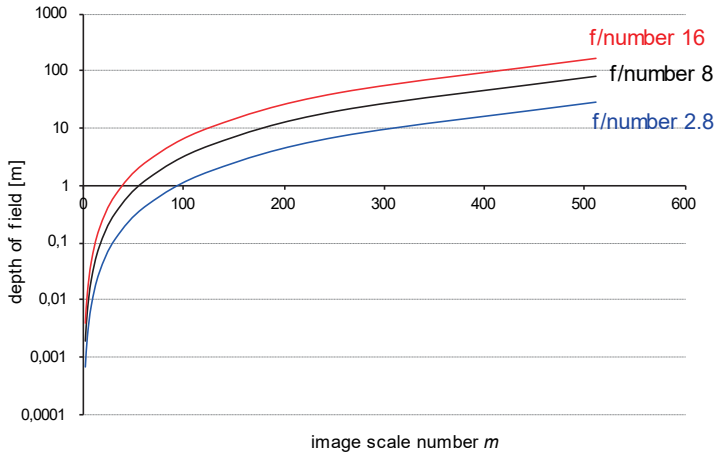
Fig. 3.14: Depth of field as a function of image scale number (u' = 20 µm).

For large object distances (3.23) can be simplified to

$$t = \frac{2u'k}{\beta'^2} \tag{3.24}$$

Example 3.2:

Given a photograph of an object at distance a = 5 m and image scale number m = 125, i.e. β' = 0.008. The depth of field for each aperture k = 2.8 and k = 11 is required.

Solution:

1. For k = 2.8 $\qquad\qquad\qquad t = \dfrac{2 \cdot u' \cdot k}{\beta'^2} = \dfrac{2 \cdot 2 \cdot 10^{-5} \cdot 2.8}{0.008^2} = 1.8$ m

2. For k = 11 $\qquad\qquad\qquad t = 6.9$ m

When imaging objects at infinity, depth of field can be optimized if the lens is not focused at infinity but to the *hyperfocal distance b*. Then the depth of field ranges from the nearest acceptably sharp point a_v to ∞:

$$b = \frac{f^2}{u'k} \qquad\qquad\qquad \text{: hyperfocal distance} \tag{3.25}$$

Sufficient depth of field must be considered especially carefully for convergent imagery, where there is variation in scale across the image, and when imaging objects with large depth variations. Depth of field can become extremely small for large image scales (small m), for example when taking images at very close ranges. Sharp focusing of obliquely imaged object planes can be achieved using the Scheimpflug condition (section 3.1.2.4).

It is worth noting that under many circumstances a slight defocusing of the image can be tolerated if radially symmetric targets are used for image measurement. On the one hand the human eye can centre a symmetrical measuring mark over a blurred circle and on the other hand the optical centroid remains unchanged for digital image measurement. Note when designing an imaging system for a particular application, the tolerable blur circle will be dependent on the features to be detected and measured. For example, a tolerable circle diameter is likely to be larger when imaging solid circular retro-target images vs applications that require the detection and measurement of natural line features.

3.1.2.4 Scheimpflug condition

For oblique imaging of a flat object, the lens eqn. (3.14) can be applied to all object points provided that the object plane, image plane and lens plane (through the lens perspective centre and orthogonal to the optical axis) are arranged such that they intersect in a common line (Fig. 3.15). This configuration is known as the Scheimpflug condition and can be realized, for example, using a *view camera* (Fig. 3.93) or special *tilt-shift lenses*, also called *perspective control lenses* (Fig. 3.105).
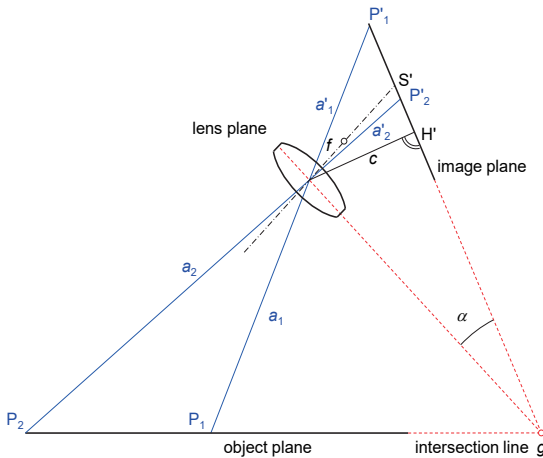
Fig. 3.15: Arrangement of object, lens and image planes according to the Scheimpflug principle.

Fig. 3.15 shows the imaging configuration which defines the Scheimpflug principle. The relative tilts of the three planes leads to the condition:

$$\frac{1}{a'_1} + \frac{1}{a_1} = \frac{1}{a'_2} + \frac{1}{a_2} \tag{3.26}$$

Compare this with eqn. (3.27) and it is clear that all points on the tilted object plane are sharply imaged. Fig. 3.16c shows an actual example of the effect.

The opposing tilts of the lens and image planes by angle α results in a separation of the principal point H' (foot of the perpendicular from perspective centre to image plane) and the point of symmetry of radial distortion S' (point of intersection of optical axis and image plane). Note that the practical maximum angles that can deliver images of appropriate quality for photogrammetric measurement will be a limited by the light transmission vs reflection characteristics of the lens elements with respect to the high incident and emergent ray angles.

a) f/number = f/3.5

b) f/number = f/16

c) f/number = f/3.5 and Scheimpflug angle α = 6.5°
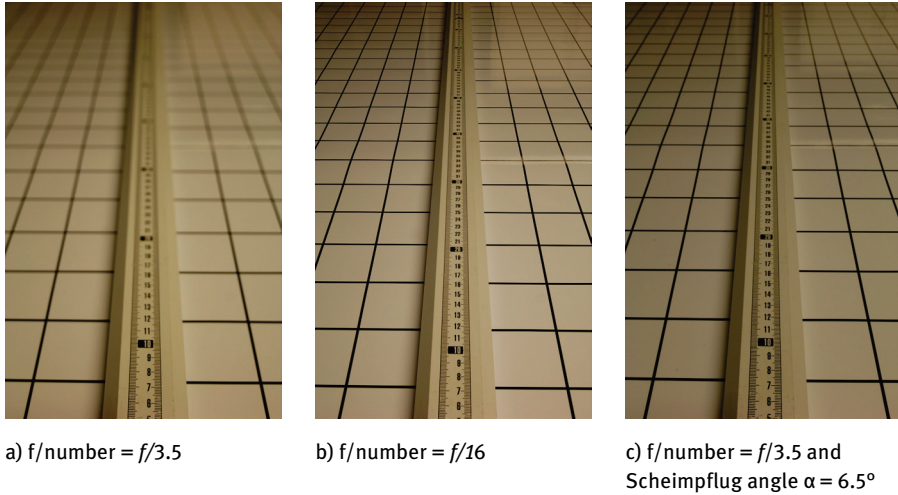
Fig. 3.16: Depth of field at different f/numbers and with the Scheimpflug condition active.

3.1.3 Aberrations

In optics, *aberrations* are deviations from an ideal imaging model, and aberrations which arise in monochromatic light are known as *geometric distortions*. Aberrations which arise through dispersion are known as *chromatic aberrations*. In photogrammetry, individual optical aberrations, with the exception of radial distortion, are not normally handled according to their cause or appearance as derived from optical principles. Instead, only their effect in the image is modelled. However, a physically-based analysis is appropriate when several aberrations combine to create complex image errors. An example would be the combination of spherical aberration, curvature of field and lateral chromatic aberration which contribute to misshapen circular retro-reflective target images at the edges of wide-angle images, see Fig. 3.20a.

3.1.3.1 Distortion

In the ideal case of Fig. 3.11 the angle of incidence τ is equal to the angle of emergence τ'. As the position of entrance pupil and exit pupil do not usually coincide with the principal planes, an incident ray enters at angle τ, and exits at a different angle τ'. This has a distorting effect Δr' in the image plane which is radially symmetric with a point of symmetry S'. In the ideal case, S' is identical with the autocollimation point H'. The sign, i.e. direction, of Δr' depends on the design combination of spherical and aspherical lens elements in the lens construction and the position of the lens aperture. For spherical lenses the distortion effect can be modelled by the Seidel series (see section 3.3.3.1).

Fig. 3.17 shows how *barrel distortion* in a spherical lens can increase when the aperture is moved towards the object such that light rays traversing the edges of the lens dominate in forming the image, and similarly *pincushion distortion* can be enhanced when the aperture is moved towards the image. When the image is free of distortion it may be called *orthoscopic*.
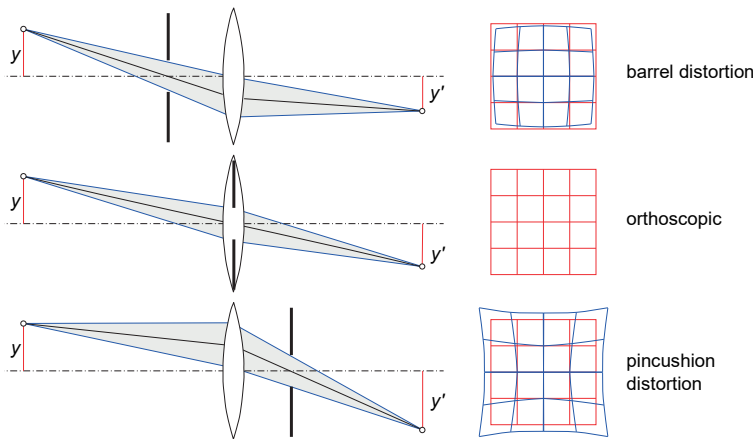
Fig. 3.17: Distortion in a spherical lens as a function of aperture position.

Radial distortion can be interpreted as a radially dependent scale change in the image. The relative distortion is defined as:

$$\frac{\Delta r'}{r'} = \frac{\Delta x'}{x'} = \frac{\Delta y'}{y'} \tag{3.27}$$

Fig. 3.18: Pincushion (left) and barrel distortion (right) in an actual image.

From the definition of image scale in (3.17) and dividing numerator and denominator in (3.27) by y, it can be seen that the relative distortion is also equivalent to the relative scale change as follows:

$$\frac{\Delta\beta'}{\beta'} = \frac{\Delta y'}{y'} = f(\tau) \tag{3.28}$$

For $\Delta y'/y' < 0$ the image is too small, i.e. the image point is shifted towards the optical axis and the result is seen as barrel distortion. Correspondingly, when $\Delta y'/y' > 0$ the result is pincushion distortion. Fig. 3.18 shows the effect in an actual image. The analytical correction of distortion is discussed in section 3.3.2.2.

3.1.3.2 Chromatic aberration

Chromatic aberration in a lens is caused by dispersion which depends on wavelength, also interpreted as colour, hence the name. *Longitudinal chromatic aberration*, also called *axial chromatic aberration*, has the consequence that every wavelength has its own focus. A white object point is therefore focused at different image distances so that an optimal focus position is not possible (see Fig. 3.19). Depending on lens quality, the effect can be reduced by using different component lenses and coatings in its construction. If the image plane is positioned for sharp imaging at mid-range wavelengths, e.g. green, then image errors will appear in the red and blue regions. In practice, the presence of chromatic aberration limits the sharpness of an image.

Lateral chromatic aberration, also called *transverse chromatic aberration*, has the consequence that an object is imaged at different scales depending on radial position in the image. For monochromatic light the effect is equivalent to radial distortion. For polychromatic light, the aberration causes a variable radial shift of colour.
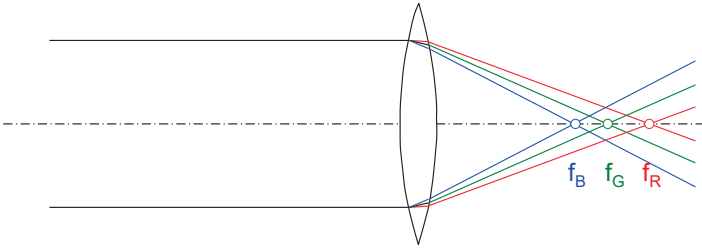
Fig. 3.19: Longitudinal chromatic aberration.

The effect can easily be seen in digital colour images. Fig. 3.20a shows white targets with poor image quality towards the edge of a colour image. The colour errors are clearly visible at the target edges, as they would occur at the edge of any other imaged object. Fig. 3.20b shows the green channel only. This has an image quality sufficient for photogrammetric point measurement. The difference image in Fig. 3.20c makes clear the difference between the red and green channels.

In black-and-white cameras chromatic aberration causes blurred edges in the image. In colour cameras, multi-coloured outlines are visible, particularly at edges with black/white transitions. In these cases, colour quality also depends on the method of colour separation (e.g. Bayer filter, see section 3.4.1.6) and the image pre-processing common in digital cameras, e.g. in the form of colour or focus correction.

a) Original RGB image   b) Green channel   c) Difference between red and green channels

Fig. 3.20: Colour shift at a black/white edge.

3.1.3.3 Spherical aberration

Spherical aberration causes rays to focus at different positions in the z' direction depending on their displacement from the optical axis. The effect is greater for off-axis object points (with bundles of rays angled to the axis).

The offset $\Delta z'$ in the longitudinal direction varies quadratically with the displacement y (Fig. 3.21). The effect can be considerably reduced by stopping down (reducing the aperture size) to the extent possible before diffraction becomes

predominant. In the image the effect causes a reduction in contrast and a softening of focus. Photogrammetrically it can be concluded that spherical aberration results in a shift of perspective centre at different field angles (off-axis locations). Good quality modern lenses optimized for the latest generation of imaging sensors, are mostly free of spherical aberration so that this source of error need not be taken into account in photogrammetric camera calibration outside of the most accurate work.
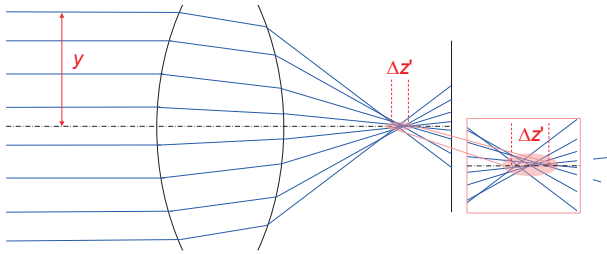
Fig. 3.21: Spherical aberration.

3.1.3.4 Astigmatism and curvature of field

Astigmatism and *curvature of field* apply to off-axis object points. They are incident at different angles to the refractive surfaces of the component lenses and have different image effects in orthogonal directions. These orthogonal directions are defined by the *meridian* and *sagittal* planes, M_1M_2 and S_1S_2 (Fig. 3.22). The meridian plane is defined by the optical axis and the *principal ray* in the imaging pencil of light, this being the ray through the centre of the entrance pupil. The two planes no longer meet at a single image point but in two image lines (Fig. 3.23). An object point therefore appears in the image as a blurred oval shape or an elliptical spot.

The curved image surfaces, which are produced as shown in Fig. 3.23, are known as *curvature of field*. If the meridian and sagittal surfaces are separated, then the effect is true astigmatism with differing curvatures through the image space. Where significant curvature is present, sensor placement within the field becomes critical in maintaining image quality across the image format.

The effects on the image of astigmatism and curvature of field are, like spherical aberration, dependant on the off-axis position of the object point (incoming ray pencils are at an angle to the axis). The imaging error can be reduced with smaller apertures and appropriately designed curvatures and combinations of component lenses. No effort is made to correct this in photogrammetry as it is assumed that lens design reduces the error below the level of measurement sensitivity. Camera manufacturer Contax has historically produced cameras with a curved photographic film image plane in order to optimize image sharpness in the presence of field curvature.
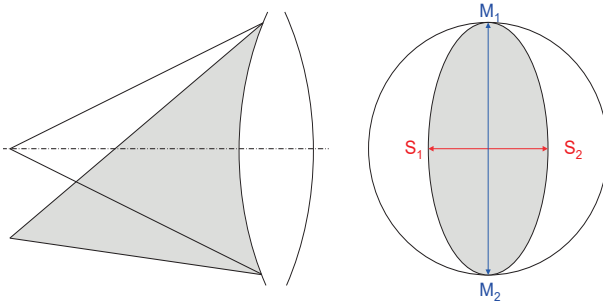
Fig. 3.22: Elliptical masking of off-axis ray bundles (after Marchesi 1985).
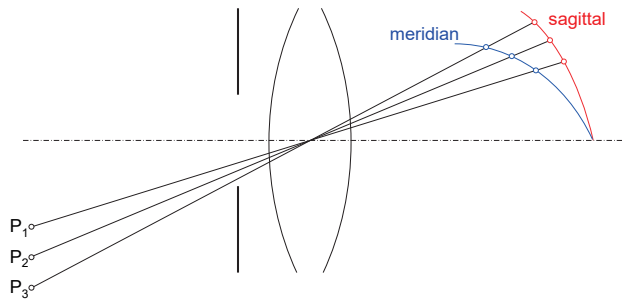
Fig. 3.23: Astigmatism (after Schröder 1990).

3.1.3.5 Light fall-off and vignetting

For conventional lenses, the luminous intensity I effective at the imaging plane is reduced with increasing field angle τ according to the $\cos^4$ law:

$$I' = I \cos^4 \tau \tag{3.29}$$

Hence the image gets darker towards its periphery (Fig. 3.24, Fig. 3.25a). The effect is particularly observable for super-wide-angle lenses where it may be necessary to use a concentrically graduated neutral density filter in the optical system to reduce the image intensity at the centre of the field of view. Fisheye lenses avoid the $\cos^4$ law through the use of different projections which reduce the fall-off in illumination at the expense of image distortions (section 3.3.7). The $\cos^4$ reduction in image intensity can be amplified if vignetting, caused by physical obstructions due to mounting parts of the lens, is taken into account (Fig. 3.25b). The fall-off in light can be compensated by analytical correction of the colour value, but at the expense of increased image noise.
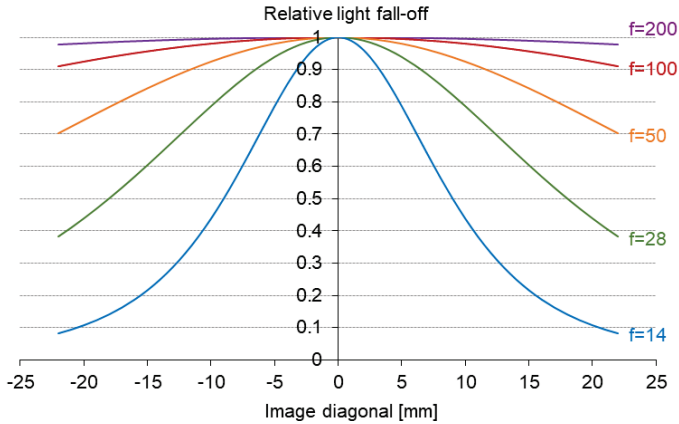
Fig. 3.24: Relative fall-off in light for small format images at different focal lengths (in mm).
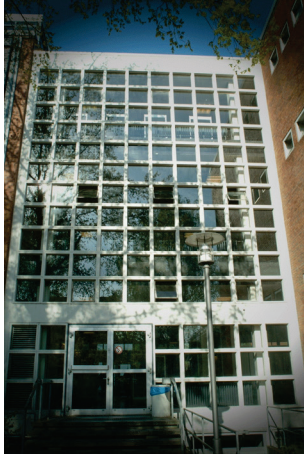
a) Light fall-off

b) Vignetting

Fig. 3.25: Images showing loss of intensity towards the edges.

3.1.4 Aspherical lenses

In the previous sections, it was implicitly assumed that lenses consist of combinations of individual spherical lenses. Improved manufacturing methods have made it cost effective for affordable lenses to deviate from the spherical shape and deliver sharper images to match increasingly smaller pixel dimensions. In general, aspherical surfaces can significantly improve lens performance in terms of sharpness, resolution and aberration compared to spherical systems.
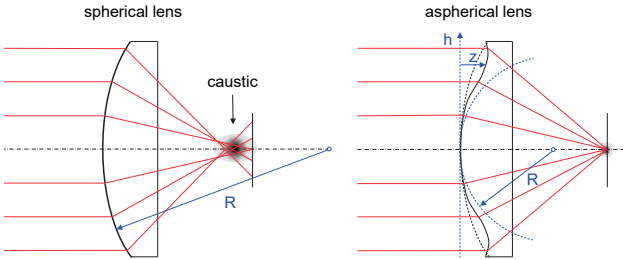
Fig. 3.26: Typical ray paths for spherical and aspherical lens.

Fig. 3.26 compares typical optical paths through a spherical and an aspherical lens element. With a spherical lens, spherical aberrations cause the image rays to intersect not at a point but in a region known as the caustic, resulting in a blurred image (see section 3.1.3.3). Aspherical lenses, on the other hand, can be constructed with a varying surface curvature optimised for example to achieve a point focus, which results in significantly sharper images.

Aspherical lenses are characterised by the fact that their surfaces can have practically any radially symmetric shape. The shape is described by the sag z as a function of the distance to the optical axis (elevation) h:

$$z(h) = \frac{\rho h^2}{1 + \sqrt{1 - (1+k)\rho h^2}} + \sum_{i=2}^{n} A_{2i} h^{2i} + \sum_{i=1}^{m} A_{2i+1} |h|^{2i+1} \qquad (3.30)$$

where

z sag
h distance to the optical axis (elevation)
ρ vertex curvature with radius $R = 1/\rho$
k conic constant (sphere: $k=0$; ellipsoid: $k>-1$; paraboloid: $k=-1$)
A_{2i} aspheric coefficients of the correction polynomial with even exponents
A_{2i+1} aspheric coefficients of the correction polynomial with odd exponents

The first summand in eqn. (3.30) describes the central spherical part of the surface. The coefficients of the correction polynomial define the aspherical curve. The odd terms describe lenses with free-form surfaces; commercially available aspherical lenses are described exclusively by the even terms.

For photogrammetry, the generally better imaging properties of aspheres are offset by the fact that, strictly speaking, the correction functions for radial-symmetrical distortion based on Seidel polynomials (section 3.3.3.1), no longer apply. The outcome in terms of radial lens distortion is specific to each lens design. A simple test is to imagine a straight line across each edge of the image and to look for variations in what would otherwise be smooth curves in a spherical lens design.

3.1.5 Resolution

3.1.5.1 Resolving power of a lens

The resolving power of a lens is limited by diffraction and aperture. For a refractive index of $n = 1$, the radius r or the diameter d of the central Airy diffraction disc is obtained from (3.13) und (3.19) as follows:

$$r = 1.22 \cdot \lambda k = 1.22 \cdot \lambda \frac{f}{d'} = d/2 \tag{3.31}$$

which corresponds to an angular resolution of

$$\delta = 1.22 \cdot \frac{\lambda}{d'} \tag{3.32}$$
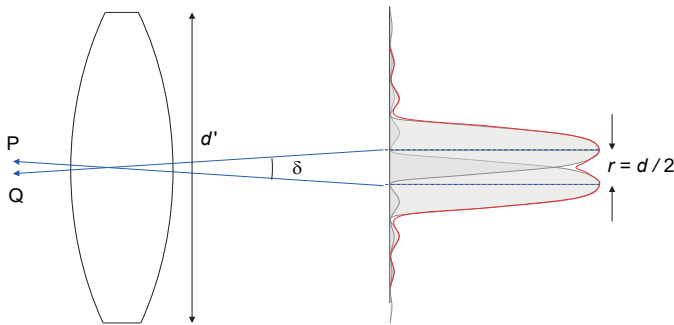
Fig. 3.27: Resolving power defined by the separation of two Airy discs (compare with Fig. 3.6).

Two neighbouring point objects can only be seen as separate images when their image separation is greater than r (Fig. 3.27). At that limiting separation the maximum of one diffraction disc lies on the minimum of the other. At greater apertures (smaller f/numbers) the resolving power increases.

Example 3.3:

The following diffraction discs with diameter d are generated by a camera for a mid-range wavelength of 550 nm (yellow light):

Aperture $f/2$: $d = 2.7\ \mu m$;　　　　　Aperture $f/11$: $d = 15\ \mu m$

A human eye with a pupil diameter of $d' = 2$ mm and a focal length of $f' = 24$ mm has an f/number of $f/12$. At a wavelength $\lambda = 550$ nm, and a refractive index in the eyeball of $n = 1.33$, a diffraction disc of diameter $d = 12\ \mu m$ is obtained. The average separation of the rod and cone cells which sense the image is 6 μm, which exactly matches the diffraction-limited resolving power.

3.1.5.2 Geometric resolving power

The geometric *resolving power* of a film or a digital imaging system defines its capability to distinguish between a number of black and white *line pairs* with equal spacing, width and contrast in the resulting image. Therefore it is a measure of the information content of an image.
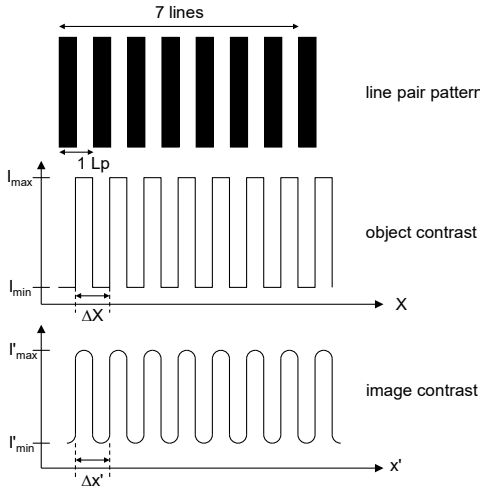
Fig. 3.28: Line pair pattern and contrast transfer.

The resolving power *RP* is measured visually as the number of *line pairs per millimetre* (Lp/mm). Alternatively, the terms *lines per millimetre* (L/mm) or *dots per inch* (dpi) may be used[2]. Such terms describe the ability of the imaging system to distinguish imaged details, with the interest usually being in the maximum distinguishable spatial frequency attainable (Fig. 3.28, see also section 3.1.6.1).

The spatial frequencies *F* in object space with respect to *f* in image space are the reciprocals of the corresponding line spacings ΔX in object space with respect to $\Delta x'$ in image space.

$$F = \frac{1}{\Delta X} \qquad\qquad f = \frac{1}{\Delta x'} \tag{3.33}$$

Resolving power can be measured by imaging a test pattern whose different spatial frequencies are known (Fig. 3.29). For example, the *Siemens star* consisting of 72 sectors (36 sector pairs) allows the maximum resolving power of the imaging system

---

2 With unit L/mm only black lines are counted, with Lp/mm black and white lines (pairs) are counted, i.e. both notions are comparable since a black line is only visible if bordered by white lines.

to be determined in Lp/mm by relating the number of sectors to the circumference (in mm) of the inner circle where the sectors are no longer distinguishable.

$$RP = \frac{36}{\pi d} \tag{3.34}$$
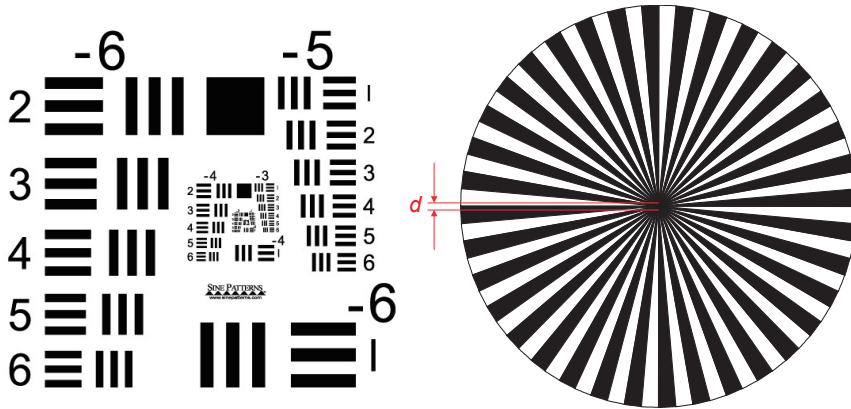
Fig. 3.29: Test chart and Siemens star for the measurement of resolving power.

The minimum resolved structure size in object space (structure resolution) ΔX is calculated from image scale and resolving power as follows:

$$\Delta X = m \cdot \Delta x' = m \cdot \frac{1}{RP} \tag{3.35}$$

The applicability of resolving power to opto-electronic sensors is discussed in section 3.4.1.7.

Example 3.4:

In the Siemens star printed above, the diameter of the unresolved circle is about 1.5 mm (observable with a magnifying glass). Thus the print resolution of this page can be computed as follows:

1. Resolving power: $\qquad RP = \dfrac{36}{\pi \cdot 1.5} \approx 8 \ \text{L/mm}$

2. Line size: $\qquad \Delta x' = 1 / RP = 0.13 \ \text{mm}$

3. Converted to dpi: $\qquad RP = \dfrac{25.4}{\Delta x'} = 194 \ \text{dpi} \approx 200 \ \text{dpi}$

3.1.5.3 Contrast and modulation transfer function

The actual resolving power of an imaging system depends on the contrast of the original object, i.e. for decreasing contrast, signal transfer performance is reduced,

particularly at higher spatial frequencies. A contrast-independent formulation of the resolving power is given by the *contrast transfer function* (CTF).

Original high contrast object pattern

Corresponding image of high contrast pattern

Sinusoidal contrast pattern in object

Corresponding image of sinusoidal contrast pattern
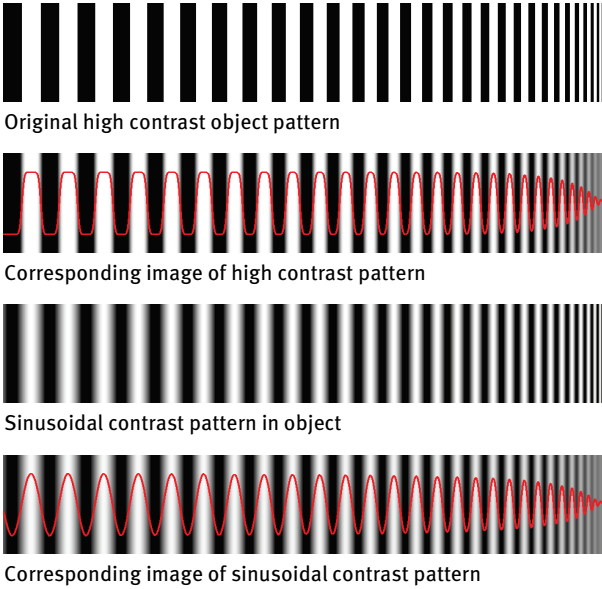
Fig. 3.30: Contrast and modulation transfer.

The object contrast K and the imaged contrast K' are functions of the minimum and maximum intensities I of the fringe pattern (Fig. 3.28, Fig. 3.30):

$$K(f) = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \qquad\qquad K'(f) = \frac{I'_{max} - I'_{min}}{I'_{max} + I'_{min}} \tag{3.36}$$

Hence the contrast transfer CT of a spatial frequency f follows:

$$CT(f) = \frac{K'(f)}{K(f)} \tag{3.37}$$

For most imaging systems, contrast transfer varies between 0 and 1. The contrast transfer function (CTF) defines the transfer characteristic as a function of the spatial frequency f (Fig. 3.31). Here the resolving power RP can be defined by the spatial frequency that is related to a given minimum value of the CTF, for example 30 % or 50 % (green line in Fig. 3.31). Alternatively, RP can be determined as the intersection point of an application-dependent threshold function of a receiver or observer that cannot resolve higher spatial frequencies. The threshold function is usually a perception-limiting function that describes the contrast-dependent resolving power of the eye with optical magnification (dashed line in Fig. 3.31).

In an analogous way, if the rectangular function of Fig. 3.28 is replaced by a sinusoidal function, the contrast transfer function is known as the *modulation transfer function* (MTF).

For an optical system an individual MTF can be defined for each system component (atmosphere, lens, developing, scanning etc.). The total system MTF is given by multiplying the individual MTFs (Fig. 3.31):

$$MTF_{total} = MTF_{imageblur} \cdot MTF_{lens} \cdot MTF_{sensor} \cdot \ldots \cdot MTF_n \tag{3.38}$$
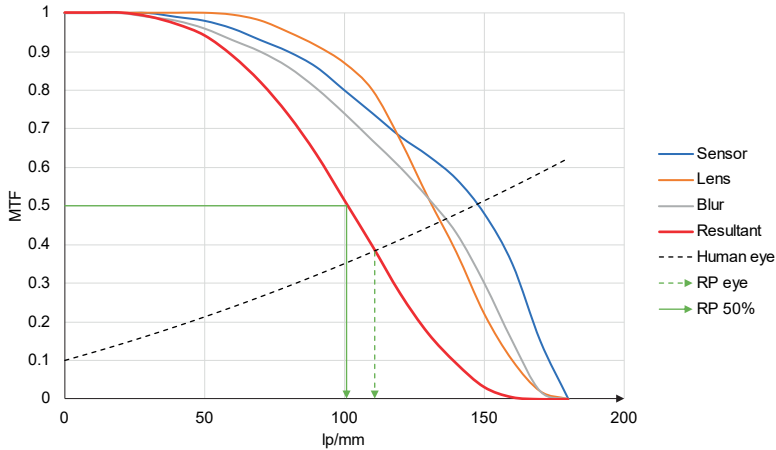
Fig. 3.31: Resulting total MTF.

3.1.6 Fundamentals of sampling theory

3.1.6.1 Sampling theorem
A continuous analogue signal is converted into a discrete signal by sampling. The amplitude of the sampled signal can then be transferred into digital values by a process known as quantization (Fig. 3.32).

If *sampling* is performed using a regular array of detector or sensor elements with spacing $\Delta s'$, then the sampling frequency f_A can be expressed as:

$$f_A = \frac{1}{\Delta s'} \qquad \text{: sampling frequency} \tag{3.39}$$

According to *Shannon's sampling theorem*, the *Nyquist frequency f_N* defines the highest spatial frequency that can be reconstructed by f_A without loss of information:

$$f_N = \frac{1}{2}f_A = \frac{1}{2\Delta s'} \qquad \text{: Nyquist frequency} \tag{3.40}$$

Spatial frequencies *f* higher than the Nyquist frequency are *undersampled,* and they are displayed as lower frequencies, an effect known as *aliasing* (see Fig. 3.33).
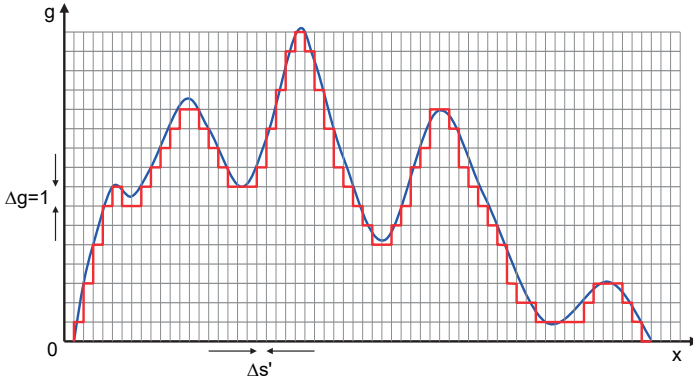
Fig. 3.32: Sampling and quantization.

To avoid sampling errors, as well as to provide a good visual reproduction of the digitized image, it is advisable to apply a higher sampling rate as follows:

$$f_A \approx 2.8 \cdot f \tag{3.41}$$

The transfer characteristic of the sampling system can be described with the modulation transfer function (MTF). With respect to a normalized frequency f/f_N the MTF falls off significantly above 1 (= Nyquist frequency). If the aliasing effects (Fig. 3.33) are to be avoided, the system must consist of a band-pass filter (anti-aliasing filter) that, in the ideal case, cuts off all frequencies above the Nyquist frequency (Fig. 3.34). As an optical low-pass filter, it is possible to use a lens with a resolving power (section 3.1.5.1) somewhat lower than the pixel pitch of the imaging sensor. There is a trend in modern DSLR camera designs utilising small pixels to forego such filters to maximize spatial resolution with lenses whose performance is matched to the sensor pixel dimensions.

The signal is quantized with a given number of integers *N* which is defined by the bit depth *K*. Bit depths $K=8$ to $K=16$ are commonly used in the camera electronics to quantize the grey values *g*. The required bit depth should be selected according to the signal-to-noise ratio of the imaging sensor (see section 3.4.1.9).

$$N = 2^K \qquad \text{where } 0 \le g \le N-1 \tag{3.42}$$
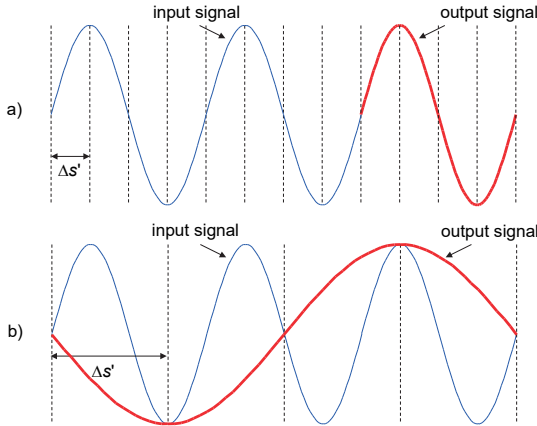
Fig. 3.33: a) Nyquist sampling; b) undersampling/aliasing.
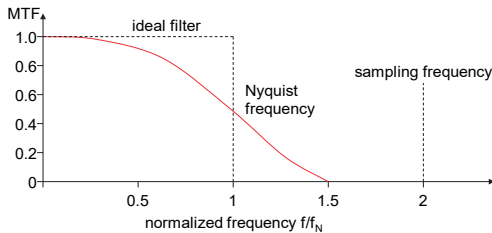
Fig. 3.34: MTF as a function of the normalized sampling frequency.

3.1.6.2 Detector characteristics

Electronic imaging devices, such as CCDs, consist of one or several detectors (sensor elements) of limited size with constant spacing with respect to each other. It is important to realize that, because of the need to place a variety of electronic devices in the sensing plane, not all of the area of each detector element is likely to be light sensitive. Sampling and transfer characteristics are therefore a function of both the size of the light-sensitive detector area (aperture size $\Delta d'$) and of the detector spacing (pixel spacing $\Delta s'$). In contrast to the sampling scheme of Fig. 3.33 real sampling integrates over the detector area.

Fig. 3.35a displays the detector output when scanning with a sensor whose light-sensitive elements are of size $\Delta d'$ and have gaps between them. This gives a detector spacing of $\Delta s' = 2\Delta d'$ (e.g. interline-transfer sensor, Fig. 3.76c). In contrast, Fig. 3.35b shows the sampling result with light sensitive regions without gaps $\Delta s' = \Delta d'$ (e.g. frame-transfer sensor, Fig. 3.76a). For the latter case, the detector signal is higher (greater light sensitivity), however, dynamic range ($u_{max} - u_{min}$) and hence modulation are reduced.
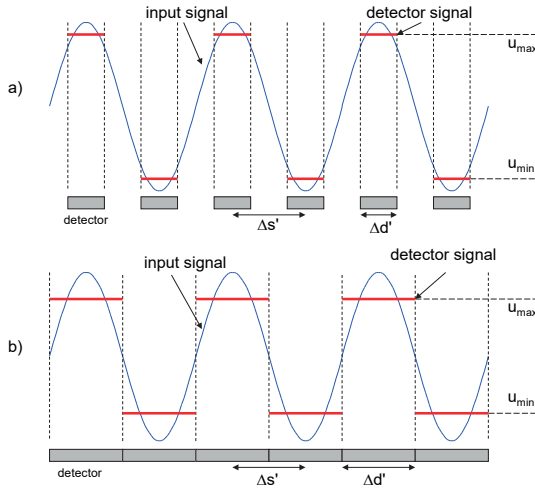
Fig. 3.35: Detector signals: a) light-sensitive size equal to half detector spacing; b) light-sensitive size equal to detector spacing.

The MTF of a sampling system consisting of sensor elements is given by:

$$MTF_{detector} = \frac{\sin(\pi \cdot \Delta d' \cdot f)}{\pi \cdot \Delta d' \cdot f} = \mathrm{sinc}(\pi \cdot \Delta d' \cdot f) \tag{3.43}$$

The sinc function was previously introduced for diffraction at aperture slits, see section 3.1.1.4. The function shows that a point signal (Dirac pulse) also generates an output at adjacent detector elements. In theory this is true even for elements at an infinite distance from the pulse. Together with possible defocusing, this gives rise to the point spread function (PSD) which, for example, causes sharp edges to have somewhat blurred grey values. Consequently both MTF and PSF can be reconstructed from an analysis of edge profiles (see Fig. 5.44).

The MTF becomes zero for $f = k / d'$ where $k = 1,2..n$. The first zero crossing ($k = 1$) is given by the frequency:

$$f_0 = \frac{1}{\Delta d'} \tag{3.44}$$

The first zero-crossing point can be regarded as a natural resolution limit. Fig. 3.36 shows a typical MTF of a detector system. Negative values correspond to reverse contrast, i.e. periodic black fringes are imaged as white patterns, and vice versa. Usually the MTF is shown up to the first zero-crossing only.
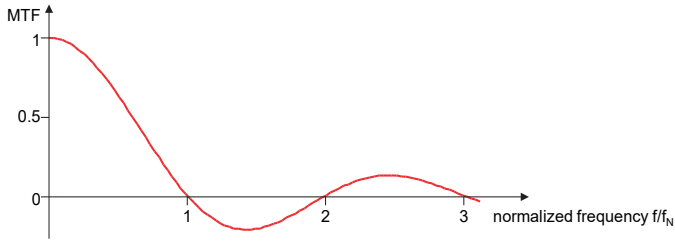
Fig. 3.36: MTF of a detector system.

3.2 Photogrammetric imaging concepts

3.2.1 Offline and online systems

With the availability of digital imaging systems, it is possible to implement a seamless data flow from image acquisition to analysis of the photogrammetric results (compare with Fig. 1.10). When image acquisition and evaluation take place in different places or at different times, this is known as *offline photogrammetry*. In contrast, when the acquired images are immediately processed and the relevant data used in a connected process, then this is known as *online photogrammetry* (Fig. 3.37). In addition, mixed forms exist. For example, intelligent cameras may use an internal camera processor to perform image processing during image acquisition, with further processing taking place offline.
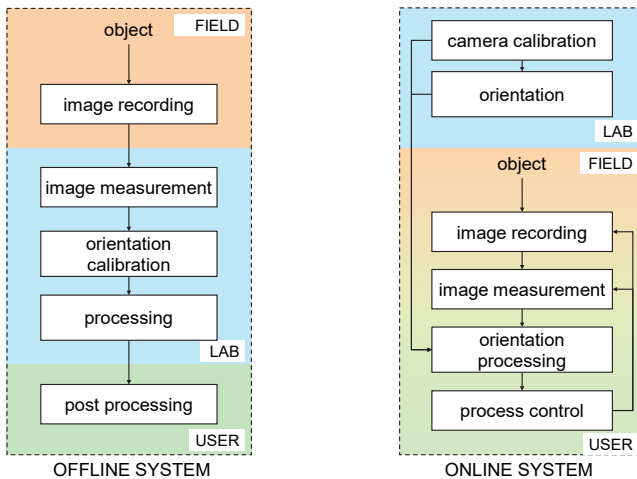
Fig. 3.37: Chain of processing in offline and online systems.

Both concepts have a direct impact on the relevant imaging technology. In many cases, cameras used offline can subsequently be calibrated from the recorded sequence of images (see section 7.3.2). Both analogue and digital cameras may be used here. In contrast, pre-calibrated digital cameras, which are assumed to be geometrically stable over longer periods of time, are typically used in online applications. By appropriate use of reference information in the object space such as control points, the orientation of online systems can be checked during measurement and, if necessary, corrected.

3.2.1.1 Offline photogrammetry

In offline applications it is typical to acquire what, in principle, can be an unlimited number of images which can then be evaluated at a later time and in a different place, possibly also by different users. If the imaging configuration is suitable (see section 7.3.2), the recording cameras can be simultaneously calibrated during the photogrammetric object reconstruction by employing a bundle adjustment (section 4.4). In this way it is possible to use lower-cost cameras, with a lower level of mechanical stability, as the photogrammetric recording system. Since object measurement and camera calibration take place simultaneously and with a high level of redundancy, the highest measurement accuracies can be achieved with offline systems. Depending on application, or technical and economic restrictions, it may also be sensible to make use of pre-calibrated cameras or metric cameras.

Examples of the use of offline photogrammetry are conventional aerial photography, the production of plans and 3D models in architecture, archaeology or facilities management, accident recording, the measurement of industrial equipment and components, as well as image acquisition from unmanned aerial vehicles (UAVs) or mobile platforms.

3.2.1.2 Online photogrammetry

Online photogrammetric systems have a limited number of cameras. There are single, stereo and multi-camera systems which, at given time intervals, deliver three-dimensional object information. Systems with image sensors integrated at fixed relative positions are generally pre-calibrated (known interior orientation) and oriented (exterior orientation). Depending on stability, this geometry remains constant over longer periods of time. Examples of mobile and stationary online systems with a fixed configuration of cameras are shown in Figs. 6.42 and 6.45. Online systems with variable camera configurations offer the option of on-site orientation which is normally achieved with the aid of reference objects. An example is shown in Fig. 6.34.

Online systems are commonly connected to further operating processes, i.e. the acquired 3D data are delivered in real time in order to control the operation of a second system. Examples here include stereo navigation systems for computer-

controlled surgery, positioning systems in car manufacture, image-guided robot positioning or production-line measurement of pipes and tubes.

3.2.2 Imaging configurations

In photogrammetry, imaging configuration is the arrangement of camera stations and viewing directions for object measurement. The following imaging configurations are typically distinguished:
– single image acquisition;
– stereo image acquisition;
– multi-image acquisition.

3.2.2.1 Single image acquisition

Three-dimensional reconstruction of an object from a single image is only possible if additional geometric information about the object is available. Single image processing is typically applied for rectification (of plane object surfaces, see section 4.2.8.1), orthophotos (involving a digital surface model, see section 4.2.8.2), as well as plane-object measurements (after prior definition of an object plane, see section 4.2.6.4) and monoplotting methods (see section 4.2.7). In addition, the use of just one image, and known object geometry, enables the measurement of relative 6DOF poses (position and orientation) between two objects in space, e.g. camera with respect to object (space resection, see section 4.2.3) or object to object (6DOF measurements, see section 4.2.5).

The achievable accuracy of object measurement depends primarily on the image scale (see section 3.3.1, Fig. 3.40 to Fig. 3.42) and the ability to distinguish those features which are to be measured within the image. In the case of oblique imaging, the image scale is defined by the minimum and maximum object distances.

3.2.2.2 Stereo image acquisition

Stereo imagery represents the minimum configuration for acquiring three-dimensional object information. It is typically employed where a visual or automated stereoscopic evaluation process is to be used. Visual processing requires near parallel camera axes similar to the normal case (Fig. 3.38a), as the human visual system can only process images which are comfortably within a limited angle of convergence. For digital stereo image processing (*stereo image matching*, see section 5.5), the prerequisites of human vision can be ignored with the result that more convergent image pairs can be used (Fig. 3.38b).

In the simplest case, three-dimensional object reconstruction using image pairs is based on the measurement of image parallax or disparity $px' = x'-x''$ (Fig. 3.38a) that can be transformed directly into a distance measure h in the viewing direction

(see section 4.3.6.2). More generally, image coordinates ($x'y'$, $x''y''$) of homologous (corresponding) image points can be measured in order to calculate 3D coordinates by spatial intersection (see section 4.4.7.1). The accuracy of the computed object coordinates in the viewing direction will generally differ from those parallel to the image plane. Differences in accuracy are a function of the intersection angle between homologous image rays, as defined by the *height-to-base ratio h/b*.

a) normal   b) convergent
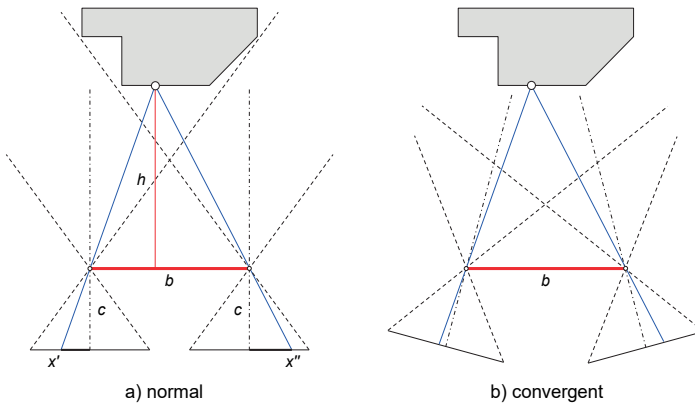
Fig. 3.38: Normal and convergent stereo image configurations.

Stereo imagery is most important for the measurement of non-signalized (targetless) object surfaces that can be registered by the visual setting of an optical floating mark (see section 4.3.6.3). Special stereometric cameras have been developed for stereo photogrammetry (see section 3.5.4). An example application in industry is the measurement of free-form surfaces where one camera can be replaced by an oriented pattern projector (see section 6.7.3).

3.2.2.3 Multi-image acquisition

Multi-image configurations (Fig. 3.39) are not restricted with respect to the selection of camera stations and viewing directions. In principle, the object is acquired by an unlimited number of images from locations chosen to enable sufficient intersecting angles of bundles of rays in object space. At least two images from different locations must record every object point to be coordinated in 3D.

Object coordinates are determined by multi-image triangulation (bundle adjustment, see section 4.4, or spatial intersection, see section 4.4.7.1). If a sufficient number and configuration of image rays (at least 3–4 images per object point) are provided, uniform object accuracies in all coordinates can be obtained (see section 3.3.1.2).

In close-range photogrammetry, multi-image configurations are the most common case. They are required in all situations where a larger number of different viewing locations are necessary due to the object structure, e.g. occlusions or the measurement of both interior and exterior surfaces, or to maintain specified accuracy requirements. Images can be arranged in strips or blocks (Fig. 4.58) or as all-around configurations (see Fig. 3.39, Fig. 4.59) but, in principle, without any restrictions (other examples in Fig. 1.6, Fig. 1.8, Fig. 8.5).
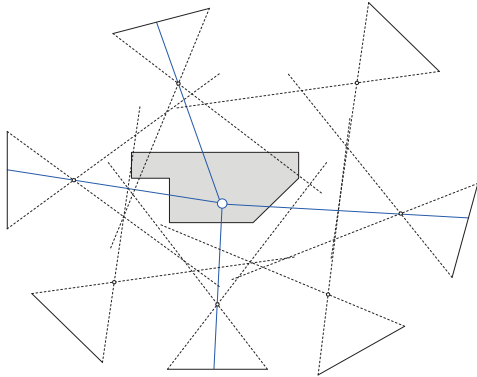
Fig. 3.39: Multi-image acquisition (all-around configuration).

Where the configuration provides a suitable geometry, multi-image configurations enable the simultaneous calibration of the camera(s) by self-calibrating bundle adjustment procedures (see sections 4.4.2.4 and 7.3.2).

3.3 Geometry of the camera as a measuring device

3.3.1 Image scale and accuracy

Image scale and achievable accuracy are the basic criteria of photogrammetric imaging and will dominate the choice of camera system and imaging configuration.

3.3.1.1 Image scale

The image scale number or magnification m is defined by the ratio of object distance h to the principal distance c (lens focal length plus additional shift to achieve sharp focus). It may also be given as the ratio of a distance in object space X to the corresponding distance in image space x', assuming X is parallel to x' (see eqn. 1.1, compare with section 3.1.2):

$$m = \frac{h}{c} = \frac{X}{x'} = \frac{1}{M} \tag{3.45}$$

In order to achieve a sufficient accuracy and detect fine detail in the scene, the selected image scale m must take into account the chosen imaging system and the surrounding environmental conditions. Fig. 3.40 illustrates the relationship between object distance, principal distance, image format and the resulting image scale.

a) Different image formats

b) Different object distances

c) Different principal distances

d) Equal image scales
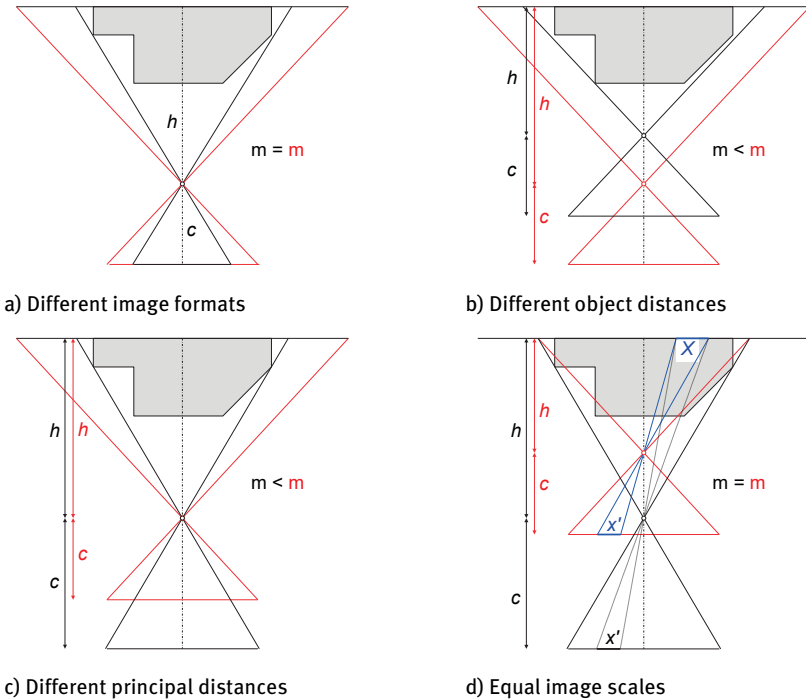
Fig. 3.40: Dependency of image scale on image format, focal length and object distance.

Using a camera with a smaller image format and the same image scale (and principal distance) at the same location, the imaged object area is reduced (Fig. 3.40a). In contrast, a larger image scale can be achieved if the object distance is reduced for the same image format (Fig. 3.40b).

For a shorter object distance (Fig. 3.40b), or a longer principal distance (Fig. 3.40c), a larger image scale will result in a correspondingly reduced imaged object area, i.e. the number of images necessary for complete coverage of the object will increase.

Fig. 3.40d shows that equal image scales can be obtained by different imaging configurations. With respect to image scale it can be concluded that the selection of imaging system and camera stations is often a compromise between contrary

requirements. Note, however, that any change in the position of the camera with respect to the object will result in a different perspective view of the object. Conversely, changing the focal length of the lens, or altering the camera format dimensions, whilst maintaining the camera position, will not alter the perspective.

An image has a uniform scale only in the case of a plane object which is viewed normally (camera axis normal to the object plane). For small deviations from the normal an average image scale number related to an average object distance can be used for further estimations. In practical imaging configurations, large deviations in image scale result mainly from

– large spatial depth of the object and/or
– extremely oblique images of a plane object.

Example 3.5:

Given a camera with image format $s' = 60$ mm and a principal distance of $c = 40$ mm (wide angle). Compute the object distance h, where an object size of 7.5 m is imaged over the full format.

1. Image scale number: $m = \dfrac{X}{x'} = \dfrac{7500}{60} = 125$ $(M = 1{:}125)$

2. Object distance: $h = m \cdot c = 125 \cdot 40 = 5000 \,\text{mm} = 5\,\text{m}$
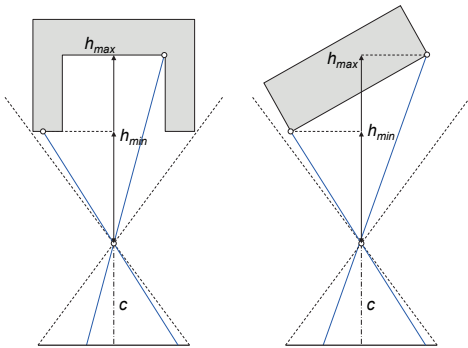
Fig. 3.41: Different image scales.
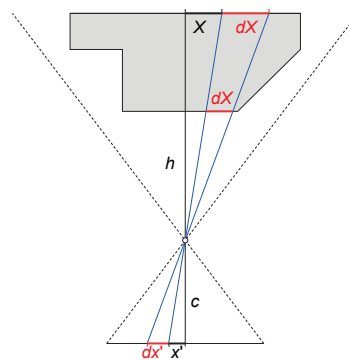
Fig. 3.42: Single image acquisition.

For these cases (see Fig. 3.41) minimum and maximum image scales must be used for project planning and accuracy estimations.

3.3.1.2 Accuracy estimation

The achievable accuracy[3] in object space of a photogrammetric measurement requires assessment against an independent external standard, but the precision of the derived coordinates can be estimated approximately according to Fig. 3.42.

Differentiation of eqn. (3.45) shows that the uncertainty of an image measurement dx' can be transferred into object space by the image scale number m:

$$dX = m \cdot dx' \tag{3.46}$$

Applying the law of error propagation, the standard deviation gives:

$$s_X = m \cdot s_{x'} \tag{3.47}$$

In many cases a relative precision, rather than an absolute value, that is related to the maximum object dimension S, or the maximum image format s', is calculated:

$$\frac{s_X}{S} = \frac{s_{x'}}{s'} \tag{3.48}$$

Eqn. (3.48) shows that a larger image format results in better measurement precision.

Example 3.6:

Given an object with object size $S = 7.5$ m photographed at an image scale of $m = 125$. What is the necessary image measuring accuracy $s_{x'}$, if an object precision of $s_X = 0.5$ mm is to be achieved?

1. According to (3.47): $\qquad s_{x'} = \dfrac{s_X}{m} = \dfrac{0.5}{125} = 0.004\,\text{mm} = 4\,\mu\text{m}$

2. Relative accuracy: $\qquad \dfrac{s_X}{S} = 1:15000 = 0.007\%$

Firstly, the achievable object precision according to (3.47) indicates the relationship between scale and resulting photogrammetric precision. Furthermore, it is a function of the imaging geometry (number of images, ray intersection angles in space) and the extent to which measured features can be identified. A statement of relative precision is only then meaningful if the measured object, processing methods and accuracy verification are effectively described.

Eqn. (3.47) must be extended by a *design factor q* that provides an appropriate weighting of the imaging configuration:

$$s_X = q \cdot m \cdot s_{x'} = \frac{q_D}{\sqrt{k}} m \cdot s_{x'} \tag{3.49}$$

---

3 Here the term "accuracy" is used as a general quality criterion. See sections 2.4.3 and 7.2 for a definition of accuracy, precision and reliability.

The design parameter q_D is related to the intersection geometry of the imaging configuration, and k defines the mean number of images per camera location. For a normal case, in practice $k = 1$ and therefore $q_D = q$.

Practical values for the design factor q vary between 0.4–0.8 for excellent imaging configurations, (e.g. all-around configuration, see Fig. 3.39), and up to 1.5–3.0 or worse for weak stereo configurations (see Fig. 3.38).

If the object is targeted (marked, or signalized, e.g. by circular high-contrast retro-reflective targets) and imaged by an all-around configuration, (3.49) provides a useful approximation for all three coordinate axes, such that $q = 0.5$ can be achieved. In cases where the object cannot be recorded from all sides, accuracies along the viewing direction can differ significantly from those in a transverse direction. As an example, the achievable precision in the viewing direction Z for a normal stereo pair (see section 4.3.6.2) can be estimated by:

$$s_Z = \frac{h^2}{b \cdot c} s_{px'} = m \frac{h}{b} s_{px'} \tag{3.50}$$

Here b defines the distance between both camera stations (stereo base) and $s_{px'}$ the measurement precision of the x-parallax; base b and principal distance c are assumed to be free of error. Measurement precision in the viewing direction depends on the image scale (h/c) and on the intersection geometry, as defined by the height-to-base ratio (h/b).

Example 3.7:

Given a stereo image pair with an image scale of $M = 1{:}125$, an object distance of $h = 5$ m and a base length of $b = 1.2$ m. Compute the achievable precision in the XY-direction (parallel to the image plane) and in the Z-direction (viewing direction) respectively, given a parallax measurement precision of $s_{px'} = 4$ μm (assume $s_{px'} = s_{x'} \sqrt{2}$).

1. Precision in Z: $\quad\quad\quad\quad\quad\quad s_Z = m \cdot \frac{h}{b} \cdot s_{px'} = 125 \cdot \frac{5}{1.2} \cdot 0.004 = 2.1 \text{mm}$

2. Precision in X,Y where $q_{XY} = 1$: $\quad s_X = s_Y = m \cdot s_{x'} = 125 \cdot 0.006 = 0.75$ mm

3. Design factor from (3.49): $\quad\quad\quad q_Z = \frac{s_Z}{m \cdot s_{x'}} = 2.8$

The example shows that the precision in the viewing direction is reduced by a factor of almost 3.

3.3.2 Interior orientation of a camera

The interior (intrinsic) orientation of a camera comprises all instrumental and mathematical elements which completely describe the imaging model within the camera. By taking proper account of the interior orientation, the real camera conforms to the pinhole camera model. It is a requirement of the model that there

exists a reproducible image coordinate system so that geometric image values, such as measured image coordinates, can be transformed into the physical-mathematical imaging model.

a) Original image   b) Compression   c) Extract
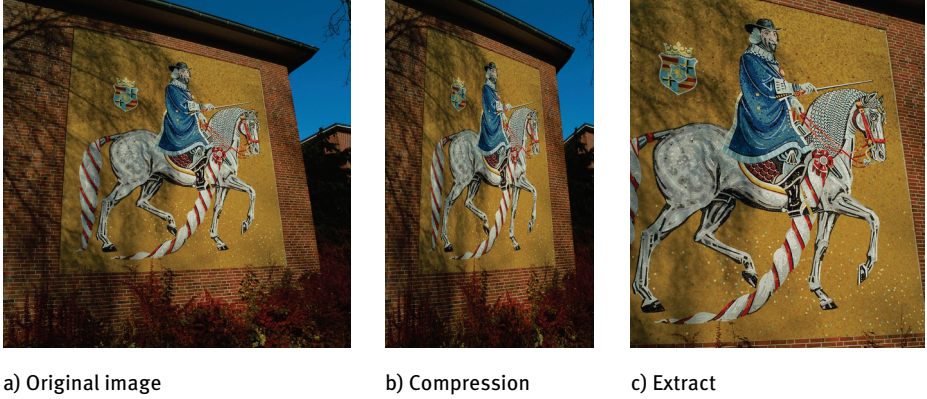
Fig. 3.43: Original and processed images.

Fig. 3.43 shows an arbitrarily selected, original digital image that has been stored in exactly this form and therefore has a reproducible relationship to the camera. The applied processing (compression and extraction) results in a loss of reference to the image coordinate system fixed in the camera, unless the geometric manipulation can be reconstructed in some way (see section 3.3.2.4).

3.3.2.1 Physical definition of the image coordinate system

The image coordinate system must not only be defined physically with respect to the camera lens, but must also be reconstructable within the image. For this purpose, analogue photogrammetric cameras used either fiducial marks at the edge of the image (metric camera) or a réseau (partial-metric camera) whose nominal coordinates were calibrated at the factory. For the digital cameras used exclusively today, artificial reference points for the image coordinate system can be omitted if there is a clear reference between the digital image and the optoelectronic image sensor (rows and columns of the sensor matrix). This is usually the case when using fixed area sensors (CCD or CMOS sensor matrix) and with direct digital image readout. For images without a given reference system (see Fig. 3.51 for an example), the direct linear transformation (DLT, section 4.2.4.1) can be used for photogrammetric orientation tasks.

To ensure a defined physical relationship between sensor and lens, a static sensor without sensor-based auto-focus, motion compensation and anti-dust vibration (see 3.4.2.3), a fixed focal length lens, the focus and aperture of which is mechanically

locked in place, are preferred for accurate photogrammetric reconstruction. In such cases, the image manipulations shown in Fig. 3.43 should be omitted. By convention the origin of these systems is usually at the top left and the coordinate system is left-handed (see section 2.1.1). The conversion of pixel coordinates into metric image coordinates is done with eqn. (2.2). Radiometric image processing (e.g. contrast changes) usually does not change the image geometry and is permissible in the photogrammetric sense.

3.3.2.2 Perspective centre and distortion

Mathematically, the perspective centre is defined by the point through which all straight-line image rays pass. For images created with a compound lens (multiple lens components) both an external and an internal perspective centre can be defined. Each is defined by the intersection point of the optical axis with the entrance pupil *EP* and the exit pupil *E'P*, respectively (see section 3.1.2). The position and size of the entrance and exit pupils are defined by the lens design and its limiting aperture (Fig. 3.11, Fig. 3.44). Hence, the position of the perspective centre depends on the chosen aperture and, due to the influence of dispersion, is additionally dependent on wavelength.
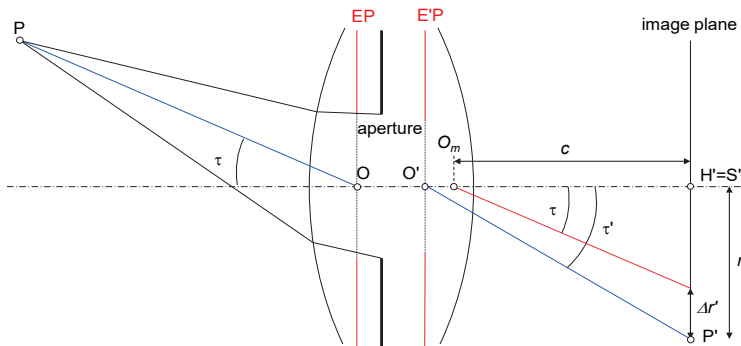
Fig. 3.44: Perspective centres O, O' and principal distance *c* (after Kraus 1994).

In the ideal case of Fig. 3.11 the angle of incidence τ is equal to the exit angle τ', and the principal distance c is equal to the image distance a' (between principal plane and image plane). As the position of entrance pupil and exit pupil do not usually coincide with the principal planes, an incident ray enters at angle τ, and exits at a different angle τ'. This effect is radially symmetric with a point of symmetry S'. Compared with the ideal case, an image point P' is shifted by an amount $\Delta r'$ that is known as *radial distortion* (see also Fig. 3.12 and Fig. 3.17).

$$\Delta r' = r' - c \cdot \tan \tau \tag{3.51}$$

In this formulation the distortion is therefore linearly dependent on the principal distance (see also Fig. 3.53).
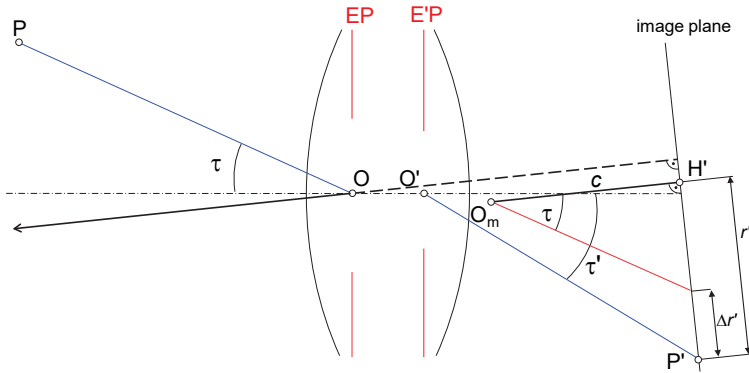
Fig. 3.45: Definition of principal point for a tilted image plane (after Kraus 2000).

The *point of autocollimation* H' is defined as the intersection point of the optical axis of the lens OO' and the image plane. The mathematical perspective centre O_m, used for photogrammetric calculations, is chosen to maximize the symmetry in the distortion in the image plane. The principal distance c is normally chosen such that the sum of the distortion components across the whole image format is minimized. If the optical axis is not normal to the image plane, O_m is not positioned on the optical axis (Fig. 3.45). H' is also known as the *principal point*. In real cameras, principal point, point of symmetry and the centre of the image can all be separate. In principle, the image coordinate system can be arbitrarily defined (Fig. 3.46).
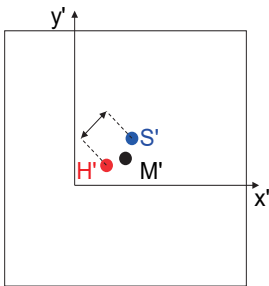
Fig. 3.46: Possible locations of principal point, point of symmetry, centre of image and origin of image coordinate system.

Hence the photogrammetric reference axis is defined by the straight line O_mH'. In object space it is given by the parallel ray passing through O^4. The image radius r' of an image point, and the residual of radial distortion $\Delta r'$, are defined with respect to the principal point H' (see section 3.3.2.3).

If a camera is typically calibrated by means of a bundle adjustment (see section 4.4.2.4), the principal distance, principal point and distortion parameters are all calculated as a result of this least-squares optimization which incorporates all participating observations. They correspond to the optically defined parameters if the imaging network is strong and the scale definition accurate.
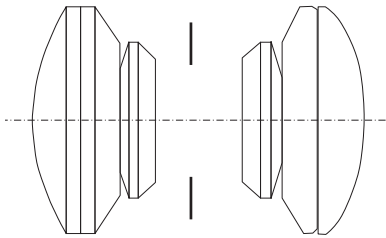
Fig. 3.47: Example of a symmetric lens design (Zeiss Lametar 8/f=200, $\Delta r'_{max} = \pm 4$ μm at $r' < 90$ mm).
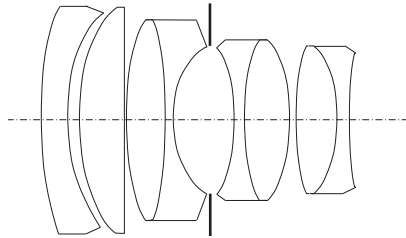
Fig. 3.48: Example of an asymmetric lens design (Leica 50/0.95 Noctilux-M, $\Delta r' = \pm 200$ μm at $r' = 20$ mm, $\Delta r' = \pm 60$ μm at $r' = 10$ mm).

In practice, every lens generates distortions. The radial distortion described above can be reduced to a level of $\Delta r' < 4$ μm for high-quality lenses designed symmetrically about the lens aperture (Fig. 3.47). In contrast, asymmetric lens designs produce significantly larger distortion values, especially towards the corners of the image. See Fig. 3.48 for an example of an asymmetric lens design.

In contrast, tangential distortions are attributable to decentring and tilt of individual lens elements within the compound lens (Fig. 3.58 shows the effect in image space). For good quality lenses, these distortions are usually 10 times smaller than radial distortion and thus can be neglected for many photogrammetric purposes. However, the simple low-cost lenses which are increasingly used have been shown to exhibit significantly larger tangential and asymmetric radial distortion values. Distortions in the range of more than 30 μm are possible and are attributable to the low cost of these lenses, combined with the small size of their individual elements.

3.3.2.3 Parameters of interior orientation

A camera can be modelled as a spatial system that consists of a planar imaging area (film or electronic sensor) and the lens with its perspective centre. The parameters of

---

4 Usually the notations O or O' are used, even when O_m is meant.

interior orientation of a camera define the spatial position of the perspective centre, the principal distance and the location of the principal point with respect to the image coordinate system defined in the camera. They also encompass deviations from the principle of central perspective to include radial and tangential distortion and often affinity and orthogonality errors in the image.
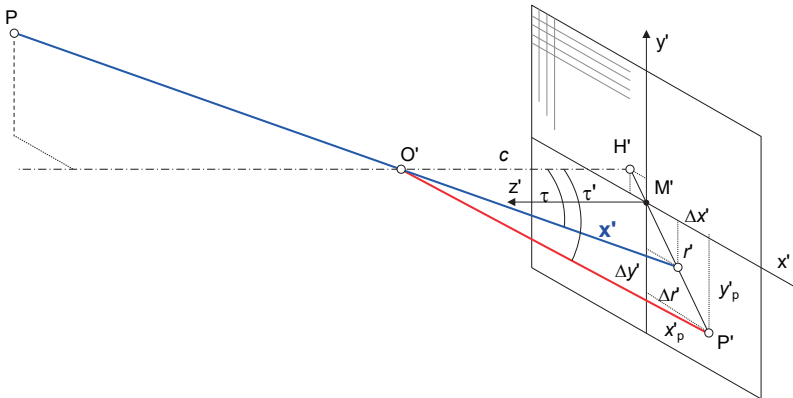
Fig. 3.49: Interior orientation.

Fig. 3.49 illustrates the schematic imaging process of a photogrammetric camera. Position and offset of the perspective centre, as well as deviations from the central perspective model, are described with respect to the image coordinate system as defined by reference or fiducial points (film-based system) or the pixel array (electronic system). The origin of the image coordinate system is located in the image plane. For the following analytical calculations, the origin of the image coordinate system is shifted to coincide with the perspective centre according to Fig. 2.2.

Hence, the parameters of interior orientation are (see section 3.3.2.1):

–   Principal point H':

    Foot of perpendicular from perspective centre to image plane, with image coordinates (x'_0, y'_0). For commonly used cameras approximately equal to the centre of the image: H' ≈ M'.

–   Principal distance c:

    Perpendicular distance to the perspective centre from the image plane in the negative z' direction. When focused at infinity, c is approximately equal to the focal length of the lens ($c \approx f'$). The principal distance is also known as the *camera constant*. For analytical computations c is a negative number ($z' = -c$), see section 2.1.2.

–   Parameters of functions describing imaging errors:
     Functions or parameters that describe deviations from the central perspective
     model are dominated by the effect of symmetric radial distortion $\Delta r'$.

If these parameters are given, the (error-free) *image vector* **x'** can be defined with
respect to the perspective centre, and hence also the principal point:

$$\mathbf{x'} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x'_p - x'_0 - \Delta x' \\ y'_p - y'_0 - \Delta y' \\ -c \end{bmatrix} \tag{3.52}$$

where
x'_p, y'_p: measured coordinates of image point P'
x'_0, y'_0:   coordinates of the principal point H'
$\Delta x', \Delta y'$:   correction values for errors in the image plane

The parameters of interior orientation are determined by camera calibration (see
sections 3.3.2.5 and 4.4.2).

3.3.2.4 Metric and semi-metric cameras

The expression *metric camera* is used for photogrammetric cameras with a stable
optical and mechanical design. For these cameras the parameters of interior
orientation can be calibrated in the factory (laboratory) and are assumed to be
constant over a long period of time. Usually metric cameras consist of a rigidly
mounted fixed-focus lens with minimal distortion. In addition, they have a flat image
plane. A *semi-metric camera* meets the above metric camera requirements only with
respect to a plane image surface and its corresponding plane image coordinate
system. These specifications are fulfilled by a réseau for analogue film cameras, and
the physical surface of the imaging sensor for digital cameras.

a) Variation of principal point (in mm)   b) Variation of principal distance (in mm)
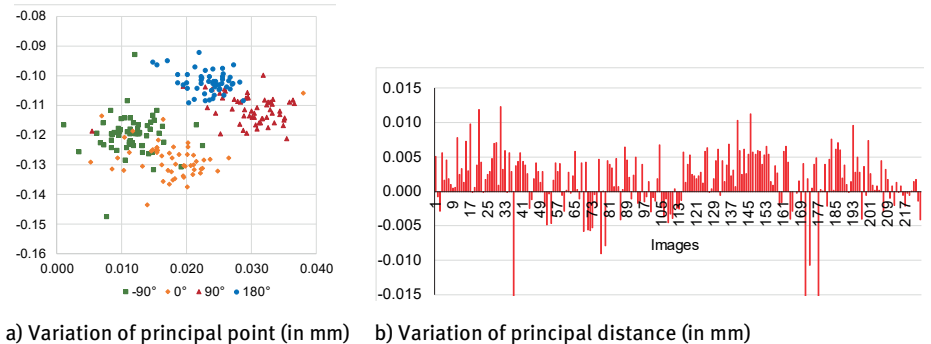
Fig. 3.50: Variation in perspective centre position across a series of images
(Nikon D4 with Zeiss 35mm lens).

For a semi-metric camera, the spatial position of the principal point is only given approximately with respect to the image coordinate system. It is not assumed to be constant over a longer period of time. Movements between perspective centre (lens) and image plane (sensor) can, for example, result from the use of variable focus lenses, or an unstable mounting of lens or imaging sensor. Fig. 3.50 illustrates this effect for a series of 220 images acquired with a Nikon D4 DSLR camera held in different orientations. The observed variations can only be handled by a calibration which varies with the image (image-variant camera calibration).
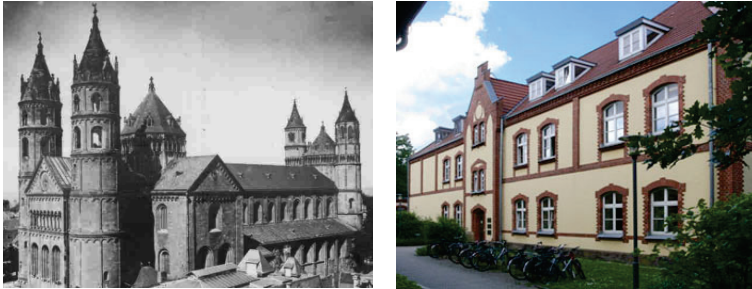
Fig. 3.51: Image examples of non-metric cameras
(left: historical postcard; right: image posted on Internet).

Cameras, such as analogue photographic cameras, without a suitable photogrammetric reference system and/or without a planar image surface, are known as *amateur* or *non-metric cameras*. If images taken by amateur cameras, for example from old postcards, or an unknown source, such as an internet platforms, are to be processed photogrammetrically, they often have no unique reference points from which the principal point can be determined. Often, even the original image corners cannot be found. In such cases, a suitable analytical method for image orientation is the direct linear transformation (DLT, see section 4.2.4.1). This method does not require an image coordinate system. For multiple image configuration, unknown images can be self-calibrated but with limited accuracy, for example using techniques such as structure-from-motion (see section 5.5.2.2 and also compare with Fig. 3.43). Fig. 3.51 shows an example of an image where neither fiducial marks nor image corners in the camera body are visible.

3.3.2.5 Determination of interior orientation (calibration)
In photogrammetry, the determination of the parameters of interior orientation is usually referred to as *calibration*. This is based on the idea that once a mechanically stable camera is calibrated, it may be moved from one image taking location to another whilst retaining its calibration. State-of-the-art close-range techniques

employ analytical calibration methods to derive the parameters of the chosen camera model indirectly from photogrammetric image coordinate observations. For this purpose the imaging function is extended by the inclusion of additional parameters that model the position of the perspective centre and image distortion effects.

Usually calibration parameters are estimated by bundle adjustment (simultaneous calibration, section 4.4.2.4). Depending on available object information (reference points, distances, constraints) suitable imaging configurations must be chosen (see section 7.3).

The necessity for the periodic calibration of a camera depends on the accuracy specifications, the mechanical construction of the camera and environmental conditions on site at the time of measurement. Consequently, the time and form of the most appropriate calibration may vary:

– One-time factory calibration:
  Imaging system: metric camera
  Method: factory or laboratory calibration
  Reference: calibrated test instruments, e.g. goniometer, comparators, reference target points
  Assumption: camera parameters are valid for the life of the camera
– Long-term checks, e.g. annual:
  Imaging system: metric camera
  Method: laboratory or test-field calibration
  Reference: calibrated test instruments, reference target points, scale bars, plumb lines
  Assumption: camera parameters are valid for a long period of time
– Calibration immediately before or after object measurement:
  Imaging system: semi-metric camera, metric camera with high stability
  Method: test field calibration, self-calibration (see section 7.3.1.1)
  Reference: reference points, reference lengths within the test field, straight lines
  Assumption: camera parameters do not alter until the time of object measurement
– Calibration integrated into object reconstruction:
  Imaging system: semi-metric camera, metric camera with moderate stability
  Method: self-calibration, on-the-job calibration (see section 7.3.1.3)
  Reference: reference points, distances on object, straight lines
  Assumption: constant interior orientation during image network acquisition
– Calibration of each individual image:
  Imaging system: semi-metric camera with limited stability
  Method: self-calibration with variable interior orientation (see section 4.4.2.4)
  Reference: reference points, distances on object, straight lines
  Assumption: only limited requirements regarding camera stability, e.g. constant distortion values

In the default case, the parameters of interior orientation are assumed to be constant for the duration of image acquisition. However, for close-range applications in it may be the case that lenses are changed or re-focused during an image sequence, and/or mechanical or thermal changes occur. Every change in camera geometry will result in a change of interior orientation which must be taken into account in the subsequent evaluation by use of an individual set of parameters for each different camera state.

From the standpoint of camera calibration, the difference between metric and semi-metric cameras becomes largely irrelevant as the stability of interior orientation depends on the required accuracy. Consequently, even metric cameras are calibrated on-the-job if required by the measuring task. In such cases their prior calibration data can be used as a check or as input observations into the new calibration. On the other hand, semi-metric digital cameras can be calibrated in advance if they are components of multi-camera online systems.

3.3.3 Standardized correction functions

Deviations from the ideal central perspective model, attributable to imaging errors, are expressed in the form of correction functions $\Delta x'$, $\Delta y'$ to the measured image coordinates. Techniques to establish these functions have been largely standardized and they capture the effects of radial, tangential and asymmetric distortion, as well as affine errors in the image coordinate system. Extended models for special lenses and imaging systems are described in section 3.3.4.

3.3.3.1 Radial distortion

According to section 3.1.3.1, distortion is related to the principal point, i.e. the measured image coordinates x'_P, y'_P must first be corrected by a shift of origin to the principal point at x'_0, y'_0:

$$x° = x'_P - x'_0$$
$$y° = y'_P - y'_0$$

: image coordinates relative to principal point

where $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (3.53)

$$r' = \sqrt{x^{°2} + y^{°2}}$$

: image radius, distance from the principal point

The correction of the image coordinates $x°$, $y°$ for distortion is then given by:

$$x' = x° - \Delta x'$$
$$y' = y° - \Delta y'$$

: corrected image coordinates $\qquad\qquad\qquad$ (3.54)

The distortion corrections $\Delta x'$, $\Delta y'$ must be calculated using the final image coordinates x', y' but must be initialized using the approximately corrected

coordinates $x°,y°$. Consequently, correction values must be applied iteratively (see section 3.3.4.6).
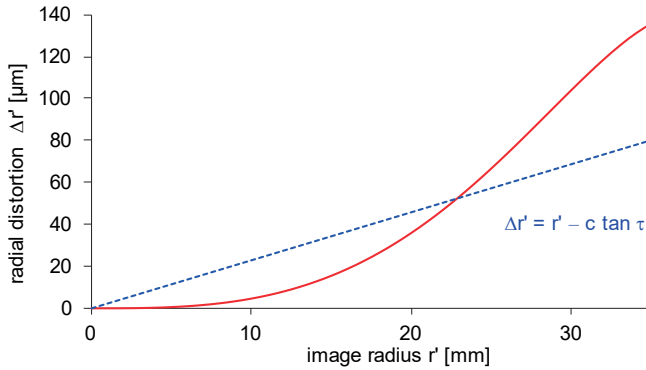
Fig. 3.52: Symmetric radial distortion.

Symmetric radial distortion, commonly known as radial distortion, constitutes the major imaging error for most camera systems. It is attributable to variations in refraction at each individual component lens within the camera's compound lens. It depends also on the wavelength, aperture setting, focal setting and the object distance at constant focus.

Fig. 3.52 shows the effect of radial distortion as a function of the radius of an imaged point. In the example, the distortion increases with distance from the principal point, and it can easily reach 1% (relative distortion $\Delta r'/r'$) or more for off-the-shelf lenses. The *distortion curve* is usually modelled with a polynomial series (Seidel series) with distortion parameters K_0 to K_n: [5]

$$\Delta r'_{rad} = K_0 r' + K_1 r'^3 + K_2 r'^5 + K_3 r'^7 \tag{3.55}$$

$$\Delta r'_{rad} = A_0 r' + A_1 r'^3 + A_2 r'^5 + A_3 r'^7 + \dots \tag{3.56}$$

For most lens types the series can be truncated after the second or third term ($i=3$) without any significant loss of accuracy. The linear term with K_0 (A_0) describes the function $\Delta r' = r' - c \cdot \tan \tau$ (see eqn. 3.51 and Fig. 3.52). Eqn. (3.51) and Fig. 3.53 show that the effect of introducing the linear term with K_0 (A_0) can also be achieved by a change in principal distance. Hence, the distortion parameters defined in (3.55) are numerically correlated with image scale or principal distance, and consequently, K_0 (A_0) and c cannot be calculated simultaneously within one system of equations.

---

5 Here and in the following paragraphs both common notations (A_i) and (K_i) are displayed in parallel.
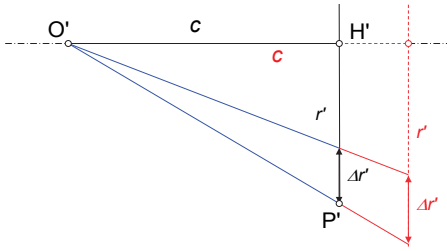
Fig. 3.53: Linear relationship between c and $\Delta r'$.

In order to avoid these correlations, the linear part of the distortion function is removed or set to a constant value. This is equivalent to a rotation of the distortion curve into the direction of the r' axis, thus resulting in a second zero-crossing. This generates a distortion function in which only the differences from the straight line $\Delta r' = r' - c \cdot \tan\tau$ (see eqn. 3.51) in Fig. 3.52 must be modelled.
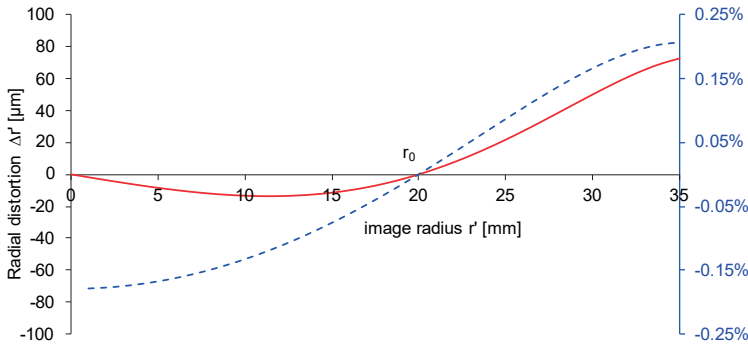
Fig. 3.54: Typical balanced lens distortion curve (Rollei Sonnar 4/150)
red: distortion effect $\Delta r'$; blue: relative distortion $\Delta r'/r'$.

Alternatively, a polynomial of the following type is used:

$$\Delta r'_{rad} = A_1 r'(r'^2 - r_0^2) + A_2 r'(r'^4 - r_0^4) + A_3 r'(r'^6 - r_0^6) \tag{3.57}$$

By simple rearrangement of (3.57) it can be shown that this has the same effect as (3.55):

$$\Delta r'_{rad} = A_1 r'^3 + A_2 r'^5 + A_3 r'^7 - r'(A_1 r_0^2 + A_2 r_0^4 + A_3 r_0^6) \tag{3.58}$$

Here the term in brackets is a constant analogous to K_0. In practice r_0 should be chosen such that minimum and maximum distortion values are more or less equal with

respect to the complete image format (*balanced radial distortion*). Usually r_0 is set to approximately 2/3 of the maximum image radius. Fig. 3.54 shows a typical radial distortion curve according to (3.57), and the relative distortion $\Delta r'/r'$.

distortion parameters:

$A_1 = 4.664 \cdot 10^{-6}$
$A_2 = -6.456 \cdot 10^{-10}$
$A_3 = 0$

$r_0 = 20.0$ mm
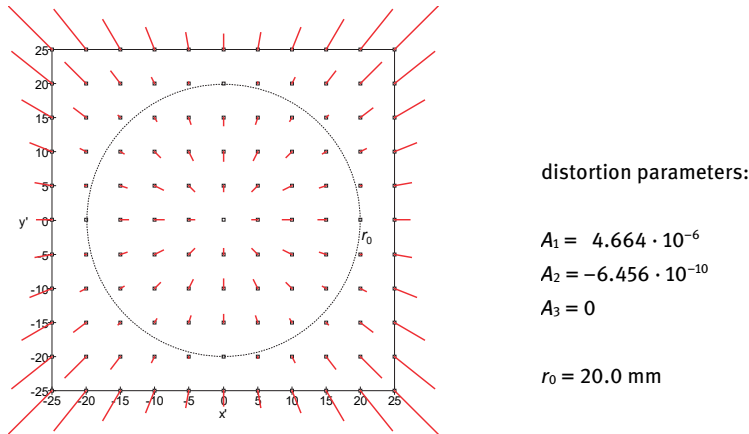
Fig. 3.55: Effect of radial distortion (data from Table 3.1).

Fig. 3.55 displays the corresponding two-dimensional effect with respect to the image format. Balanced radial lens distortion is only necessary when the distortion correction is applied using analogue methods, for example by specially shaped cams within the mechanical space rod assembly. The term r_0 was introduced for analogue systems where mechanical cameras were used to apply mechanical corrections but for purely digital calibration it no longer has any practical meaning and can, without restriction, be set to zero (corresponding to $K_0 = 0$).

Table 3.1: Correction table for distortion (see Fig. 3.54, Fig. 3.55, all values in mm).

r'	Δr'	r'	Δr'	r'	Δr'	r'	Δr'
0	0.0000	9	-0.0125	18	-0.0057	27	0.0350
1	-0.0018	10	-0.0130	19	-0.0031	28	0.0419
2	-0.0035	11	-0.0133	20	0.0000	29	0.0494
3	-0.0052	12	-0.0132	21	0.0035	30	0.0574
4	-0.0068	13	-0.0129	22	0.0076	31	0.0658
5	-0.0082	14	-0.0122	23	0.0121	32	0.0748
6	-0.0096	15	-0.0112	24	0.0170	33	0.0842
7	-0.0107	16	-0.0098	25	0.0225	34	0.0941
8	-0.0117	17	-0.0080	26	0.0285	35	0.1044

A table of correction values for radial distortion (lens map function) is often derived from a camera calibration (Table 3.1). It can then be used by a real-time processing system to provide correction values for image coordinates by linear interpolation.

Finally, the image coordinates are corrected proportionally:

$$\Delta x'_{rad} = x' \frac{\Delta r'_{rad}}{r'} \qquad\qquad \Delta y'_{rad} = y' \frac{\Delta r'_{rad}}{r'} \qquad\qquad (3.59)$$

The meaning of the distortion parameters

The sign of parameter A_1 (K_1) determines the form of the distortion as either barrel ($A_1 < 0$) or pincushion ($A_1 > 0$), see Fig. 3.56. A_2 and A_3 (K_2 and K_3) model deviations of the distortion curve from the cubic parabolic form with each term introducing an additional inflection. A_2 normally has its primary effect towards the edges of the image (Fig. 3.57). The introduction of the term A_3 makes it possible to model lenses with large distortion values at the image edges (Fig. 3.57).
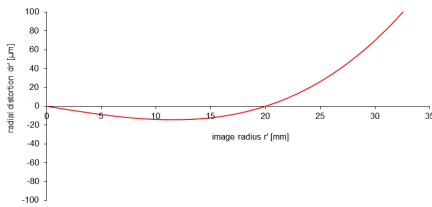
Parameters:
$A_1 = 4.664 \cdot 10^{-6}$ $\qquad A_2 = 0 \qquad A_3 = 0$
$r_0 = 20.0$ mm
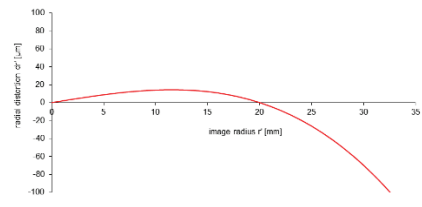
Parameters:
$A_1 = -4.664 \cdot 10^{-6}$ $\quad A_2 = 0 \qquad A_3 = 0$
$r_0 = 20.0$ mm

Fig. 3.56: Effect of changing the sign of A_1.
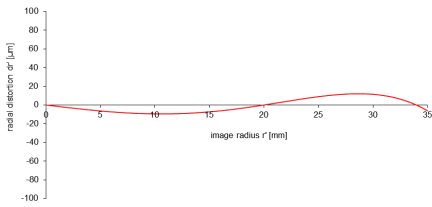
Parameters:
$A_1 = 4.664 \cdot 10^{-6}$ $\qquad A_2 = -3.0 \cdot 10^{-9} \qquad A_3 = 0$
$r_0 = 20.0$ mm
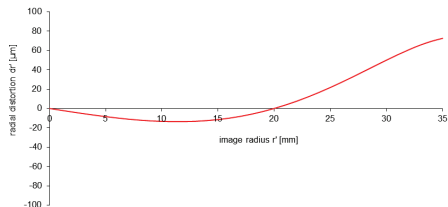
Parameters:
$A_1 = -4.664 \cdot 10^{-6}$ $\quad A_2 = 0 \qquad A_3 = -1.0 \cdot 10^{-12}$
$r_0 = 20.0$ mm

Fig. 3.57: Effects of parameters A_2 and A_3.

3.3.3.2 Tangential distortion

Tangential distortion (Fig. 3.58), also known as decentring distortion, is attributable to decentring and misalignment of individual lens elements from the ideal central axis of the camera's compound lens. It can be described by the following functions with the parameters P_1, P_2 or alternatively B_1, B_2:

$$\Delta x'_{tan} = P_1(r'^2 + 2x'^2) + 2P_2 x'y' \qquad \Delta x'_{tan} = B_1(r'^2 + 2x'^2) + 2B_2 x'y'$$
$$\Delta y'_{tan} = P_2(r'^2 + 2y'^2) + 2P_1 x'y' \qquad \Delta y'_{tan} = B_2(r'^2 + 2y'^2) + 2B_1 x'y' \qquad (3.60)$$
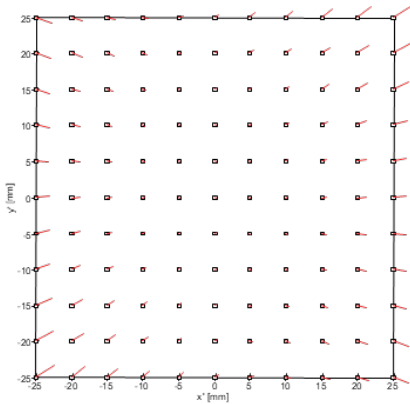
Fig. 3.58: Effect of tangential distortion.   **Fig. 3.59:** Effect of affinity and shear.
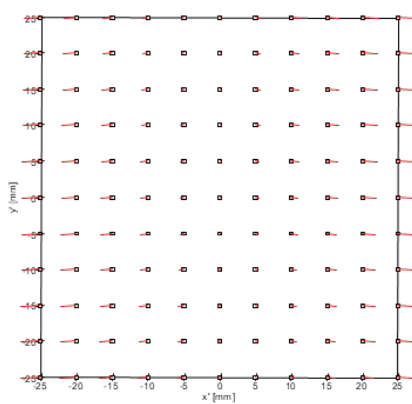
According to Conrady (1919) and Brown (1966), the complete approach for modelling asymmetric tangential distortion also consists of a coupled radially symmetric component described by the coefficients P_3 and P_4:

$$\Delta x'_{tan} = [P_1(r'^2 + 2x'^2) + 2P_2 x'y'](1 + P_3 r'^2 + P_4 r'^4)$$
$$\Delta y'_{tan} = [P_2(r'^2 + 2y'^2) + 2P_1 x'y'](1 + P_3 r'^2 + P_4 r'^4) \qquad (3.61)$$

P_3 and P_4 describe a non-linear effect that may have a larger impact at the image borders. Fig. 3.60 illustrates the effect for two image points P', Q', in different image positions. The imaged points are displaced asymmetrically by a radial and tangential component (first term in eqn. 3.61) which is itself scaled by another radial component (second term in eqn. 3.61). The radial term in eqn. (3.61) is not considered in most software solutions since it has a minor impact in good quality lenses and is otherwise highly correlated with the standard parameters for radial distortion.
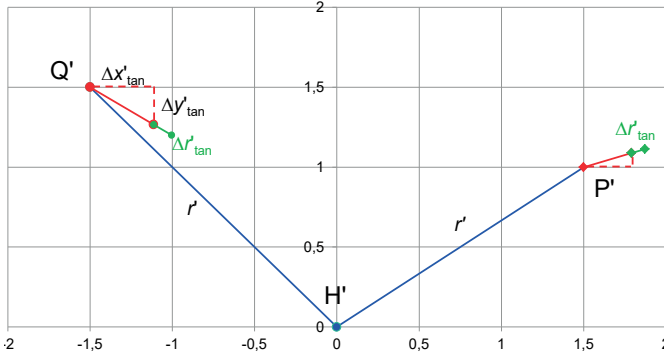
Fig. 3.60: Example of tangential distortion with a radially symmetric component.

For most quality lenses the effects of decentring distortion are significantly smaller than for radial distortion and are often only determined where accuracy demands are high. If low-cost lenses are used, as is often the case in surveillance or webcam systems, significant decentring distortion can be present.

3.3.3.3 Affinity and shear

Affinity and shear are used to describe deviations of the image coordinate system with respect to orthogonality and uniform scale of the coordinate axes (Fig. 3.59). The effect can have several causes:

–   sensors which are not aligned orthogonally to the optical axis of the camera lens;
–   sensor elements which are not distributed on a regular grid or which are not square;
–   video images which are transferred in analogue form before being digitized by a frame grabber;
–   imaging sensors with rolling shutters (see section 3.4.2.2);
–   scanned analogue films.

The following function can be used to provide an appropriate correction:

$$\Delta x'_{aff} = C_1 x' + C_2 y' \qquad\qquad \Delta y'_{aff} = 0 \qquad\qquad (3.62)$$

A similar effect is modelled by the parameters m' and s' of the calibration matrix in eqn. (3.66). Affinity and shear are not associated with the lens but are distortions occurring within the sensor and image planes. If relevant, these should be processed in the order shown below for total distortion correction. The C parameters are close to zero in many modern digital cameras (sensors) and need only be calibrated for very high-accuracy applications.

3.3.3.4 Total correction

The simplified, but widely used, definition of total distortion correction is obtained by simple summing of the individual correction terms:

$$\Delta x' = \Delta x'_{rad} + \Delta x'_{tan} + \Delta x'_{aff} \qquad \Delta y' = \Delta y'_{rad} + \Delta y'_{tan} + \Delta y'_{aff}$$
$$x' = x^\circ - \Delta x' = x'_P - x'_0 - \Delta x \qquad y' = y^\circ - \Delta y' = y'_P - y'_0 - \Delta y' \qquad (3.63)$$

Strictly speaking, the measured image coordinates must initially be corrected by the sensor-related affinity and shear terms. Subsequently, these pre-corrected coordinates are used to calculate the corresponding corrections for lens-related radial and decentring distortion. The total correction then becomes:

$$x'' = x' - \Delta x'_{aff} \qquad y'' = y' - \Delta y'_{aff}$$
$$\Delta x' = \Delta x'_{rad}(x'',y'') + \Delta x'_{tan}(x'',y'') \qquad \Delta y' = \Delta y'_{rad}(x'',y'') + \Delta y'_{tan}(x'',y'') \qquad (3.64)$$
$$x' = x^\circ - \Delta x' \qquad y' = y^\circ - \Delta y'$$

Which of the two approaches should be used to correct measured image coordinates depends on the software used for camera calibration. It should also be noted that the sign of the correction terms and the assumption as to the pixel origin, e.g. top left, centre or bottom right, may vary between different programs. Careful reading of software documentation is therefore strongly recommended if data is to be exchanged between different software packages.

Example 3.8:

A calibration of a Canon EOS1000 digital camera with lens $f = 18$ mm gives the following set of correction parameters:

$A_1 = -4.387 \cdot 10^{-4}$	$A_2 = 1.214 \cdot 10^{-6}$	$A_3 = -8.200 \cdot 10^{-10}$
$B_1 = 5.572 \cdot 10^{-6}$	$B_2 = -2.893 \cdot 10^{-6}$	
$C_1 = -1.841 \cdot 10^{-5}$	$C_2 = 4.655 \cdot 10^{-5}$	$r_0 = 8.325$ mm

Compute the effects of individual distortion terms for two image points, the first located in the centre of the image and the second in one corner:

	x'_1	y'_1	x'_2	y'_2
x',y'	1.500	1.500	11.100	7.400 mm
A_1, A_2, A_3	34.4	34.4	−215.4	−143.6 µm
B_1, B_2	0.0	0.0	1.9	0.2 µm
C_1, C_2	0.0	0.0	0.1	0.0 µm
Total	34.4	34.4	−213.4	−143.4 µm

This example indicates that the effect of radial distortion predominates. However, if the accuracy potential of this camera of about 0.2 µm ($1/50^{th}$ pixel) is to be reached, the other sources of image errors must be taken into account.

3.3.4 Alternative correction formulations

3.3.4.1 Simplified models
In Tsai's approach, only one parameter k is used in addition to principal distance, principal point and two scaling parameters. The distorted image coordinates, with origin at the principal point, are given by:

$$x° = \frac{2x'}{1+\sqrt{1-4kr'^2}} \qquad\qquad y° = \frac{2y'}{1+\sqrt{1-4kr'^2}} \qquad (3.65)$$

This widely used model does not include any high-order distortion effects and is not therefore equivalent to a Seidel polynomial with two or three parameters (A_1, A_2, A_3). Typical lenses with large distortion components are not fully modelled by this approach and the disadvantage increases towards the image edges. In addition, asymmetric and tangential distortions are not taken into account at all.

In computer vision the parameters of interior orientation are typically expressed by a *calibration matrix* **K** which consists of five degrees of freedom (principal distance c, principal point x'_0, y'_0, shear s and differential scale m' between the axes).

$$\mathbf{K} = \begin{bmatrix} 1 & s' & x'_0 \\ 0 & 1+m' & y'_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} c & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} c & cs' & x'_0 \\ 0 & c(1+m') & y'_0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (3.66)$$

K is part of a general 3x4 transformation matrix from object space into image space by homogeneous coordinates (see section 2.2.3.1). Lens distortion cannot directly be integrated in this method and must be modelled by a position-dependent correction matrix $\mathbf{dK}(x', y')$.

In computer vision, camera parameters are often based on pixel units. The transformation of pixel-based calibration parameters into metric parameters is described in section 3.3.6.

3.3.4.2 Additional parameters
An approach by Brown (1971) was developed specifically for large format analogue aerial cameras, but can also be applied to digital cameras:

$$\begin{aligned}
\Delta x'_{Brown} &= D_1 x' + D_2 y' + D_3 x'y' + D_4 y'^2 + D_5 x'^2 y' + D_6 x'y'^2 + D_7 x'^2 y'^2 + \\
&\quad [E_1(x'^2 - y'^2) + E_2 x'^2 y'^2 + E_3(x'^4 - y'^4)]x'/c + \\
&\quad [E_4(x'^2 + y'^2) + E_5(x'^2 + y'^2)^2 + E_6(x'^2 + y'^2)^3]x' + E_7 + E_9 x'/c \\
\Delta y'_{Brown} &= D_8 x'y' + D_9 x'^2 + D_{10} x'^2 y' + D_{11} x'y'^2 + D_{12} x'^2 y'^2 + \\
&\quad [E_1(x'^2 - y'^2) + E_2 x'^2 y'^2 + E_3(x'^4 - y'^4)]y'/c + \\
&\quad [E_4(x'^2 + y'^2) + E_5(x'^2 + y'^2)^2 + E_6(x'^2 + y'^2)^3]y' + E_8 + E_9 y'/c
\end{aligned} \qquad (3.67)$$

In addition to parameters for modelling deformations in the image plane (D_1 to D_{12}) terms for the compensation of lack of flatness in film or sensor are included (E_1 to E_6). These are formulated either as a function of the radial distance or of the tangent of the imaging angle (x'/c or y'/c).

Additional correction models, based on polynomials, were created by Ebner (1978) and Grün (1978) for the calibration of analogue aerial cameras:

$$\Delta x'_{Ebner} = E_1 x' + E_2 y' - E_3 2x'^2 + E_4 x' y' + E_5 y'^2 + E_7 x' y'^2 + E_9 x'^2 y' + E_{11} x'^2 y'^2$$
$$\Delta y'_{Ebner} = -E_1 y' + E_2 x' + E_3 x' y' - E_4 2y'^2 + E_6 x'^2 + E_8 y' x'^2 + E_{10} y'^2 x' + E_{12} x'^2 y'^2$$

(3.68)

$$
\begin{aligned}
\Delta x'_{Grün} = {} & G_1 x' + G_2 y' + G_3 x' y' + G_4 y'^2 - G_6 (10/7) x'^2 + G_7 x'^3 + G_8 x'^2 y' \\
& + G_9 x' y'^2 + G_{10} y'^3 + G_{15} x'^4 + G_{16} x'^3 y' + G_{17} x'^2 y'^2 + G_{18} x' y'^3 \\
& + G_{19} y'^4 + G_{25} y' x'^4 + G_{26} x'^3 y'^2 + G_{27} x'^2 y'^3 + G_{28} x' y'^4 + G_{33} x'^3 y'^3 \\
& + G_{35} x'^2 y'^4 + G_{39} x'^4 y'^3 + G_{40} x'^3 y'^4 + G_{43} x'^4 y'^4 \\
\Delta y'_{Grün} = {} & -G_1 y' + G_2 x' - G_3 (10/7) y'^2 + G_5 x'^2 + G_6 x' y' + G_{11} x'^3 + G_{12} x'^2 y' \\
& + G_{13} x' y'^2 + G_{14} y'^3 + G_{20} x'^4 + G_{21} x'^3 y' + G_{22} x'^2 y'^2 + G_{23} x' y'^3 \\
& + G_{24} y'^4 + G_{29} y' x'^4 + G_{30} x'^3 y'^2 + G_{31} x'^2 y'^3 + G_{32} x' y'^4 + G_{37} x'^3 y'^3 \\
& + G_{38} x'^2 y'^4 + G_{41} x'^4 y'^3 + G_{42} x'^3 y'^4 + G_{44} x'^4 y'^4
\end{aligned}
$$

(3.69)

However, these polynomials are geometrically difficult to interpret. Other approaches to camera modelling based on Legendre or Chebyshev polynomials show similar characteristics to high-order polynomials. Fourier series can also be used in which coefficients are inherently uncorrelated and independent in x and y. Here again, however, there are a large number of parameters which are difficult to handle and interpret.

In summary, the above approaches are more difficult to interpret geometrically and can easily be over-parameterised, leading to a dependency or correlation between individual parameters. As a result, these approaches are not often used in photogrammetry.

For a self-calibrating bundle adjustment, individual parameters should be tested for their significance and correlation with respect to each other. Any parameters which fail such tests should be eliminated, starting with the weakest first.

An example of an extension to the parameter set to accommodate digital cameras is given below:

$$
\begin{aligned}
\Delta x'_{Beyer} = {} & \Delta x'_0 - \Delta c\, x'/c + K_1 x' r'^2 + K_2 x' r'^4 + K_3 x' r'^6 \\
& + P_1 (r'^2 + 2x'^2) + 2P_2 x' y' - C_1 x' + C_2 y' \\
\Delta y'_{Beyer} = {} & \Delta y'_0 - \Delta c\, y'/c + K_1 y' r'^2 + K_2 y' r'^4 + K_3 y' r'^6 \\
& + 2P_1 x' y' + P_2 (r'^2 + 2y'^2) + C_2 x'
\end{aligned}
$$

(3.70)

Here the parameters K describe radial distortion, P describes tangential (decentring) distortion and C models affinity and shear. The parameters $\Delta x'_0$, $\Delta y'_0$ and Δc are used for small corrections to the spatial position of the perspective centre. Together with the factors x'/c and y'/c they have an effect similar to r_0 in eqn. (3.57).

The correction of radial distortion according to (3.55) can be formulated as a function of image angle rather than image radius:

$$\Delta r = W_1\theta^2 + W_2\theta^3 + W_3\theta^4 + \ldots\ldots + W_i\theta^{i+1} \tag{3.71}$$

where

$$\theta = \arctan\left(\frac{1}{c}\sqrt{x'^2 + y'^2}\right) \qquad \text{: angle between optical axis and object point}$$

x', y':   corrected image coordinates, projection of the object point in the image

c: principal distance
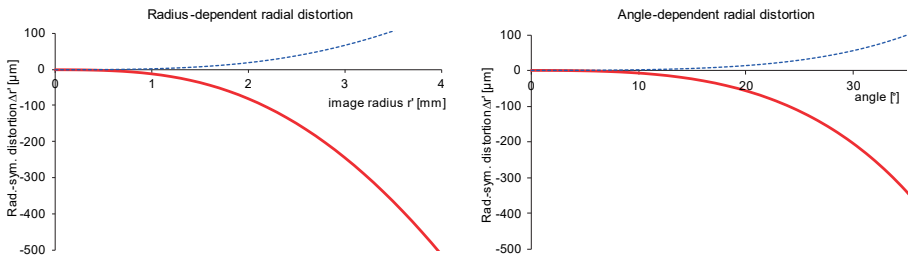
Fig. 3.61: Radial distortion (red) as a function of radius (left) and image angle (right) for a lens with large distortion (Basler, see Table 3.2), with the uncertainty of the distortion curve for each case shown in blue.

A distortion model based on angle is particularly suitable for central perspective optics with large symmetric radial distortion. In this case the distortion curve has a lower slope into the corners (Fig. 3.61).

The increasing uncertainty is in general due to a reduced density of imaged points with increasing image radius.

In aspheric wide-angle lenses with aspherical components, the radial lens distortion gradients can be sharper and more varied than for spherical lenses. The conventional odd-power polynomial model (section 3.3.3.1) can lead to significant systematic errors in image distortion corrections (see Fig. 7.37). In such cases it can be useful to apply additional even-power terms to the classical Seidel polynomial.

3.3.4.3 Correction of distortion as a function of object distance

Strictly speaking, the above approaches for the correction of lens distortion are valid only for points on an object plane that is parallel to the image plane, and focused,

according to the lens eqn. (3.14), at a distance a' or with respect to the interior perspective centre. Imaging rays of points outside this object plane pass through the lens along a different optical path and hence are subject to different distortion effects.

This effect can be taken into account by a correction dependent on distance. This requires introduction of a scaling factor:

$$\gamma_{SS'} = \frac{c_{S'}}{c_S} = \frac{S'}{S} \cdot \frac{(S-c)}{(S'-c)} \tag{3.72}$$

where c_S: principal distance (image distance) of object distance S
                     $c_{S'}$: principal distance (image distance) of object distance S'

For a given set of distortion parameters $K_{1S'}$, $K_{2S'}$, $K_{3S'}$ applying to an object plane at distance S', according to (3.55), the correction of radial distortion for object points at a focused distance S can be calculated as follows:

$$\Delta r'_{SS'} = \gamma^2_{SS'} K_{1S'} r'^3 + \gamma^4_{SS'} K_{2S'} r'^5 + \gamma^6_{SS'} K_{3S'} r'^7 \tag{3.73}$$

This model is suitable for high-precision measurements made at large scales ($m < 30$) and lenses with relatively steep distortion curves. As the effect of distortion dependent on distance increases with image scale (decreasing object distance), an empirically estimated correction factor $g_{SS'}$ can be introduced that can be determined for each individual lens:

$$\Delta r'_{SS'} = \Delta r'_S + g_{SS'} (\Delta r'_{S'} - \Delta r'_S) \tag{3.74}$$

As an example, Fig. 3.62 shows the distortion curves of object points with different image scales, i.e. at different object distances.
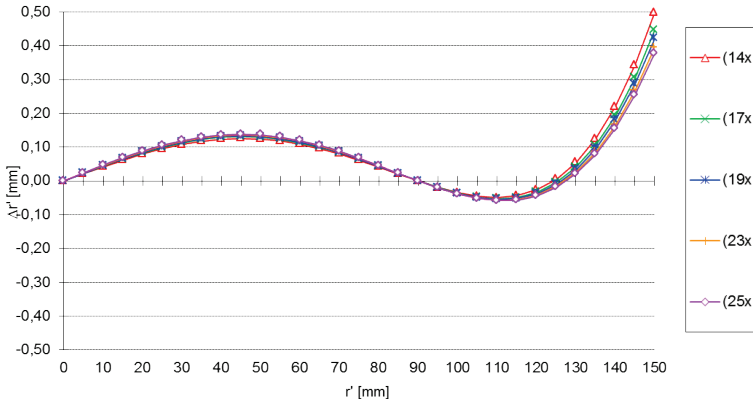
Fig. 3.62: Lens distortion curves for different image scales (Dold 1997).

Tangential distortion can also be formulated as a function of the focused distance S:

$$\Delta x'_S = \left(1 - \frac{c}{S}\right)\left[P_1(r'^2 + 2x') + 2P_2 x'y'\right]$$
$$\Delta y'_S = \left(1 - \frac{c}{S}\right)\left[P_2(r'^2 + 2y') + 2P_1 x'y'\right]$$

$$(3.75)$$

In contrast to the above approaches, the following set of parameters for the correction of distance-dependent distortion can be estimated completely within a self-calibrating bundle adjustment. However, it should be noted that incorporation of such corrections within a self-calibration require very strong networks that contain many images taken at each distance setting in order to provide a robust and reliable parameter set.

$$\Delta r'_{dist} = \frac{1}{Z^*}\left[D_1 r'(r'^2 - r_0^2) + D_2 r'(r'^4 - r_0^4) + D_3 r'(r'^6 - r_0^6)\right]$$

$$(3.76)$$

where Z^*: denominator of collinearity equations (4.9) $\approx S$ (object distance)

Extension of (3.63) leads to the following total correction of imaging errors:

$$\Delta x' = \Delta x'_{rad} + \Delta x'_{tan} + \Delta x'_{aff} + \Delta x'_{dist} \qquad \Delta y' = \Delta y'_{rad} + \Delta y'_{tan} + \Delta y'_{aff} + \Delta y'_{dist} \qquad (3.77)$$

Usually, distance-dependant distortion does not exceed more than 1μm at the edge of the image. Hence, it must only be considered for high-accuracy measurement tasks where sub-micron image measuring accuracies are required. This is relevant to large-scale industrial applications with analogue large-format cameras, but is particularly relevant to high-resolution digital cameras that provide an accuracy potential of better than 0.5 μm. The need will be most prevalent where there is a large range of depth over the scene to be recorded or if the scene is highly linear and requires wide angle oblique views from close ranges. An example of this last case might be an aircraft wing jig located in the confined working environment of a factory.

3.3.4.4 Image-variant calibration

Cameras and lenses, which have so little mechanical stability that the geometry of the imaging system can vary within a sequence of images, can be modelled by an image-variant process. Here individual parameters defining the perspective centre (principal distance and principal point) are determined individually for every image j. In contrast, distortion parameters are normally assumed to be constant for the entire image sequence. Image coordinate adjustment is then done with correction terms which are calculated for every image as a function of the perspective centre's location.

$$\Delta x'_{var} = \Delta x'_{\{\Delta c, \Delta x'_0, \Delta y'_0\}_j} \qquad\qquad \Delta y'_{var} = \Delta y'_{\{\Delta c, \Delta x'_0, \Delta y'_0\}_j} \qquad (3.78)$$

Fig. 3.50 shows an example for a digital camera, where principal distance and principal point vary due to mechanical handling and the effects of gravity. The calculation of image-variant camera parameters is done using an extended bundle adjustment.

3.3.4.5 Correction of local image deformations

Local image deformations are imaging errors which appear only in specific areas of the image format and are not therefore covered by the global correction functions described above.
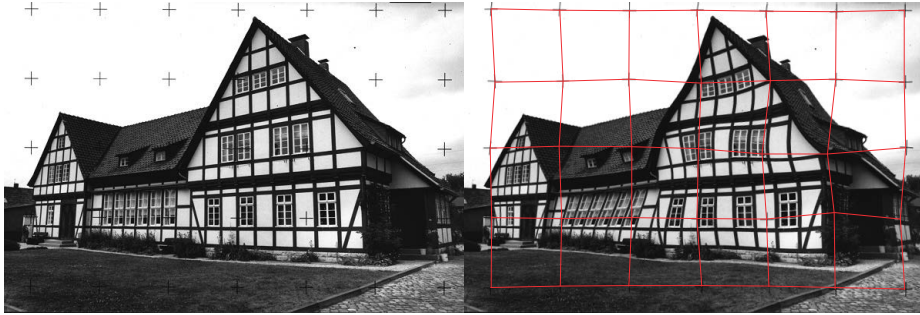
Fig. 3.63: Example of simulated film deformation on an image with a réseau.

Local deformations are due, for example, to local lack of flatness in the image plane. The lack of flatness often seen in analogue film cameras can be corrected by application of the réseau method, i.e. by local geometric transformation of the image onto a reference grid (Fig. 3.63).

A similar analytical formulation can be adopted to describe the lack of flatness of digital image sensors. A finite element approach, based on localized correction nodes, is used. This requires a grid of two-dimensional correction vectors at the intersection points of the grid lines. Corrections within the grid are computed using an interpolation process, typically according to the linear process in the following equation (Fig. 3.64):

$$
\begin{aligned}
\Delta x'_{corr} &= (1 - x_l - y_l + x_l \cdot y_l) \cdot k_{x[i,j]} & \Delta y'_{corr} &= (1 - x_l - y_l + x_l \cdot y_l) \cdot k_{y[i,j]} \\
&\quad + (x_l - x_l \cdot y_l) \cdot k_{x[i+1,j]} & &\quad + (x_l - x_l \cdot y_l) \cdot k_{y[i+1,j]} \\
&\quad + (y_l - x_l \cdot y_l) \cdot k_{x[i,j+1]} & &\quad + (y_l - x_l \cdot y_l) \cdot k_{y[i,j+1]} \\
&\quad + x_l \cdot y_l \cdot k_{x[i+1,j+1]} & &\quad + x_l \cdot y_l \cdot k_{y[i+1,j+1]}
\end{aligned}
\tag{3.79}
$$
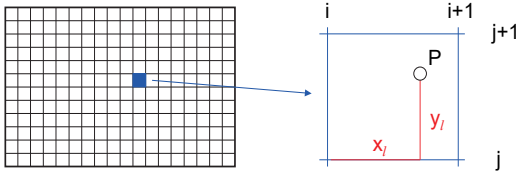
Fig. 3.64: Interpolation within the correction grid.

Here x_{corr} is the corrected value of the image coordinate x, coordinates x_l, y_l represent the local position of the image point within the grid element and $k_{x[i,j]}$, $k_{x[i+1,j]}$, $k_{x[i,j+1]}$, $k_{x[i+1,j+1]}$ are the correction vector components at the corresponding grid intersection points. An analogous correction applies to image coordinate y. The collinearity equations (4.9) are extended by the above formulation so that the grid parameters can be estimated by a bundle adjustment (see section 4.4).

In order to separate out the noise component, i.e. the random measurement error in the image point, from the sensor deformation error and other lens imaging errors not otherwise considered, deformation conditions at the nodes are introduced as pseudo-equations:

$$
\begin{aligned}
0 &= (k_{x[i,j-1]} - k_{x[i,j]}) - (k_{x[i,j]} - k_{x[i,j+1]}) \\
0 &= (k_{x[i-1,j]} - k_{x[i,j]}) - (k_{x[i,j]} - k_{x[i+1,j]}) \\
0 &= (k_{y[i,j-1]} - k_{y[i,j]}) - (k_{y[i,j]} - k_{y[i,j+1]}) \\
0 &= (k_{y[i-1,j]} - k_{y[i,j]}) - (k_{y[i,j]} - k_{y[i+1,j]})
\end{aligned}
\tag{3.80}
$$

These nodal conditions are applied in both x and y directions in the image plane. Within the system of equations, this leads to a new group of observations. The accuracy with which the equations are introduced depends on the expected "roughness" of the unflatness parameters in the correction grid, as well as the number of measured points in the imaging network which appear on each grid element. In addition, the equations protect against potential singularities within the complete set of adjustment equations when there are grid elements which contain no image points.

As an example, Fig. 3.65 shows a calculated correction grid for a digital camera (Canon EOS1000) fitted with a zoom lens. The grids show similar trends despite setting the lens at two different zoom settings. In addition to the simultaneously determined distortion parameters, effects are modelled in image space which have similar characteristics in both cases. However, these are not necessarily pure sensor deformations such as lack of flatness. The finite element grid compensates for all remaining residual errors in the image sequence.

f = 18 mm   f = 28 mm
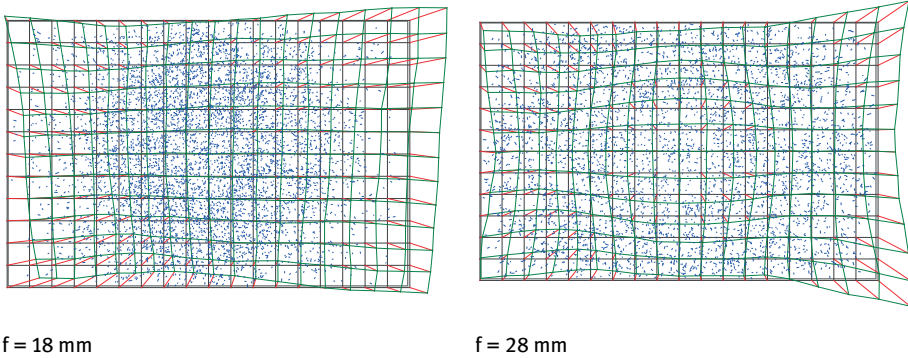
Fig. 3.65: Computed correction grids for the same camera body with a zoom lens at different settings.

The finite element method of calibration is also suitable for modelling ray paths in complex optics, for example when using a stereo-mirror attachment (see Fig. 3.110).

3.3.4.6 Chromatic aberration

In principle, all image distortions caused by the lens depend on wavelength. Strictly speaking, RGB colour cameras or multi-spectral cameras should be calibrated separately for each channel.

Red – Green   Blue – Green
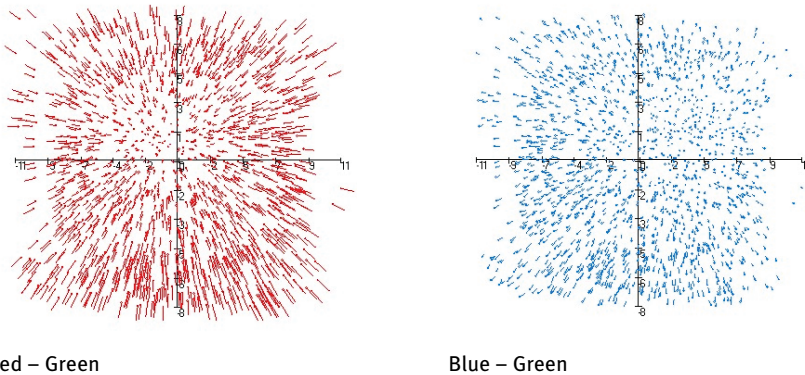
Fig. 3.66: Image measurements compared in separate colour channels (Fuji S2 Pro, f = 20 mm).

Fig. 3.66 shows the difference vectors between measured image points from the separate RGB channels of a true-colour image. The radially symmetric distortion, corresponding to the transverse chromatic aberration (section 3.1.3.2), can be clearly observed.

For self-calibration, at every imaging location each colour channel is treated as a separate camera with the additional constraint that the corresponding exterior orientations are identical. The adjusted parameters of interior orientation then represent the physical shift in focus due to colour, i.e. the principal distance of the green channel lies between those of the red and blue channel (see example in Table 3.2). In contrast, the position of the optical axis (principal point x'_0 und y'_0) does not change significantly.

Table 3.2: Principal point position [mm] after calibration using separate colour channels (Fuji S2 Pro, f = 20 mm).

Parameter	Red	Green	Blue
c	−20.5739	−20.5557	−20.5468
x'_0	0.2812	0.2818	0.2812
y'_0	−0.2080	−0.2095	−0.2105

3.3.5 Iterative correction of imaging errors

Measured image points can be corrected a priori if the parameters of interior orientation are known. Example cases are the calculation of object coordinates by space intersection, or the resampling of distortion-free images. However, there are often misunderstandings about the sequence and sign of corrections to image errors. For clarity it is necessary to know the derivation of the correction values and details about the software implementation, which is often not available in practice.

The correction model described in (3.63) assumes that the error corrections are calculated by self-calibrating bundle adjustment. This incorporates the collinearity model (4.9), which enables image points to be calculated for points defined in object space. Any possible distortion values are then added to these image coordinates.

$$x' = f(X,Y,Z,X_0,Y_0,Z_0,\omega,\varphi,\kappa,c,x'_0) + \Delta x'(x',y')$$
$$y' = f(X,Y,Z,X_0,Y_0,Z_0,\omega,\varphi,\kappa,c,y'_0) + \Delta y'(x',y')$$

(3.81)

After adjustment, the correction parameters relate to adjusted object coordinates and orientation parameters or, expressed differently, to "error-free" image coordinates. Consequently, these correction values are directly applicable to image coordinates calculated by applying the collinearity equations from object space to image space. However, since the corrections $\Delta x'$, $\Delta y'$ depend on the current values of image coordinates, in the case where image coordinates, measured in a distorted image, must be corrected for previously established distortion parameters, then corrections must be applied iteratively. In this process, the currently corrected image positions

are the starting point for the subsequent calculation of corrections. The process continues until the computed corrections are insignificant.

a) Behaviour for lenses with normal distortion

b) Behaviour for lenses with high levels of distortion
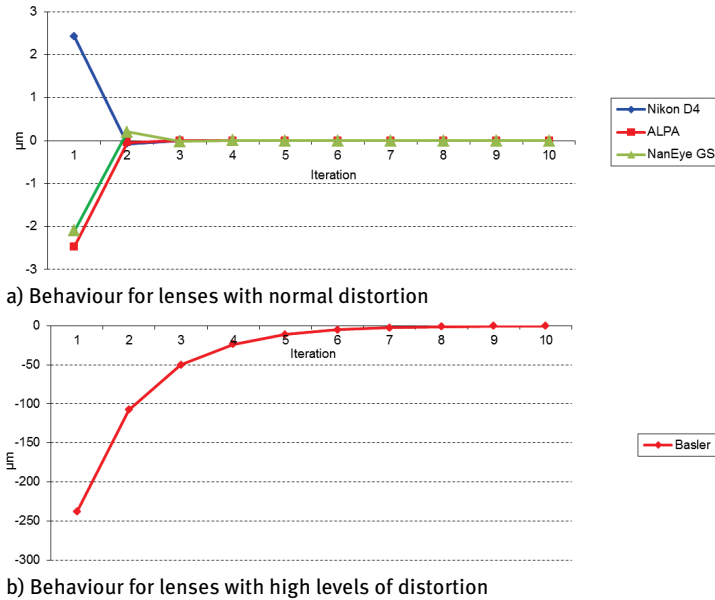
Fig. 3.67: Iterative correction of distortion.

The iterative correction of image coordinates is of particular practical importance when lenses with high levels of distortion are used. Table 3.3 summarizes the effect of iterative correction of image coordinates for typical camera systems. The row labelled "Total correction" shows the effect of iterative correction. It is obvious that cameras with wide angles of view (short focal lengths) and high distortion parameters (see row labelled "corrections due to further iterations") have distortion errors which cannot be ignored without iterative correction (ALPA: −2.5μm, Nikon D4: 2.4μm). Fig. 3.67a shows that 3-4 iterations are necessary for a final correction. Cameras with low-distortion lenses or small angles of view can be corrected without iteration (JAI: 0.2 μm). Relatively large values occur for the example micro camera where iterative correction is also necessary (NanEye GS: −2.1μm).

If a high-distortion wide-angle lens is evaluated, e.g. the Basler camera with a 4.8mm Pentax lens (see sample images in Fig. 3.68), then correction values have two features of note. Firstly, they are high, in this case 238 μm. Secondly, they are generated in a slow convergence process (Fig. 3.67b) which may even diverge in the corners.

Table 3.3: Effects of iterative distortion correction for typical camera systems (mm).

Camera	ALPA 12 FPS	Nikon D4	JAI	Basler	NanEye GS
Focal length (mm)	40	24	8	4.8	6
Sensor format (mm)	44 x 33	36 x 24	6.5 x 4.8	9.0 x 6.7	2.3 x 2.3
c	−41.9600	−35.6657	−8.2364	−4.3100	−6.006917
x'_0	-0.50598	−0.0870	0.2477	−0.0405	-0.02647737
y'_0	-0.21847	0.4020	0.0839	−0.0011	-0.1640278
A_1	−2.4799E-05	−9.0439E-05	−1.9016E-03	−1.4700E-02	2.1644E-03
A_2	1.3858E-08	6.3340E-08	3.1632E-05	3.0396E-04	1.2617E-03
A_3	−2.4341E-12	−1.5294E-11	0	3.4731E-06	2.8721E-04
B_1	−2.4965E-06	2.5979E-06	−2.6175E-05	1.1358E-05	3.1584E-05
B_2	2.8628E-06	−7.2654E-07	−7.3739E-05	−4.5268E-05	6.5409E-05
C_1	2.6503E-05	1.0836E-04	9.8413E-03	1.9833E-05	6.6634E-05
C_2	2.3514E-05	1.0964E-05	−7.2607E-06	8.9119E-05	−6.1250E-05
x'	21.942	18.000	3.200	4.400	1,152
y'	16.451	12.000	2.400	3.300	1,152
Imaging angle	33.68°	41.89°	24.49°	52.08°	16.39°
Total correction	−0.1351	−0.0796	0.0045	−0.5674	0.0206
Corrections due to further iterations	−0.0025	0.0024	−0.0002	−0.2385	−0.0021
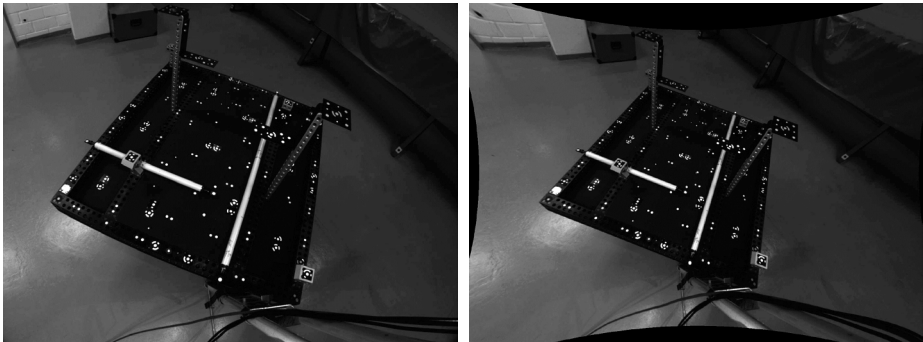
Fig. 3.68: Example of a significantly distorted image (left) and its corrected version (right).

3.3.6 Transformation of interior orientation parameters

Different software systems allow for camera calibration based either on pixel coordinates or metric coordinates. If those parameters are used for further processing

in another program, they can usually be transformed from one unit to the other as required.

Table 3.4: Transformation of camera parameters from pixel to metric units.

Parameter	metric	pixel-based	Transformation	Unit
Principal distance	c	f	$c = f \cdot \Delta s'$	mm
Principal point	x'_0	u_0	$x'_0 = u_0 \cdot \Delta s'$	mm
	y'_0	v_0	$y'_0 = -v_0 \cdot \Delta s'$	
Radial distortion	A_1	K_1	$A_1 = K_1 / c^2$	$1/mm^{-2}$
	A_2	K_2	$A_2 = K_2 / c^4$	$1/mm^{-4}$
	A_3	K_3	$A_3 = K_3 / c^6$	$1/mm^{-6}$
Decentring distortion	B_1	P_1	$B_1 = P_1 / c$	$1/mm$
	B_2	P_2	$B_2 = -P_2 / c$	$1/mm$
Affinity and shear	C_1	B_1	$C_1 = B_1 / f$	
	C_2	B_2	$C_2 = -B_2 / f$	

Table 3.4 summarizes the transformation equations from pixel to metric units. It is assumed that the origin of the pixel coordinate system is located in the image centre and that the pixels are quadratic, i.e. $\Delta s' = \Delta s'_{x'} = \Delta s'_{y'}$, (see section 3.3.2.1). For absolute metric parameters, the physical pixel size $\Delta s'$ must be known. There is a corresponding reverse transformation from metric to pixel units.

3.3.7 Fisheye projections

As the angle of view of an imaging system increases beyond about 110°, optical performance decreases rapidly. Degradation in image quality is seen in the capability to correctly image straight lines in the object as straight lines in the image and in reduced illumination towards the extremes of the image format following the cos^4 law (section 3.1.3.5). As an example a 110° angle of view would be given by a 15 mm focal length lens on a full frame, FX format digital SLR sensor (Fig. 3.78). Such lenses are available, but extremely expensive.

A solution is to change the imaging geometry from the central perspective projection where incident and exit angles τ and τ' are equal, to one where the incident angle τ from a point P in object space is greater than the exit angle τ' in image space (Fig. 3.69). Fisheye designs allow the projection of a half hemisphere onto the image plane with the optical axis coinciding with the centre of the resultant circular image.

If the image format is larger than the resultant image circle, the camera is termed a fisheye system. Conversely, if the format is smaller than the circle, such that the image diagonal is 180°, a quasi-fisheye system is produced.
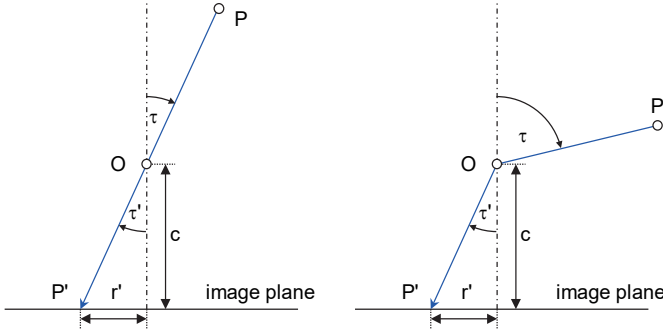
Fig. 3.69: Central perspective projection (left) and a generic fisheye projection (right) (after Schneider 2008).

Three fisheye projections are in optical usage: stereographic, equidistant and orthographic. They are defined by the following equations.

$$r' = c \cdot \tan \tau \qquad \qquad \text{: central perspective} \qquad \qquad (3.82)$$

$$r' = 2c \cdot \tan \tau / 2 \qquad \qquad \text{: stereographic} \qquad \qquad (3.83)$$

$$r' = c \cdot \tau \qquad \qquad \text{: equidistant} \qquad \qquad (3.84)$$

$$r' = c \cdot \sin \tau \qquad \qquad \text{: orthographic} \qquad \qquad (3.85)$$

When modelling the distortions in a fisheye lens, conventional radial lens distortion corrections (eqn. 3.55) are mathematically unstable beyond the central region of the image format where the gradient of the distortion curve describing the departure from the central perspective case is low. For sensors that capture a significant area of the fisheye image circle, for example a DX sensor with a 10.5 mm fisheye lens or an FX sensor with a 16 mm fisheye lens, it is necessary to apply the appropriate fisheye lens model before using a radial distortion model such as eqn. (3.55), in order to account for any remaining radial departures from the lens's fisheye projection. Alternatively, a pre-correction can be calculated by interpolation using a lens map function which may be available from the lens manufacturer (see section 3.3.3.1).

Fisheye cameras and their applications are presented in sections 3.4.3.4 and 3.5.6.3. Aspects of fisheye lens calibration are discussed in section 7.3.3.2.
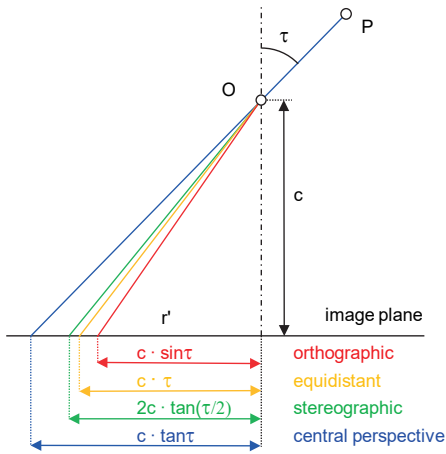
Fig. 3.70: Central projection and spherical projections.

3.4 System components

Electronic imaging systems use opto-electronic sensors for image acquisition instead of photographic emulsions. They directly provide an electronic image that can be digitized by suitable electronic components and transferred to a local processor or host computer for measurement and analysis. Hence the term *electronic imaging system* summarizes all system components involved in the generation of a digital image (Fig. 3.71).
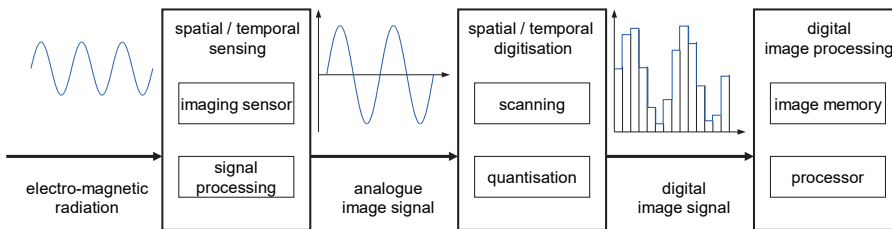
Fig. 3.71: Electronic imaging system.

The electro-magnetic radiation (light) emitted or reflected by the object is imaged by a sensor as a function of time (exposure time, integration time) and space (linear or area sensing). After signal enhancement and processing an analogue image signal, in the form of an electric voltage proportional to the amount of light falling on the sensor, is produced. In a second stage, this signal is sampled by means of an analogue

to digital converter in order to produce a digital image consisting of a series of discrete numerical values for each light sensitive cell or pixel in the image. This digital image can then be used for further processing such as discrete point measurement, or edge detection.

As far as photogrammetry is concerned, where geometrically quantifiable images are required, the development of digital imaging technology is closely related to the technology of CCD image sensors (charge coupled devices). Invented at the beginning of the 1970s, they dominated the digital imaging market and photogrammetric applications for two decades and provided solutions to many imaging challenges. More recently, CMOS opto-electronic imaging sensors have become the dominant technology (see section 3.4.1.3). Manufactured with silicon fabrication processes shared with CPU and memory chips, they are more affordably mass produced than CCDs and dominate mobile phone and consumer imaging devices.

Although camera products are developing into a product continuum, the following definitions of imaging systems are useful to characterise those typically used in photogrammetry (status 2023):

– Mobile phone cameras:
  Miniaturised consumer cameras integrated into mobile phones and delivering still frames in standard image formats (JPEG, TIFF etc.) or digital video signals in real-time (25 to 30 frames per second, fps) in standard video formats (AVI, MPEG etc.). The number of pixels varies between ca. 780 x 580 pixels (corresponding to older video standards such as PAL or NTSC), 1280 x 780 pixels (HD), 1980 x 1080 Pixel (Full HD) up to ca. 4096 x 2160 pixels (4K standard) and 8192 x 6144 (8K standard). These systems are often equipped with multiple sensors and lenses in the same device to give wide, normal and narrow angle fields of view.

– Action cameras:
  Physically rugged and often waterproof cameras for the consumer market delivering images between ca. 3000 x 2000 (6 Mpixel) and ca. 6000 x 4000 pixels (24 Mpixel). Possible frame rates vary between ca. 10 fps and 0.5 fps depending on sensor resolution and internal memory design. These range from ruggedized fixed lens devices to units able to take interchangeable lenses. Examples include products from GoPro, Nikon, Olympus and Canon.

– High-resolution digital cameras:
  Cameras for professional photographic and technical applications with pixel numbers between ca. 4000 x 3000 (12 Mpixel) and ca. 12000 x 8000 pixel (100 Mpixel). Most products use either imaging sensors in DX format (18 mm x 12 mm) or FX full format (36 mm x 24 mm). New developments already achieve resolutions of 150 Mpixel up to 250 Mpixel for DSLR cameras. Specialized cameras for digital aerial images can now deliver 350 Mpixel with a single sensor.

– High-speed cameras:
  Digital high-speed cameras provide typical frame rates between 500 and 2000 fps. The number of pixels ranges from ca. 1000 x 1000 to ca. 2000 x 2000 pixels.

The total recording time is limited by internal or external storage media. For special cases very high frame rates of more than 1 million fps can be reached with highly reduced image sizes.

– Scanning cameras:

Imaging systems which increase the pixel density or image format by sequential scanning using a movable imaging sensor. Different principles (see section 3.5.5) enable pixel numbers to be increased from ca. 3000 x 2300 pixels up to ca. 20000 x 20000 pixels. These are useful only for stationary imaging situations. Scanning panorama cameras with CCD line sensors provide image sizes up to ca. 50000 x 10000 pixels.

3.4.1 Opto-electronic imaging sensors

3.4.1.1 Principle of CCD sensor

Solid-state imaging sensors are exclusively used in digital photogrammetric systems. Solid state imaging sensors consist of a large number of light-sensitive detector elements that are arranged as lines or arrays on semi-conductor modules (linear or area sensor). Each detector element (sensor element) generates an electric charge that is proportional to the amount of incident illumination falling on it. The sensor is designed such that the charge at each individual element can be read out, processed and digitized.

Fig. 3.72 illustrates the principle of a single sensor element. Incident light, in the form of photons, is absorbed in a semi-conducting layer where it generates pairs of electron holes (charged particles). The ability of a sensor element to create a number n_E of charged particles from a number n_P of incident photons is expressed by the external quantum efficiency η. The quantum efficiency depends on the sensor material and wavelength of the incident light.

$$\eta = \frac{n_E}{n_P} \tag{3.86}$$

The negatively charged particles are attracted by a positive electrode. Charges are accumulated in proportion to the amount of incident light until saturation or overflow of charge is achieved. The positive electric field of the electrode is generated by a potential well that collects the negative charge. In CCD sensors the detector elements are formed from MOS capacitors (*metal-oxide semiconductor*).

Sensor elements can be arranged in lines or two-dimensional arrays. Fig. 3.73 shows the simplified layout of a CCD line sensor. Each active sensor element is directly connected to a serial read-out register that is used to output the generated charge. In contrast, bilinear CCD lines can be resampling into what is effectively a

single line to provide increased resolution if the sensor elements are coupled in an alternating manner with two read-out registers.
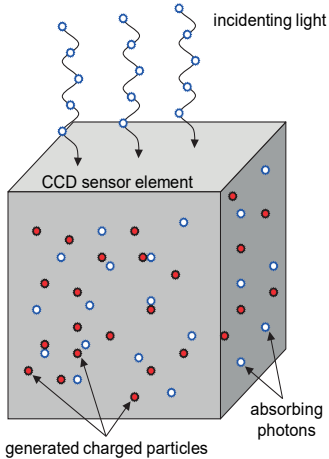
Fig. 3.72: Conversion of photons into charged particles.
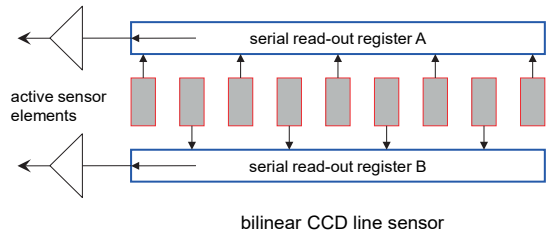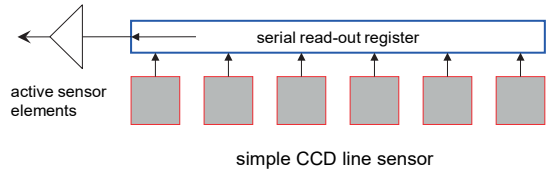
Fig. 3.73: Principle of simple and bilinear CCD line sensors.

The core problem for such sensor arrangements is the transportation of the charge stored in the sensor element to an output. Fig. 3.74 illustrates a typical solution for a linear arrangement of sensor elements. Here, the individual sensor elements have three electrodes, each connected to a different voltage phase. (In practice, most sensors use a 4-phase technique.) The cumulated charge at electrode 1 cannot discharge at time t_1, since the voltage is high on electrode 1 and low on electrode 2. At time t_2 the voltages on electrodes 1 and 2 are equal, forcing a portion of the charge under electrode 1 to flow to electrode 2. At time t_3 the voltages of electrode 1 and 3 have a low value, i.e. the complete charge is shifted under electrode 2. The result of the sequence is that the charge has shifted one electrode width to the right.

This process is continued until the charge reaches the read-out register at the end of a line. There the charges are read out and transformed into electrical voltage signals. The process is usually known as the CCD principle (*charge coupled device*), or bucket-brigade principle. In addition to the CCD principle, the CMOS principle for solid-state area sensors has also become well established (see section 3.4.1.3).

CCD line sensors can consist of more than 12 000 sensor elements. Given a sensor spacing of ca. 4 μm to 20 μm, the length of line sensors can be more than 100 mm. Line sensors are used in a wide variety of devices such as line cameras, fax machines, photo scanners or digital copiers.
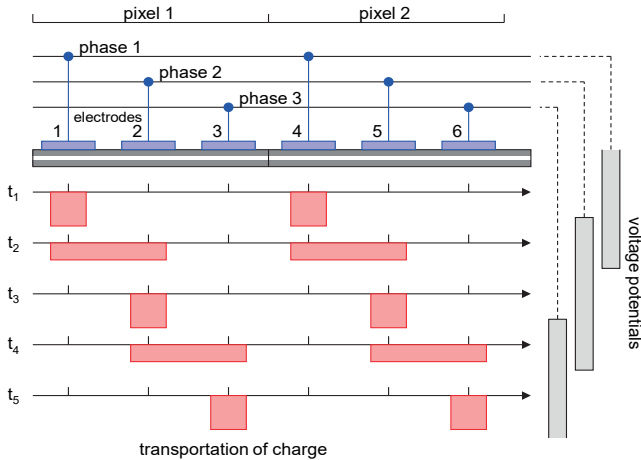
Fig. 3.74: Principle of CCD charge transportation (red-rectangles = bucket-brigade device).

3.4.1.2 CCD area sensors

Area sensors, which have their sensor elements arranged in a two-dimensional matrix, are almost exclusively used for photogrammetric image acquisition. In comparison with line sensors, the construction of matrix sensors is more complicated since the read-out process must be accomplished in two dimensions. Examples are shown in Fig. 3.75.
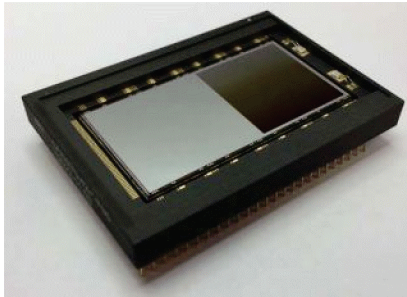
a) Frame-transfer sensor with imaging zone and storage zone (Teledyne Imaging)
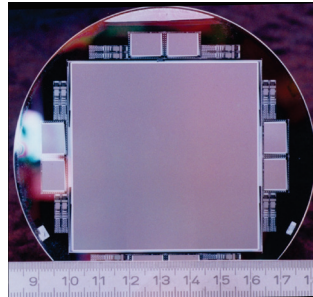
b) A full-frame transfer sensor with 4096 x 4096 elements on a single silicon wafer

Fig. 3.75: CCD matrix sensors.

There are three different arrangements of CCD matrix sensors that differ in layout and read-out process: frame transfer, full-frame transfer and interline transfer.

Frame-transfer sensors (FT) consist of a light-sensitive, image-recording zone and an equally sized, opaque storage zone (Fig. 3.75a). Each contains a parallel array

of linear CCD sensors. After exposure, charges are moved along the arrays from the imaging zone into the storage zone. From there they are rapidly shifted line by line into the read-out register (Fig. 3.76a). Charge transfer from the imaging to the storage zone can be carried out very rapidly, allowing high frame rates to be achieved since the imaging zone can be exposed again whilst the previous image is written out of the camera from the storage zone. Because imaging and storage zones are completely separate areas, the elements in the imaging zone can be manufactured with almost no gaps between them.

A simpler variation is given by the full-frame transfer sensor (FFT, Fig. 3.75b). It consists only of an imaging zone from where charges are directly transferred into the read-out register (Fig. 3.76b). During read-out the sensor may not be exposed. In contrast to FT sensors, FFT sensors tend to show greater linear smearing effects since longer transfer times are required. The simpler layout enables the construction of very large sensor areas[6] with very small sensor elements (6–9 µm size). Such layouts are used for high-resolution digital cameras with typically more than 1000 x 1000 sensor elements (manufacturers: e.g. Thomson, Kodak, Fairchild, and Dalsa). Note that the number of FFT sensor elements is often based on integer powers of 2 (512 x 512, 1024 x 1024, 4096 x 4096).
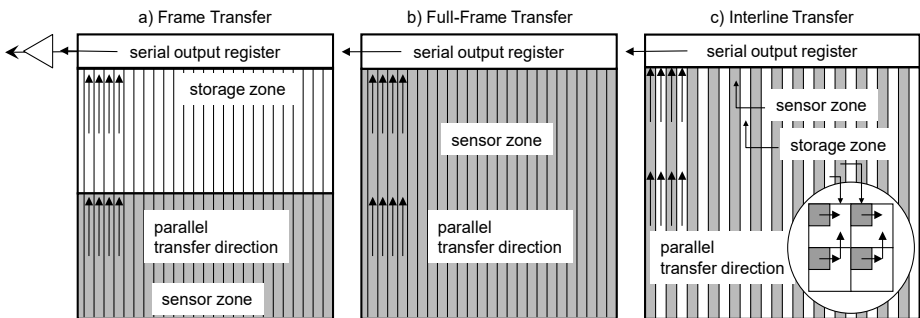
Fig. 3.76: Charge transfer for CCD sensors.

In contrast, interline-transfer sensors (IL) have a completely different layout. Here linear CCD arrays, which are exposed to light, alternate with linear CCD arrays which are opaque to light. Following exposure, charges are first shifted sideways into the opaque arrays which act as transfer columns. Then they are shifted along the columns to the read-out registers (Fig. 3.76c). The light sensitive area of the detector covers only about 25% of the total sensor area, compared with 90 to 100% for FT sensors, i.e.

---

6 The production of very large CCD sensors is limited mainly by economic restrictions (production numbers, quality) rather than technological restrictions.

IL sensors are less light-sensitive. IL sensors with standard pixel numbers of about 780 x 580 pixels are mainly used for CCD video and TV cameras (especially colour cameras). High-resolution IL sensors have up to 1900 x 1000 pixels. Since the whole area of a CCD array sensor is exposed at once, the image acquisition can be regarded as exposure by a global shutter (section 3.4.2.2).

3.4.1.3 CMOS matrix sensors

CMOS technology (*complementary metal oxide semi-conductor*) is a widely used technique for the design of computer processors and memory chips. It is increasingly used in the manufacture of opto-electronic imaging sensors, since it has significant advantages over CCD technology:
– only 1/10 to 1/3 power consumption;
– lower manufacturing costs;
– directly addressable sensor elements;
– acquisition of arbitrary image windows;
– frame rates of more than 2000 frames per second (see section 3.5.3);
– can be provided with on-chip processing, e.g. for sensor control or image processing;
– high dynamic range and low image noise (see section 3.4.1.9).
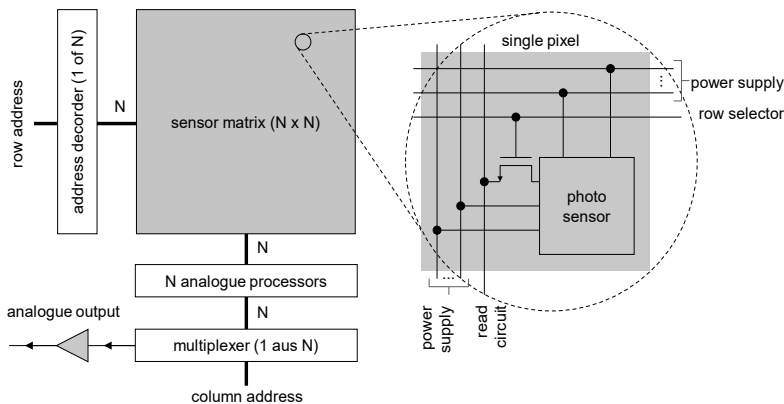
Fig. 3.77: Architecture of a simple 2D CMOS sensor (after Hauschild 1999).

CMOS imaging sensors are currently available with up to 150 million sensor elements. In contrast to CCD sensor elements comprising linked Metal Oxide Semiconductor (MOS) capacitors, CMOS detectors are based on photo diodes or transistor elements which are widely used in the construction of both integrated circuit chips and analogue detectors. They are therefore less costly to produce and can benefit from advances in fabrication of the diodes and transistors.

In a CMOS sensor, the charge generated by the incident light is directly processed by an integrated amplifier and digitizer unit attached to the pixel element. Individual sensor elements can be directly addressed for both readout and processing. This then avoids the need for sequential internal charge transfer which ensures that the sensor is less sensitive to blooming and transfer loss. Fig. 3.77 illustrates the basic architecture of a CMOS matrix sensor. Due to the presence of additional electronic components, the CMOS sensor's ratio of light sensitive surfaces to total surface area (fill factor) is smaller than for FFT CCD sensors. CMOS sensors are therefore generally equipped with microlenses (section 3.4.1.5). Sensor structure is layered with the latest generation of sensors having light sensitive elements in the lowest fabrication layer. The sensor is then flipped over in the camera so that light from the lens directly illuminates the back side of the sensor. Known as back illumination, this enables higher fill factors and efficiency because the electronic circuits do not then interfere the incoming light. The layer structure can also be used to provide multiple sensing layers, each responsive to a selected spectral band (see section 3.4.1.6).

CMOS sensor differences include, among others, the exposure and readout processes. Here it is usual to distinguish between global shutter and rolling shutter principles which function in a similar way to inter-lens and focal-plane shutters. Different shutter details are given in section 3.4.2.2.

3.4.1.4 Sensor formats

According to eqn. (2.2) the physical format of an imaging sensor is derived from the number of pixels and the spacing between them. From a photogrammetric or photographic point of view, larger sensor formats are preferred since they generate less noise and are more sensitive to light. In addition, larger formats are useful for generating larger image scales (see section 3.3.1.1).

Table 3.5: Typical image formats of video cameras.

Sensor size (typical) [mm]	Diagonal (tube) [inch]	Diagonal sensor [mm]	Number of pixels (typical)	Size of sensor element [μm]	Example
3.7 x 2.7	1/4"	4.6	656 x 490	5.6 x 5.6	Baumer VLG02
4.9 x 3.6	1/3"	6.1	1296 x 966	3.75 x 3.75	Basler ace A1300
6.1 x 4.9	1/2"	7.8	1280 x 1024	4.8 x 4.8	IDS UI3140
8.5 x 7.1	2/3"	11.1	2452 x 2056	3.45 x 3.45	AVT Manta G-505
9 x 6.7	2/3"	11.2	1392 x 1040	6.45 x 6.45	Ximea ICX285
12.5 x 9.97	1"	16	3376 x 2704	3.69 x 3.69	Grasshopper3 9.1
18.1 x 13.6	4/3"	22.7	3296 x 2472	5.5 x 5.5	AVT Prosilica GT 3300

Traditionally, sensor formats for video cameras are given in inches. With respect to old vidicon (tube) cameras, the sensor formats vary between ¼" and 1". For the image formats in Table 3.5, pixel sizes range from around 4 µm (¼" cameras) to around 16 µm (1" cameras) and with sensors having rectangular or square pixels.

For digital consumer cameras the following are common sensor formats (Fig. 3.78):

– Full format (FX):

Sensor formats which are close to the size of analogue 35mm film (36mm x 24mm) and with aspect ratio 4:3. In full format, lenses with standard focal lengths generate imaging angles that are similar to 35mm cameras, e.g. f = 50 mm for a normal lens (see section 3.4.3.2).

– APS and DX formats:

APS formats are defined with respect to earlier APS film formats which were 21 mm x 14 mm up to 24 mm x 16 mm for APS-C (also known as DX) and up to 28.7 mm x 19.1 mm for APS-H, with aspect ratio 3:2. Here crop factors are of the order of 1.5 to 1.6x when comparing the same lens used on an FX format sensor.

– Four Thirds:

The Four-Thirds format refers to sensors with a diagonal of around 4/3" and dimensions of around 18 mm x 13 mm with aspect ratio 4:3, i.e. they are equivalent to half-frame sensors. The newer Micro-Four-Thirds standard has the same format but a smaller flange focal distance (see section 3.4.2.4), hence these kind of sensors are especially suited for high-resolution compact cameras.

– HD and other formats:

In addition to the standard formats above, many new digital cameras offer variable image formats such as a 16:9 aspect ratio. With a sensor size of, for example, 18.7 mm x 10.5 mm, these correspond to the HD and Full-HD formats offered by video cameras. In addition, various special forms are available, e.g. with square sensor formats.
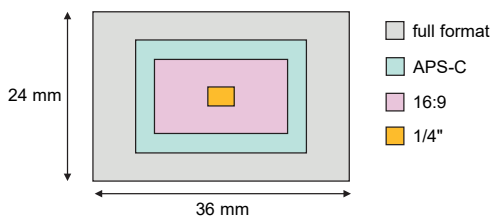
Fig. 3.78: Image formats of array sensors (selection).

Technical developments in recent years indicate that future CCD and CMOS array sensors will have an increasing number of pixels and larger formats and will deliver

improved signal to noise ratios as mass-market technology delivers consumer-driven improvements in image quality.

3.4.1.5 Microlenses

Current sensor designs typically employ microlens arrays in order to increase the fill factor of each pixel. A microlens array consists of a series of lens elements, each of which is designed to collect the light falling on a region approximating to the area of a single pixel and to direct that light to the smaller light-sensitive region of the actual sensor element (Fig. 3.79, Fig. 3.80). Whilst microlenses significantly enhance pixel fill factor, they typically have the limitation of only being able to receive light over a ±30-degree range of angles. This performance can limit the use of such arrays for extreme wide angle recording unless special optics are used. Modern lens designs for DSLR cameras reflect this requirement utilising optical configurations that produce less divergent emergent rays than earlier designs which were developed to optimize performance with legacy photographic film technologies.
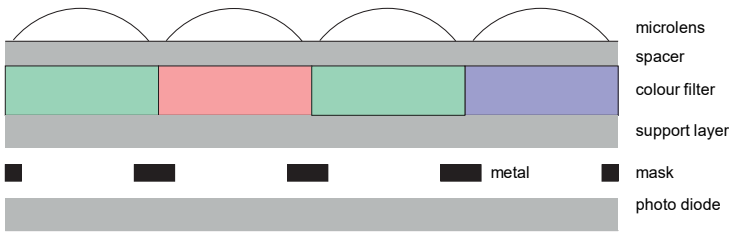
Fig. 3.79: Example of a microlens structure.
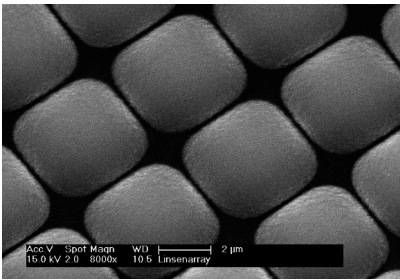
Fig. 3.80: Microlens array (raster electron microscope, 8000x magnification).

3.4.1.6 Colour cameras

In order to create true colour images, incident light must be separated into three spectral bands, typically red, green and blue.

Table 3.6: Features of colour separating methods.

	3-chip camera	RGB filter	Colour (mosaic) mask	True colour sensor
Number of sensors	3	1	1	1
Number of images	1	3	1	1
Number of video signals	3	1	1	1
Dynamic scenes	yes	no	yes	yes
Resolution	full	full	half	full
Colour convergency	adjustment	yes	interpolation	yes

Separation can be performed by four common methods (summarized in Table 3.6):

– Parallel or 3-chip method:
A prism system is used to project incident light simultaneously onto three CCD sensors of the same design. Each sensor is located behind a different colour filter so that each registers the intensity of only one colour channel (Fig. 3.81). Full sensor resolution is retained but exact alignment is required in order to avoid colour shifts. The camera delivers three separated analogue image signals that must be temporally synchronized and digitized in parallel. The principle is used for professional colour cameras, most of which are used either in TV studios or by mobile film crews.
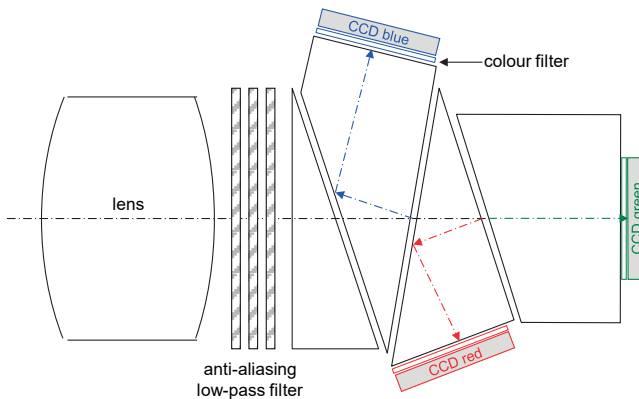
Fig. 3.81: Schematic optical diagram of a 3-CCD or 3-chip camera.

– Time-multiplex or RGB filter method:
Colour separation is performed by the sequential recording of one sensor whereby for each image a primary colour filter is introduced into the path of light. Full sensor resolution is preserved, but dynamic scenes cannot be imaged. The

method can be applied to both matrix sensor cameras and scanning cameras. The camera delivers a single image signal which must be filtered temporally in order to generate a digital RGB combination from the single colour bands.

— Space-multiplex or colour-mask methods:
A filter mask is mounted in front of the CCD matrix so that individual sensor elements react to only one colour. Strip or mosaic masks are used with the Bayer pattern mosaic mask being the most common. In comparison with the previous methods, geometric resolution will decrease since the output of three or more (typically four) sensor elements are combined to form each colour pixel (Fig. 3.82). The principle enables the recording of moving objects and its cost effectiveness means that it used for practically all consumer digital camera systems.
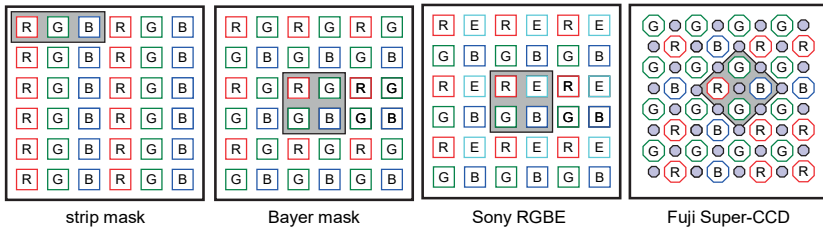
Fig. 3.82: Strip and mosaic masks.

When using a Bayer pattern, RGB colour values are calculated by interpolating neighbouring grey values. Because there are twice as many green as red or blue pixels, different calculation methods are used as follows:
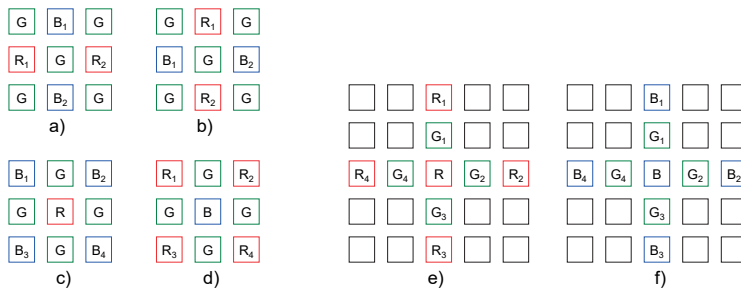
Fig. 3.83: Different combinations used in Bayer colour interpolation.

Calculation of red and blue components for a green pixel (Fig. 3.83a und b):

$$R_{green} = \frac{R_1 + R_2}{2} \qquad\qquad B_{green} = \frac{B_1 + B_2}{2} \qquad (3.87)$$

Calculation of the blue component for a red pixel and the red component for a blue pixel (Fig. 3.83c und d):

$$B_{red} = \frac{B_1 + B_2 + B_3 + B_4}{4} \qquad\qquad R_{red} = \frac{R_1 + R_2 + R_3 + R_4}{4} \qquad (3.88)$$

Calculation of the green component for red and blue pixels (Fig. 3.83e and f):

$$
\begin{aligned}
G_{red} &= \begin{cases} (G_1 + G_3)/2 & \text{for } |R_1 - R_3| < |R_2 - R_4| \\ (G_2 + G_4)/2 & \text{for } |R_1 - R_3| > |R_2 - R_4| \\ (G_1 + G_2 + G_3 + G_4)/2 & \text{for } |R_1 - R_3| = |R_2 - R_4| \end{cases} \\[2mm]
G_{red} &= \begin{cases} (G_1 + G_3)/2 & \text{for } |B_1 - B_3| < |B_2 - B_4| \\ (G_2 + G_4)/2 & \text{for } |B_1 - B_3| > |B_2 - B_4| \\ (G_1 + G_2 + G_3 + G_4)/2 & \text{for } |B_1 - B_3| = |B_2 - B_4| \end{cases}
\end{aligned} \qquad (3.89)
$$

There are additional interpolation functions which differ in their operation with regard to edge sharpness, colour fringes, noise correction and speed of operation.

– True colour sensor:

Foveon manufactures a CMOS-based, high-resolution, single-chip, true colour sensor consisting of three layers that are each sensitive to one primary colour (Fig. 3.84). It utilizes the property of silicon that light of different wavelengths penetrates to different depths. Hence, this sensor provides the full resolution of a usual CMOS sensor with true-colour registration capability.
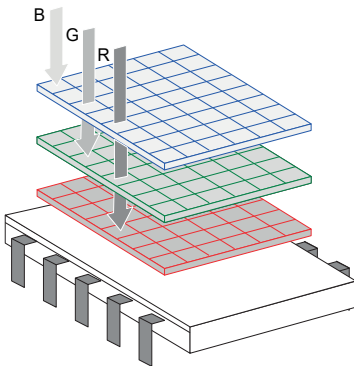
Fig. 3.84: Simplified structure of the Foveon X3 RGB sensor.

3.4.1.7 Sensor architecture

A typical imaging sensor has a structure comprising separate layers which are sandwiched together (Fig. 3.85). The lowest level is a layer of light-sensitive elements (photodiodes). The Bayer colour mask is attached on top of this, followed by the microlens array. Next comes a low-pass filter to supress the aliasing effect (section 3.1.5.1). The top layer consists of a filter for blocking near infrared radiation (section 3.4.1.9).
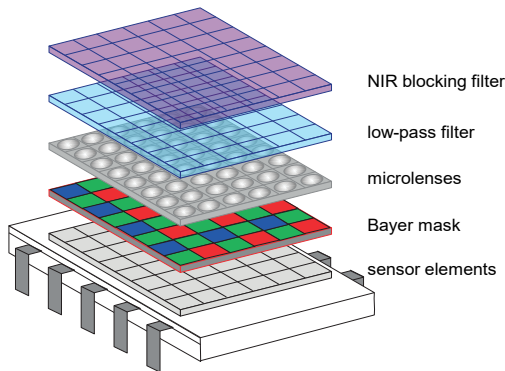
Fig. 3.85: Layered construction of a typical imaging sensor.

In principle, each layer is an optical component that affects the imaging rays and hence image generation. It is therefore possible that this causes local geometric imaging errors which cannot be modelled by standard methods of interior orientation.

3.4.1.8 Geometric properties
3.4.1.8.1 Resolving power

The theoretical resolving power of monochrome imaging sensors is limited by two factors:

–   detector spacing $\Delta s'$ (distance between sensor elements) and scanning theorem (Nyquist frequency f_N);
–   detector size $\Delta d'$ (aperture size) and MTF (limiting frequency f_0).

According to section 3.1.6.2, there are different theoretical resolution limits for FT, FFT and IL sensors due to their different arrangements of sensor elements. Furthermore, for all types, image quality can differ in both x and y directions where rectangular, rather than square, light sensitive regions have been used (Table 3.7).

For FFT sensors, or progressive-scan sensors (no interlaced mode), with square detector elements, resolving power can be expected to be equal in both directions.

IL sensors have approximately the same detector spacing as FT sensors. However, each pixel is split into a light sensitive detector and a shift register. Hence, the resulting theoretical resolving power is four times higher than the Nyquist frequency, and about two times higher than for FT and FFT sensors.

In practice, theoretical resolving power cannot be achieved unless the sampling interval is small enough to push the Nyquist frequency up beyond the effective resolution limit (cut-off) for that MTF. Normally the best that can be achieved is a full-fill system, for example the Kodak FFT in Table 3.7, where the Nyquist frequency is about half the theoretical resolution. In practical systems, frequencies higher than the Nyquist frequency are filtered out in order to avoid aliasing and micro lenses are used in order to provide pixel fill factors close to unity. Micro scanning systems (section 3.5.5.1) are able to achieve closer to theoretical resolving power since they subsample by moving the detector in fractions of a pixel between images.

New sensors with very small detector spacing (e.g. the Kodak KAC-05020 used in mobile phone cameras, see last column in Table 3.7) have a theoretically very high resolving power. However, these sensors also have a lower sensitivity to light and higher image noise. In order to match the pixel dimensions with an appropriate optical resolution, very high-quality lenses are required (see section 3.1.5.1 for comparison). The advantage of the small detector size therefore lies principally in the small dimensions of the imaging sensor.

Table 3.7: Resolving power of different CCD sensors.

		FT Valvo NXA	FFT Kodak	IL Sony	CMOS Kodak
Detector spacing in x	$\Delta s'_x$ [µm]	10.0	9.0	11.0	1.4
Detector spacing in y	$\Delta s'_y$ [µm]	7.8	9.0	11.0	1.4
Detector size in x	$\Delta d'_x$ [µm]	10.0	9.0	5.5	1.4
Detector size in y	$\Delta d'_y$ [µm]	15.6	9.0	5.5	1.4
Nyquist frequency in x	f_{Nx} [lp/mm]	50	55	45	357
Nyquist- frequency in y	f_{Ny} [lp/mm]	64	55	45	357
Theoretical resolution in x	f_{0x} [lp/mm]	100	111	180	714
Theoretical resolution in y	f_{0y} [lp/mm]	64	111	180	714

In comparison to photographic emulsions, recent opto-electronic sensors have equal or even better resolutions but, at the same time, usually much smaller image formats. A comparable resolution is achieved with sensor element sizes of about 7 µm or less.

In the digital photographic industry, alternative image quality measures are currently in use. The value *MTF50* defines the spatial frequency in lp/mm where the MTF is equal to 50%. With *line widths per picture height* (LW/PH) digital cameras are

classified as a function of line width instead of line pairs. LW/PH is equal to 2 x lp/mm x (picture height in mm). The term *cycles or line pairs per pixel* (c/p or lp/p) is used to give an indicator of the performance of a pixel.

3.4.1.8.2 Geometric accuracy

The geometric accuracy of matrix sensors is mainly influenced by the precision of the position of sensor elements. Due to the lithographic process used to manufacture semi-conductors, CCD matrix sensors have regular detector positions of better than 0.1–0.2 μm, corresponding to 1/60 to 1/100 of the size of a sensor element. This does not mean that the resulting image can be evaluated to this accuracy. Several electronic processing steps are performed between image acquisition and digital storing that may degrade image geometry and contrast.

An additional effect is given by the possible lack of flatness of the sensor surface. For sensor areas of 1500 x 1000 pixels departures of up to 10 μm from a plane surface have been demonstrated. Depending on the viewing angle, a perpendicular displacement of a sensor element causes a corresponding lateral shift in the image (Fig. 3.86). If there is a non-systematic lack of flatness in sensor surface, the usual approaches to distortion correction fail (additional parameters). A suitable correction model based on finite elements has been presented in section 3.3.4.4. Fig. 3.87 shows deformations of an imaging sensors computed by this approach.

$$\Delta r' = \Delta h' \frac{r'}{z'} = \Delta h' \tan \tau' \tag{3.90}$$
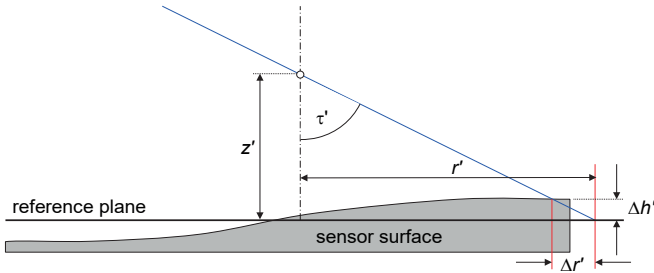
Fig. 3.86: Lateral displacement for an uneven sensor surface.

According to (3.90), image displacement is greater given large incident ray angles to the sensor from shorter focal lengths z' (wide angle lenses), increasing distance r' from the optical axis and greater lack of flatness $\Delta h'$.
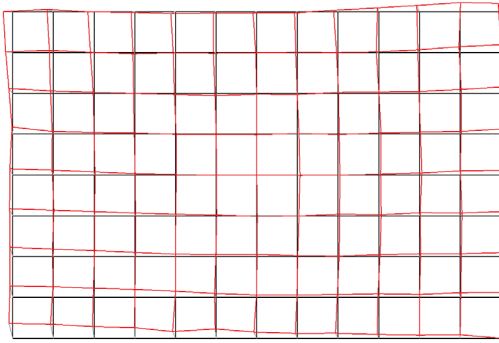
Fig. 3.87: Sensor deformations (red) after finite element calibration (Kodak DCS460).

3.4.1.8.3 Warm-up effects

Imaging sensors and on-board electronics require a warm-up period of up to 2 hours. During warm up, image coordinate displacements of several tenths of a pixel have been shown to occur. Fig. 3.88 shows an example of drift measured under controlled laboratory conditions. The figure shows coordinate differences extracted from five images of a warm-up time series. In the x direction, small shifts only can be observed which, after 10 minutes, are below the measuring uncertainty of around ±0.02 pixel. In contrast, in the y direction a drift of about 0.15 pixel can be observed within the first 20 minutes. Part of this effect is caused by temperature increases on mechanical components, but electronic devices inside the camera also contribute as they warm up. High performance photogrammetric systems therefore require both calibration and image mage acquisition once a steady thermal and mechanical state has been reached.
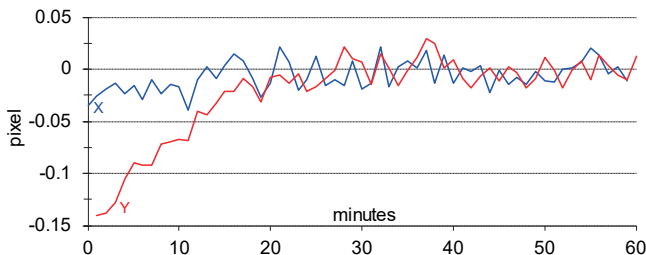
Fig. 3.88: Drift of measured sensor coordinates during warm-up.

3.4.1.9 Radiometric properties

Light falling onto an imaging sensor is either reflected on the sensor surface, absorbed within the semi-conducting layer, or transmitted if high-energy photons are

present. Absorption happens if the wave length of light is shorter than a threshold wave length λ_g which is a property of the radiated material and defined according to:

$$\lambda_g = \frac{h \cdot c_0}{E_g} \tag{3.91}$$

where

c_0: velocity of propagation $= 3 \cdot 10^8$ m/s

h: Planck's constant $= 6.62 \cdot 10^{-34}$ Js

E_g: energy between conduction and valence band
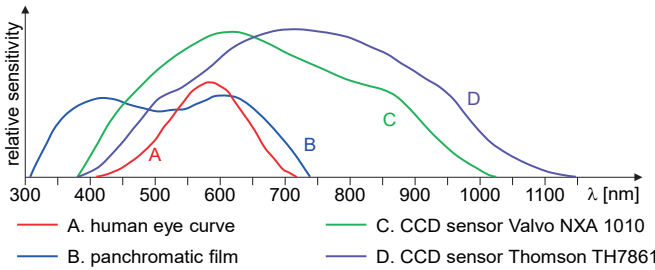
Fig. 3.89: Spectral sensitivity of different sensors.

Silicon used in the manufacture of imaging sensors has an E_g value of 1.12 eV, thus a threshold wavelength of $\lambda_g = 1097$nm which is in the near infrared. In comparison with film and the human eye, imaging sensors show a significantly wider spectral sensitivity (Fig. 3.89). Optionally, infrared absorbing filters can be attached to the sensor in order to restrict incident radiation to visible wavelengths. The spectral sensitivity of imaging sensors with Bayer colour masks (see section 3.4.1.6) varies between the individual colour channels. Typically, the blue channel has the lowest sensitivity (Fig. 3.90).
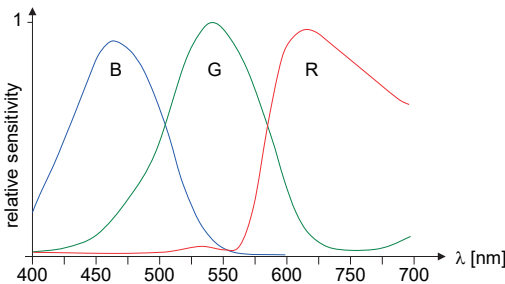
Fig. 3.90: Spectral sensitivity of a colour CCD sensor (Sony ICX 098BQ).

Independently of the image-forming incident light, thermal effects in the semi-conducting layer can generate small charges which appear as background noise in the sensor signal. This background noise is known as dark current since it occurs independently of any image illumination and can be observed in total darkness.

The intensity signal g, created by the imaging sensor, depends on the dark current signal g_0, an electronic amplification factor K, the quantum efficiency η and the number of collected photons n_P (see eqn. 3.86).

$$g = g_0 + K \cdot \eta \cdot n_P \tag{3.92}$$

The noise (variance) of a digitized grey value is given by:

$$\sigma_g^2 = K^2 \sigma_d^2 + \sigma_q^2 + K(g - g_0) \tag{3.93}$$

where

σ_d^2 : system noise of camera electronics

σ_q^2 : quantization noise

The noise depends on the intensity, i.e. dark image areas show higher noise than bright areas. By cooling the sensor, the noise level can be reduced in order to improve the radiometric dynamic range. Cooling a sensor by 5–10° C readily reduces noise by a factor 2. Cooled sensors are mainly used in low light applications with long integration times. e.g. imaging sensors for astronomy.

The radiometric dynamic range is defined by the signal-to-noise ratio (SNR):

$$SNR = \frac{g - g_0}{\sigma_g} = 20 \log \frac{g - g_0}{\sigma_g} \text{[dB]} \tag{3.94}$$

For matrix sensors typical SNR lies somewhere between 1000:1 (approx. 60 dB) and 5000:1 (approx. 74 dB). If an additional noise for the subsequent A/D conversion is assumed, the sensor signal should be digitized with at least 10–12 bits per pixel. In low-light conditions or where exposure times are very short (e.g. high frequency cameras, section 3.5.3), the SNR has a significant role. For this purpose scientific CMOS sensors (sCMOS) which have a very low SNR are available. In static situations the dynamic range can be increased by multiple exposures with different apertures and exposure settings, as well as analytical combination of individual images (High Dynamic Range Imaging, HDR photography).

For very bright areas in the scene (hot spots) the CCD sensor may be subject to saturation or overflow effects where the imaged signal flows into adjacent pixels. Such effects are caused by movement of charged particles into neighbouring sensor elements (*blooming*, see Fig. 3.91a,b), and by continuous charge integration during the read-out process (*smear*, see Fig. 3.91c,d). In photogrammetric applications these effects can be observed for brightly illuminated retro-reflective targets resulting in an incorrect determination of the target centre. Blooming and smear can also occur

within images of natural scenes where large variations in lighting and contrast are present. This manifestation of blooming and smear does not occur in CMOS sensors. However, over-exposure can lead to pixel blooming (parasitic light sensitivity, shutter efficiency) or to bright pixels appearing as dark (black sun effect).
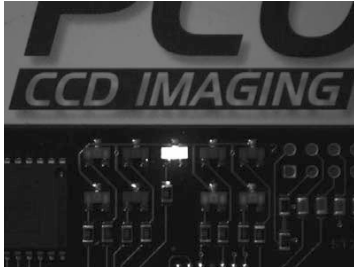
a) Normal exposure (no blooming)
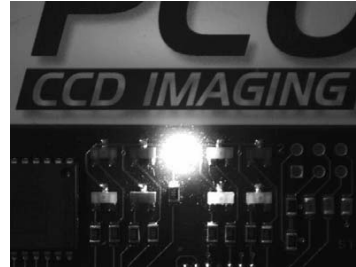
b) Blooming due to flash illumination
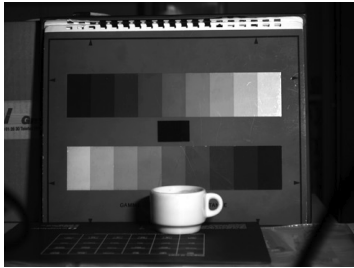
c) Blooming through over-exposure
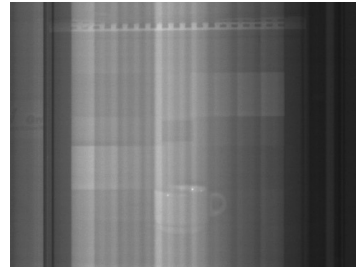
d) Continuous illumination

Fig. 3.91: Blooming effect and smear (PCO).

3.4.2 Camera technology

3.4.2.1 Camera types

Three generic camera designs can be identified among the wide variety of camera systems that are available to the photographer:

– Cameras without viewfinder:

  Numerous industrial or surveillance cameras (examples in Fig. 3.112) consist only of a sensor unit with electronics and a lens, hence no integrated optical or digital viewfinder. However, live images can be transferred via a digital interface to an external processor or computer which can be used to display the live image. The low cost and availability of low-cost network interfaces, such as Gigabit Ethernet, expand the capabilities of these systems to include rapid processes such as on-line lens focusing, target image detection and feature processing. These cameras tend to be very small for their imaging capability and can provide an excellent

solution where space is restricted or a multitude of cameras are needed to simultaneously observe a scene.

– Viewfinder or compact cameras:

Historically a viewfinder camera (Fig. 3.92 left, example in Fig. 3.114a) uses a viewing lens that is separated from the actual image taking lens. Such camera designs are generally very light in weight as the optical systems and image formats are compact. However, direct observation of the image is not possible for focusing, depth of field control or lens interchange. More seriously for close-range work, the difference in content between the viewfinder image and the recorded image give rise to parallax errors. An overlay in the viewfinder can indicate the area imaged by the camera, but it is not possible to fully correct the difference in perspective. The viewfinder displays an upright image as would be seen by the eye.

Digital camera technology has replaced the optical viewfinder with a live display of the camera image on an LCD screen integrated into the camera. The screen may be located on the back of the camera or on hinge mechanism to allow comfortable viewing for high and low vantage points. Many consumer compact digital cameras and mobile phone cameras are of this type (see Fig. 3.114 and Fig. 3.119) and are capable of 3D scene reconstruction using public domain apps and open source software.

– Single-lens reflex camera:

In a single-lens reflex (SLR) camera, viewing is done directly through the camera lens by means of a plane mirror which deflects the path of rays into a viewfinder (Fig. 3.92 right). Before exposure, the mirror is flipped out of the optical path. Cameras employing this principle with a digital sensor are termed Digital SLRs or DSLRs (see Fig. 3.115 to Fig. 3.118).

camera with
separated viewfinder

single-lens reflex camera
with pentagonal prism
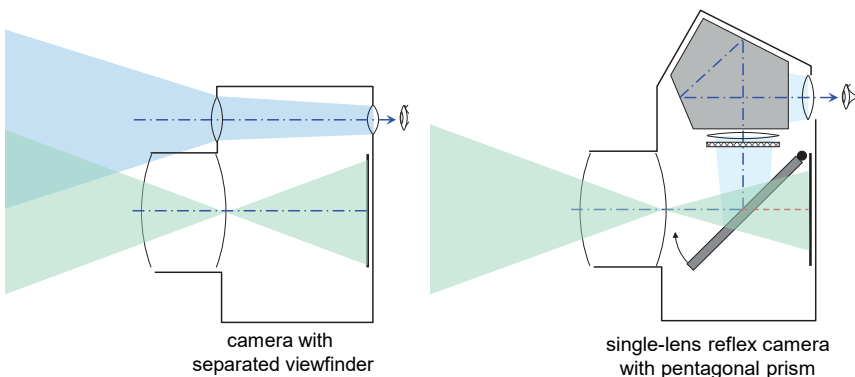
Fig. 3.92: Single lens reflex cameras (after Marchesi 1985).

Fig. 3.115 to Fig. 3.118 show examples of digital single-lens reflex cameras. Here it is also normally possible to view the live camera image on the LCD display. In this case, the mirror must be flipped up for permanent exposure of the sensor. This mode can cause significant warm-up of the camera.

Most modern DSLR cameras maintain the mirror and ground glass viewing system, but support direct live viewing on an LCD screen at the back of the camera in a similar manner to the consumer compact camera. Often termed "live view", this method requires the mirror to be in the up position so that light can reach the CCD or more commonly CMOS sensor. Maximum framerates for such systems are governed by the mechanical requirements of the mirror and shutter system to deliver between 3 to 10 frames per second (fps).

– Mirrorless cameras:

Mirrorless cameras (example in Fig. 3.114b) do not have an optical viewfinder but the image of the sensor is directly transferred to an integrated display that can be observed either through an ocular or by a small monitor. Mirrorless cameras are equipped with interchangeable lens systems which are also available with fixed focal length. They also have short sensor to lens-flange distances enabling small high-quality wide-angle lens designs to be used. As battery power needed to illuminate their real-time electronic viewfinders improves, this type of camera is being marketed to replace DSLR systems by the major camera manufactures. High frame rates (30 to 60fps being possible), portability and high quality construction make such designs well-suited to photogrammetric purposes provided features such as image stabilisation (see section 3.4.2.3) and image plane focussing can be switched off as these can compromise the physical stability between sensor and lens required for accurate and reliable photogrammetric calibration.

A camera design, known as a bridge camera, removes the interchangeability of the lens system and replaces it with a large-ratio zoom lens that can provide wide angle and very narrow angle views suitable for all types of photography from landscapes to recording distant wildlife.

– Studio camera:

Studio cameras (Fig. 3.93) allow for the individual translation and rotation of lens and image planes. By tilting the lens, special focus settings can be enabled, e.g. Scheimpflug condition (section 3.1.2.4) which maximizes depth of field in a particular plane. When the lens is shifted the perspective imaging properties are changed, e.g. to avoid convergent lines for vertical architectural pictures. Studio cameras represent the ultimate in photographic quality by virtue of their large image sizes (5"x7" and 10"x8" being common). Due to their bulky nature and relatively cumbersome deployment they are typically used in professional studio applications or for landscape and architectural work.
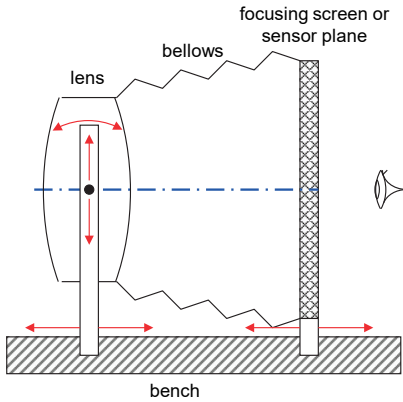
Fig. 3.93: Studio camera (Rollei).

3.4.2.2 Shutter
3.4.2.2.1 Mechanical shutters

The shutter is used to open the optical path for the duration of time necessary for correct exposure. Mechanical shutters are also used in digital cameras in order to control the readout process with an electronic shutter (see below). For conventional camera systems two basic types are used, the focal-plane shutter and the inter-lens or leaf shutter.
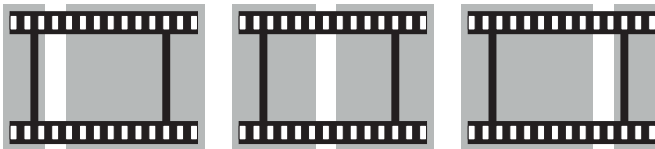
Fig. 3.94: Principle of focal plane shutter.

The majority of 35mm single lens reflex cameras use a mechanical focal-plane shutter that is mounted directly in front of the image plane and, during exposure, moves a small slit across it (Fig. 3.94). Variation in the size of the slit allows a variety of short duration shutter settings. Focal plane shutters are easy to design and provide shutter times of less than 1/8000 s. If the camera moves parallel to the shutter movement, e.g. photos from a moving platform, imaging positions are displaced, i.e. at the different slit positions the image has a different exterior orientation (see example in Fig. 3.97).

Inter-lens shutters are typically mounted between the lens elements and close to the aperture stop. Because of this, each lens must have its own shutter. Mechanically sprung blades are used to open the shutter radially (Fig. 3.95). In the figure, the circles

represent the diameter of the shutter as it opens and closes and the corresponding tone in the image shows the level of light reaching the image plane. In practice the opening and closing of the shutter can be regarded as instantaneous, which means that the complete image is exposed simultaneously with the same projection even if the camera platform moves. Inter-lens shutters are mechanically more complex than focal-plane shutters since the individual elements must be sprung open and then closed. Consequently, shortest shutter times are restricted to about 1/1000 s. In photogrammetry they are usually encountered in low-cost viewfinder cameras and in professional medium and large format cameras.
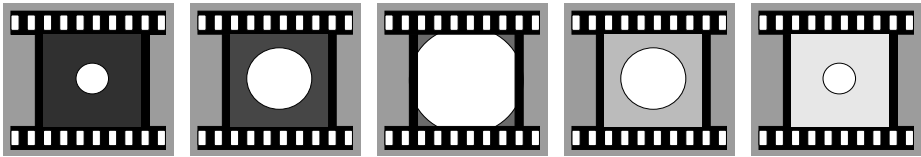
Fig. 3.95: Principle of the inter-lens shutter.

3.4.2.2.2 Electronic shutters

The process of integration and readout in digital imaging sensors is described as an electronic shutter. Depending on the camera and sensor type (CCD, CMOS), it is combined with a mechanical shutter in order to control the incoming light and protect active sensor elements against additional exposure.

Using a reset command, the exposed sensor elements are discharged and the exposure process can be restarted. The time required for exposure is called the integration time. After integration, the sensor elements are read out. For CCD and CMOS sensors, the reset is implemented using different timings.

Several technologies are in use today:

– Interline odd/even line:

   All even lines in the image are captured followed by all odd lines. The two sets of lines are then presented as a pair of half images to form a full frame consecutively. The initial purpose of this standard was to provide a high frame rate (60 Hz) in early TV transmission with limited given bandwidth.

– Global Shutter:

   All sensor elements are discharged simultaneously in a global reset process and the entire sensor is then exposed for the next image (all pixels simultaneously exposed). After the integration period, the charges are shifted to a covered sensor area so that there is no further exposure of the pixels during the readout process. The chosen exposure time t_{image} equals the integration time t_{int}. Fig. 3.96 (left) shows the principle of the exposure and readout scheme. The ratio between integration time t_{int} and readout time t_{read} is variable. Standard CMOS sensors do

not have a light-protected storage zone, hence a true global shutter can only be implemented by additional storage cells as, for instance, employed in high-speed cameras (section 3.5.3). The global shutter corresponds to a mechanical inter-lens shutter and is preferred for dynamic applications.

An approximation is possible whereby integration starts for all pixels at the same time and the data is read out line by line so that the first lines receive less integration time than the later lines. This is often used in progressive scan cameras equipped with a trigger output so that an electronic flash can be used to freeze motion.

– Rolling Shutter:

Lower cost CMOS sensor use the rolling shutter or progressive scan principle (Fig. 3.96 right). Here lines are progressively read out and exposed line by line as an electronic shutter is moved across the array with a time offset which relates to the readout time t_{line}. This process is analogous to the higher speed focal plane SLR shutters which sweep two blinds with a variable gap across the sensor. Each pixel in the array receives the same exposure duration, but the scan over the hole image is temporally variant in the direction of the progressive scan. The total exposure time t_{image} of an image is therefore longer than the chosen exposure time.

As with mechanical focal plane shutters, a relative movement between rolling shutter camera and object leads to image distortion (Fig. 3.97 right). In addition, each image line has an individual exterior orientation if the camera platform is moving. This effect can be modelled mathematically for systematic camera motion and can be compensated within an extended bundle adjustment.
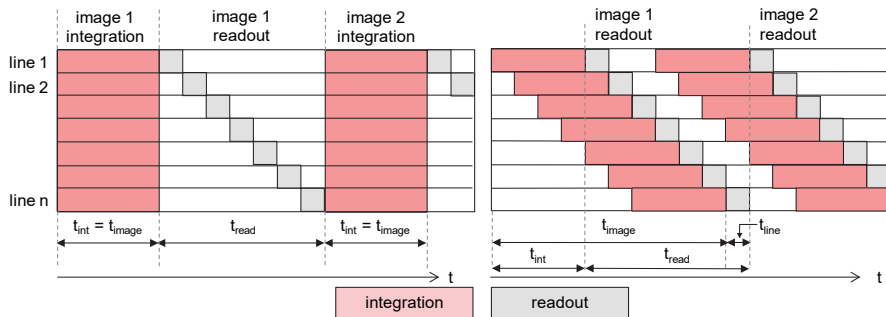
Fig. 3.96: Integration and readout for global shutter (left) and rolling shutter (right).

DSLR cameras with CMOS sensors are often combined with a mechanical shutter. This is faster than the rolling shutter and can therefore reduce image distortion caused by the rolling shutter effect.
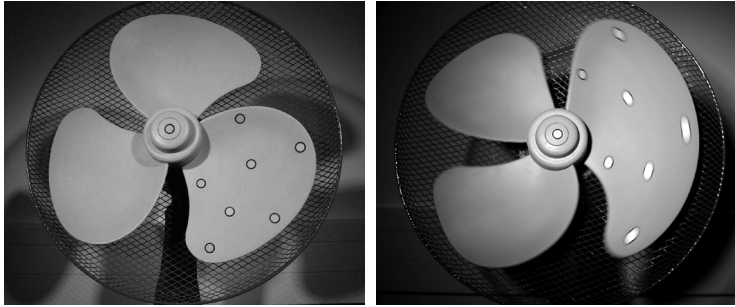
Fig. 3.97: Recording of a moving object at the same shutter duration with global shutter (left) and rolling shutter (right).

3.4.2.3 Image stabilization and cleaning

Modern lenses and digital cameras can be equipped with electronic image stabilization which ensures that blurring caused by hand-held operation is reduced. The vibrations, for which this compensates, lie approximately in the range 1 to 10 Hz and are corrected either in the camera body or within the lens.

Image stabilizers in the camera body have a movement sensor which detects the accelerations acting on the camera and immediately applies a counteracting movement of the imaging sensor (Fig. 3.98a). The compensating movement of the imaging sensor is implemented with piezoelectric or electromagnetic elements which can generate shifts of up to three millimetres. Cameras with built-in image stabilizers can utilize this option with any lens provided resulting image movements remain within the limits of the sensor mechanical motion capabilities.

Image stabilization at the lens (Fig. 3.98b) is implemented with a correction lens (or lenses) moved by piezoelectric elements. This lens group alters the imaging path in a way which compensates for the blurring movement. To make use of this feature the lens requires a digital interface which receives a signal from the camera when exposure takes place.

a) Image stabiliser in camera   b) Image stabiliser in lens
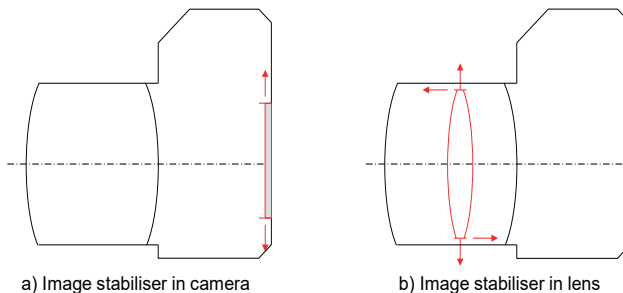
Fig. 3.98: Techniques for automatic image stabilization.

Numerous DSLR cameras are also now equipped with automatic sensor cleaning. This is done by piezoelectric vibration of the sensor which removes dust particles. Image cleaning techniques change the sensor location and can introduce systematic changes in the interior orientation of the camera between groups of images. The use of this feature requires adoption of calibration methods suited to group variations between image sets.

Image stabilization techniques are more complex to accommodate. First advice is to turn these off, however if the resultant image quality is then too poor for the identification and subsequent measurement of imaged features, an image-variant camera calibration (see section 3.3.4.4) needs to be adopted.

3.4.2.4 Sensor to lens flange distance

The sensor to lens-flange distance is the distance between the lens mounting flange and imaging plane of the sensor. For compatibility, camera bodies and lenses are each designed with a specified set of mounting tolerances. This enables the full range of focus setting to be used, from infinity to the minimum focus distance. Advantages of short distances are simpler wider-angle lens designs and smaller more portable camera systems. If the flange distance between camera and lens is mismatched, focusing errors will arise. Where the flange distance in the camera is shorter, an adapter may be used to attach a lens designed for a longer flange distance.

Table 3.8 lists some typical lens mount systems and their and their respective flange distances.

Table 3.8: Example flange distances (FD) for different lens mounts.

Camera type	Lens adapter	FD	Mount
Industrial camera	C-Mount	17.526 mm	screw thread
Industrial camera	CS-Mount	12.50 mm	screw thread
Mirrorless camera	Micro Four Thirds	19.25 mm	bayonet catch
DSLR	Nikon F-Mount	46.50 mm	bayonet catch
DSLR	Canon EF	44.00 mm	bayonet catch

3.4.3 Lenses

3.4.3.1 Relative aperture and f/number

The light gathering capacity of a lens is measured by the relative aperture which is the ratio of the iris diameter d' of the entrance pupil to the focal length f (see Fig. 3.99):

$$\text{Relative aperture} = \frac{d'}{f} \tag{3.95}$$

For a lens with an entrance pupil of $d' = 20$ mm and a focal length of $f = 80$ mm, the relative aperture is 1:4.

The f/number is given by the reciprocal of relative aperture:

$$\text{f/number} = \frac{f}{d'} \tag{3.96}$$

A higher f/number corresponds to a smaller relative aperture, i.e. the light gathering capacity is reduced. Changes in f/number follow a progression whereby the area of aperture (hence the amount of gathered light) changes by a factor of 2 from step to step (Table 3.9). The f/number also alters the depth of field (see section 3.1.2.3).

Table 3.9: Standard f/number sequence.

1	1.4	2	2.8	4	5.6	8	11	16	22	32

3.4.3.2 Field of view

The field of view (format angle) $2\alpha'$ of a lens is defined by its focal length and the diameter of the entrance pupil EP (Fig. 3.99):

$$\tan \alpha' = \frac{d'}{2f} \tag{3.97}$$
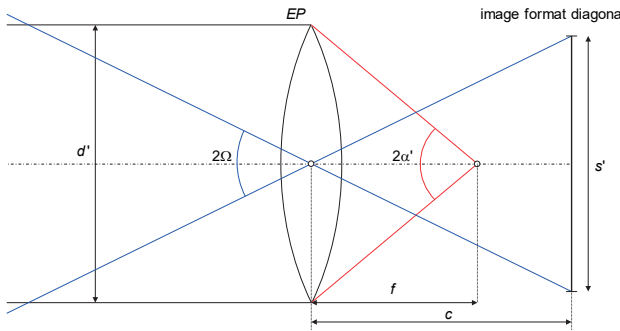
Fig. 3.99: Field of view and format angle.

In contrast, the format angle (field angle) 2Ω is given by the maximum usable image angle with respect to the diagonal of a given image format s' and the principal distance c:

$$\tan \Omega = \frac{s'}{2c} \qquad\qquad (3.98)$$

Table 3.10: Corresponding focal lengths for typical lens types and image formats.

Lens type	Format angle (2Ω)	Mobile phone [3 x 2 mm²]	Video [8 x 6 mm²]	Full format [36 x 24 mm²]	Medium format [44 x 33 mm²]
Image diagonal		3.6 mm	10 mm	43 mm	55 mm
Telephoto (small angle)	15–25°	>8 mm	> 22 mm	> 80 mm	> 95 mm
normal	40–50°	4–5 mm	11–14 mm	45–55 mm	55–70 mm
wide angle	60–100°	2–3 mm	5–8 mm	24–35 mm	35 – 60 mm
fisheye	> 100°	<2 mm	<4 mm	< 24 mm	< 30 mm

Format angle is a convenient method of distinguishing between different basic lens types (see Table 3.10). As a rule of thumb, the focal length of a normal lens is approximately equal to the diagonal of the image format. Small image formats (found in video cameras, mobile phones and low-cost digital cameras) require short focal lengths in order to produce wide angles of view.

3.4.3.3 Image circle and sensor format

A lens generates a circular image area which is different from the rectangular area of a typical sensor. Fig. 3.100 illustrates the difference in the case of two different sensor sizes.
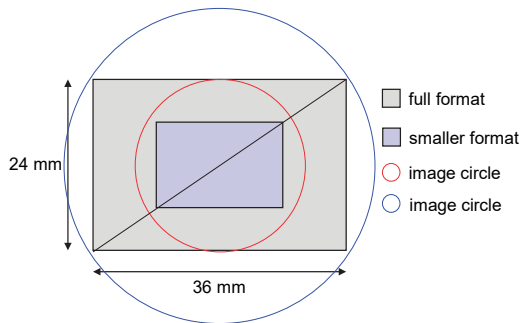
Fig. 3.100: Sensor formats and image circles.

If the diameter of the image circle is equal to, or larger than the sensor diagonal (blue circle), the complete sensor area is illuminated without vignetting. Alternatively, if

the image circle is smaller than the sensor diagonal (red circle), the sensor is only partially illuminated. This causes darkening at the image borders and corners. However, if the sensor format is reduced from the larger to the smaller format, full sensor illumination is restored.

With increasing distance from the centre, the optical quality of a lens generally degrades and distortion generally increases. From metrological considerations it is therefore recommended to use lenses with larger image circles to give better quality over the sensor area and to stop down the aperture to reduce light fall off (vignetting) caused by the use of large apertures.

3.4.3.4 Super wide-angle and fisheye lenses

In close-range photogrammetry, short focal length lenses with wide fields of view are often selected because they allow for shorter object distances, greater object coverage and, consequently, more favourable ray intersection angles. Wide-angle lenses have fields of view typically in the range 60–75°. Lenses with larger fields of view designed to maintain the central perspective projection are known as super wide-angle lenses (approx. 80–120°). Whilst they have increased image aberrations and illumination fall-off towards the extremes of the image format (section 3.1.3.5), their use in photogrammetry is common as it supports working in cluttered environments with short taking distances. Fisheye lenses utilize a different optical design that departs from the central perspective imaging model to produce image circles of up to 180° (section 3.3.7).

a) f = 14 mm   b) f = 15 mm   c) f = 20 mm

Fig. 3.101: Imaging with a) super-wide-angle; b) fisheye; c) wide angle lenses.

In general, the effect of radial distortion increases with larger fields of view. Fig. 3.101 shows examples of images taken with lenses of different focal lengths and fields of view. For some lenses the distortion is clearly visible and here the 15 mm quasi-fisheye lens shows the greatest distortion, but the illumination fall-off identifiable in the 14mm super wide-angle lens image is noticeably absent. In addition, towards the lens perimeter there is reduced image sharpness, increased chromatic aberration and greater fall-off in light according to the $\cos^4$ law (see section 3.1.3.5).

Fisheye lenses must be modelled using the appropriate fisheye projection model (see section 3.3.7) as the polynomials described in section 3.3.3.1 will be unstable due to the very high distortion gradients. Conventional test fields with circular targets often cannot be used for modelling purposes because the large distortion parameters deform the images so much that the targets cannot be measured with standard methods. Solutions for the calibration of fisheye lenses are discussed in section 7.3.3.2.

The last 15 years have seen a step change in mass-market wide-angle lens design. Here, the use of high-dispersion glass, aspheric lens elements and precision plastic construction delivers increased optical performance over the complete field of view, along with a one to two stops increase in lens aperture. A prime example is the change made by Nikon from its AF-D to AF-S lens designs where f/1.8 wide-angle lenses in the 20mm to 28mm focal length range are of high quality and readily affordable.

3.4.3.5 Zoom lenses

Zoom or varifocal lenses enable a varying focal length to be produced from a single lens system. Designs may also permit constant focusing and maintenance of the same relative aperture as focal length is changed. Fig. 3.102 illustrates the principle of a zoom lens where moving a central lens group gives a change in focal length whilst motion of a second group of lenses provides focus compensation.
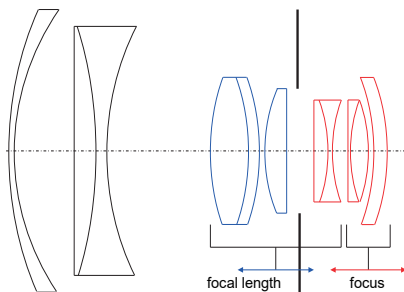
Fig. 3.102: Principle of a zoom lens (after Marchesi 1985).
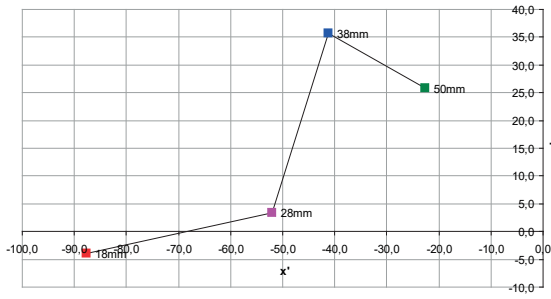
Fig. 3.103: Variation of the principal point in μm for the different focal lengths of a zoom lens (Canon EOS 1000D with Sigma DC 18−50 mm).

According to section 3.3.2.2, a change in focal length results in a new interior orientation. Due to the zoom lens construction, not only the spatial position of the perspective centre will change (Fig. 3.103) but parameters of radial and tangential distortion will also change (Fig. 3.104). Whilst zoom lenses can be calibrated photogrammetrically, off-the-shelf designs cannot be assumed to provide high mechanical stability. Thus, whilst they provide great flexibility, they are seldom used in practice for accurate work.
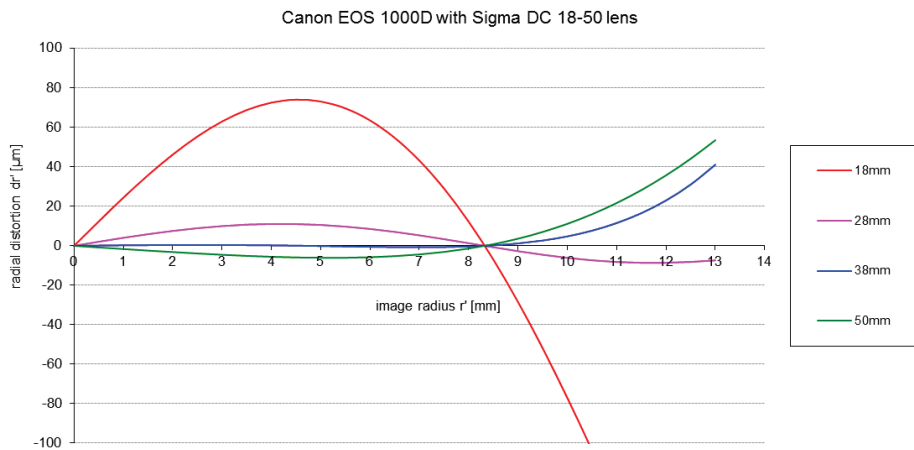
Fig. 3.104: Variation in radial distortion for different focal settings of a zoom lens.

3.4.3.6 Tilt-shift lenses

Tilt-shift lenses are special lenses which permit a lateral shift and tilt of the lens relative to the image sensor (Fig. 3.105). Using the shift function, the optical axis can be displaced in order to eliminate converging lines in oblique views. As an example,

Fig. 3.106 shows the image of a building façade in which the converging effect of perspective has been eliminated by shifting the lens in the vertical direction. In both cases the camera is set to point horizontally.

Fig. 3.105: A digital SLR camera fitted with a tilt-shift lens.

a) Conventional image
b) Image using shift function

Fig. 3.106: Correction of converging lines by the use of a tilt-shift lens.

The lateral shift of a lens can also be used to extend the stereoscopic view of a stereo camera pair without the need for convergent viewing of the two cameras, i.e. their axes remain parallel.
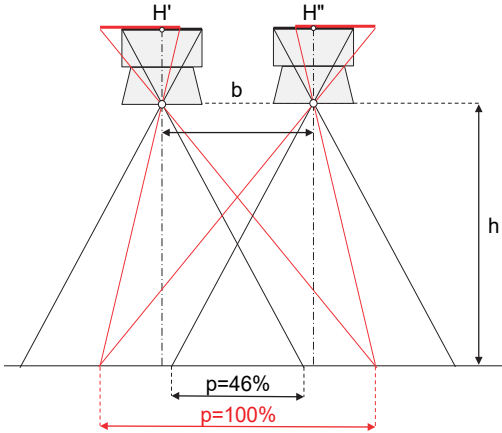
Fig. 3.107: Lateral sensor shift for the stereo case.

In Fig. 3.107, a pair of cameras is configured for the normal case of stereo photogrammetry with an overlap of $p = 46\%$. By shifting the sensor with respect to the lens a theoretical overlap of $p = 100\%$ can be achieved. However, care must be taken that the image circles of the lenses are large enough to cover their respective sensor areas after shifting.

The tilt function permits the use of the Scheimpflug condition (section 3.1.2.4) which enables sharp imaging of an obliquely imaged object plane.

3.4.3.7 Telecentric lenses

Telecentric lenses are designed such that all object points are imaged at equal image scales regardless of their distance. A two-stage telecentric lens consists of two lens groups where the object-side principal plane of the second (right) system coincides with the focal plane of the first (left) system (Fig. 3.108). The aperture stop is also positioned at this location. An object point P located within one focal length of the first system is virtually projected into P'. The second system projects P' sharply into the image plane at P". Since all points in object space lie on a parallel axis (hence distance-independent), light rays are projected through identical principal rays and are therefore imaged at the same position in the focal plane.

Limits to the bundles of rays place limits on the size of the object which can be seen in the image plane. The limiting factor will be either the maximum diameter of the aperture stop or the lens. Telecentric lenses are mainly used for imaging small objects ($\varnothing < 100$ mm) in the field of two-dimensional optical metrology. The imaging model does not correspond to a central projection but can be modelled as a parallel projection.
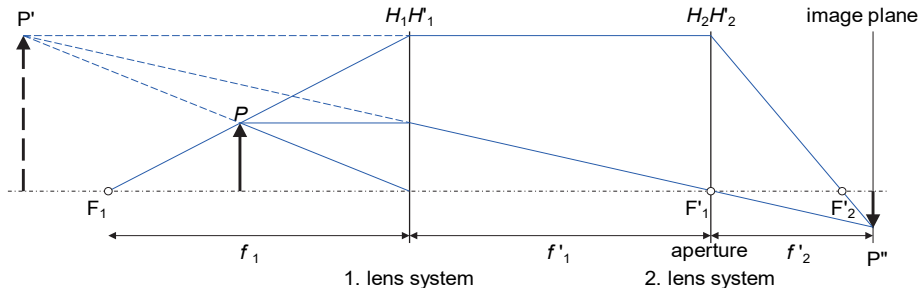
Fig. 3.108: Two-stage telecentric lens (after Schmidt 1993).

3.4.3.8 Reflective optics

Reflective optics refers to lenses where curved mirror surfaces, normally paraboloids, are used to focus the imaging rays. Parabolic mirrors have a focal point and a corresponding focal length. Since imaging does not involve optically refracting media, there is no chromatic aberration when imaging with mirrors. Reflective optics are therefore suitable for imaging across the whole wavelength spectrum.

Fig. 3.109 shows the imaging principle of a concentric arrangement of mirrors (mirror telescope) and an asymmetric arrangement (Schiefspiegler, also called tilted-component telescope). The concentric design has an area in the centre which cannot be imaged. In contrast, the asymmetric design with off-axis mirrors can acquire a complete image.
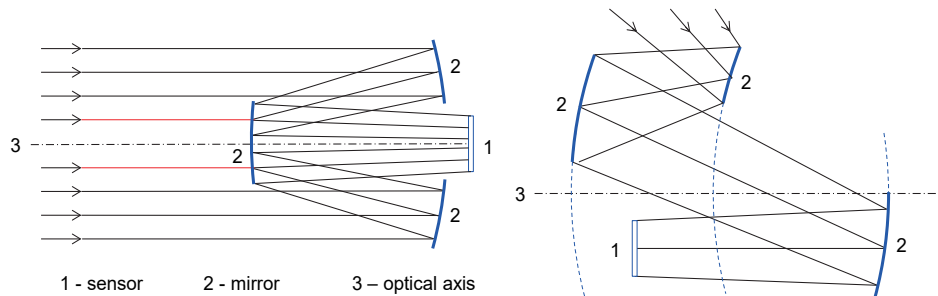
Fig. 3.109: Reflective optics (after Richter et al. 2013).

The geometric calibration of reflective optics is complex, especially for asymmetric designs. These generate large asymmetric distortions for which standard models of radial and decentring distortion are insufficient. In addition, local deviations in the mirror surfaces lead to local errors in the image which are much larger than those created by local deviations on spherical lens surfaces. The calibration can be solved

by using extended polynomials (section 3.3.4.2) or corrective grids based on finite elements (section 3.3.4.5).

3.4.3.9 Stereo image splitting

With the aid of beam splitters and mirrors it is possible, using only a single camera, to make stereo and multiple image recordings with only one exposure. This possibility is of particular interest for recording dynamic processes where synchronization of camera imaging is mandatory. A single-camera, split-image system is intrinsically synchronized and is therefore suitable for recording fast-changing events.

Fig. 3.110 shows the principle of a stereo mirror attachment with a central lens. With this arrangement, it is possible to use an existing camera/lens combination or the camera in combination with other lenses. Assuming that the mirror and beam splitting surfaces are planar, the arrangement generates the equivalent of two virtual cameras with perspective centres O' and O", each of which provides a central perspective image onto half of the image sensor area.
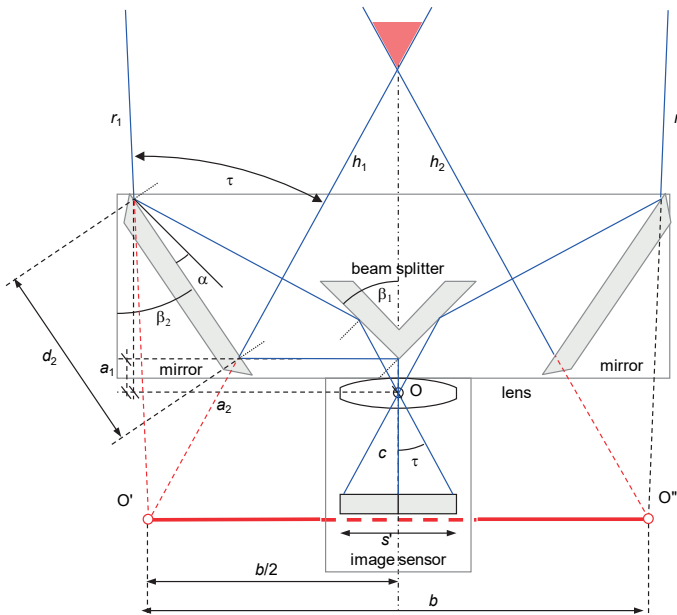
Fig. 3.110: Principle of the stereo-mirror attachment.

The stereo base b is the separation between the points O' and O". The photogrammetric exterior orientation is defined at these points, i.e. a change in the tilt or position of the mirror attachment causes a change in the imaging geometry analogous

to a conventional measurement with separate cameras. The reference point for the interior orientation of the camera is the perspective centre O located within the lens. The field of view τ for the half image is derived from half of the horizontal format $s/2$ and the principal distance c. The inclination angle of the inner mirror must meet the requirement $\beta_1 > \tau$ so that the outer ray at the edge of the image can still be reflected by the mirror. The outer mirror has an inclination angle β_2 and, relative to the inner mirror, a rotation angle of α. Angle β_2 must be smaller than β_1 so that optical axes h_1 and h_2 converge, thus ensuring stereoscopic coverage of the object space. At the same time, β_2 should be set with respect to τ such that the edge rays r_1 and r_2 diverge in order to capture an object space wider than the base b.

The size of the outer mirror is primarily a function of the field of view of the lens, as well as its rotation angle. Mirror size increases linearly with decreasing focal length and increasing mirror offset. The resulting horizontal (stereoscopic) measurement range is limited by the inner and outer imaging rays. The vertical measurement range continues to be defined by the field of view and vertical image format.

High demands are made of the planarity of the mirror surfaces. Departures from planarity result in non-linear image deformations which can only be removed by a significant calibration effort.

3.4.4 Filters

Various filters are employed in analogue and digital imaging procedures. These absorb or transmit different parts of the optical spectrum. The following types of filter are typically used. Fig. 3.111 illustrates their transmission properties:

– Ultraviolet blocking filter:
  UV blocking filters are mostly used where the light conditions have a strong UV component, e.g. in snow. They absorb all radiation under about 380 nm.
– Infrared blocking filter:
  IR blocking filters are generally employed to suppress the natural sensitivity of digital image sensors to light in the infrared part of the spectrum (see section 3.4.1.9). They work from about 720 nm and are an integrated filter layer in many digital cameras.
– Band-pass filter:
  Band-pass filters are designed to transmit a limited range of wavelengths. All wavelengths outside the band defined by the central wavelength and the half-power width are suppressed. Band-pass filters are used in optical systems where a monochromatic or spectrally narrow illumination is used and extraneous light outside this illumination band must be excluded from reaching the imaging sensor. Their use with colour imaging systems deploying matrix filters (see section 3.4.1.6) has to be carefully considered as their capability to completely

block light to some areas of the matrix pattern will impact on the colour image reconstruction process.
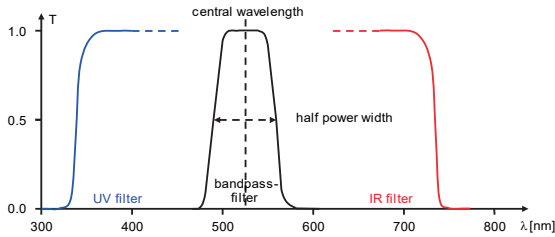
Fig. 3.111: Transmission properties of UV and IR blocking filters.

–   Polarizing filter:
    Polarizing filters (also known as pol filters) restrict the oscillation of light to a certain direction perpendicular to the direction of propagation (see section 3.1.1.5). Amongst other purposes, polarising filters can be used to reduce reflections from smooth and polished object surfaces. Fig. 3.9 shows an object with smooth surface which generates a bright hot spot from a light source which can almost be eliminated by application of cross-polarization. In this technique, one polarizing filter is attached to the light source (flash). Polarizing filters reduce the amount of light energy and often lead to more intensive colours.

3.5 Imaging systems

Imaging systems closely follow the consumer market with significant advances being made in design and capability. Photogrammetric system development parallels these advances, from making use of the very first analogue CCD camera systems to adopting lessons learned for the accelerated development of systems based on low-cost webcams and mass-market DSLR systems.

3.5.1 Industrial cameras

The term *industrial camera* describes all digital cameras which comprise only a lens, the imaging sensor and on-board electronics, and therefore have no viewfinder or manual controls (see also section 3.4.2.1). These include surveillance cameras, webcams, miniature cameras, cameras for driver-assistance systems or cameras for industrial inspection. Examples are illustrated in Fig. 3.112.
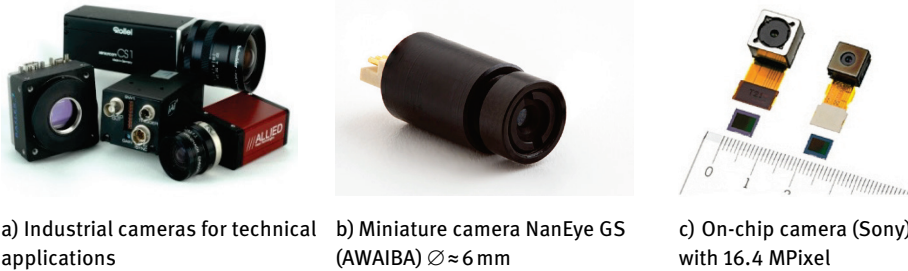
a) Industrial cameras for technical applications

b) Miniature camera NanEye GS (AWAIBA) $\varnothing \approx 6\,mm$

c) On-chip camera (Sony) with 16.4 MPixel

Fig. 3.112: Examples of digital industrial cameras.

These cameras usually use a C-mount or CS-mount lens adapter with screw thread (see also section 3.4.2.4). Whilst high quality optics are available, for off-the-shelf cameras it cannot be assumed that the individual optical components making up the lens are sufficiently well aligned to the optical axis, nor have a homogeneity that meets photogrammetric requirements. In addition, many video camera lenses have large radial and tangential distortions that must be modelled (optical distortions of tens of pixels being common). Provided that the opto-mechanical properties of the camera are physically stable, modelling is a routine process.
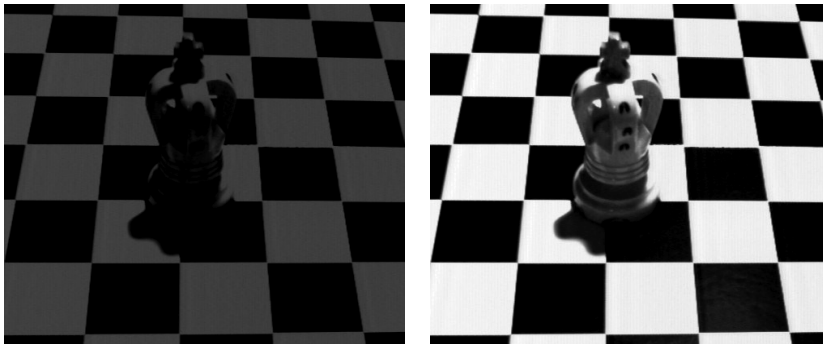
Fig. 3.113: Digital image without (left) and with (right) automatic gain control at constant illumination.

After sensor read-out the image signal is amplified and pre-processed. Subsequently its contrast and brightness may be automatically adjusted through the use of an *automatic gain control* whereby the analogue signal is amplified to the maximum amplitude. The level of amplification is controlled by the brightest and the darkest locations in the image. Many cameras allow gain control to be controlled externally (example in Fig. 3.113). Since automatic gain control is a post-processing step, loss of geometric quality and undesired brightness changes cannot be excluded. In addition,

and most noticeably under poor lighting conditions, gain control can increase image noise. Automatic gain control should be switched off under conditions of sufficient and constant illumination.

Both image sensor and video signal require high-frequency synchronization signals that are delivered by a pulse generator (*sensor clock*). Synchronization can also be controlled by an external signal, e.g. for the synchronized image acquisition of several cameras, or for the synchronization of a video frame grabber for the precise digitization of a video signal.

Modern industrial cameras store and transfer image data in digital formats. Power supply and data transfer are usually made via cable connection. In addition to the acquisition of single frames, most also allow for the recording of image sequences in standard video and audio formats, e.g. MPEG-2, MPEG-4 or AVI. Typical frame rates range from 20 to 100 fps. Standardized data transfer protocols and interfaces are available, e.g. USB, Firewire, CameraLink, Gigabit Ethernet. Wireless connections are also possible.

3.5.2 Digital cameras

Following the first developments at the end of the 1980s, such as the Kodak Megaplus, 1320 x 1035 pixel FT sensor, a growing number of high-resolution imaging sensors have become available at economic cost. In this field, consumer-orientated developments in digital photography are the driving force in that they offer a wide range of imaging systems for both amateurs (low-resolution examples in Fig. 3.114), and professional photographers (high-resolution monochrome and colour camera examples in Fig. 3.115 and Fig. 3.116).

a) Compact camera Fujifilm X100V   b) Mirrorless camera Sony Alpha 6700

Fig. 3.114: Examples of digital cameras.

High-resolution digital cameras with up to 100 Mpixel are becoming readily available for photogrammetric practice, examples including still-video cameras, scanning

cameras, digital camera backs or specialized metric cameras (examples in Fig. 3.117, Fig. 3.118). In combination with established photographic camera technologies, they provide powerful and user-friendly systems. Data transfer is either performed offline by means of an internal storage device (e.g. SD Card) or online with a connected computer (Wifi, USB, Firewire, Gigabit Ethernet). In addition, there are imaging systems with integrated processors where a variety of image processing functions are performed directly inside the camera.

Fig. 3.115: Nikon D6.

Fig. 3.116: Leica SL.

There are several trends in the current development of digital cameras which are tuned to the consumer market drive for smaller devices and improved pictorial imaging performance. Any new system needs to be carefully tested as improvements in one area can be outweighed by new metric imaging challenges from other innovations:

– smaller pixel and sensor sizes for compact and mobile phone cameras;
– larger image resolutions at economic cost in SLR image formats;
– sensor back illumination whereby light passes the shortest distance through the sensor to arrive at the light sensitive part of each pixel improving light sensitivity and reducing sensor noise;
– combination camera systems able to selectively provide still and video images from the same sensor;
– formats up to 54 mm x 40 mm in digital camera backs or medium format cameras;
– smaller choice of cameras designed specifically for optical measurement;
– high-resolution optics, usually equipped with aspherical elements to match the capabilities of smaller pixel dimensions with sharper better corrected optical images;
– a move from DSLR to high specification mirrorless cameras as improvements in battery capacity and display technologies replace the need for optical viewing systems;

- actuated sensor movement to compensate for camera motion during exposure and potentially utilise sensor motion for autofocus and focus stacking. Note that whilst desirable pictorially, mechanical instabilities in the relationship between lens and sensor from image to image are likely to impact photogrammetric performance;
- onboard manufacturer's lens distortion correction which can benefit from knowledge of the specific lens design. Challenges are that corrections are typically not applied on a per-lens basis and can be irreversibly built into all output images making lens distortion correction for more accurate applications challenging.

In close-range photogrammetry, digital compact cameras are only used for measurements with lower accuracy requirements, for example applications such as accident recording, UAV applications or texturing of 3D city models. Digital SLRs are used in many photogrammetric applications due to their favourable price/performance ratio. Specialized SLRs with robust metal housings and high-quality lenses can achieve high measurement accuracies with appropriate calibration.

Fig. 3.117: Leica S.

Fig. 3.118: Hasselblad H4D-60.

In principle, digital cameras with a small image format (examples in Fig. 3.115 and Fig. 3.116) permit the use of standard lenses for this type of camera so that a wide choice of lenses is available, particularly for wide-angle use. However, cameras with very small pixel sizes have greater noise in the image and are less sensitive to light. In addition, the corresponding lenses are subject to high demands with respect to resolution which are often not fulfilled, resulting in a lower image quality than available from cameras with a lower number of pixels.

Medium-format digital cameras (Fig. 3.117, Fig. 3.118) are mainly used in high-end applications requiring a maximum number of pixels (up to 100 Mpixel) and large

image format (up to 54 mm x 40 mm). Due to their high cost, these medium-format cameras do not have a significant presence in close-range photogrammetry.

Purpose-built digital metric cameras, specially designed for measuring purposes, guarantee high geometric stability due to their opto-mechanical design. In particular, lens and image sensor are rigidly connected in the camera housing. An integrated ring flash is used to illuminate retro-reflecting targets. The DynaMo series of metric cameras from Geodetic Systems Inc. (GSI), one of which is shown in Fig. 3.119, have resolutions ranging from 5 Mpixel to 45 Mpixel, frame rates from 15 Hz to 100Hz and accuracy to 5μm + 4μm/m (RMS-1σ). A maximum length-measuring error of around 25μm can be anticipated for the 12 and 45 Mpixel cameras, according to VDI/VDE 2634/1 (see section 7.2.3). The DynaMo cameras are designed for high-accuracy industrial metrology applications, and they incorporate special features such as built-in image processing, proprietary very fast image compression, long-life LED flash systems and power-over-ethernet cabling.

The Hexagon C1 (Fig. 3.120) is a metric camera utilizing a full-format CMOS imaging sensor with 50.6 Mpixel which is installed together with a 28mm lens in a particularly dust- and shock-proof housing. A projection element displays the object area visible from the camera. The photogrammetric accuracy has been evaluated at around 2μm + 5μm/m (RMS-1σ) with a maximum length-measuring error of 15μm+ 15μm/m according to VDI/VDE 2634/1 (see section 7.2.3).

Fig. 3.119: GSI DynaMo/D45.

Fig. 3.120: Hexagon DPA Industrial C1 camera.

The SingleCam metric camera from AXIOS 3D (Fig. 3.156) is a mechanically stable video camera with an optimized assembly of lens, sensor and housing. It has an integrated, monochromatic, LED ring flash for recording retro-reflecting targets. The camera is shock-resistant to 50 times the acceleration of gravity without any change in interior orientation.

3.5.3 High-speed cameras

High-speed cameras allow the recording of fast-changing scenes with image frequencies much higher than for standard video cameras. Today, high-speed cameras exist that can record more than 1200 images per second at resolutions of the order of 2000 x 2000 pixel. They are mainly used for monitoring dynamic production processes, but are also used for analysing object movement in industrial, medical and research applications. For special applications frame rates of more than 1 million fps can be reached at highly reduced pixel resolutions.

In general, high speed imaging systems can be characterized as follows:

– CMOS sensors:

As an alternative to progressive-scan sensors, CMOS imaging sensors are available that provide exposure times of a few microseconds using an electronic shutter which is integrated in the sensor. These sensors facilitate direct addressing of individual pixels, but have the disadvantage of increased image noise when compared with CCD sensors.

– Electronic shutter:

Electronic shutters control the integration time of the imaging sensor with exposure times of down to 50 μs, corresponding to 1/20000 s.

– Solid-state image memory:

Solid-state image memory modules permit immediate storage of the image sequence either in the camera or nearby. The maximum number of images, and the maximum recording period, depend on the image memory capacity and the (selectable) sensor resolution.

– External trigger signals:

External trigger signals serve to control the image acquisition process by allowing the camera to be synchronized with an external event, e.g. a particular point on a machine production cycle or the simultaneous acquisition of image sequence from multiple cameras.

– Recording of additional information:

Storing of additional information (image number, data rate etc.) allows for subsequent image sequence analysis.

– Onboard processing:

CMOS sensors, combined with specialized Field Programmable Gate Array (FPGA) processors, can process incoming images in real-time. For example, the AICON MoveInspect HF4 is a specialized photogrammetric high-speed camera (Fig. 3.121b). It has a 4 Mpixel CMOS image sensor and an FPGA capable of automatically measuring up to 10000 bright targets at an image frequency of up to 500 Hz with full sensor resolution. Only the measured image coordinates are transferred to the externally connected computer and the current image is then immediately overwritten. In this way, recording sequences can be of any

duration. The camera is used, for example, to measure wheel movements on a moving vehicle (see section 6.10.2.2).

High-speed cameras are offered by different suppliers, including Excelitas PCO, IDT and Microtron. For very demanding environmental conditions, such as on-board recording of crash tests, special cameras are available which can tolerate high levels of impact and acceleration and still maintain a very stable interior orientation. Fig. 3.121a shows two such cameras, with control unit, from Excelitas PCO. As an application illustration, Fig. 3.122 shows a subset of an image sequence recorded by the pco.dimax (1920 x 1080 pixels, 1100 frames per second). The duration of the recording sequence is limited by the internal or external storage. This camera type can be of the cabled or self-contained variety.
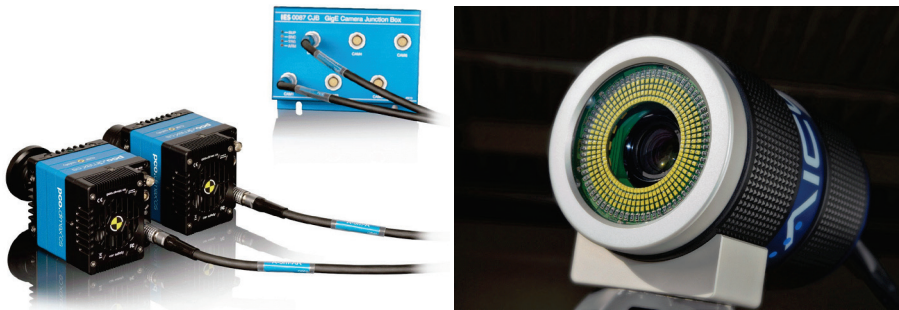
a) High-speed camera pco.dimax cs   b) AICON MoveInspect HF4

Fig. 3.121: High-speed cameras.

Synchronization of several high-speed cameras is a demanding requirement since sensor exposure must be guaranteed even at very high frame rates. Usually the cameras are synchronized by an external trigger signal, or in master-slave mode where one camera triggers the other. Departures from synchronization lead to photogrammetric measurement errors (see section 6.10.1). Where cameras of the same product type are used, synchronization errors should be less than 50 µs. Synchronization can be tested optically, e.g. using the Synchronometer from manufacturer IES (Fig. 3.123). This device displays a high-speed sequence of LED codes which must be imaged in the same sequence by a set of synchronized cameras.

Synchronized high-speed stereo images can also be recorded using a stereo mirror attachment. (see section 3.4.3.8). Fig. 3.124 shows a high-speed camera with the stereo mirror attached. There are demanding requirements in respect of planarity and alignment of the mirrors, and significant asymmetric distortion effects if these are not met. Fig. 3.125 shows radial and decentring distortion for the left (red) and the

right (green) half of an image taken with a stereo mirror. It can easily be seen that the mirrors cause both image halves to display completely different characteristics.
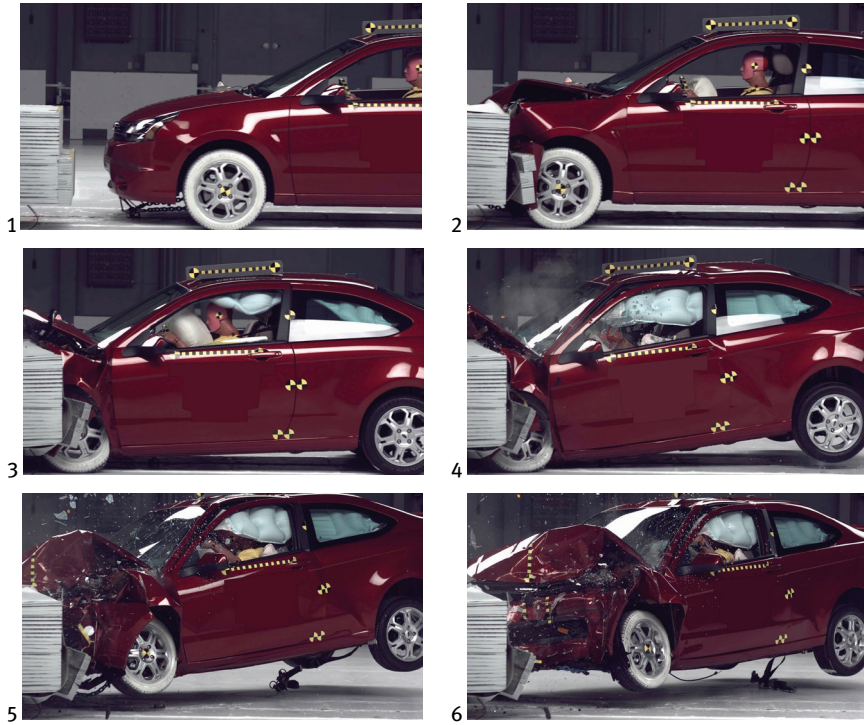
Fig. 3.122: Subset of an image sequence (PCO).
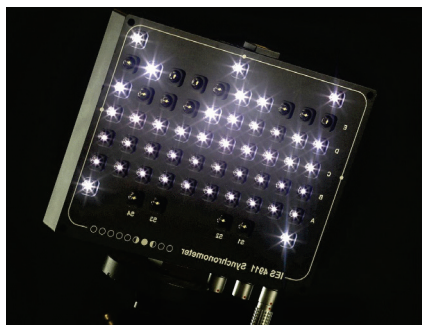
Fig. 3.123: Synchronometer (IES).
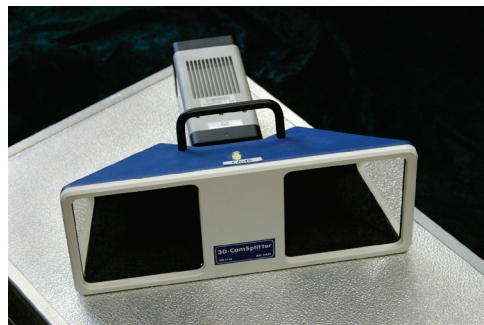
Fig. 3.124: High-speed camera with stereo mirror attachment (High Speed Vision).
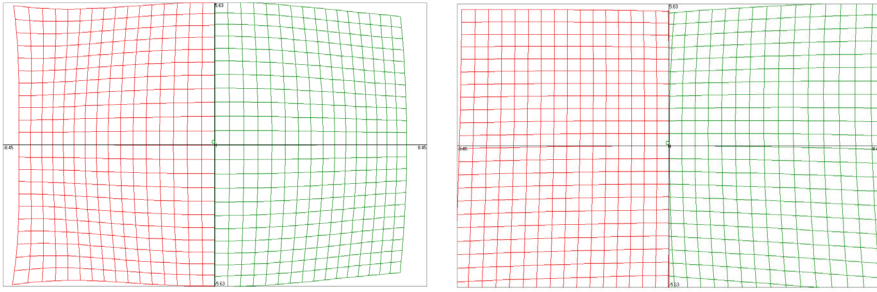
Fig. 3.125: Radial (10-times enlarged, left) and decentring distortion (100-times enlarged, right) in a camera with a stereo mirror attachment (Weinberger Visario, focal length 12.5 mm).

Fig. 3.126 shows two images from a stereo recording sequence. Only half the sensor format is available for the left- and right-hand parts of the image. Between each half image there is a narrow strip with no useful image data.
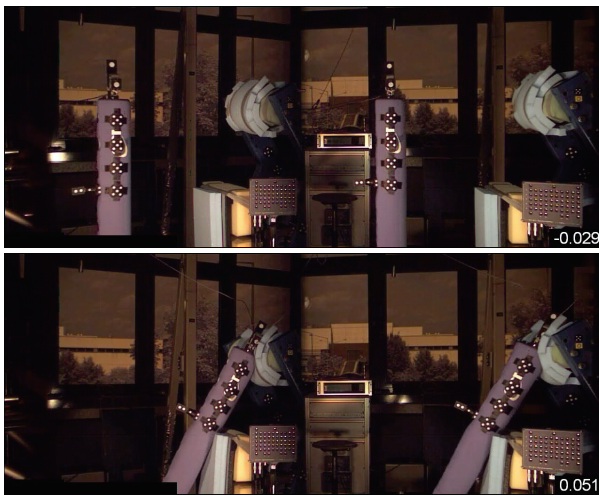
Fig. 3.126: High-speed stereo images taken with the stereo mirror attachment (Volkswagen).

3.5.4 Stereo and multi-camera systems

Stereo and multi-camera systems, in which the cameras are rigidly mounted in a single housing, are suitable for 3D applications requiring the synchronous recording of at least two images (Fig. 3.127). These systems permit the imaging of object points with simultaneous calculation of their 3D coordinates by the technique of spatial intersection (see section 4.4.7.1). The cameras have largely constant orientation

parameters and are normally calibrated in the factory. For example, the CamBar stereo camera system from AXIOS 3D (Fig. 3.127a) is shock-resistant to 75 g and specified to operate in the range 15-30° C, without any change in interior orientation of an individual camera or the relative orientation between cameras in the housing.

a) Stereo camera for use in medical navigation applications (AXIOS 3D)

b) 3-camera system for monitoring dynamic processes (AICON)

Fig. 3.127: Stereo and multi-camera systems.
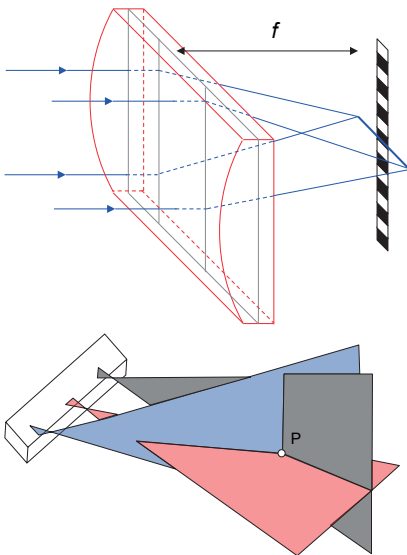
a) Three-line measurement principle

b) Three-line camera (Nikon Metrology)

Fig. 3.128: Principle and example of three-line measurement.

Another form of multi-camera system is represented by three-line systems which have three linear imaging elements. Each element has a cylindrical lens which creates a line image of a LED target point, with the image being recorded by a linear CCD array

set in the image plane at right angles to the image line (Fig. 3.128a). This effectively defines a plane in space on which the target point lies. By orienting the central element at right angles to the outer elements, the target point can be located by the intersection of three planes in space. Linear array technology permits a measurement frequency of up to 3000 Hz which is temporally split across the number of active targets being imaged. Fig. 3.128b shows a camera system based on this principle.

Stereo and multi-camera systems are often used for navigation tasks in which objects must be absolutely tracked or positioned in space, or located relative to other objects. One of the most common application areas is in medical navigation. Here the systems are employed during operations to track the movements of surgical instruments. Typical industrial applications include the positioning of parts and robots (see also section 6.9.3).

3.5.5 Micro and macro-scanning cameras

Since the resolution of digital cameras was a limiting factor some years ago, scanning cameras were developed in order to increase pixel resolution by sequential scanning of an object or scene. Two basic scanning principles can be distinguished: *micro scanning* where small sub-pixel steps of the sensor are made sequentially within the same image format area to increase spatial resolution and *macro scanning* where the sensor is sequentially stepped by a significant portion of its total dimension to expand the image format. Such systems can only deliver high quality images if there is no relative movement between camera and object. Depending on implementation, image sizes of the order of 3000 x 2300 pixels to 20 000 x 20 000 pixels can be obtained.

3.5.5.1 Micro scanning

In the case of micro-scanning cameras, piezo elements are used to translate an interline-transfer sensor horizontally and vertically, in fractions of the sensor element size (micro-scan factor) (Fig. 3.129). A high-resolution output image is created by alternating the storage of single images. The final image has a geometric resolution that is increased by the micro-scan factor in both directions. From a photogrammetric standpoint, the micro-scanning principle results in a reduction of effective pixel size whilst maintaining the usable image format of the camera.

Cameras based on this principle are the RJM JenScan (4608 x 3072 pixel), Kontron ProgRes 3012 (4608 x 3480 pixel) and Jenoptik Eyelike (6144 x 6144 pixel, Fig. 3.131). Even higher resolutions can be achieved if micro scanning is combined with the principle of macro scanning (see below).
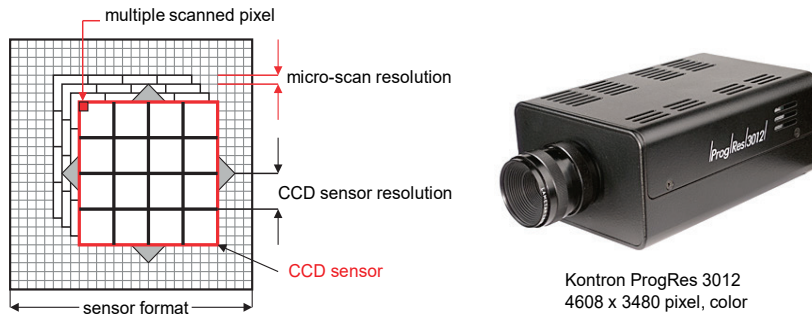
Fig. 3.129: Principle and example of a micro-scanning camera.

3.5.5.2 Macro scanning

Macro-scanning systems shift a linear or area imaging sensor in large steps over a large image format. The position of the separate partial images is determined either mechanically or by an optical location technique. This enables the images to be combined into one complete, large-format image.
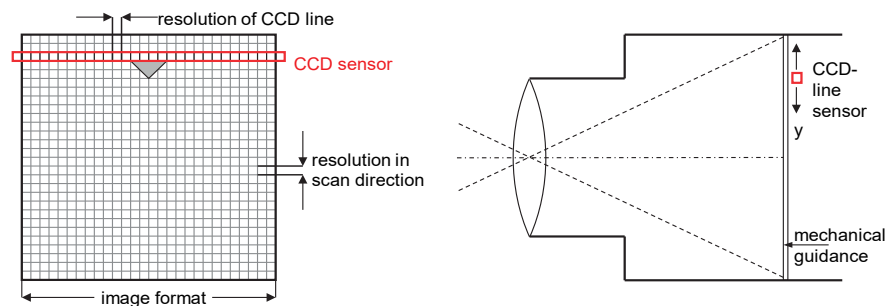
Fig. 3.130: Principle of a line-scanning camera.

Fig. 3.130 illustrates the principle of a line-scanning camera. A single linear CCD array, set parallel to one axis of the image, is mechanically moved along the other image axis in order to scan the entire image area. Image resolution is therefore given in one direction by the resolution of the array and in the other direction by the step resolution of the mechanical guide.

An example of a high-resolution, line-scan camera is the PentaconScan 6000 shown in Fig. 3.132. This camera has a linear RBG CCD sensor which is scanned across an image format of 40 mm x 40 mm and, at its highest resolution, can deliver images with 10 000 x 10 000 pixel for each colour channel. Due to limitations in the precision which is achievable in the mechanical guidance mechanism, these systems are

principally designed for professional still photography and not for photogrammetric applications.

Fig. 3.131: Jenoptik Eyelike.

Fig. 3.132: Pentacon Scan 6000.
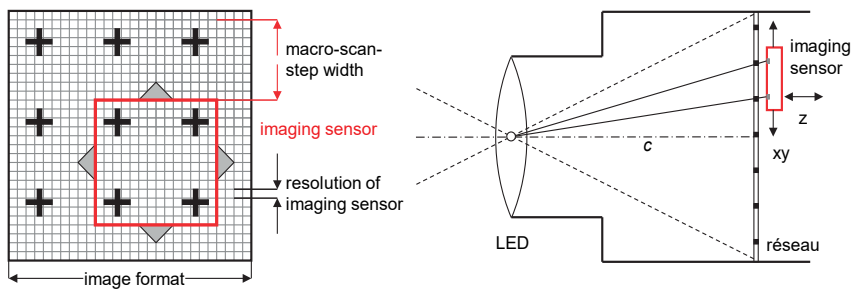
Fig. 3.133: Principle of the réseau-scanning camera.

The macro-scanning principle can also be employed with area sensors. The two-dimensional mechanical positioning technique which is required in this case is considerably more complex and therefore more expensive to produce. The réseau-scanning principle offers an alternative solution. Here the individual sensor locations are determined by measuring the images of réseau crosses in the sub-image delivered by the sensor. This technique does not require an accurate mechanical positioning mechanism (Fig. 3.133). In addition to movement in the xy direction, the sensor can also be shifted in the z direction, parallel to the optical axis. This makes it possible to focus without altering the parameters of interior orientation. An example of a réseau-scanning camera is the Rollei RSC which is no longer in production (see Fig. 1.43).

3.5.6 Panoramic cameras

3.5.6.1 Line scanners

Digital panoramic cameras with a scanning line sensor form a special case of the macro-scan technique. A vertically mounted sensor line is rotated about a vertical axis, thereby imaging the surrounding object area (rotating line scanner). The mathematical projection model is a central perspective projection in the vertical direction and mapped by rotation angle to a cylindrical projection surface in the horizontal direction. Some systems internally convert the camera output to a spherical projection. Providing the particular panoramic imaging geometry is taken into account, photogrammetric methods such as bundle adjustment and spatial intersection can be applied in an analogous way to conventional images. Scanning panoramic cameras based on CCD line sensors can achieve image sizes of the order of 50 000 x 10 000 pixels, often known as gigapixel images.
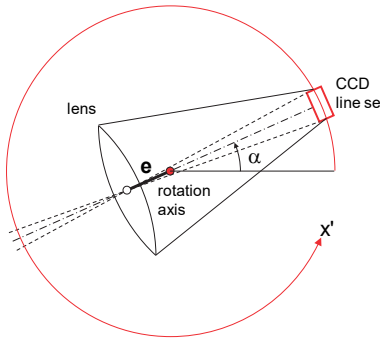
Fig. 3.134: Principle of a panoramic scanning camera.

Fig. 3.135: Digital panorama camera SpheroCam HDR (Spheron).

An eccentricity vector **e** between perspective centre and rotation axis must be determined by camera calibration (Fig. 3.134). For photogrammetric processing of panoramic images, the mathematical model is extended for the x' direction whereby the image coordinate x' (column position in image) is defined as a function of the rotation angle α and the eccentricity vector **e**. A number of panoramic cameras are currently available on the market, e.g. the SpheroCam HDR from Spheron (Fig. 3.135, 50 Mpixels) or the Panoscan Mark III (max. 9000 x 6000 pixels). Most cameras provide 360° images and can also produce smaller, user-definable sections.

Panoramic images can also be generated from single frame images if the individual images have a sufficient overlap. A geometrically exact stitching of such

panoramas can only be performed if the single images are acquired around a common rotation axis, and if the interior orientation parameters are known.

3.5.6.2 Panorama stitching

A panoramic image can be generated from a number of central-perspective, single images which overlap horizontally or vertically (see example in Fig. 3.138). The individual images can be merged together in various ways:

– Manual merging:
  The overlapping images are interactively positioned relative to one another until there is a good match in the areas of overlap.
– Automatic merging:
  Procedures for automatically creating the panoramic image are based on detecting corresponding regions in the overlaps using feature extraction or correlation methods (see section 5.4.2). Once these common locations are found, a simple geometric transformation is used to make a best fit between them in the overlaps. The technique makes no use of a global geometric model, such as a cylindrical projection, and this gives rise to residual errors where neighbouring images are connected.
– Photogrammetric merging:
  The photogrammetric construction of a panorama takes into account the outer and inner orientation of the individual images. In an analogous way to orthophoto creation (section 4.2.8.2), the resultant panorama is generated on a cylindrical surface with the colour values calculated by back-projection into the original images.

Capabilities are included in cameras of all levels from mobile phones to DSLRs.

The geometrically correct generation of a panoramic image requires that the rotation axis of the camera passes through the perspective centre. Only then do all imaged rays pass through the same point, as would be the case in a perfect panoramic image.

To achieve this condition it is advantageous to connect the camera to an adjustable panoramic adapter, one which permits a shift along the camera axis (Fig. 3.136). This adjustment can be achieved by simple means, e.g. the imaging of two points in line with the rotation axis (Fig. 3.137). When correctly adjusted, both object points overlap in the same image point. When the ideal configuration is achieved and the rotation axis passes through the perspective centre, then the resulting panoramic image has the same central-perspective properties as a normal image, although in this case the horizontal field of view is 360°. If there is a residual displacement error, then the camera's perspective centre moves on a circle around the rotation axis. Note that in panoramic photography the perspective centre is also known as the nodal point. Fig. 3.138 shows an example panoramic image created by the camera in Fig. 3.136.

Fig. 3.136: Digital camera on a panoramic adapter.
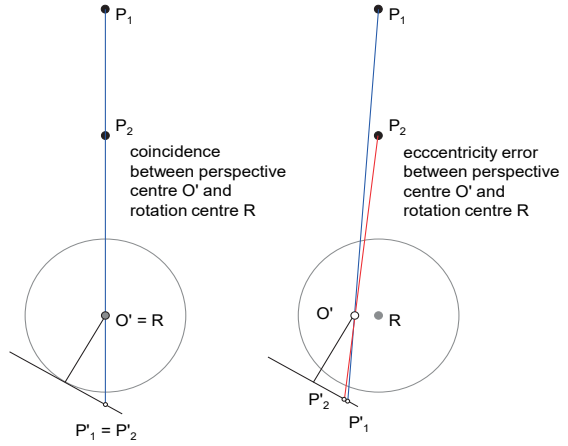
Fig. 3.137: Alignment of perspective centre with rotation centre by observing collinear points.

Fig. 3.138: A digital cylindrical panorama created by stitching together overlapping camera images.

As an alternative to the use of a camera, a panoramic image can also be generated by a video theodolite, or imaging total station with integrated camera (see section 6.3.2.1). A panoramic camera with integrated laser distance meter is presented in section 6.8.3.

3.5.6.3 Panoramas from fisheye lenses

360° panoramas can also be created by combining two images taken with fisheye lenses (see section 3.3.7) which are pointing in opposite directions. Fig. 3.139 shows two original fisheye images and Fig. 3.140 the resultant panorama. The source of the images is a system designed for mobile mapping (see section 6.11.1). A wide range of action and panoramic cameras are now available which have been designed for leisure and web applications rather than for professional metrology (Fig. 3.141). They are constructed from multiple individual cameras which enables panoramic images and videos to be acquired and processed directly.

a) Camera view to front

b) Camera view to rear

Fig. 3.139: Images from two cameras with fisheye lenses which are pointing in opposite directions (Cyclomedia).

Fig. 3.140: Digital panorama created from the images in Fig. 3.139 (Cyclomedia).

a) 360° camera (Giroptic) b) Ball with 32 cameras (Panono)

Fig. 3.141: Panoramic cameras with multiple individual cameras.

3.5.7 Endoscopes

A distinction is made between rigid and flexible endoscopes. Rigid endoscopes (borescopes) usually have a camera attached to the end, whereby light and image transmission are usually achieved with a rod lens system. Optical fibres are typically used for both light and image transmission in flexible endoscopes. However, the number of individual optical fibres in the connecting coherent fibre bundle will limit image resolution. Due to ever smaller image sensors and LED light sources, it is now possible to build one or more cameras into the tip of the endoscope (chip-on-the-tip) along with the necessary illumination. Fig. 3.142 shows examples of endoscopes.
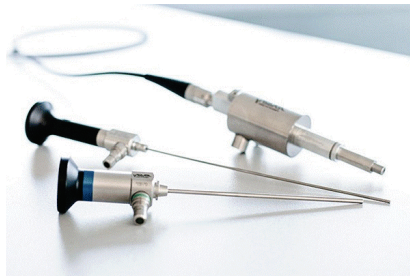
a) Rigid endoscopes without camera b) Stereo endoscope with two miniature cameras in the tip of the endoscope

Fig. 3.142: Examples of endoscopes (Schölly).

Photogrammetric calibration of endoscope cameras can cause particular problems due their small field of view and depth of field. Flat checkerboard test fields are often used for this purpose, but they tend to create high correlations between interior and

exterior orientation parameters (see sections 4.4.5.5 and 7.3.2). However, it is also possible to use miniaturised 3D test fields equipped with conventional targets. Suitable arrays can be 3D printed.

In addition to monocular endoscopes, there are also stereo endoscopes (Fig. 3.142b). These were designed to give the observer a visual 3D impression, but they can also be used photogrammetrically and offer advantages in the matching of image points through the use of epipolar geometry (section 5.5.4). Fig. 5.98 shows the experimental setup of a trinocular endoscope that avoids ambiguities in image assignment. The third camera could also be replaced by another image sensor, e.g. an infrared camera or range sensor.

3.5.8 Thermal imaging cameras

Thermal imaging cameras (thermographic cameras) are used for analytical tasks in building research and materials science. They deliver a thermal image using wavelengths in the range 2.5 to 14 µm (medium to near infrared part of the spectrum, see Fig. 3.2). Typical thermal sensitivity of the cameras lies approximately in the range −40° C to +2500° C. Until now, thermal imaging cameras have been rarely used for geometric and photogrammetric analyses because they are costly and, with sensor resolutions in the range 320 x 240 pixel to 1024 x 768 pixel, have a relatively poor resolution compared with modern CCD and CMOS sensors.

Fig. 3.143: Thermal imaging camera (InfraTec).
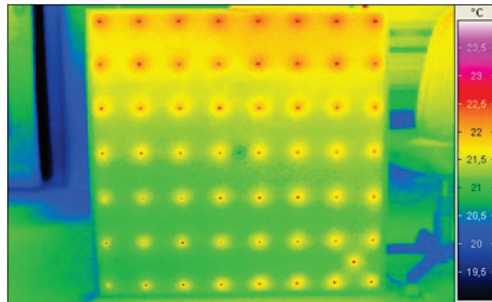
Fig. 3.144: Thermographic test-field image.

Thermal imaging cameras (example shown in Fig. 3.143) work either with thermal detectors or quantum detectors. Thermal detectors (micro-bolometers, pyro-electric detectors) measure the temperature directly incident on the sensor, independently of wavelength. The cameras do not need to be cooled, only stabilized at their own constant temperature. In contrast, quantum detectors absorb incident infrared

photons and operate in a similar way to CCD and CMOS sensors. They must be cooled to very low temperatures (60-140 K) and are functionally dependent on wavelength.

Typical sizes of detector elements (pixel sizes) lie in the range 17 µm to 50 µm. According to eqn. (3.13), the diffraction-limited resolution of thermal imaging cameras is around 12 µm for a wavelength of 5 µm and an f/number of 1. Because of the requirements imposed by the refractive indices of thermal radiation, the lenses which can be used are not constructed from normal glass but from germanium. They are expensive to produce and therefore represent the principal component cost of a thermal imaging camera.

Like conventional cameras, thermal imaging cameras can be geometrically calibrated. Test fields for this purpose have target points which radiate in the thermal region, for example active infrared LEDs, heated or cooled metal discs, or retro-reflecting targets exposed to a suitable infrared illumination source. Fig. 3.144 shows a thermal image of such a photogrammetric test field.

3.5.9 Multi-spectral and hyperspectral cameras

Multi-spectral or hyperspectral cameras are designed to capture images over specified wavelength ranges either as a continuum of colour samples across a designated area of the spectrum or in discrete bands. Sensors may be silicon based, for example CMOS or CCD, which are capable of acquiring images over wavelengths in the range of ca 350 to 1000 nm covering near UV (NUV), visible and near IR (NIR) parts of the spectrum, (see Fig. 3.1 and section 3.4.1.8). Beyond the visible spectrum, different sensor and recording principles are needed. For example, InGaAs-sensors can be used to collect data in the short-wave infrared band (SWIR) between 1000 nm and 2500 nm. Lens design in these cases must be carefully considered as optical glasses and materials have different spectral transmission properties and need to be optimised to achieve consistent image quality over the selected wavelength range.

3.5.9.1 Multi-spectral cameras

Multi-spectral cameras are designed to simultaneously acquire images across defined broad spectral bands. Typically, single cameras are combined such that each sensor is sensitive to a specific spectral band by incorporating a band pass filter in its optical path (section 3.4.4). Fig. 3.145a shows a camera with five 1280 x 960 pixel sensors which record images in the blue (459-491 nm), green (546-574 nm), red (660-676 nm), "red edge" (711-723 nm) and near IR (814-870 nm) bands. Cameras with up to 10 spectral channels are also available. The red to near IR channels are often selected for vegetation detection and classification, e.g. by band combinations such as the normalized difference vegetation index (NDVI) which calculates the ratio (NIR-Red)/(NIR+Red) based on the red-edge phenomena.

The calibration of interior orientation must be performed for each camera individually. Because the combined sensor will have a series of physically separate optical centres, it is useful to apply a mathematical geometric constraint for fixed relative orientation between the cameras in order to stabilize the calibration process (see section 4.4.2.3).

Fig. 3.145b shows example multi-spectral images acquired by the camera in Fig. 3.145a. It can be seen that different object types have significantly different reflection properties in each band. The false-colour composite infrared image (CIR) displayed is derived from combining red, green and NIR bands. Due to the different positions of the lenses of each sensor, a clear displacement of colours can be observed. Multispectral images typically have fewer than 20 bands with each band covering a wider wavelength range. Cameras of this type are relatively new for close-range purposes, but are long established in aerial mapping.

a) Sensor head (MicaSense)

b) Example multi-spectral bands (top: green and red; bottom: infrared and CIR image) (Hochschule München)
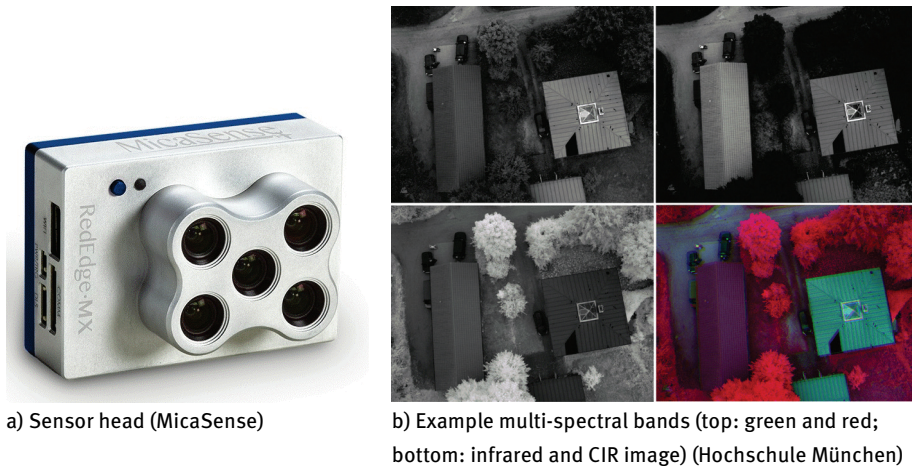
Fig. 3.145: Multispectral camera for five spectral bands.

3.5.9.2 Hyperspectral cameras

Hyperspectral cameras generate images with a spectral profile per pixel. Spectral profiles consist of between 20 to several hundred narrow spectral bands, each band being 10 nm or less in width. Ideally spectral bands will be continuous over the wavelength range of the imaging system. There are three common hyperspectral camera designs: spatial scanning, spectral scanning and snapshot cameras.

For hyperspectral spatial scanning cameras, the typical optical separation of the spectral ranges is done with the help of a dispersion element (prism, grating, cf. Fig. 3.5), whose image is recorded with an area sensor having n x k pixels. Fig. 3.146 shows the imaging principle within the camera. The object is scanned line by line with n pixels. Each pixel of this line is expanded via several optical elements to form a

hyperspectral line, which is then recorded in the sensor column with k lines. Area images can only be captured by sequentially scanning the object.

Hyperspectral spatial scanning cameras are rarely used for photogrammetric tasks because they are expensive and can only record areas if the exterior orientation of each individual image can be determined with sufficient accuracy. This can be realized, for example, by a motorized shift or rotation of the camera or a known object movement (e.g. conveyor belt). Their advantage lies in their high radiometric resolution, which can be used for material or object classification.
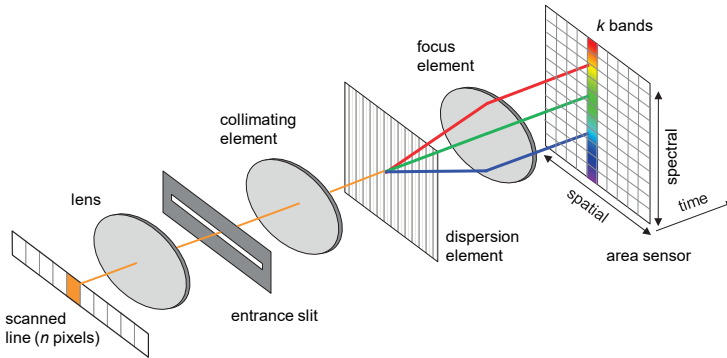
Fig. 3.146: Imaging principle of a line-scanning hyperspectral camera.

Hyperspectral scanning cameras capture spectral images over time. To do this they rely on sequential exposures utilising different narrow band filters each of which delivers a spectral band. This gives the advantage of maintaining spatial resolution at the disadvantage of requiring the relationship between the camera and the scene to remain constant whilst successive exposures are made. A typical implementation is to use a filter wheel or optically tuneable filter in front of the lens.

Hyperspectral snapshot cameras use a frame sensor which is covered by a filter mask where each pixel in a defined neighbourhood (e.g. 4x4 or 5x5 pixels) is sensitive to a specific narrow spectral band. Comparable to RGB mosaic filters such as the Bayer system (see section 3.4.1.6), spectral information needs to be spatially interpreted across the image using demosaicing. Therefore, the actual spatial resolution is reduced compared with the line scanning principle. As an example, a snapshot camera with a 2048 x 1088 array sensor would logically yield an effective resolution of 409 x 216 using a 5x5 filter grid (25 spectral bands).

3.6 Reflection and illumination

3.6.1 Reflection models

This section summarises some essential aspects of the reflection of light at material surfaces. Basic knowledge of reflection types is important for understanding the intensity and colour of a captured image signal, but also has significance for the design of targets. In addition, the generation of synthetic images is based on the consideration of surface and material properties, light sources and recording positions, which essentially determine the reflectivity.

3.6.1.1 Reflection types

The *ambient* reflection model is based on the idea that a proportion k_a of the incident light I_i reflects uniformly in all directions. The reflected radiation I_r is then given by:

$$I_r(\lambda) = k_a \cdot I_i(\lambda) \qquad \text{where } 0 \leq k_a \leq 1 \tag{3.99}$$

This model does not consider the spatial orientation of the surface. The factor k_a results in a uniform intensity variation of the whole scene.
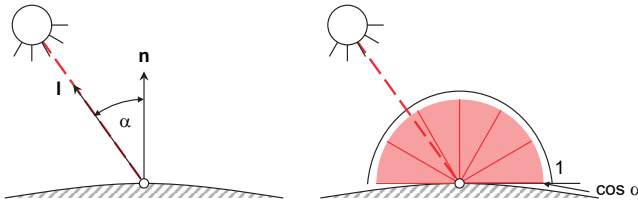
Fig. 3.147: Diffuse reflection.

Diffuse reflection (Fig. 3.147) is based on Lambert's law. The intensity of the reflected light reduces as a function of the cosine of the angle α between the surface normal **n** and the direction to the light source **l** and is expressed by the scalar triple product of the two vectors:

$$I_r(\lambda) = k_d(\lambda) \cdot \max((\mathbf{n} \cdot \mathbf{l}), 0) \cdot I_i(\lambda) \tag{3.100}$$

The term $k_d(\lambda)$ is material dependent and defined as a function of wavelength. It is therefore dependent on the colour of the illuminant, the surface and the colour sensitivity of the sensor or observer. The reflected component is independent of the observer's position. Techniques like Reflectance Transformation Imaging (RTI, see section 3.6.3.5), photometric stereo and shape-from-shading calculate normals from reflection.

The basic idea of specular or *mirror-like* reflection (Fig. 3.148) is the principle that angle of incidence = angle of reflection. For a perfect specular reflector or mirror the light is reflected in direction **r**. Consequently, in viewing direction **b** the intensity is given by:

$$I_r(\lambda) = \begin{cases} I_i(\lambda) & \text{if } \mathbf{b} \cdot \mathbf{r} = 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.101}$$
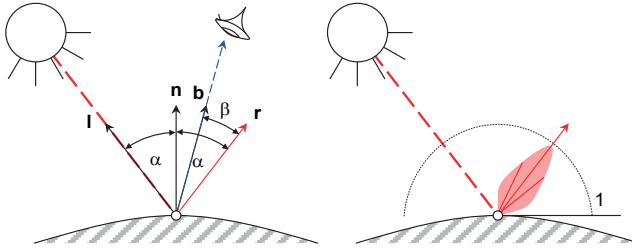
Fig. 3.148: Specular reflection.

Since perfect mirrors rarely exist in reality, and as they would generate a signal only in one direction **b** = **r** (β = 0), some spread around this direction is permitted:

$$I_r(\lambda) = k_s(\lambda) \cdot \max((\mathbf{b} \cdot \mathbf{r}), 0)^m \cdot I_i(\lambda) \tag{3.102}$$

Here the exponent m defines the spread characteristic which is dependent on material and term $k_s(\lambda)$ specifies the spectral reflective properties. Large values of m describe polished surfaces (metals, mirrors), small values specify matt, non-metallic surfaces. In general, the roughness of mirror surfaces must always be smaller than half of the incoming wavelength.

Natural surfaces often have a mixed reflection, with ambient, diffuse and specular components overlapping. The radiation received by an observer or a camera depends on the following influencing variables:

– spectral characteristics of light sources,
– spatial position of light sources,
– properties of the atmosphere (media),
– spectral characteristics of surface materials (reflection, absorption, transmission),
– surface structure (roughness),
– location and orientation of surfaces,
– optical and radiometric characteristics of the camera or sensor,
– location and orientation of the camera or observer.

With advancing computational capability, illumination and reflection models play an increasingly important role in photogrammetry, e.g. in multispectral image analysis, in the simulation of synthetic images, in the design of a suitable illumination situation or in evaluation methods that explicitly include position, direction and spectral properties of light sources (e.g. Photometric Stereo or RTI, see section 3.6.2.5). Methods of optical 3D metrology, computational photography and material sciences flow together here.

3.6.1.2 Retro-reflection

Retro-reflectors reflect incident light back in the direction of the light source over a relatively large range of incident angles, i.e. the reflector is not required to point accurately back towards the source of illumination. The technical realisation is achieved with retro-reflective spheres or with triple mirrors (corner cubes, see Fig. 3.149 showing two of the mirrors). Retro-reflective film, for example, has a dense array of spheres of approx. 30-50 µm diameter embedded in its surface, and is suitable for the production of retro-reflective targets (see section 6.2). When used with a camera equipped with a light source close to the objective, typically a ring light, the strong reflection back to the recording camera produces high-contrast imaging of targets under a wide variety of ambient light levels (see Fig. 3.155 or Fig. 3.156).

a) Reflection by a spherical mirror   b) Reflection by mirrors with normals at 90°
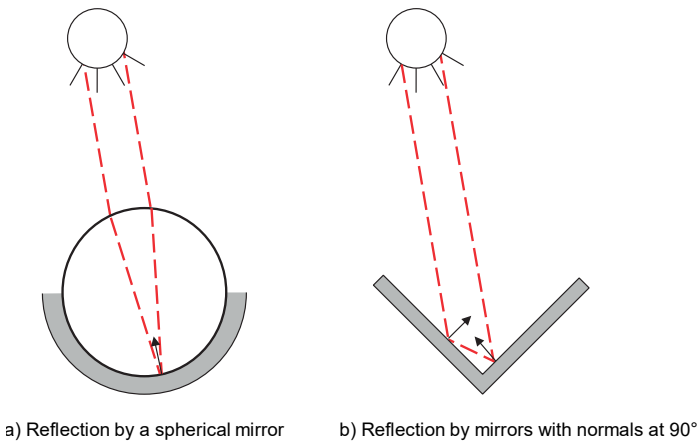
Fig. 3.149: Retro-reflection.

For retro-reflective spherical mirrors, the viewing angle is limited to ±45°. Triple-mirrors allow larger viewing angles. One alternative concept makes use of glass spheres with a refractive index of 2 at the wavelength of the illuminating light. This achieves a total internal retro-reflection without the need for a reflecting surface seen in Fig. 3.149a. However, only a small proportion of light is returned by this design.

In optical metrology large retro-reflectors based on triple mirrors or prisms are mainly used with laser trackers and totalstations (see section 6.3.2).

3.6.2 High contrast photogrammetric targeting

Applications that require accurate photogrammetric measurement typically make use of purpose engineered targets whose optical centres can be reliably and accurately determined from multiple camera views. Specific designs and target configurations are presented in section 6.2.

3.6.2.1 Retro-reflective targets

Retro-reflective targets (retro-targets) are widely used in practice, particularly in industrial applications. For typical photogrammetric applications, they consist of a retro-reflective material that is either covered by a black pattern mask, or is stamped into the target shape. Usually retro-targets are circular in shape, but they can be manufactured in any size and shape (examples in Fig. 3.150). In addition, the retro-reflective material can be mounted on spheres (retro-spheres) which can be viewed over a wider range of directions (see section 6.2.1.2). The retro-reflective material consists either of a dense arrangement of small reflective balls ($\varnothing \approx 80\,\mu m$), or an array of micro-prisms (with triple-mirror corners).
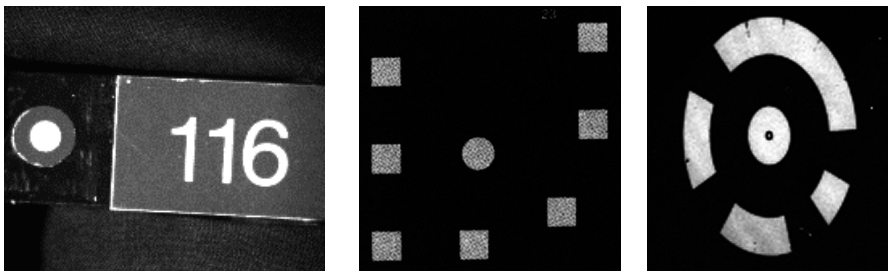
a) Circular target with point number

b) Circular target with area code

c) Circular target with ring code

Fig. 3.150: Examples of retro-reflective targets.

In order to achieve high contrast target images, retro-reflective targets must be illuminated from a position close to the camera axis (e.g. by a ring flash, see Fig. 3.155). The resulting images can be simply and fully automatically measured since, for all practical purposes, only the high-contrast measurement locations, without any background information, are imaged (examples in Fig. 7.30).

The microscopic balls or prisms of a retro-target are attached to a base material. Incident light is reflected internally within each ball and returns parallel to its incident path. Optionally, the material may have a protective plastic coating to allow the surface to function under wet conditions, but at the expense of a reduced light return. For masked targets, balls can partially be occluded at the edge of the target, so that the measured centre is laterally displaced. The shift depends on the viewing direction and gives rise to a 3D target location which is above the physical centre of the target if a triangulation method is applied (Fig. 3.151). The opposite effect occurs for stamped targets and leads to a triangulated 3D point below physical target level. Both effects can result in a shift of about 50 μm and should be corrected by the processing software where a best estimate of the mechanical plane of the target is required.

M_L

incident light from the left

M_R

incident light from the right

M

M_R M_L

triangulated point for
masked retro-targets

M

M_L M_R

triangulated point for
stamped retro-targets
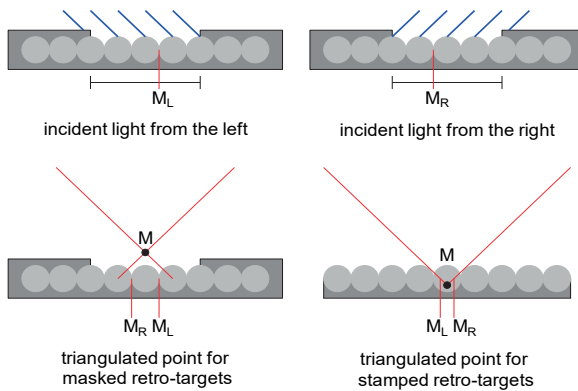
Fig. 3.151: Position of virtual centre of circular retro-targets (after Dold 1997).
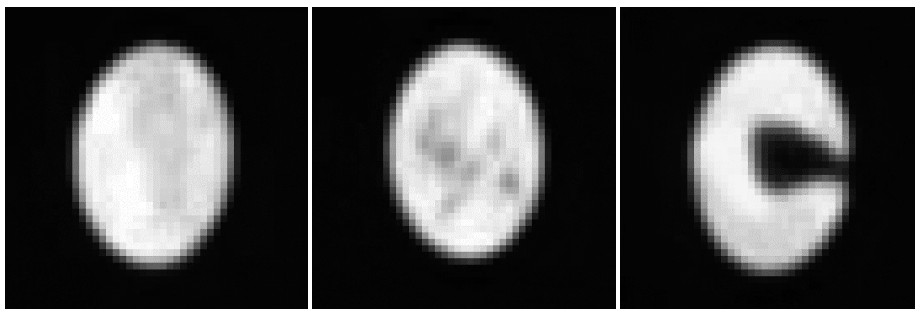
a) Finger marks

b) Scratches

c) Water drops

Fig. 3.152: Degraded retro-targets.

Retro-reflective target materials are sensitive to surface marking caused by fingerprints, dust, liquids and mechanical abrasion, as well as humidity and aging. In these cases, reflectivity and target outline are affected, which may lead to degradation in 3D measurement quality. This particularly affects measurement techniques which determine the target centre from the grey value distribution across the target image, e.g. centroid methods and template matching, see section 5.4.2. Fig. 3.152 shows examples of degraded retro-targets. In such cases the plastic-coated versions may be of more practical use, since the ability to wipe these clean can more than offset their reduced light return and viewing angle.

Further drawbacks of retro-targets are caused by:

–   relatively high manufacturing costs, particularly for large or coded targets;
–   finite target dimensions which restrict the target density which can be applied to the object surface;
–   self-adhesive targets can generally only be used once, therefore requiring a new set of targets for each object to be measured;
–   method of illumination produces images which show only the targets and no other object information.

3.6.2.2 Other target materials

The drawbacks of retro-reflective targets mentioned above demands the use of alternative materials or targeting techniques for a number of applications.

a) Target with a bright plastic centre ($\varnothing = 50$ mm)   b) Target printed on paper ($\varnothing = 15$ mm)
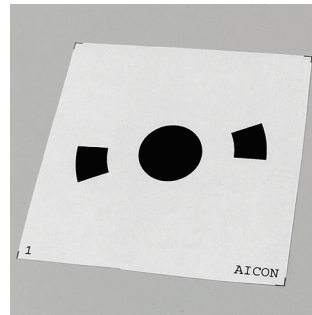
Fig. 3.153: Examples of artificial circular targets.

Circular plastic targets with a central target point are suitable for long-term targeting of outdoor objects such as buildings and bridges, and are also useful for geodetic measurements. They can be produced in almost any size. Fig. 3.153a shows an example of a circular plastic target.

Targets printed on paper (Fig. 3.153b) or adhesive film are simple to manufacture, for example with a laser printer. Self-adhesive targets can be used for temporary

targeting of locations where no long-term reproducibility in measurement is required. At minimal cost, this type of targeting offers an almost unlimited range of target designs with sharp edges and additional information. Luminous targets, where illumination at one wavelength is used to emit a different wavelength from a fluorescent material, can be very advantageous since the illumination wavelength can be filtered out at the camera removing reflections from shiny surfaces. Since printed targets are not retro-reflective, image contrast and target detectability strongly depend on high quality photography and sophisticated feature extraction algorithms during post-processing. The adhesive must also be selected with regard to ease of removal, potential damage to the object surface and resistance to heat.
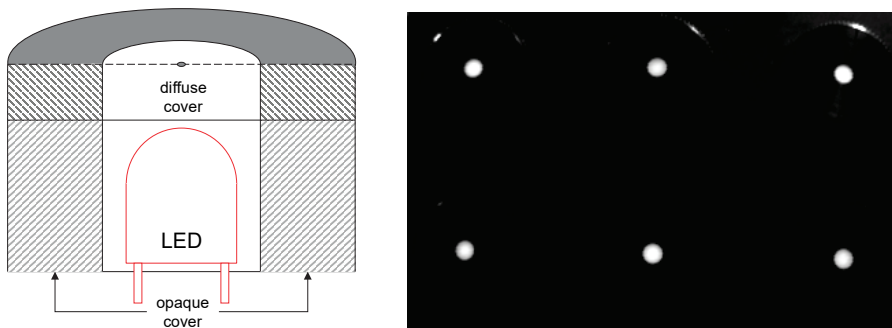
Fig. 3.154: Conceptual design for a luminous target.

Self-luminous (active) targets are more complex. They are used in those applications where no artificial illumination is possible or where a recording or measurement process is controlled by switching active targets on or off. Self-luminous targets can, for example, be designed with an LED (light emitting diode) that is mounted behind a semi-transparent plastic cover. These targets provide optimal contrast and sharp edges (Fig. 3.154). As examples, LEDs are incorporated into manual probes used in online measuring systems (Fig. 6.9).

3.6.3 Illumination and projection techniques

3.6.3.1 Electronic flash
Electronic flash systems can be attached and synchronized to almost all modern digital camera systems. Whilst the electronic flash output may occur in 1/70 000 s or less, camera shutters use synchronization times ranging between 1/30 s to 1/250 s. Electronic flash performance is characterized by the *guide number Z*. The guide number indicates the ability of the flash to illuminate an object at a given object

distance a in metres and a given f/number k. Higher guide numbers indicate greater light output.

$$a = \frac{Z}{k} \qquad\qquad (3.103)$$

Ring flashes, which are circular flash tubes mounted concentrically around the lens, are of special importance for photogrammetry (see Fig. 3.155). They are mainly used for illuminating objects to which retro-reflective targets are attached. Their light output, which is concentric with, and close to, the lens axis, ensures the retro-reflective target images are well separated from the background. It is important to note that electronic flash units with automatic exposure options utilize light measurement systems that are calibrated for general-purpose photography. Correctly illuminated retro-targets do not require as much light output and it is generally necessary to use manual controls in order to obtain consistent image quality. When illumination is controlled effectively, it is possible to make the background disappear almost completely in the image such that only reflecting target images remain (example in Fig. 7.30). Such a situation is ideal for rapid, automated target detection and measurement.

Fig. 3.155: Digital SLR camera with ring flash.

Fig. 3.156: Digital metric camera with integrated LED ring flash (AXIOS 3D).

Modern LED technology makes it possible to construct almost any form of illumination geometry, for example on a ring or flat field. LED illumination is available in the colours blue, green, yellow, red and white, as well as infrared, which permits the generation of almost any colour mix. They are extremely robust, durable and have low energy requirements. If LEDs are used for high-voltage flash illumination, very high illumination levels can be achieved. Fig. 3.156 shows a digital metric camera with an integrated LED ring flash operating in the infrared.

3.6.3.2 Pattern projection

Arbitrary light patterns, for example lines, fringes or random patterns, can be projected onto object surfaces by means of analogue slide projectors or computer-controlled digital projectors. These devices are typically used if the object does not provide sufficient natural texture to enable a photogrammetric surface reconstruction. They are also known as structured light projectors, particularly if the projected image is designed to form part of the surface reconstruction algorithm. Fig. 3.157 shows examples of projectors used for industrial photogrammetry. The GSI Pro-Spot system can project up to 22 000 circular dots onto an object. These can be measured like conventional targets to produce a dense object surface measurement.
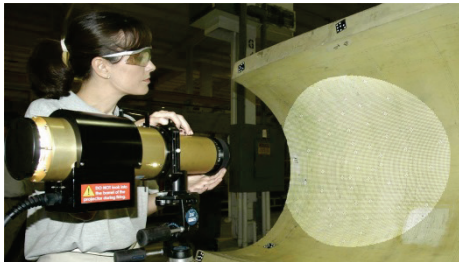
a) Point pattern projector (GSI)

c) Fringe projection sensor (Hexagon)
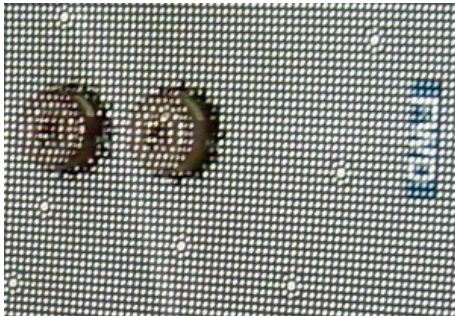
b) Component surface showing projected point pattern (GSI)
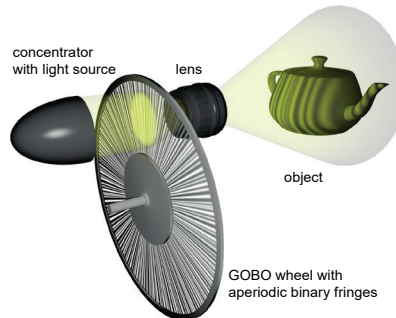
d) GOBO principle (Fraunhofer IOF)

Fig. 3.157: Surface pattern projection.

Digital projectors use either LCD (liquid crystal display), LCOS (liquid crystal on silicon) or DLP technologies (digital light processing) with micro-mirror devices (digital micro-mirror device, DMD). For LCD and LCOS projectors, reflective condenser optics are used to illuminate an LCD array where the programmed line pattern is projected through a lens onto the object surface. LCD video projectors are also available which can project video and computer images in colour. Current projectors have resolutions up to 1600 x 1400 lines at frame rates of up to 60 Hz.

Higher light efficiency and image contrasts can be achieved by digital micro-mirror devices (DMD). Movable micro-mirrors of about 5–16 μm size are attached to a semiconductor. They can be controlled digitally in order to reflect or absorb incident light. Micro-mirror chips are available with up to 4090 x 2160 points or more at frame rates of up to 100 Hz.

GOBO projectors (GOes Before Optics) have a high-speed rotating disk supporting a pattern, e.g. aperiodic fringes (see section 6.7.3.4), which is projected via a lens onto the object surface of interest (Fig. 3.157d). Using this method, patterns can be projected with high radiant power (>100 W) and at high projection frame rates of more than 10000 Hz, without being restricted in lateral resolution by the pixel structure of a digital light projector.

If a projector is calibrated and oriented it can be treated as a camera in the photogrammetric process. In such a situation the position of a projection point within the slide or digital projection plane is analogous to an image coordinate measurement and the collinearity equations can be used to compute its direction. This approach is used for projectors integrated into structured light systems for optical 3D metrology. There is also an increasing demand in the areas of virtual and augmented reality, e.g. for the superimposition of CAD data onto the real object (section 6.12.3).

3.6.3.3 Laser projectors

Laser projectors can be used to create a structured pattern on an object surface, e.g. with points or lines. In contrast to other light sources, high illumination powers can be achieved even for eye-safe laser classes (up to Class 2).

Laser projectors can be classified into three groups:

– Point projection:
   A single laser spot is projected onto the object, for example with a laser pointer. The resulting pattern is not a perfect circle and oblique projection angles make the spot elliptical. Interference between the coherent laser light and the roughness of the incident surface gives rise to speckle effects and generates an inhomogeneous intensity distribution within the laser spot which depends on viewpoint and any relative motion between laser, surface and viewpoint. As a result the optical centroid does not correspond to the geometric centre (Fig. 3.158). For these reasons, laser point projection is seldom used for high accuracy photogrammetric targeting.
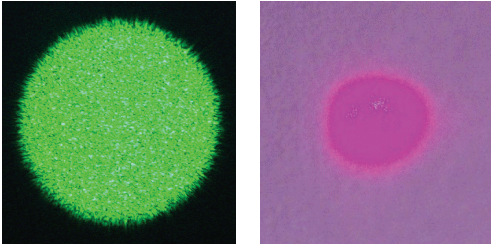
Fig. 3.158: Magnified image of laser points (left λ=511nm; right λ=785nm).

– 1D line projection:
Laser lines can be projected by means of a cylindrical lens (see Fig. 3.128) mounted in front of the laser source. Projected lines are used for triangulation with light-section methods (sections 6.4.1 and 8.4.2).
– 2D pattern projection:
Two-dimensional laser patterns can be created by special lenses and diffractive elements mounted in front of the laser source (Fig. 3.159). This is used, for example, in low-cost 3D sensors and 3D cameras (see section 6.7.6).
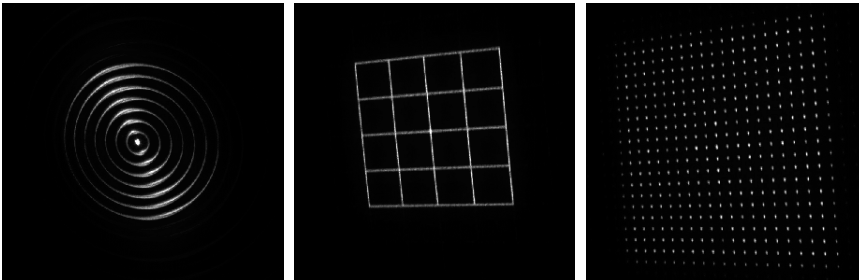
Fig. 3.159: Laser projection of two-dimensional patterns by front lenses.

Arbitrary patterns can be generated by fast 2D laser scanners. A laser beam is reflected by two galvanometer mirrors which rotate at high frequencies (up to 15 MHz). If the projection frequency is higher than the integration time of the camera (or the human eye), a continuous two-dimensional pattern is visible in the image. Such systems can be used in industrial measurement to set out information on a surface.

3.6.3.4 Directional lighting
Directional lighting techniques are of major importance since the measurement of particular object areas, for example edges, can be enhanced by the controlled

generation of shadows. This applies, for example, to the illumination of object edges which, without further targeting, are to be located by image edge detection (see section 4.4.7.2). Under such situations the physical edge of the object and the imaged edge must be identical.

Depending on the relative position of object surface and camera, directional lighting must be chosen such that one surface at the object edge reflects light towards the camera, while the other surface is shadowed, or does not reflect into the camera (Fig. 3.160). In the image, contrast characteristics and edge structure depend on the reflective properties of the two surfaces, and on the sharpness of the actual object edge. For example, a right-angled edge can be expected to give better results than a curved or bevelled edge.
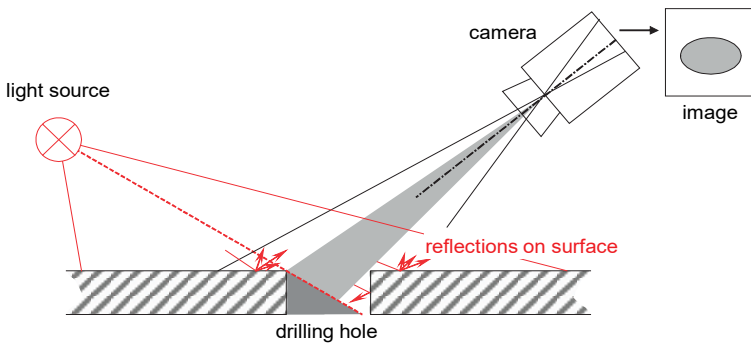
Fig. 3.160: Directional lighting for edge measurement.

a) Directly from the camera   b) According to Fig. 3.160   c) Diffuse from below
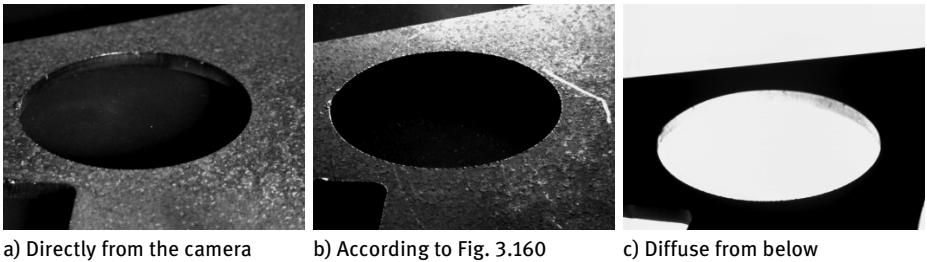
Fig. 3.161: Illumination of a drill hole in a metal component.

Fig. 3.161 shows the influence of different types of illumination on the images of object edges on a metal workpiece. Illumination from the viewing direction of the camera (a) results, as expected, in a poor image of the edge. If light is incident from the opposite direction (b), the edge of the drilled hole is correctly located in the object surface. Conditions permitting, a diffuse illumination source located below the

drilled hole (c) will also provide a suitable image of the edge, as required for example in the optical measurement of pipes and tubes (section 6.6.2.1).

3.6.3.5 Reflectance Transformation Imaging

RTI techniques (reflectance transformation imaging) are first used to record an object with a sequence of images from a camera position which is stationary with respect to the object of interest. Each individual image in the sequence is illuminated from a different direction using either a large number of permanently installed lamps (Fig. 3.162a) or a light source moving around the object (Fig. 3.162b). In this way, shadows are cast even on the smallest object details. The position of a light source in relation to the camera can be determined by means of the reflection on a bare metal sphere, which must be in the camera's field of view. This allows conclusions to be drawn about the surface texture, thus enabling the computation of surface normals for every pixel from sequentially acquired images using triplets of light sources. The surface normal solution is obtained independently of the surrounding pixels and without the need for a 3D surface model preserving the full spatial image resolution. RTI was developed as a tool for visualisation. For example, the method allows advanced texturing methods such as polynomial texture mapping to be used in order to determine the colour values of the surface under varying illumination conditions.

a) Illumination using a light dome   b) Illumination using a moving light source
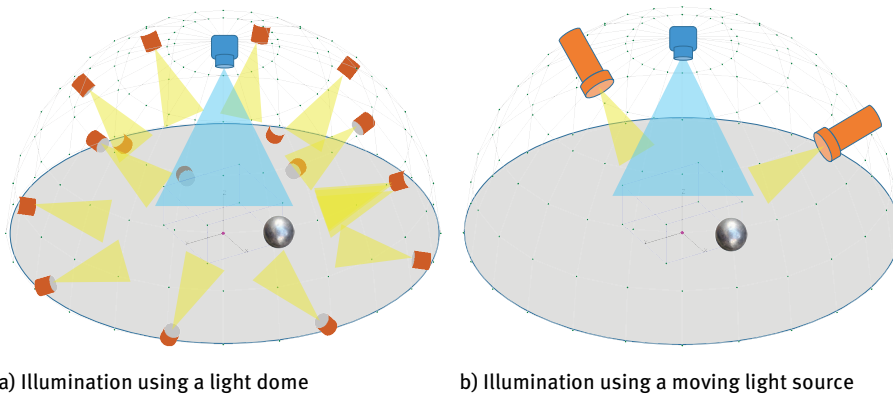
Fig. 3.162: Arrangement of light sources for RTI.

In the example of Fig. 3.163, local carving detail on a prehistoric bone shows how RTI generates shadows in the minimal carvings which are not revealed in the original RGB image due to its limited camera resolution. In associated software solutions, a fictitious light source can be moved interactively and continuously so that the viewer selects a suitable lighting direction in which the desired surface structures and reflectance become visible.

a) Original RGB image   b) Visualisation using RTI

Fig. 3.163: Visualisation of small structures by RTI.

The technique can be extended to metric surface recording if the spatial relationship between the camera and each illumination location is established. Where the lamps are mounted statically in a dome, a network of photogrammetric images of the constellation of lights taken with a second camera can be used to compute each lamp location. If a motion system is used to move a single lamp, the motion system can be calibrated much like the end effector on a robot arm. When the camera and illumination information are available, shading-based 3D methods such as Shape from Shading or Photometric Stereo can be used to determine surface normals.

RTI systems work exceptionally well for diffuse reflecting surfaces (section 3.6.1.1). However, for specular surfaces, small rotations of the light-source geometry with respect to the surface geometry are needed to optimise the light returned to the camera. The objective is to avoid direct specular returns which will locally saturate the sensor, whilst optimising if possible, illumination geometries that capture near-specular returns just below sensor saturation levels. In such cases it is also possible to record and determine the colour differences between specular and diffuse surfaces. This can be important in understanding metallic heritage objects.

4 Analytical methods

4.1 Overview

This chapter deals with the analytical methods which are essential for the calculation of image orientation parameters and object information (coordinates, geometric elements). The methods are based on measured image coordinates derived from both manual image measurement and digital image processing.

Due to their differing importance in practice, the analytical methods of calculation are classified according to the number of images involved. It is common to all methods that the relationship between image information and object geometry is established by the parameters of interior and exterior orientation. Procedures usually have the following stages:

– provision of object information (reference points, distances, geometric elements);
– measurement and matching of image points for orientation (image coordinates);
– calculation of orientation parameters (interior and exterior orientations);
– object reconstruction from oriented images (new points, dense point clouds, geometric elements, maps, orthophotos).

Depending on method and application, these stages are performed either sequentially, simultaneously or iteratively in a repeated number of processing steps.
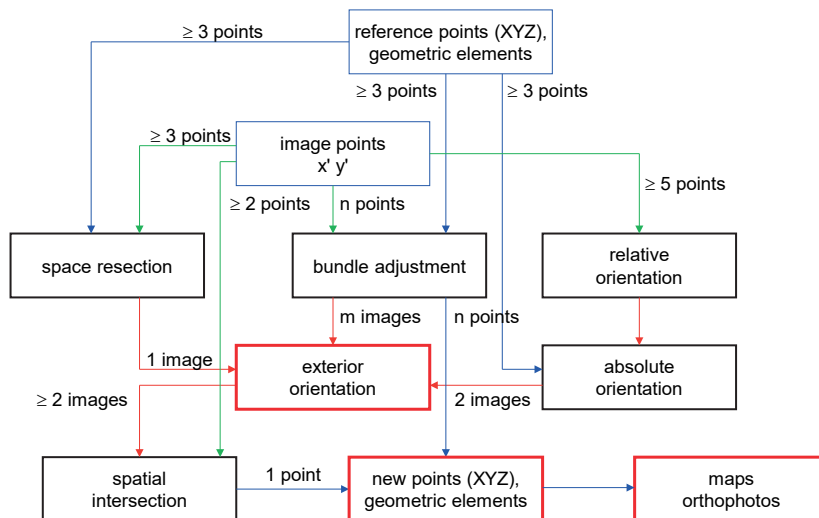
Fig. 4.1: Methods and data flow for orientation and point determination.

Fig. 4.1 is a simplified illustration of typical methods for the determination of orientation parameters (exterior only) and for the reconstruction of 3D object geometry (new points, geometric elements). Both procedures are based on measured image coordinates and known object information. It is obvious that 3D point determination for object reconstruction cannot be performed without the parameters of exterior orientation, i.e. position and orientation of an image in space.

If orientation data and coordinates of reference points or measured object points are available, further object reconstruction can be performed according to one or more of the following methods:

– Numerical generation of points and geometric elements:
  Using images of known orientation for the determination of additional object coordinates and geometric elements by, for example, spatial intersection.
– Graphical object reconstruction:
  Extraction of graphical and geometric information to create maps, drawings or CAD models.
– Rectification or orthophoto production:
  Using various projections, the transformation of the measurement imagery into image-based products such as photo maps, image mosaics and 3D animations.

With regard to the number of images involved, the following methods can be identified:

– Single image analysis (see section 4.2):
  Analysis of single images which takes into account additional geometric information and constraints in object space (straight lines, planes, surface models etc.). Here a distinction is made between the calculation of object coordinates and the production of rectified images and orthophotos (Fig. 4.2), see also section 4.2.7 (Monoplotting).
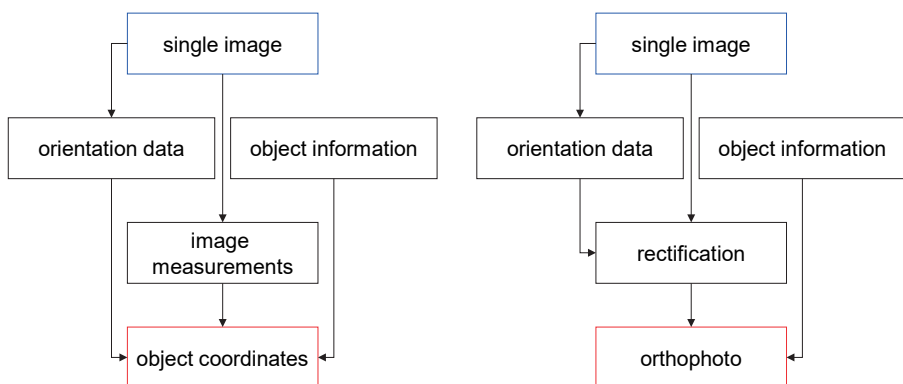
Fig. 4.2: Methods of single image processing.

– Stereo image processing (see section 4.3):
  Visual or digital processing of image pairs based on the principles of stereoscopic image viewing and analysis, in particular for the measurement of natural features (non-targeted points) and the capture of free-form surfaces (Fig. 4.3).
– Multi-image processing (see section 4.4):
  Simultaneous evaluation of an unlimited number of images of an object, e.g. for camera calibration or the reconstruction of complex object geometries.
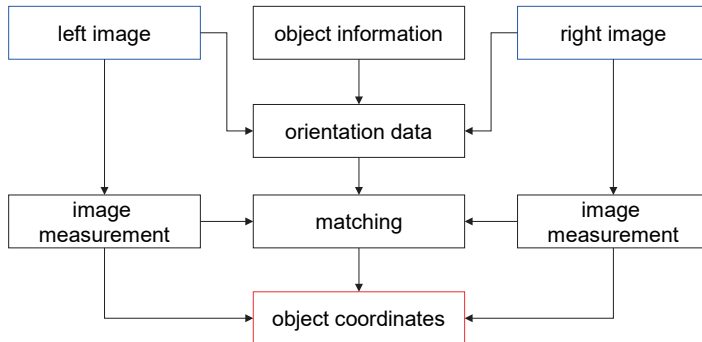
Fig. 4.3: Methods of stereo image processing.

If the geometric imaging model is appropriately modified, the techniques of orientation and 3D object reconstruction can be extended to panoramic applications (section 4.5) and multi-media photogrammetry (section 4.6).

4.2 Processing of single images

4.2.1 Exterior orientation

4.2.1.1 Standard case

The *exterior orientation* consists of six parameters which describe the spatial position and orientation of the camera coordinate system with respect to the global object coordinate system (Fig. 4.4). The terms pose or 6DOF are also used to denote the six orientation parameters. Fig. 4.5 illustrates the projective relationship between an object point P and the exterior orientation of an image. The standard case in aerial photography of a horizontal image plane is also used as the basic model in close-range photogrammetry, i.e. for a non-rotated image the optical axis is approximately aligned with the Z axis of the object coordinate system. Terrestrial photogrammetry is covered in the next section as a special case.

The camera coordinate system has its origin at the perspective centre of the image (see section 2.1.2). It is further defined by reference features fixed in the camera (fiducial marks, réseau, sensor system). It can therefore be reconstructed from the image and related to an image measuring device (*interior orientation*, section 3.3.2).
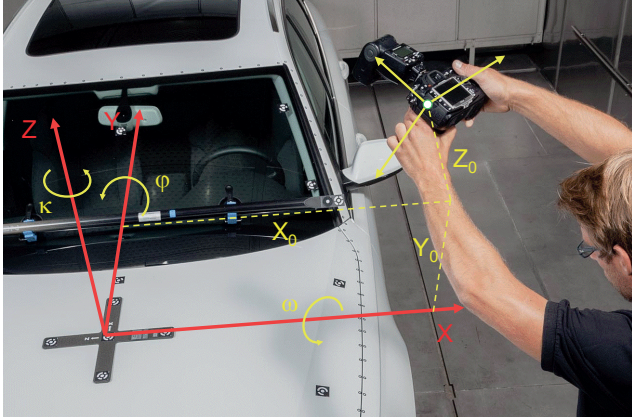
Fig. 4.4: Exterior orientation of a camera in object coordinate system XYZ.
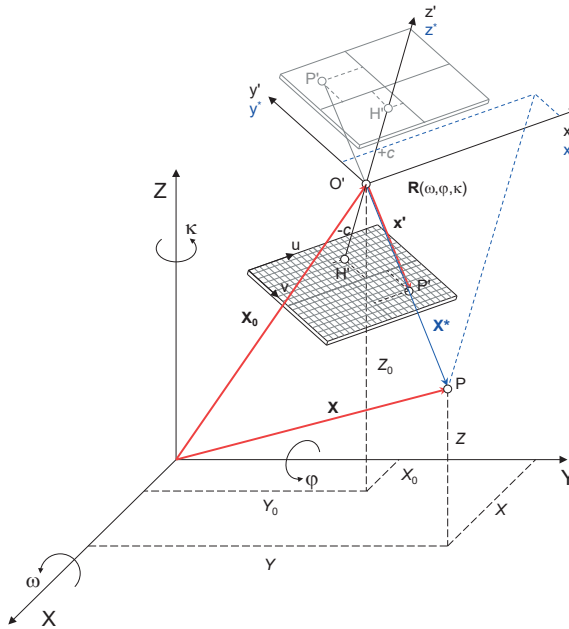
Fig. 4.5: Exterior orientation and projective imaging.

The spatial position of the image coordinate system is defined by the vector $\mathbf{X}_0$ from the origin to the perspective centre O'. The orthogonal rotation matrix $\mathbf{R}$ defines the angular orientation in space. It is the combination of three independent rotations ω,φ,κ about the coordinate axes X,Y,Z (see section 2.2.2).

$$\mathbf{X}_0 = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} \qquad \text{: position of perspective centre} \qquad (4.1)$$

$$\mathbf{R} = \mathbf{R}_\omega \cdot \mathbf{R}_\varphi \cdot \mathbf{R}_\kappa$$

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \qquad \text{: rotation matrix} \qquad (4.2)$$

The elements of the rotation matrix r_{ij} can be defined either as trigonometric functions of the three rotation angles or as functions of quaternions (section 2.2.2.2).

With given parameters of exterior orientation, the direction from the perspective centre O' to the image point P' (image vector $\mathbf{x}'$) can be transformed into an absolutely oriented spatial ray from the perspective centre to the object point P.

The exterior orientation further describes the spatial transformation (rotation and shift) from camera coordinates x^*,y^*,z^* into object coordinates X,Y,Z (see Fig. 4.5):

$$\begin{aligned} \mathbf{X} &= \mathbf{X}_0 + \mathbf{R} \cdot \mathbf{x}^* = \mathbf{X}_0 + \mathbf{X}^* \\ X &= X_0 + r_{11}x^* + r_{12}y^* + r_{13}z^* \\ Y &= Y_0 + r_{21}x^* + r_{22}y^* + r_{23}z^* \\ Z &= Z_0 + r_{31}x^* + r_{32}y^* + r_{33}z^* \end{aligned} \qquad (4.3)$$

4.2.1.2 Special case of terrestrial photogrammetry

For the special case of conventional terrestrial photogrammetry, the camera axis is approximately horizontal. In order to avoid singularities in trigonometric functions, the rotation sequence must either be re-ordered (see section 2.2.2.1) or the image coordinate system must be defined by axes x' and z' (instead of x' and y', see section 2.1.2). In this case, image acquisition systems which provide angle measurements, e.g. video theodolites, can use rotation angles ω (tilt about horizontal axis), κ (roll around optical axis) and φ or α (azimuth) instead of the standard sequence ω,φ,κ (Fig. 4.6). It must be remembered here that geodetic angles are positive clockwise.

The modified rotation order ω,φ,κ leads to the following rotation matrix (compare with eqn. 2.37):

$$\mathbf{R}_{terr.} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \tag{4.4}$$

$$= \begin{bmatrix} \cos\varphi\cos\kappa - \sin\varphi\sin\omega\sin\kappa & -\sin\varphi\cos\omega & \cos\varphi\sin\kappa + \sin\varphi\sin\omega\cos\kappa \\ \sin\varphi\cos\kappa + \cos\varphi\sin\omega\sin\kappa & \cos\varphi\cos\omega & \sin\varphi\sin\kappa - \cos\varphi\sin\omega\cos\kappa \\ -\cos\omega\sin\kappa & \sin\omega & \cos\omega\cos\kappa \end{bmatrix}$$
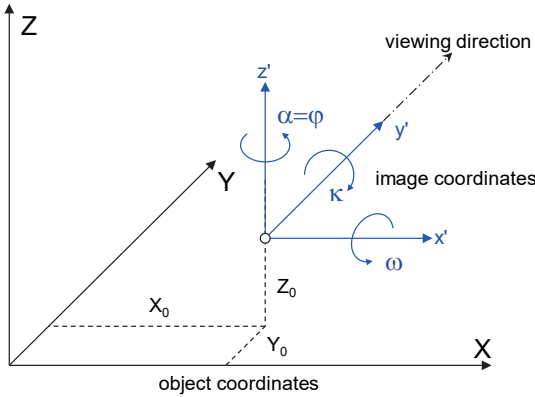
Fig. 4.6: Exterior orientation for terrestrial photogrammetry.

4.2.2 Collinearity equations

The central projection in space is at the heart of many photogrammetric calculations. Thus, the coordinates of an object point P can be derived from the position vector to the perspective centre $\mathbf{X}_0$ and the vector from the perspective centre to the object point $\mathbf{X}^*$ (Fig. 4.5):

$$\mathbf{X} = \mathbf{X}_0 + \mathbf{X}^* \tag{4.5}$$

The vector $\mathbf{X}^*$ is given in the object coordinate system. The image vector $\mathbf{x'}$ may be transformed into object space by rotation matrix $\mathbf{R}$ and a scaling factor m. Then, since it is in the same direction as $\mathbf{X}^*$ (with $z' = -c$):

$$\mathbf{X}^* = m \cdot \mathbf{R} \cdot \mathbf{x'} \tag{4.6}$$

Hence, the projection of an image point into a corresponding object point is given by:

$$\mathbf{X} = \mathbf{X}_0 + m \cdot \mathbf{R} \cdot \mathbf{x'}$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + m \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \tag{4.7}$$

The scale factor m is an unknown value which varies for each object point. If only one image is available, then only the direction to an object point P can be determined but not its absolute position in space. The 3D coordinates of P can only be computed if this spatial direction intersects another geometrically known element, e.g. intersection with a second ray from another image or intersection with a given surface in space. Range-measuring 3D cameras represent an exception because they directly deliver polar coordinates (section 6.7.6).

By inverting eqn. (4.7), adding the principal point $H'(x'_0, y'_0)$ and introducing correction terms $\mathbf{\Delta x'}$ (image distortion parameters), the image coordinates are given by (see also eqn. 3.52):

$$\mathbf{x'} = \mathbf{x'}_P - \mathbf{x'_0} - \mathbf{\Delta x'} = \frac{1}{m} \mathbf{R}^{-1}(\mathbf{X} - \mathbf{X}_0)$$

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x'_P - x'_0 - \Delta x' \\ y'_P - y'_0 - \Delta y' \\ -c \end{bmatrix} = \frac{1}{m} \begin{bmatrix} r_{11} & r_{21} & r_{31} \\ r_{12} & r_{22} & r_{32} \\ r_{13} & r_{23} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{bmatrix} \tag{4.8}$$

Note that the inverse rotation matrix is equal to its transpose. By dividing the first and second equations by the third equation, the unknown scaling factor m is eliminated:

$$\frac{x'_P - x'_0 - \Delta x'}{z'} = \frac{r_{11}(X - X_0) + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)}$$

$$\frac{y'_P - y'_0 - \Delta y'}{z'} = \frac{r_{12}(X - X_0) + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} \tag{4.9}$$

As a check, this reduces to the simple non-rotated case ($\mathbf{R} = \mathbf{I}$) shown in Fig. 4.7 where, for example, the image-side ratio $x':z'$ is equal to the object-side ratio $(X-X_0):(Z-Z_0)$. After further rearrangement, the *collinearity equations* follow:

$$x' = x'_0 + z' \cdot \frac{r_{11} \cdot (X - X_0) + r_{21} \cdot (Y - Y_0) + r_{31} \cdot (Z - Z_0)}{r_{13} \cdot (X - X_0) + r_{23} \cdot (Y - Y_0) + r_{33} \cdot (Z - Z_0)} + \Delta x'$$

$$y' = y'_0 + z' \cdot \frac{r_{12} \cdot (X - X_0) + r_{22} \cdot (Y - Y_0) + r_{32} \cdot (Z - Z_0)}{r_{13} \cdot (X - X_0) + r_{23} \cdot (Y - Y_0) + r_{33} \cdot (Z - Z_0)} + \Delta y' \tag{4.10}$$

These equations describe the transformation of object coordinates (X,Y,Z) into corresponding image coordinates (x',y') as functions of the interior orientation parameters $(x'_0, y'_0, z' = -c, \Delta x', \Delta y')$ and exterior orientation parameters $(X_0, Y_0, Z_0, \mathbf{R})$ of

one image. Here (x'_P, y'_P) denote the measured (distorted) image coordinates whilst (x', y') are the non-distorted image coordinates of a point P which are reduced to the principal point (see eqn. 3.52).
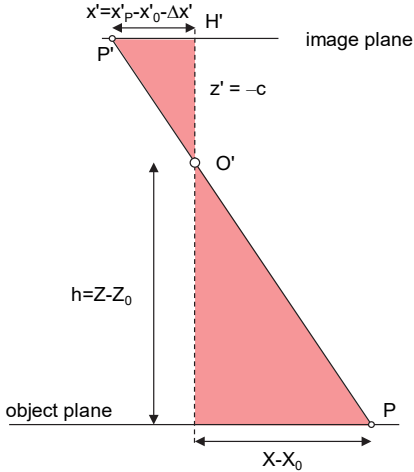
Fig. 4.7: Image and object-side proportions for the case **R = I**.

An alternative form is given if the object coordinate system is transformed by shifting to the perspective centre and orienting parallel to the image coordinate system. Within the local coordinate system which results, object coordinates are denoted by x^*, y^*, z^* (see Fig. 4.5):

$$\begin{bmatrix} x^* \\ y^* \\ z^* \end{bmatrix} = \mathbf{R}^{-1} \begin{bmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{bmatrix} \qquad \text{: spatial translation and rotation} \qquad (4.11)$$

Multiplying out and substitution results in the following transformation equation:

$$\begin{aligned} \mathbf{x}^* &= \mathbf{R}^{-1} \cdot \left(\mathbf{X} - \mathbf{X}_0 \right) \\ &= \mathbf{R}^{-1} \cdot \mathbf{X} - \mathbf{R}^{-1} \cdot \mathbf{X}_0 \qquad \text{: spatial translation and rotation} \qquad (4.12) \\ &= \mathbf{X}_0^* + \mathbf{R}^* \cdot \mathbf{X} \end{aligned}$$

where $\mathbf{R}^* = -\mathbf{R}^{-1}$.

By introducing the image scale $1/m$ and corrections for shift of principal point and image errors, the collinearity equations for image coordinates are again obtained:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{z'}{z^*} \begin{bmatrix} x^* \\ y^* \end{bmatrix}$$

$$m = \frac{z^*}{z'}$$

: projection into image plane   (4.13)

In the following, principal point and image error corrections are ignored. By multiplying out eqn. (4.13) and multiplying through with $1/Z^*_0$, the following spatial projection equations are obtained:

$$x' = \frac{\dfrac{z'r_{11}}{Z^*_0} X + \dfrac{z'r_{21}}{Z^*_0} Y + \dfrac{z'r_{31}}{Z^*_0} Z + \dfrac{z'}{Z^*_0} X^*_0}{\dfrac{r_{13}}{Z^*_0} X + \dfrac{r_{23}}{Z^*_0} Y + \dfrac{r_{33}}{Z^*_0} Z + \dfrac{Z^*_0}{Z^*_0}}$$

$$y' = \frac{\dfrac{z'r_{12}}{Z^*_0} X + \dfrac{z'r_{22}}{Z^*_0} Y + \dfrac{z'r_{32}}{Z^*_0} Z + \dfrac{z'}{Z^*_0} Y^*_0}{\dfrac{r_{13}}{Z^*_0} X + \dfrac{r_{23}}{Z^*_0} Y + \dfrac{r_{33}}{Z^*_0} Z + \dfrac{Z^*_0}{Z^*_0}}$$

(4.14)

X_0^*, Y_0^*, Z_0^* are the components of vectors $\mathbf{X}_0^*$ from eqn. (4.12). If the fractional terms are replaced by coefficients a',b',c', the three dimensional image equations are obtained as follows (see for comparison eqn. 2.21 and eqn. 4.24):

$$x' = \frac{a'_0 + a'_1 X + a'_2 Y + a'_3 Z}{1 + c'_1 X + c'_2 Y + c'_3 Z} \qquad\qquad y' = \frac{b'_0 + b'_1 X + b'_2 Y + b'_3 Z}{1 + c'_1 X + c'_2 Y + c'_3 Z}$$

(4.15)

The collinearity equations demonstrate clearly that each object point is projected into a unique image point, if it is not occluded by other object points. The equations effectively describe image generation inside a camera by the geometry of a central projection.
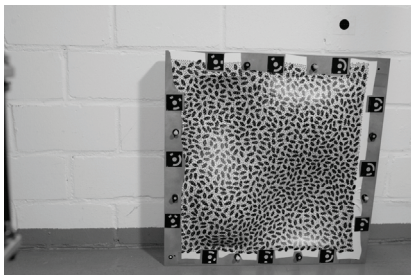
a) Original image (taken by Nikon D2X, f = 24 mm)
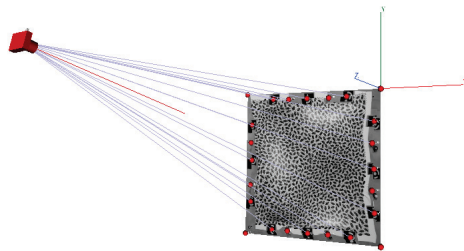
b) Position of the camera (and image) in the object coordinate system

Fig. 4.8: Exterior orientation of an image.

The equations (4.10) form the *fundamental equations of photogrammetry*. It is important to note that, since the observed measurements stand alone on the left-hand side, these equations are suitable for direct use as observation equations in an over-determined least-squares adjustment (see section 2.4.2). For example, the collinearity equations are used to set up the equation system for spatial intersection (section 4.4.7.1), space resection (section 4.2.3) and bundle triangulation (section 4.4). Additionally, they offer the mathematical basis for the generation of orthophotos (section 4.2.8.2) and the principle of stereo plotting systems.

Example 4.1:

The following data are available for the image illustrated in Fig. 4.8:

Interior orientation:	$c = -24.2236$ mm	$x'_0 = 0.0494$ mm	$y'_0 = -0.2215$ mm
Exterior orientation:	$X_0 = -471.89$ mm	$Y_0 = 11.03$ mm	$Z_0 = 931.07$ mm
	$\omega = -13.059°$	$\varphi = -4.440°$	$\kappa = 0.778°$
Object coordinates:	$X_1 = -390.93$ mm	$Y_1 = -477.52$ mm	$Z_1 = 0.07$ mm
	$X_2 = -101.54$ mm	$Y_2 = -479.19$ mm	$Z_2 = 0.10$ mm

Image coordinates of the object points are required in accordance with equations (2.32) and (4.10).

Image coordinates:	$x'_1 = 0.0104$ mm	$y'_1 = -6.5248$ mm
	$x'_2 = 6.7086$ mm	$y'_2 = -6.5162$ mm

From the object and image separations of both points, an approximate image scale is calculated as $m \approx 43$.

4.2.3 Space resection

Orientation of single images is taken to mean, in the first instance, the process of calculating the parameters of exterior orientation. Since direct determination, for example by angle or distance measurement, is not usually possible or not accurate enough, methods of indirect orientation are employed. These make use of XYZ reference points whose image coordinates may be measured in the image. Common calculation procedures can be divided into two groups:

1. Calculation of exterior orientation based on collinearity equations:
   The method of space resection provides a non-linear solution that requires a minimum of three XYZ reference points in object space and approximate values for the unknown orientation parameters (see also section 4.2.3.1).
2. Calculation of exterior orientation based on projective relations:
   The most popular method in this group is the Direct Linear Transformation (DLT). It requires a minimum of six XYZ reference points, but provides a direct solution without the need for approximate values (see also section 4.2.4). Linear methods in projective geometry function in a similar way.

4.2.3.1 **Space resection with known interior orientation**

Space resection is used to compute the exterior orientation of a single image. The procedure requires known *XYZ* coordinates of at least three object points P_i which do not lie on a common straight line. The bundle of rays through the perspective centre from the reference points can fit the corresponding points in image plane P'_i in only one, unique, position and orientation of the image (Fig. 4.9).[1] Fig. 4.10 shows the respective intersection figure of image plane and control point tetrahedron for two images with different exterior orientations.
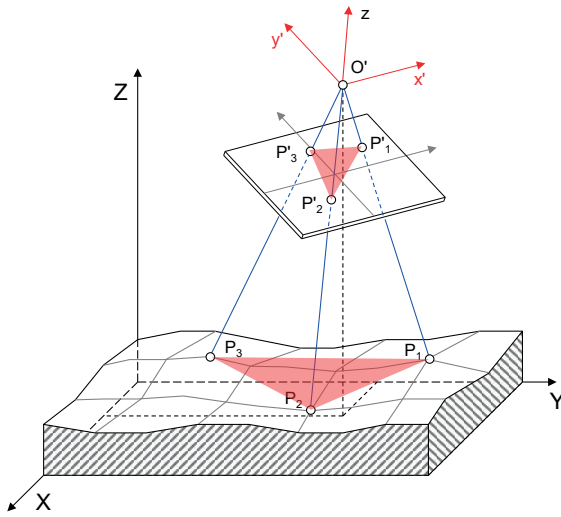
Fig. 4.9: Space resection.
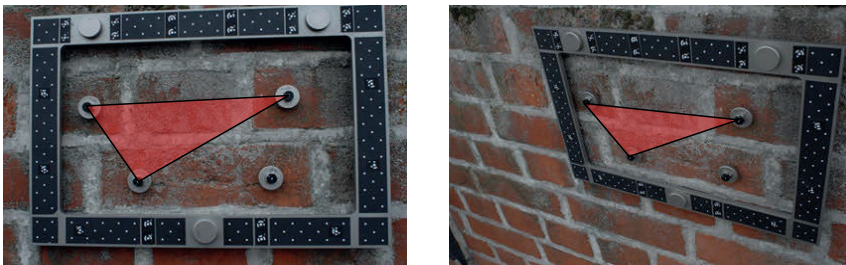
Fig. 4.10: Intersection figures for two images with different exterior orientations.

---

1 Strictly speaking, multiple solutions exist with only three reference points but a single unique solution is possible if at least one more non coplanar reference point is added (further information in section 4.2.3.4).

Using the measured image coordinates of the reference points, corrected for distortion and reduced to the principal point, and with known interior orientation, the following system of correction equations can be derived from the collinearity equations (4.10):

$$x' + vx' = F(X_0, Y_0, Z_0, \omega, \varphi, \kappa, z', X, Y, Z)$$
$$y' + vy' = F(X_0, Y_0, Z_0, \omega, \varphi, \kappa, z', X, Y, Z)$$

(4.16)

Function F is a representation for equations (4.10) in which the red values are introduced as unknowns. This system can be linearized at approximate values by Taylor-series expansion and solved by least-squares adjustment.

Each of the measured image points provides two linearized correction equations:

$$vx'_i = \left(\frac{\partial x'}{\partial X_0}\right)^0 dX_0 + \left(\frac{\partial x'}{\partial Y_0}\right)^0 dY_0 + \left(\frac{\partial x'}{\partial Z_0}\right)^0 dZ_0$$
$$+ \left(\frac{\partial x'}{\partial \omega}\right)^0 d\omega + \left(\frac{\partial x'}{\partial \varphi}\right)^0 d\phi + \left(\frac{\partial x'}{\partial \kappa}\right)^0 d\kappa - (x'_i - x'^{\,0}_i)$$

$$vy'_i = \left(\frac{\partial y'}{\partial X_0}\right)^0 dX_0 + \left(\frac{\partial y'}{\partial Y_0}\right)^0 dY_0 + \left(\frac{\partial y'}{\partial Z_0}\right)^0 dZ_0$$
$$+ \left(\frac{\partial y'}{\partial \omega}\right)^0 d\omega + \left(\frac{\partial y'}{\partial \varphi}\right)^0 d\phi + \left(\frac{\partial y'}{\partial \kappa}\right)^0 d\kappa - (y'_i - y'^{\,0}_i)$$

(4.17)

Here x_i' and y_i' are the corrected image coordinates, x'^0_i and y'^0_i are the image coordinates which correspond to the approximate orientation parameters. If quaternions are used to define rotations (see eqn. 2.38) then the partial derivatives with respect to ω, φ, κ, must be replaced by derivatives with respect to the algebraic parameters a, b, c, d.

Simplification of the collinearity equations (4.10), by substituting k_x and k_y for the numerators and N for the denominator, leads to

$$x' = x'_0 + z' \frac{k_X}{N} + \Delta x' \qquad\qquad y' = y'_0 + z' \frac{k_Y}{N} + \Delta y'$$

(4.18)

from which the derivatives of (4.17) are given by:

$$\frac{\partial x'}{\partial X_0} = \frac{z'}{N^2} \cdot (r_{13} k_X - r_{11} N) \qquad \frac{\partial x'}{\partial Y_0} = \frac{z'}{N^2} \cdot (r_{23} k_X - r_{21} N) \qquad \frac{\partial x'}{\partial Z_0} = \frac{z'}{N^2} \cdot (r_{33} k_X - r_{31} N)$$

$$\frac{\partial x'}{\partial \omega} = \frac{z'}{N} \cdot \left\{ \frac{k_X}{N} \cdot \left[r_{33}(Y - Y_0) - r_{23}(Z - Z_0) \right] - r_{31}(Y - Y_0) + r_{21}(Z - Z_0) \right\}$$

(4.19a)

$$\frac{\partial x'}{\partial \varphi} = \frac{z'}{N} \cdot \left\{ \frac{k_X}{N} \cdot \left[k_Y \cdot \sin\kappa - k_X \cdot \cos\kappa \right] - N \cdot \cos\kappa \right\} \qquad\qquad \frac{\partial x'}{\partial \kappa} = \frac{z'}{N} \cdot k_Y$$

$$\frac{\partial y'}{\partial X_0} = \frac{z'}{N^2} \cdot (r_{13}k_Y - r_{12}N) \qquad \frac{\partial y'}{\partial Y_0} = \frac{z'}{N^2} \cdot (r_{23}k_Y - r_{22}N) \qquad \frac{\partial y'}{\partial Z_0} = \frac{z'}{N^2} \cdot (r_{33}k_Y - r_{32}N)$$

$$\frac{\partial y'}{\partial \omega} = \frac{z'}{N} \cdot \left\{ \frac{k_Y}{N} \cdot \left[r_{33}(Y - Y_0) - r_{23}(Z - Z_0) \right] - r_{32}(Y - Y_0) + r_{22}(Z - Z_0) \right\} \qquad (4.19b)$$

$$\frac{\partial y'}{\partial \varphi} = \frac{z'}{N} \left\{ \frac{k_Y}{N} \left[k_Y \sin\kappa - k_X \cos\kappa \right] + N \sin\kappa \right\} \qquad\qquad \frac{\partial y'}{\partial \kappa} = -\frac{z'}{N} \cdot k_X$$

The coefficients r_{ij} can be derived from the rotation matrix **R** according to eqn. (2.32).

Example 4.2:

The following data are available for the space resection of the image in Fig. 4.8:

Interior orientation: $z' = -24.2236$ mm

Reference points:		
$X_1 = -390.93$ mm	$Y_1 = -477.52$ mm	$Z_1 = 0.07$ mm
$X_2 = -101.5$ mm	$Y_2 = -479.19$ mm	$Z_2 = 0.11$ mm
$X_3 = -23.22$ mm	$Y_3 = -256.85$ mm	$Z_3 = -0.09$ mm
$X_4 = -116.88$ mm	$Y_4 = -21.03$ mm	$Z_4 = 0.07$ mm
$X_5 = -392.17$ mm	$Y_5 = -21.46$ mm	$Z_5 = 0.11$ mm
$X_6 = -477.54$ mm	$Y_6 = -237.65$ mm	$Z_6 = -0.16$ mm

Corrected image coordinates:	
$x'_1 = -0.0395$ mm	$y'_1 = -6.3033$ mm
$x'_2 = 6.6590$ mm	$y'_2 = -6.2948$ mm
$x'_3 = 9.0086$ mm	$y'_3 = -1.3473$ mm
$x'_4 = 7.3672$ mm	$y'_4 = 4.5216$ mm
$x'_5 = 0.2936$ mm	$y'_5 = 4.7133$ mm
$x'_6 = -2.0348$ mm	$y'_6 = -0.7755$ mm

The parameters of exterior orientation are required:

Exterior orientation: $\qquad X_0 = -471.859$ mm $\qquad Y_0 = 11.021$ mm $\qquad Z_0 = 931.128$ mm

$$\mathbf{R} = \begin{bmatrix} 0.996908 & -0.013556 & -0.077392 \\ 0.030729 & 0.973815 & 0.225255 \\ 0.072312 & -0.226937 & 0.971221 \end{bmatrix}$$

$\omega = -13.0577° \qquad\qquad \varphi = -4.4387° \qquad\qquad \kappa = 0.7791°$

4.2.3.2 Space resection with unknown interior orientation

For images from cameras with unknown parameters of interior orientation, such as amateur cameras, the number of unknown parameters, ignoring distortion in the first instance, increases by 3 (c, x'_0, y'_0) to a total of 9. Two more additional reference points, providing 4 additional image observation equations, are required for the solution (a minimum of 5 reference points in total).

If all reference points lie approximately on a plane, then the normal system of equations for the resection is singular, since the problem can be solved by an 8-parameter projective transformation between image and object planes (see equations

2.21 and 4.45). However, if one of the unknown parameters, such as the principal distance, is fixed to an arbitrary value, a unique solution can be computed.

If a suitable spatial distribution of object points is available, the space resection approach can be used to calibrate the parameters of interior orientation c, x'_0, y'_0, $\Delta x'$, $\Delta y'$ from only one image. The number of elements to be determined increases to a total of 11 if the parameters A_1 and A_2 for radial distortion are introduced (6 for exterior and 5 for interior orientation). In this case a minimum of 6 spatially distributed XYZ reference points are required.

4.2.3.3 Approximate values for resection

In some situations, approximate values for the unknown parameters of exterior orientation can be readily determined by one of the following methods:

– Approximate values by direct measurement:

Approximate orientation values can possibly be measured on site, for example by geodetic methods or the use of GNSS and INS sensors. The camera position can often be determined by simple means such as estimation off a site plan. Even rotation angles can be estimated sufficiently well in simple configurations, for example without oblique or rolled views.

– Small image rotations:

If the coordinate axes of the image system are approximately parallel to the object coordinate system, initial rotation angles can be approximated by the value zero. The parameters for translation X0 and Y0 can be estimated from the centroid of the reference points; the object distance (Z–Z0) can be determined from the principal distance and approximately known image scale.

– General solution:

A general solution for approximate values of the six parameters of outer orientation can be derived with the aid of three reference points as the next method explains.

4.2.3.4 Resection with minimum object information

In the general case of arbitrary position and orientation of an image, approximate values can no longer be easily estimated but must be computed. This can be performed according to the following scheme which may also be called space resection with minimum object information. It is noted that other solutions, including algebraic solutions, exist but are not discussed here.

Given a minimum of three XYZ reference points, the position of the perspective centre O' is first determined. As shown in Fig. 4.11 a tetrahedron can be formed by the perspective centre and the three object points. From the simple properties of triangles, the angles α, β, γ of the tetrahedrons $O'P'_1P'_2P'_3$ and $O'P_1P_2P_3$ can be calculated from the measured image coordinates of P'_1, P'_2 and P'_3 and the principal

distance c. However, the side lengths of the tetrahedron, d_i, $i = 1...3$, necessary to complete the location of O', remain unknown at this stage.

If each of the tetrahedron sides is rotated into the plane of the three object points, the configuration of Fig. 4.12 is obtained. Taking as an example the triangle formed by object points P_1, P_2 and the required perspective centre O', it is obvious that a circle K_1 can be constructed from known distance s_{12} and angle α calculated previously, and that O' lies on this circle. Similarly, circle K_2 can be constructed from s_{23} and circle K_3 from s_{13}. Three spindle tori shapes are formed when the three circles are rotated about their corresponding chords P_1P_2, P_2P_3, and P_3P_1 and their intersection points provide possible solutions for the spatial position of O'.
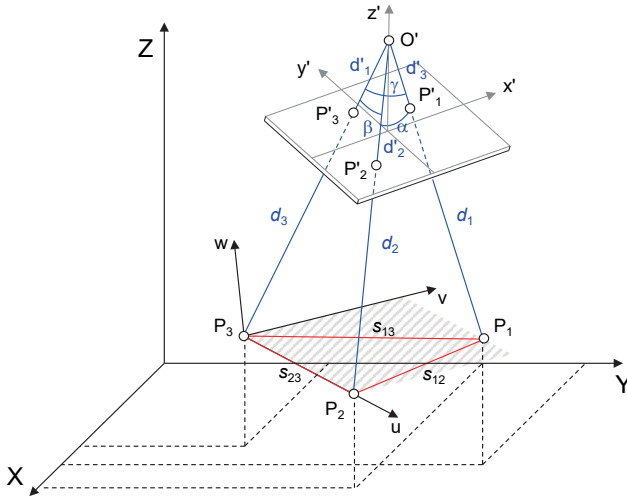
Fig. 4.11: Tetrahedron for space resection.

However, rather than attempting to compute the intersection of these relatively complex shapes, the following iterative search strategy estimates the unknown side lengths of the tetrahedron and from these the perspective centre can be more simply calculated as described below. The search is made as follows. Starting near P_1 in circle K_1, a test point R is stepped around the circle and distances d_{R1} and d_{R2} calculated at test positions. Distance d_{R1} is an estimate of side length d_1 and d_{R2} an estimate of side length d_2. At each test position, distance d_{R2} is transferred into circle K_2 where a corresponding value d_{R3} can be calculated from angle β and the known distance s_{23}. Finally, d_{R3} is transferred into circle K_3 where, in a similar way, distance d_1 is again estimated as value d'_{R1}. At the end of the loop a difference $\Delta d_1 = d_{R1} - d'_{R1}$ results. If R is stepped further around the circle until it approaches P_2, the sign of the difference at some position will change. At this position, sufficiently good approximate values for the side lengths d_i are available.
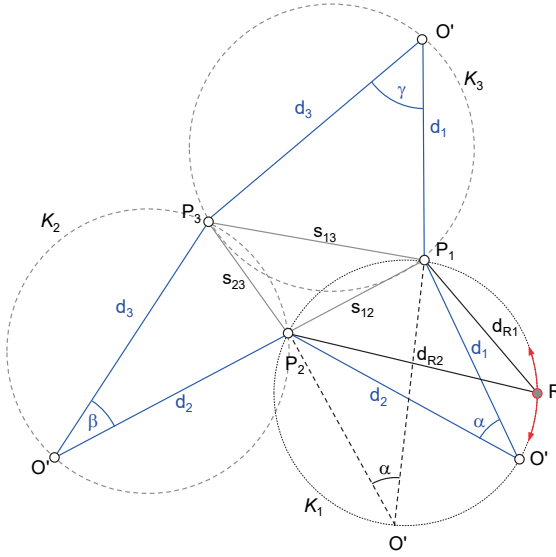
Fig. 4.12: Approximate values for space resection.

When transferred into circle K_2, d_{R2} in general creates two possible positions for O' and therefore two possible values of distance d_{R3}. When transferred into circle K_3 each value of d_{R3} generates two possible positions for O' which therefore leads to a total number of four possible solutions for the position of the perspective centre. To ensure that all possible solutions are investigated, and correct solutions with $\Delta d_1 = 0$ are found, the circles must be searched in order of increasing size, starting with the smallest.

Now the rotation and translation of the image coordinate system with respect to the object coordinate system can be determined. Using the side lengths estimated above for one of the solutions, the coordinates of reference points P_i, $i = 1\dots3$, are calculated in the camera coordinate system xyz, with origin O', which coincides with the image coordinate system x'y'z' (Fig. 4.11):

$$\mathbf{x}_i = \frac{d_i}{d_i'} \cdot \mathbf{x}'_i \qquad \text{where } d' = \sqrt{x'^2 + y'^2 + z'^2} \tag{4.20}$$

$$x_i = d_i \frac{x'_i}{d'_i} \qquad\qquad y_i = d_i \frac{y'_i}{d'_i} \qquad\qquad z_i = d_i \frac{z'_i}{d'_i}$$

Approximate values for the rotation parameters can then be derived by a spatial similarity transformation (see section 2.2.4), since the coordinates of P_1, P_2, P_3 are known in both the xyz and the XYZ systems (Fig. 4.9).

$$\mathbf{X}_i = \mathbf{X}_0 + m \cdot \mathbf{R} \cdot \mathbf{x}_i \quad \text{for} \quad i = 1,2,3 \tag{4.21}$$

Since the distances, P_iP_{i+1} are the same in both systems, the scale factor, m, will be unity. The rotation matrix **R** rotates the xyz coordinate system parallel to the XYZ system. Finally, the required translation, $\mathbf{X}_0$, may be found using the xyz coordinates $\mathbf{x}_i$ of a reference point, P_i, and rotated by **R**. For example, using P_1:

$$\mathbf{X}_0 = \mathbf{X}_1 - \mathbf{R} \cdot \mathbf{x}_1 \tag{4.22}$$

For the reasons given above, the spatial position of O', and the corresponding rotation matrix, are not unique. With the correct solution, O' lies on the visible side of the object. With the wrong solution, the point is mirrored in the plane defined by the object points. To solve the ambiguity, the distances of points P_3 and P'_3 to the plane $O'P_1P_2$ are calculated with eqn. (2.151). With the correct solution, both distances have the same sign, hence lie on the same side of the object plane. Alternatively, the ambiguity can be resolved by use of a fourth reference point.

The approximate values found in this way can then initiate a linearized solution to generate optimized values. The normal system of equations is set up and corrective additions added to the solutions in a sequence of k iterative calculations until these additions lie below a threshold value:

$$
\begin{aligned}
X_0^{k+1} &= X_0^k + dX_0^k & \omega^{k+1} &= \omega^k + d\omega^k \\
Y_0^{k+1} &= Y_0^k + dY_0^k & \varphi^{k+1} &= \varphi^k + d\varphi^k \\
Z_0^{k+1} &= Z_0^k + dZ_0^k & \kappa^{k+1} &= \kappa^k + d\kappa^k
\end{aligned}
\tag{4.23}
$$

Amongst other applications, space resection is used to compute initial orientation values from approximate object point coordinates (see section 4.4.4.1). If the mirror effect mentioned above is not taken into account, then wrong approximations can result for plane object surfaces.

4.2.3.5 Quality measures

In addition to the accuracy of image coordinate measurement, the quality of the resection depends on the number and distribution of reference points. Measured image points should ideally fill the image format. If all the reference points are located on or close to a common straight line, the normal system of equations becomes singular or numerically weak. Similarly, there is no solution if object points and perspective centre are all located on the same *danger surface* such as a cylinder (see also section 4.3.3.6).

As in other adjustment problems, the a posteriori standard deviation of unit weight s_0 can be used as a quality criterion for the resection. It represents the internal accuracy of the observations, in this case the measured image coordinates. In addition, the standard deviations of the estimated orientation parameters can be analysed. They can be derived from the covariance matrix and hence depend on s_0.

4.2.4 Linear orientation methods

4.2.4.1 Direct linear transformation (DLT)

Using the *Direct Linear Transformation* (DLT) it is possible, by solving a linear system of equations, to determine the orientation of an image without the need for approximate initial values. The method is based on the collinearity equations, extended by an affine transformation of the image coordinates. No image coordinate system fixed in the camera is required.

The transformation equation of the DLT is given by:

$$x' = \frac{L_1 X + L_2 Y + L_3 Z + L_4}{L_9 X + L_{10} Y + L_{11} Z + 1} \qquad\qquad y' = \frac{L_5 X + L_6 Y + L_7 Z + L_8}{L_9 X + L_{10} Y + L_{11} Z + 1} \qquad (4.24)$$

Here x and y are the measured comparator or image coordinates and X,Y,Z are the 3D coordinates of the reference points. The coefficients L_1 to L_{11} are the DLT parameters to be estimated and from these the parameters of interior orientation (3) and exterior orientation (6) can be derived. The two remaining elements describe shearing and scaling of the affine transformation. By re-arrangement of eqn. (4.24) the following linear system is obtained:

$$L_1 X + L_2 Y + L_3 Z + L_4 - x' L_9 X - x' L_{10} Y - x' L_{11} Z - x' = 0$$
$$L_5 X + L_6 Y + L_7 Z + L_8 - y' L_9 X - y' L_{10} Y - y' L_{11} Z - y' = 0 \qquad (4.25)$$

Determination of the 11 DLT parameters requires a minimum of 6 reference points, each of which provides two equations, resulting in minimum total of 12 equations. The solution is calculated according to the usual model

$$\mathbf{v} = \mathbf{A} \cdot \hat{\mathbf{x}} - \mathbf{l}$$

for which the design matrix $\mathbf{A}$ is set up as follows:

$$\mathbf{A}_{n,u} = \begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -x'_1 X_1 & -x'_1 Y_1 & -x'_1 Z_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -y'_1 X_1 & -y'_1 Y_1 & -y'_1 Z_1 \\ X_2 & Y_2 & Z_2 & 1 & 0 & 0 & 0 & 0 & -x'_2 X_2 & -x'_2 Y_2 & -x'_2 Z_2 \\ 0 & 0 & 0 & 0 & X_2 & Y_2 & Z_2 & 1 & -y'_2 X_2 & -y'_2 Y_2 & -y'_2 Z_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_n & Y_n & Z_n & 1 & 0 & 0 & 0 & 0 & -x'_n X_n & -x'_n Y_n & -x'_n Z_n \\ 0 & 0 & 0 & 0 & X_n & Y_n & Z_n & 1 & -y'_n X_n & -y'_n Y_n & -y'_n Z_n \end{bmatrix} \qquad (4.26)$$

Since equations (4.26) are linear in the unknowns, L_i, no approximate values of the unknown parameters are required. Because of the affine transformation applied to the measured image coordinates, there is no need for an image coordinate system defined by reference points fixed in the camera, such as fiducial marks. Instead it is possible to make direct use of measured comparator coordinates and, in general,

coordinates from an arbitrary image measuring device with non-orthogonal axes and different axial scale factors, e.g. pixel coordinates. Images from non-metric cameras, which have no image coordinate system or have an unknown interior orientation, can therefore be evaluated by this method.

Because of its robust linear form, the DLT is also used for the calculation of approximate exterior orientation parameters prior to a bundle adjustment. The more familiar orientation parameters can be derived from the DLT parameters as follows:

With

$$L = \frac{-1}{\sqrt{L_9^2 + L_{10}^2 + L_{11}^2}}$$

the parameters of interior orientation are obtained as follows:

$$
\begin{aligned}
x'_0 &= L^2 (L_1 L_9 + L_2 L_{10} + L_3 L_{11}) \\
y'_0 &= L^2 (L_5 L_9 + L_6 L_{10} + L_7 L_{11})
\end{aligned}
\qquad : \text{coordinates of principal point} \qquad (4.27)
$$

$$
\begin{aligned}
c_x &= \sqrt{L^2 (L_1^2 + L_2^2 + L_3^2) - x'^2_0} \\
c_y &= \sqrt{L^2 (L_5^2 + L_6^2 + L_7^2) - y'^2_0}
\end{aligned}
\qquad
\begin{aligned}
&: \text{principal distance} \\
&\ \ (\text{different scales in x and y})
\end{aligned}
\qquad (4.28)
$$

The parameters of exterior orientation, as defined by the elements of rotation matrix **R,** are given by

$$r_{11} = \frac{L \cdot (x'_0 \cdot L_9 - L_1)}{c_x} \qquad r_{12} = \frac{L \cdot (y'_0 \cdot L_9 - L_5)}{c_y} \qquad r_{13} = L \cdot L_9$$

$$r_{21} = \frac{L \cdot (x'_0 \cdot L_{10} - L_2)}{c_x} \qquad r_{22} = \frac{L \cdot (y'_0 \cdot L_{10} - L_6)}{c_y} \qquad r_{23} = L \cdot L_{10} \qquad (4.29)$$

$$r_{31} = \frac{L \cdot (x'_0 \cdot L_{11} - L_3)}{c_x} \qquad r_{32} = \frac{L \cdot (y'_0 \cdot L_{11} - L_7)}{c_y} \qquad r_{33} = L \cdot L_{11}$$

The position of the perspective centre is given by

$$
\begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} = - \begin{bmatrix} L_1 & L_2 & L_3 \\ L_5 & L_6 & L_7 \\ L_9 & L_{10} & L_{11} \end{bmatrix}^{-1} \cdot \begin{bmatrix} L_4 \\ L_8 \\ 1 \end{bmatrix}
\qquad (4.30)
$$

It is sometimes possible in this process that the determinant of the rotation matrix **R** is negative. This must be checked and, if necessary, the matrix must be multiplied by −1 to make its determinant positive. If the determinant of the rotation matrix in (4.29) is negative, L must be multiplied by −1.

To avoid possible numerical uncertainties, the elements of the rotation matrix must be normalized to an orthonormal matrix (see section 2.2.2.1). The individual rotation angles can be further derived according to eqn. (2.36), with due regard to the ambiguities indicated. The DLT model can be extended by correction terms for radial distortion.

Example 4.3:

Using the six image and reference points from example 4.2, a DLT is calculated:

DLT parameters:

$$L_1 = 0.025599 \qquad L_2 = 0.000789 \qquad L_3 = 0.000914$$
$$L_4 = 10.341817 \qquad L_5 = -0.000347 \qquad L_6 = 0.025008$$
$$L_7 = -0.008166 \qquad L_8 = 4.986816 \qquad L_9 = 0.000082$$
$$L_{10} = -0.000239 \qquad L_{11} = -0.000551$$

Applying eqns. (4.27ff), the following orientation parameters are then derived:

Interior orientation:

$$c_x = 42.0688 \text{ mm} \qquad c_y = 43.1748 \text{ mm}$$
$$x'_0 = 3.8202 \text{ mm} \qquad y'_0 = -4.0620 \text{ mm}$$

Exterior orientation:

$$X_0 = -471.12 \text{ mm} \qquad Y_0 = 318.14 \text{ mm} \qquad Z_0 = 1604.96 \text{ mm}$$

$$\mathbf{R} = \begin{bmatrix} 0.990725 & -0.000546 & -0.135153 \\ 0.066680 & 0.917685 & 0.393592 \\ 0.118394 & -0.397306 & 0.909295 \end{bmatrix}$$

$$\omega = -23.4056° \qquad \varphi = -7.7675° \qquad \kappa = 0.0316°$$

The DLT result departs significantly from the space resection calculation in example 4.2. The reason is that in this example the object points lie close to a plane. In this case, the DLT generates values which are not suitable for either orientation or calibration.

Together with the benefits of the DLT mentioned above, there are some drawbacks. If the parameters of interior orientation are known, the DLT has an excess of parameters. In addition, singular or weakly conditioned systems of equations arise if all reference points are located on a common plane, or if the denominator in equations (4.24) is close to zero. Measurement errors in the image coordinates, and errors in reference point coordinates, cannot be detected by the DLT and this results in false parameters. Finally, the minimum number of 6 reference points cannot always be provided in real applications.

4.2.4.2 Perspective projection matrix

The creation of an image point from an object point can also be formulated using the methods of projective geometry. Here object and image coordinates are expressed as homogeneous vectors. Object coordinates X are transformed into image space with image coordinates x' using a 3x4 projection matrix $\mathbf{P}$.

$$\boldsymbol{x}' = \mathbf{P} \cdot \boldsymbol{X} \tag{4.31}$$

where

$$\mathbf{P} = \mathbf{K} \cdot \mathbf{R} \cdot \left[\boldsymbol{I} \mid -\boldsymbol{X_o} \right] = \mathbf{K} \cdot \left[\mathbf{R} \mid -\mathbf{R} \cdot \boldsymbol{X_o} \right] \tag{4.32}$$

P is based on the parameters of exterior orientation **R** and $\boldsymbol{X_o}$, as well as a calibration matrix **K** which contains five parameters of interior orientation.

$$\mathbf{K} = \begin{bmatrix} 1 & s' & x'_0 \\ 0 & 1+m' & y'_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} c & cs' & x'_0 \\ 0 & c(1+m') & y'_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.33}$$

In addition to the coordinates of the perspective centre (x'_0, y'_0, c), s' and m' describe the shear and differential scales of the image coordinate axes.

With the six parameters of exterior orientation, **P** therefore has 11 independent transformation parameters. They can be solved in a linear system of equations and transformed into the 11 DLT parameters.

Written in full, (4.32) gives the following system of equations:

$$\mathbf{P} = \begin{bmatrix} c & cs' & x'_0 \\ 0 & c(1+m') & y'_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -X_0 \\ 0 & 1 & 0 & -Y_0 \\ 0 & 0 & 1 & -Z_0 \end{bmatrix} \tag{4.34}$$

(4.31) can therefore be written as follows:

$$\boldsymbol{x}' = \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \begin{bmatrix} c & cs' & x'_0 \\ 0 & c(1+m') & y'_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -X_0 \\ 0 & 1 & 0 & -Y_0 \\ 0 & 0 & 1 & -Z_0 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{4.35}$$

From the right-hand side, (4.35) expresses a translation of the perspective centre, followed by a rotation **R**, and finally a perspective scaling and correction to the principal point, in image space. With $x'=u'/w'$ and $y'=v'/w'$ the classical collinearity equations can be derived. Distortion effects are taken into account in the above model by the addition of correction values $\Delta x', \Delta y'$:

$$\mathbf{K}' = \begin{bmatrix} c & cs' & x'_0 + \Delta x' \\ 0 & c(1+m') & y'_0 + \Delta y' \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta x' \\ 0 & 1 & \Delta y' \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c & cs' & x'_0 \\ 0 & c(1+m') & y'_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.36}$$

4.2.5 Object position and orientation (pose) by inverse resection

4.2.5.1 Position and orientation of an object with respect to a camera

An inverse space resection enables the determination of the spatial position and orientation of an object, also called pose, with respect to the 3D camera coordinate system. The procedure is sometimes known as the 6 DOF (degrees of freedom) calculation for the target object.

Based on the parameters of exterior orientation described in equations 4.1 and 4.2, the relationship between coordinates of a point, $\mathbf{X}$ in the object system and $\mathbf{x}^*$ in the camera system (see Fig. 4.5), is given by:

$$\mathbf{X} = \mathbf{X}_0 + \mathbf{R} \cdot \mathbf{x}^\star \tag{4.37}$$

Rearranging (4.37) gives:

$$\mathbf{x}^\star = \mathbf{R}^{-1} \cdot \left(\mathbf{X} - \mathbf{X}_0 \right) \tag{4.38}$$

where $\mathbf{x}^\star$ gives the coordinates with respect to the camera system of a point on the object.

When expanded, eqn. (4.38) gives:

$$\mathbf{x}^\star = -(\mathbf{R}^{-1} \cdot \mathbf{X}_0) + \mathbf{R}^{-1} \cdot \mathbf{X} = \mathbf{X'}_0 + \mathbf{R'} \cdot \mathbf{X} \tag{4.39}$$

This has the same form as eqn. (4.37) and can be interpreted as an inverse operation, i.e. object coordinates in the camera system rather than vice versa. If $\mathbf{X_0}$ and $\mathbf{R}$ are obtained from a space resection, then $\mathbf{X'_0}$ and $\mathbf{R'}$ are the inverse resection values. Typically, this is applied to the monitoring of a known object. If the camera remains fixed during an imaging sequence, the spatial motion (pose) of a known object can be fully determined by repeated inverse space resections.

4.2.5.2 Position and orientation of one object relative to another

The spatial relationship between two objects can also be calculated by space resection, provided that both objects appear in the same image. In this case a reference object is defined with its own local coordinate system, XYZ, in which the position and orientation of a second object is to be determined. This second object, here called the probe or locator, has a separate coordinate system, xyz (Fig. 4.13). Two space resections are calculated using control points given in each of these two coordinate systems (see also eqn. 4.3):

$$\mathbf{X} = \mathbf{X}_0 + m \cdot \mathbf{R}_R \cdot \mathbf{x}^\star \qquad \text{: resection on reference points} \tag{4.40}$$

$$\mathbf{x} = \mathbf{x}_0 + m \cdot \mathbf{R}_L \cdot \mathbf{x}^\star \qquad \text{: resection on locator points} \tag{4.41}$$

Rearranging (4.41) gives:

$$\mathbf{x}^{\star} = \frac{1}{m} \cdot \mathbf{R}_L^{-1} \cdot (\mathbf{x} - \mathbf{x}_0) \tag{4.42}$$

Substituting for $\mathbf{x}^{\star}$ in (4.42) from (4.40) eliminates the scale factor m to give:

$$\mathbf{X} = \mathbf{X}_0 + \mathbf{R}_R \cdot \mathbf{R}_L^{-1} \cdot (\mathbf{x} - \mathbf{x}_0)$$
$$\mathbf{X} = \mathbf{X}_0 + \mathbf{R} \cdot (\mathbf{x} - \mathbf{x}_0) \tag{4.43}$$

in which

x: position of a locator point P, e.g. a probing tip (see Fig. 4.13), within its xyz system

X: position of the same point P in the XYZ reference system (which is the result required)

$\mathbf{x}_0$, $\mathbf{X}_0$:   coordinates of the projection centre within locator xyz and reference XYZ systems respectively

$\mathbf{R}_L$, $\mathbf{R}_R$:   rotation matrices of camera axes with respect to locator xyz and reference XYZ systems respectively

R: rotation matrix of the locator xyz axes with respect to the XYZ axes

Fig. 4.14 illustrates an application in which a hand-held measurement probe is monitored by a single camera. The 6DOF calculation is made in real time at a rate of 25 Hz. Since the reference object is also simultaneously located, any movement of the reference object or camera has no effect on the calculation of relative 6DOF between reference and probe. Fig. 6.94 shows another industrial example of the same principle, used to measure the spatial movements of a wheel on a moving car.
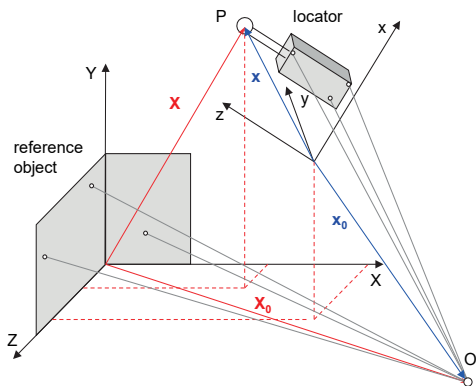
Fig. 4.13: 6DOF relation between two objects and a camera.
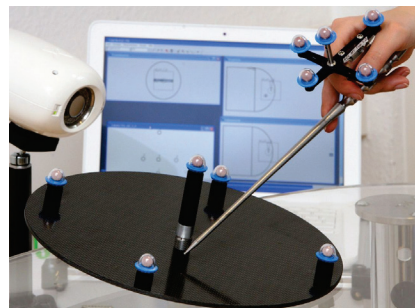
Fig. 4.14: Example of a 6DOF application with reference object and locator (AXIOS 3D).

The accuracy of the calculated 6DOF values depends on a number of parameters, including:
– focal length of camera;
– physical extent of the observed targets;
– distance of targets from camera;
– accuracy of the reference target locations and the image measurement.

a) Translation error as a function of focal length

b) Rotation error as a function of focal length

c) Translation error as a function of locator size

d) Rotation error as a function of locator size

e) Translation error as a function of image measurement accuracy

f) Rotation error as a function of image measurement accuracy
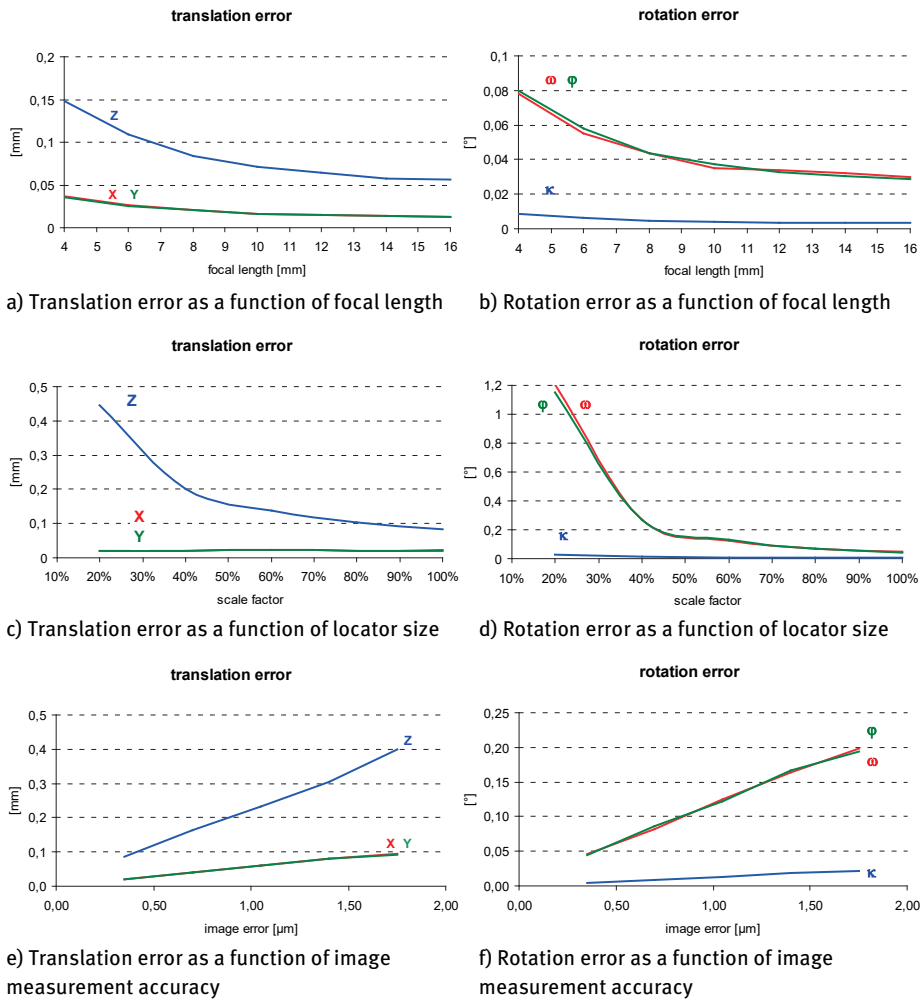
Fig. 4.15: Simulated errors in 6DOF calculation.

In general, rotations of either body around an axis perpendicular to the image plane (roll angle), as well as lateral shifts (parallel to the image plane), can be determined very accurately. Determinations of movement in the camera's viewing direction (Z),

and around the two remaining axes of rotation, is less accurate and significantly more sensitive to measurement and errors in camera orientation and calibration. Fig. 4.15 shows examples of position and rotation errors in a 6DOF calculation relating to the scenario illustrated by the application in Fig. 4.14. This problem also applies to 6DOF targets (section 6.2.2.4), whose orientation in space is captured with only one camera.

4.2.6 Projective transformation of a plane

4.2.6.1 Mathematical model

A special case is the reconstruction of plane object surfaces. The central perspective projection of an object plane onto the image plane is described by the projective transformation (see section 2.2.1.6), also known as homography.

By setting $Z = 0$ in eqn. (4.15), the transformation equations for the central projection of a plane are obtained:

$$x' = \frac{a'_0 + a'_1 X + a'_2 Y}{1 + c'_1 X + c'_2 Y} \qquad\qquad y' = \frac{b'_0 + b'_1 X + b'_2 Y}{1 + c'_1 X + c'_2 Y} \qquad\qquad (4.44)$$

or alternatively in the inverse form:

$$X = \frac{a_0 + a_1 x' + a_2 y'}{c_1 x' + c_2 y' + 1} \qquad\qquad Y = \frac{b_0 + b_1 x' + b_2 y'}{c_1 x' + c_2 y' + 1} \qquad\qquad (4.45)$$

In order to determine the eight parameters of eqn. (4.45), at least four reference points are required on the plane, no three of which may lie on a common straight line.

The (over-determined) system of equations is solved by an iterative least-squares adjustment. Using the approximate initial values:

$$a_1 = b_2 = 1 \qquad\qquad\qquad a_0 = a_2 = b_0 = b_1 = c_1 = c_2 = 0$$

and the substitution

$$N_i = c_1 x'_i + c_2 y'_i + 1$$

the following observation equations are derived:

$$v_{Xi} = \frac{1}{N_i} da_0 + \frac{x'_i}{N_i} da_1 + \frac{y'_i}{N_i} da_2 + \frac{x'_i X_i}{N_i} dc_1 + \frac{y'_i X_i}{N_i} dc_2 - \frac{lx_i}{N_i}$$

$$v_{Yi} = \frac{1}{N_i} db_0 + \frac{x'_i}{N_i} db_1 + \frac{y'_i}{N_i} db_2 + \frac{x'_i Y_i}{N_i} dc_1 + \frac{y'_i Y_i}{N_i} dc_2 - \frac{ly_i}{N_i} \qquad (4.46)$$

where

$$-lx_i = a_0 + a_1 x'_i + a_2 y'_i - c_1 x'_i X_i - c_2 y' X_i - X_i$$
$$-ly_i = b_0 + b_1 x'_i + b_2 y'_i - c_1 x'_i Y_i - c_2 y' Y_i - Y_i$$

The approximate solution values at iteration k are adjusted by the computed corrections to the unknowns and the process repeated until the changes are no longer significant, i.e.:

$$a_0^{k+1} = a_0^k + da_0^k$$
and so on

Object coordinates   Image coordinates   Coefficients

$(X_0 = 7500, Y_0 = 4000, Z_0 = 10000,$
$\omega = 0°, \varphi = 0°, \kappa = 30°)$


```
a0=7500.0000
a1= 433.0127
a2=-250.0000
b0=4000.0000
b1= 250.0000
b2= 433.0127
c1=   0.0000
c2=   0.0000
```

Image coordinates   Coefficients

$(X_0 = 15000, Y_0 = 4000, Z_0 = 10000,$
$\omega = -20°, \varphi = 40°, \kappa = 30°)$

Pt	X	Y
1	0	0
2	0	3500
3	7500	8000
4	15000	3500
5	15000	0


```
a0=6070.4880
a1=1183.9872
a2=-272.1012
b0= 360.2976
b1= 386.9559
b2= 639.9948
c1=   0.0482
c2=  -0.0004
```

Fig. 4.16: Projective transformation of a plane pentangle.

Equations 4.45 are non-linear in the coefficients $a,b,$c. A direct, non-iterative calculation of the unknown parameters is possible if linear equations (2.22) are used. This results in the following equations

$$a_0 + a_1 x'_i + a_2 y'_i - c_1 x'_i \, X_i - c_2 y'_i \, X_i = X_i$$
$$b_0 + b_1 x'_i + b_2 y'_i - c_1 x'_i \, Y_i - c_2 y'_i \, Y_i = Y_i$$

(4.47)

which can be solved directly according to the scheme:

$$\mathbf{A} \cdot \hat{\mathbf{x}} = \mathbf{l}$$

(4.48)

Using the transformation parameters of (4.45) further image coordinates can be transformed into object coordinates.

Fig. 4.16 shows the position of five plane object points and the corresponding image coordinates with their related transformation parameters. For the special case where $c_1 = 0$ and $c_2 = 0$, the projective transformation (4.45) reduces to an affine transformation (eqn. 2.12). For the further case of parallel object and image planes (Fig. 4.16 top right), eqn. (4.45) can be replaced by the plane similarity transformation (2.5).

By the use of homogeneous coordinates, the projective transformation of a plane can be formulated as a homogeneous transformation.

$$X = H \cdot x$$
$$\begin{bmatrix} w \cdot X \\ w \cdot Y \\ w \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$$

(4.49)

This formulation is known as *homography*. Since the matrix H can be arbitrarily scaled without altering its projective properties, this shows that there are eight degrees of freedom in the transformation, as there are in eqn. (4.45). Thus, the matrix could also be normalized to $h_{33}=1$. From this a set of linear equations is obtained which can be solved directly.

$$w \cdot X = h_{11} \cdot x' + h_{12} \cdot y' + h_{13}$$
$$w \cdot Y = h_{21} \cdot x' + h_{22} \cdot y' + h_{23}$$
$$w = h_{31} \cdot x' + h_{32} \cdot y' + h_{33}$$

(4.50)

If the last equation is inserted into the first two and all terms containing lower case coordinates are moved to the right of the equation, we obtain:

$$h_{33} \cdot X = h_{11} \cdot x' + h_{12} \cdot y' + h_{13} - h_{31} \cdot x' \cdot X - h_{32} \cdot y' \cdot X$$
$$h_{33} \cdot Y = h_{21} \cdot x' + h_{22} \cdot y' + h_{23} - h_{31} \cdot x' \cdot Y - h_{32} \cdot y' \cdot Y$$

(4.51)

If the normalization to $h_{33} = 1$ is applied, the same equations as (4.47) are obtained. Alternatively, the term to the left of the equal sign can be moved to the right side and the equation set to zero. In this case an eigenvalue decomposition will provide the solution. The eigenvector corresponding to the smallest eigenvalue contains the desired transformation parameters.

4.2.6.2 Influence of interior orientation

The model of central projection described above assumes straight line rays through the perspective centre. Although the spatial position of the perspective centre (principal distance c, principal point x'_0, y'_0) is modelled by the parameters of the projective transformation, it is not possible to compensate here for lens distortion.

The image coordinates must therefore be corrected for lens distortion before applying the projective transformation. For optimal accuracy, when applying distortion parameters derived from a separate process, care should be taken in case they are correlated with the interior and exterior orientation parameters. Further usage of these parameters might involve a different mathematical model where, strictly speaking, the distortion parameters should be applied with their full variance-covariance matrix from the bundle adjustment. However, this transfer of correlation information is rarely done. A determination of distortion parameters which is numerically independent of the orientation parameters can, for example, be done using the plumb-line method (see section 7.3.1.2).

4.2.6.3 Influence of non-coplanar object points

Object points which do not lie in the plane of the reference points have a positional error in image space after projective transformation. This shift depends on the height above the reference plane and on the position of the point in the image. The image plane is here assumed to be parallel to the reference plane. Since Fig. 4.17 represents a vertical plane containing the projection centre, the shift $\Delta r'$ is radial with respect to the principal point:

$$\Delta r' = \frac{r'}{h} \Delta h \tag{4.52}$$

Multiplying by the image scale gives the corresponding shift in the reference plane:

$$\Delta r = \frac{h}{c} \Delta r' = m \cdot \Delta r' \tag{4.53}$$

The equations above can also be used to determine the maximum height of a point above the reference plane which will not exceed a specified shift in image space.

Example 4.4:

For the measurement of a flat plate, a digital video camera is used with following specifications: $c = 8$ mm, $h = 2.5$ m, $r'_{max} = 5$ mm. The maximum offset above the object plane must be calculated which ensures that the resulting shift in object space Δr is less than 1mm.

1. Image scale: $m = \dfrac{h}{c} = \dfrac{2.5 \text{ m}}{0.008 \text{ m}} = 312$

2. From (4.52) and (4.53) $\Delta h = \Delta r \dfrac{c}{r'_{max}} = 1.0 \text{ mm} \cdot \dfrac{8 \text{ mm}}{5 \text{ mm}} = 1.6 \text{ mm}$
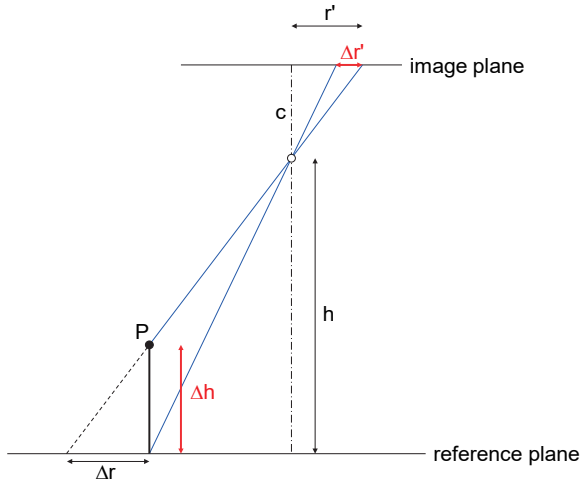
Fig. 4.17: Shift in image space caused by height differences.

4.2.6.4 Measurement of flat objects

If the projective transformation between any object plane and the image is known, measured values such as coordinates, lengths, areas or angles can be determined directly in the original image. Fig. 4.18 illustrates the measurement of flat panels on a roof. The parameters of the projective transformation are determined beforehand with the aid of a reference cross visible in the middle of the image. Every pixel or coordinate in the image can therefore be directly transformed into object coordinates with respect to the reference plane (and vice versa).

Fig. 4.18: Image showing overlay results from in-plane measurement.
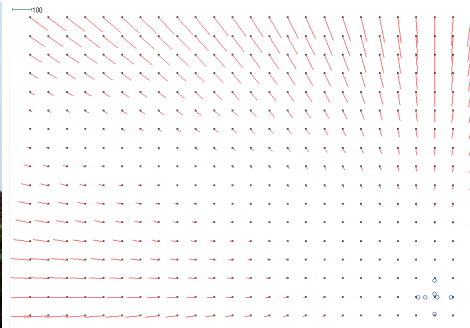
Fig. 4.19: Simulated errors of object points measured using an oblique image of a flat roof.

Measurement accuracy depends strongly on the perspective of the image. As usual, it reduces with increasing object distance, which results in measurements of very variable accuracy in the image. Fig. 4.19 shows simulated error vectors which would be expected in object points measured by oblique imaging of a flat roof. In this case the control points (blue) are in the lower right corner. The systematic and, in object space asymmetric, error distribution is obvious.

4.2.7 Monoplotting

The spatial ray defined by an image point can be intersected with the object surface, if interior and exterior orientation of the image are known and a geometric model of the surface exists. The object model can be defined by a known mathematical element (straight line, plane, cylinder, etc.), or by a dense grid of points (digital surface or elevation model). The intersection point of the ray and the object model is the desired 3D object coordinate. The procedure is known as *monoplotting* and is explained in the next two sections in more detail.

4.2.7.1 Standard geometric shapes (geometric primitives)

4.2.7.1.1 Plane

The principle of monoplotting within an object plane is illustrated in Fig. 4.20. The 3D coordinates of the object point P result from the intersection of the object surface and the ray defined by the image point P' and the perspective centre O'. The object plane can be defined, for example, by prior photogrammetric measurement of three points. The parameters of the plane can be calculated from a least-squares best-fit adjustment (see section 2.3.2.2).
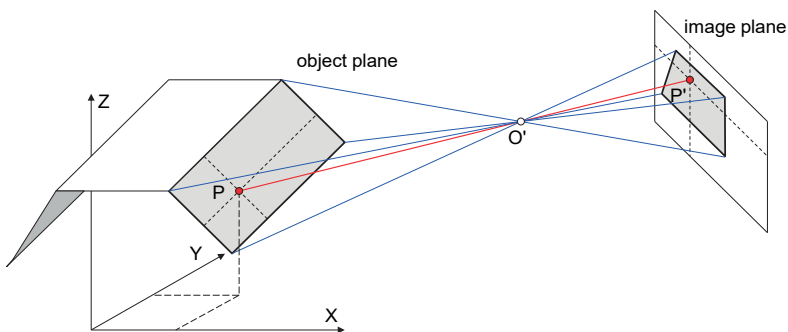
Fig. 4.20: Single image evaluation within an object plane.

A variation of plane monoplotting is represented by a triangular mesh where each individual triangle of the mesh defines a unique plane (eqn. 2.150). In addition to calculation of the intersection point of ray and plane, it is also necessary to test if the point lies with the triangular area.

4.2.7.1.2 Sphere and cylinder

In principle, any geometric element which can be described analytically can be intersected with an image ray. Spheres and cylinders are useful examples.

The intersection point of a straight line with a sphere can be solved by introducing the line in parameterized form (eqn. 2.139) into the sphere equation (2.158). This may result in two solutions (line within sphere), one solution (line tangential to sphere) or no solution (line outside sphere).

For a cylinder the same solutions exist as for a sphere. Additionally, it is necessary to check if the intersection point lies beyond the ends of the cylinder, which has finite length.

4.2.7.2 Digital surface models

An arbitrary or free-form object surface can be approximated by a suitably dense grid of 3D points (digital surface model, DSM, see section 2.3.3.1). The point grid can be regular, e.g. $\Delta X = \Delta Y = $ const., or irregular. Object edges (breaklines) can be modelled by special point codes or by additional vector data (polygons). Inside the object model further points can be interpolated.

Fig. 4.21 shows the principle of spatial point determination from a single image using a digital surface model (DSM). The spatial direction defined by the measured image coordinates x', y' and c in the image coordinate system (image vector $\mathbf{x'}$) is transformed into the spatial vector $\mathbf{X^*}$ using the exterior orientation parameters (similarity transform with arbitrary scale, e.g. $m = 1$). This ray intersects the DSM at point P using a local surface plane defined by the four adjacent points.

In order to calculate the point of intersection, the straight line g is constructed between the intersection point S of $\mathbf{X^*}$ and the XY plane, and the foot of the perpendicular O$_{XY}$ from the perspective centre to the XY plane (Fig. 4.22). A search for point P starts along this line at O$_{XY}$ until its interpolated height Z lies within two Z values of adjacent profile points.

If the DSM is defined by a triangular mesh instead of a rectangular grid structure, the intersection point is calculated as for a plane (see section 4.2.7.1). With complex 3D surfaces (see section 2.3.3) multiple triangles may be intersected, hence the point closest to the image must be selected.

Monoplotting is not very popular in close-range photogrammetry, but is gaining importance in CAD technology and through the combination of cameras with 3D point cloud recording, e.g. by laser scanners. Fig. 4.23 shows an example of monoplotting using a 3D point cloud from a terrestrial laser scanner which is equipped with a high-resolution camera (see section 6.8.1). Here geometrical primitives such as cylinders

are extracted from the measurement data. The evaluation is semi-automatic, requiring visual interpretation of the image. The enlargement of the image extract in Fig. 4.23b makes clear that a fully automatic evaluation of the complex point cloud is not practicable without additional image information.
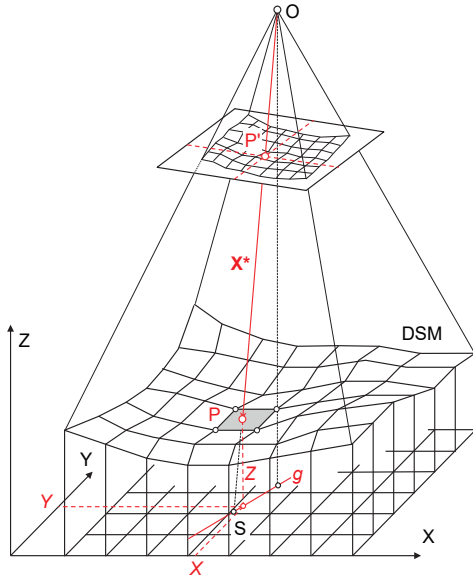
Fig. 4.21: Point determination in a digital surface model (DSM).
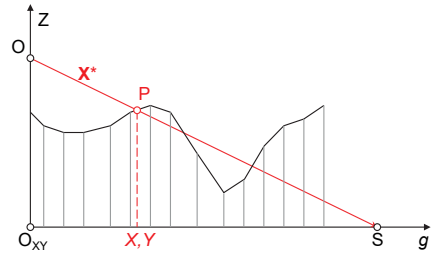
Fig. 4.22: Search and interpolation of object point P(XYZ) within the height profile above g.

a) Measurement image overlaid with simulated cylinder in object space

b) Image detail with overlaid 3D point cloud

Fig. 4.23: Monoplotting applied to the measurement of pipework by intersecting image rays with the 3D point cloud from a terrestrial laser scanner (Riegl, Phocad).

4.2.8 Rectification

4.2.8.1 Plane rectification

In addition to coordinate determination, the projective transformation is also used as the mathematical basis for digital image rectification. The aim is the transformation of an image into a new geometric (graphical) projection, in most cases a parallel projection.

For rectification of a plane, the complete image format, or a defined area of interest, is transformed into the reference system (target system) by the parameters of a single projective transformation, i.e. the same transformation coefficients are applied to every point in the source image area. In contrast, for non-planar objects every image point must be rectified as a function of its corresponding XYZ coordinates (differential rectification or orthophoto production, see section 4.2.8.2).

The digital projective rectification of a rectangular object plane (here: XY) is calculated in the following steps:

1. Measurement of reference and image points:
   A plane projective transformation requires at least four XY control points lying in a common plane and their corresponding image coordinates must be measured.

2. Calculation of transformation parameters:
   From the measured image and reference point coordinates, the coefficients a', b' and c' of the projective transformation are calculated (see section 4.2.6), by least-squares adjustment if required.

$$x'_i = \frac{a'_0 + a'_1 X_i + a'_2 Y_i}{1 + c'_1 X_i + c'_2 Y_i} \qquad\qquad y'_i = \frac{b'_0 + b'_1 X_i + b'_2 Y_i}{1 + c'_1 X_i + c'_2 Y_i} \qquad\qquad (4.54)$$

3. Definition of the rectangular object area to be rectified:
   Lower left corner: X_1, Y_1
   Upper right corner: X_2, Y_2

4. Definition of output scale or print resolution:
   The output scale (map scale) m_k and/or the desired output resolution (print, screen) $\Delta x_k, \Delta y_k$ of the rectification, define the step width (point distance) in object space. Often it will be defined according to the pixel size in object space of the original image (GSD, see section 1.2.3).

5. Definition of object resolution $\Delta X, \Delta Y$, which is used to "scan" the object:
   $$\Delta X = m_k \Delta x_k \qquad\qquad \Delta Y = m_k \Delta y_k$$

6. For each point on the object plane $(X,Y)_i$ the corresponding image or pixel coordinates $(x',y')_i$ are calculated using the parameters of the projective transformation.

7. At $(x',y')_i$ the colour value of the digital image is determined. This is usually done by interpolation since $(x',y')_i$ are real numbers which do not correspond to the integer position of a pixel (indirect rectification, see section 5.3.2):

$$g'_i = g(x',y')_i$$

8. The interpolated colour value g'_i is stored in the output image at pixel position $(x',y')_i$:

$$x_i = x_1 + j \cdot \Delta x_k \qquad\qquad g(x_i, y_i) = g'_i$$
$$y_i = y_1 + k \cdot \Delta y_k$$

In close-range photogrammetry, digital rectification has gained in importance, e.g. for the production of rectified image mosaics of building façades, or for the superposition of natural textures onto CAD elements (see section 5.3.3). Fig. 4.24 shows an example of the plane rectification of a façade. Object areas outside the reference plane (marked in red) are distorted. The curved image edge is created by considering the distortion of the input image. The digital image processing procedure for creating the new image is presented in section 5.3.1.

a) Original image   b) Rectification onto façade plane

Fig. 4.24: Example of plane rectification.

4.2.8.2 Differential rectification (orthophotos)

For the production of an orthophoto by differential rectification, each point (infinitesimally small object patch) is projected individually according to its XYZ coordinates in order to create a new image which is a parallel projection of the object surface. Compared with the process described in section 4.2.8.1, a digital orthophoto displaying a rectangular area of the ground plane XY is determined in the following steps (Fig. 4.25a):

1. Determination of the interior and exterior orientation of the image (or all images): Calibration and orientation data are determined according to the methods described in section 3.3.2 and 4.4.

2. Definition of the rectangular area of interest and resolution on the object:
3. For each grid point $(X,Y)_i$ the corresponding Z_i value is interpolated in the given surface model, e.g. by bilinear interpolation (see section 2.2.1.5):

$$X_i = X_1 + j \cdot \Delta X \qquad\qquad Z_i = Z(X_i, Y_i)$$
$$Y_i = Y_1 + k \cdot \Delta Y$$

4. Using the collinearity equations (4.10), and the given parameters of interior and exterior orientation, the image coordinate $(x',y')_i$ corresponding to $(X,Y,Z)_i$ is calculated:

$$x'_i = F(X_0, Y_0, Z_0, \omega, \varphi, \kappa, x'_0, c, \Delta x', X_i, Y_i, Z_i)$$
$$y'_i = F(X_0, Y_0, Z_0, \omega, \varphi, \kappa, x'_0, c, \Delta y', X_i, Y_i, Z_i)$$

5. At position $(x',y')_i$ the stored colour value is extracted from the digital image. The colour value is usually interpolated, since $(x',y')_i$ are floating point numbers which do not match the integer pixel raster of the image (see section 5.3.2):

$$g'_i = g(x', y')_i$$

6. The interpolated colour value g'_i is transferred into the output image at position $(x,y)_i$.

The method described above transforms each point of the orthophoto into the original image. In order to reduce the computational effort, the *anchor-point method* can be used. This transforms only a coarse grid of points from which local colour values can be linearly interpolated in the image. Fig. 4.25b shows an orthophoto of the image from Fig. 4.8 created using the corresponding surface model. In comparison, Fig. 4.25c shows the rectification calculated by plane projective transformation. At first glance, this looks identical to the orthophoto. However, the difference between the images, shown in Fig. 4.25d, demonstrates clear differences between the methods at all points where there are significant height differences from the reference plane (black: no difference, white: large difference). Object areas which are not described by the surface model are therefore shifted and distorted in image space.

Occluded object areas can only be projected in the orthophoto if additional images from different stations are available and integrated into the orthophoto process. The method however requires either extensive *ray tracing* or simultaneous multi-image processing, for example using object-based multi-image matching (see section 5.5.6).

The resulting orthophoto is geometrically correct only when *all* visible object elements are located to a corresponding accuracy in the digital surface model (DSM). Suitable surface models can be generated by point clouds which are created, for instance, by laser scanning, 3D cameras, or dense stereo or multi-image matching. Fig. 4.26 shows an example of a rectified façade generated with both a complete and incomplete surface model from laser scanned data. It can clearly be seen that object elements not part of the DSM are laterally shifted due to perspective imaging (see also

point B in Fig. 4.27). The complete, geometrically error-free rectification is known as a *true orthophoto*.

a) Principle of orthophoto calculation

b) Orthophoto calculated using surface model

c) Rectification by plane transformation

d) Difference image between b) and c)
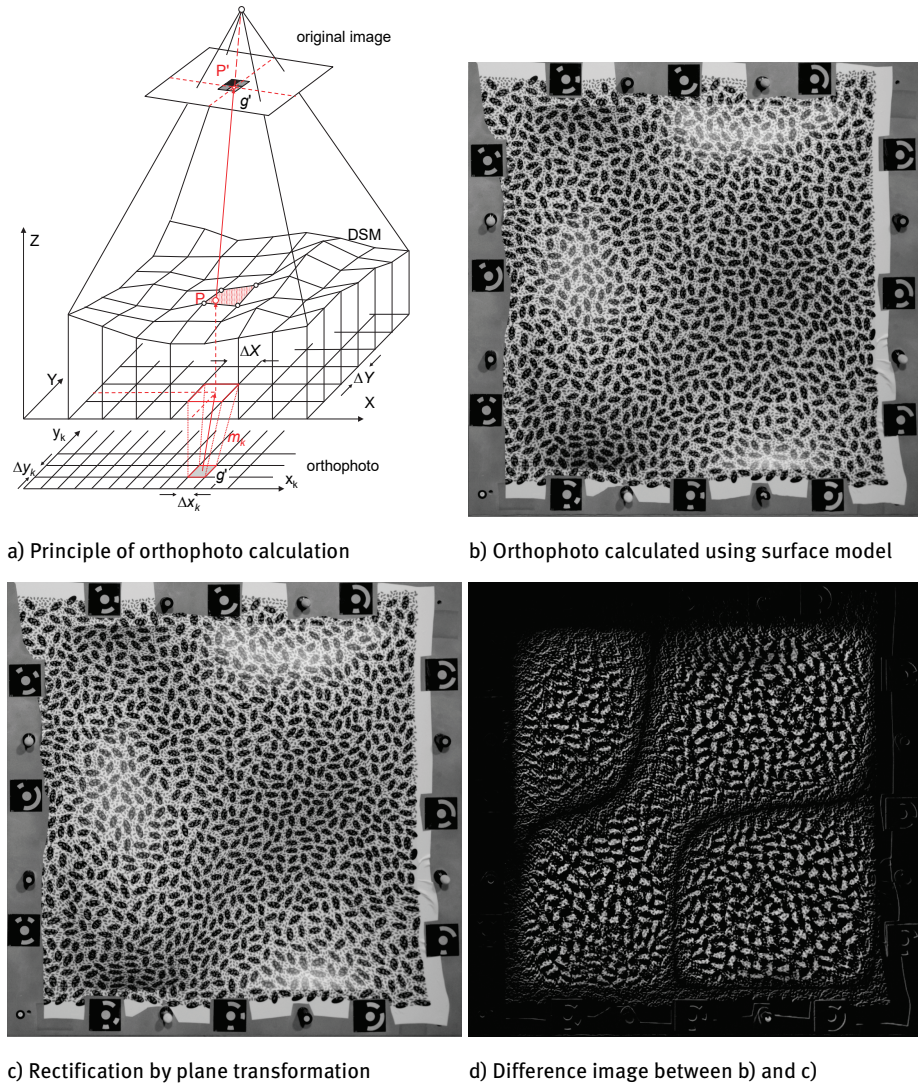
Fig. 4.25: Creation of orthophotos.

a) Orthophoto with incomplete surface model

b) True orthophoto with complete surface model

Fig. 4.26: Orthophoto generation with different surface models.

Even in true orthophotos, which are based on complete surface models, geometric errors may occur. One typical effect is known as *double mapping* (ghosting) where object areas appear twice in the final image as explained by Fig. 4.27. Point A is correctly projected because terrain and surface model are identical and occluding objects do not interrupt the imaging ray. Point B illustrates the lateral displacement which can occur when a point is re-projected when using a terrain model (B_T) instead a surface model (B_S). As a result, the image of the left façade will appear in the orthophoto. It should, of course, not appear at all since it is a vertical structure. In contrast, point C is an example of double mapping. Although C is represented by the surface model, the projected ray intersects an object, e.g. a house with roof. In the orthophoto, this part of the roof will therefore be visible at C as well as near B.
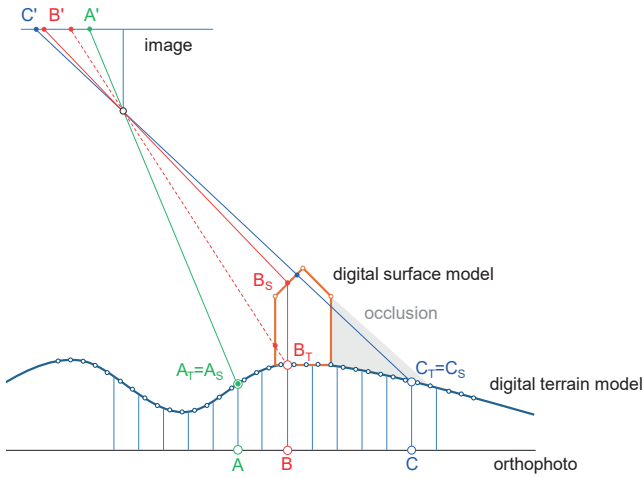
Fig. 4.27: Lateral displacement (B) and double mapping (B,C) in orthophoto generation.

Double mapping can be avoided if the rectification software uses ray tracing to detect occluding objects. Then the occluded object regions may either be drawn as empty (black) areas, or additional images from other positions may be used to fill these regions with correct colours.

a) With double mapping

b) Without double mapping

Fig. 4.28: True orthophoto generation.

Fig. 4.28a shows an example of a true orthophoto where double mapping is visible (windows on red roof appear twice). Fig. 4.28b illustrates the result without double mapping. The same digital surface model is behind both images but only the software

used to create the image of Fig. 4.28b takes ray tracing and multi-image processing into account.

Differential rectification is a general approach which, in principle, can be applied to any projections and object surfaces. In such cases it is necessary only to replace the transformations of steps 3 and 4 with functions relating the 2D coordinates of the original image to the 3D object coordinates. This method can handle, for example, cylindrical or spherical projections as well as the plane projective transformation described above (example in Fig. 4.29).

Fig. 4.29: Original image and cylindrical projection (monastery church at Gernrode (Fokus Leipzig).

4.2.8.3 Image mosaics

An orthophoto can also be generated by combining a number of partial views or original images into a common image mosaic. This procedure, used in the production of an aerial orthophoto mosaic, can be employed in close-range photogrammetry for the rectification of façades, or for texture mapping of 3D models (Fig. 4.30, section 5.3.3).

Fig. 4.30: 3D model with rectified image areas and superimposed textures
(Lurdji monastery, Tbilisi, Georgia, see also Figs. 6.111 and 6.112).

Image mosaics are essentially calculated according to the procedures above. The process can utilize either re-projection of the mosaic image points into the original images by projective or differential rectification, or previously rectified images can be combined. If more than one image is available to provide the colour value of the mosaic point, then one of the following decision rules can be applied:

- colour value is taken from the image where the angle between optical axis and surface normal is minimal;
- colour value is taken from the image where the image point is closest to the image centre;
- colour value is averaged over all images;
- minimum or maximum colour value from all related images;
- colour adjustment over all images to minimize colour edges.

Fig. 4.31 shows the example of an image mosaic generated from 9 images. The camera has been calibrated in advance, hence image distortion can be compensated. All images have been rectified using reference points known in a common plane coordinate system on the façade. For mosaicking, the average colour value of overlapping images has been used. An area-based colour adjustment has not been calculated, hence colour edges are still visible between the single images.
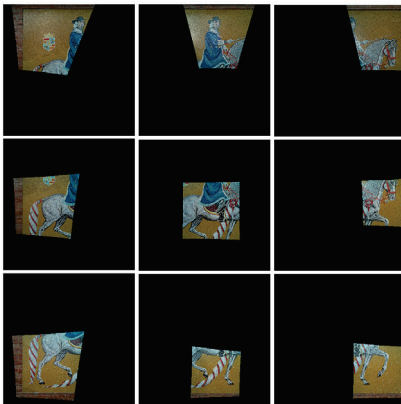
a) Rectified input images   b) Orthophoto mosaic

Fig. 4.31: Generation of an image mosaic.

4.3 Processing of stereo images

4.3.1 Stereoscopic principle

4.3.1.1 Stereoscopic matching

The photogrammetric processing of image pairs has, in many ways, a particular significance. For a start, two images of an object, taken from different positions, represent the minimum condition necessary for a 3D object measurement. Then there is the fact that human vision is highly developed and enables the stereoscopic viewing and analysis of image pairs.
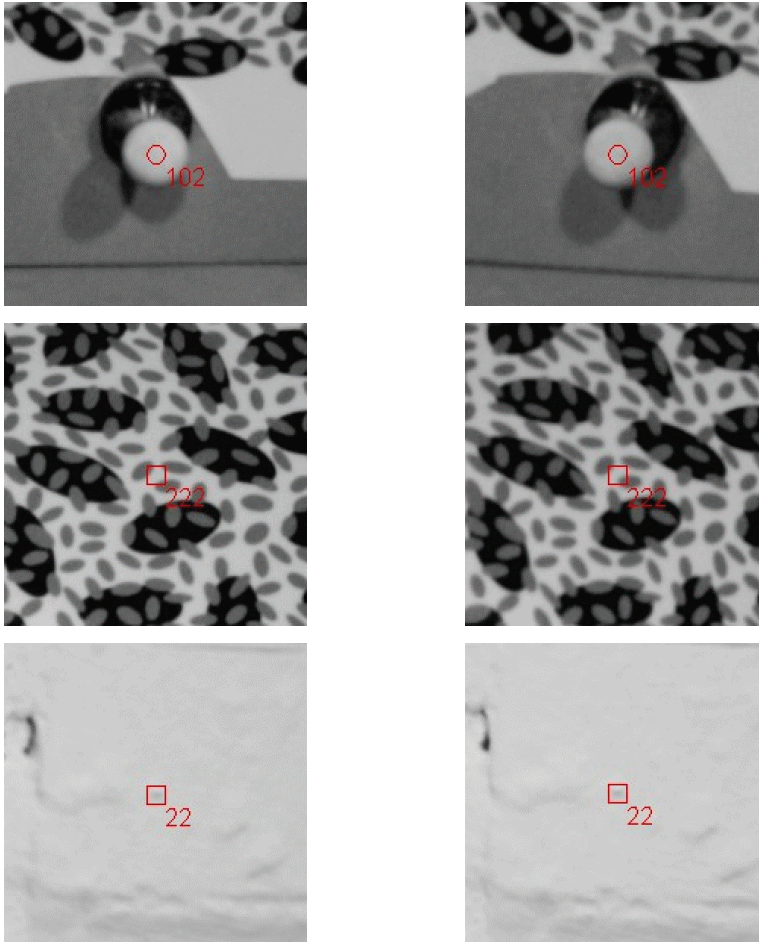
Fig. 4.32: Correspondence problem resulting from different object surface patterns in a stereo image.

The process of evaluating stereo images is typically divided into three steps:

1. Determination of homologous image features, e.g. corresponding points in both images;
2. Orientation of the image pair, potentially also with calibration of the camera(s);
3. 3D object measurement, e.g. measurement of free-form surfaces.

Depending on application, the steps can be executed in the sequence given, or in combination, or in a different order.

The essential task in stereo and multi-image processing is the solution of the correspondence problem, i.e. the matching of identical (homologous) image points. This depends first of all on the visible object texture which influences the accuracy and uniqueness of the association. Fig. 4.32 shows an example of varying textures, patterns and targets (extracts from the image pair in Fig. 4.38), which make clear that unique or ambiguous matches are possible, depending on pattern.

4.3.1.2 Tie points

Tie points are measured, homologous points in the images, i.e. they represent the same object point. They assist in the geometric connection between two or more images and need not be reference points. More specifically, they must be selected to cover a sufficient area in image and object space in order to provide a robust connection between the images.
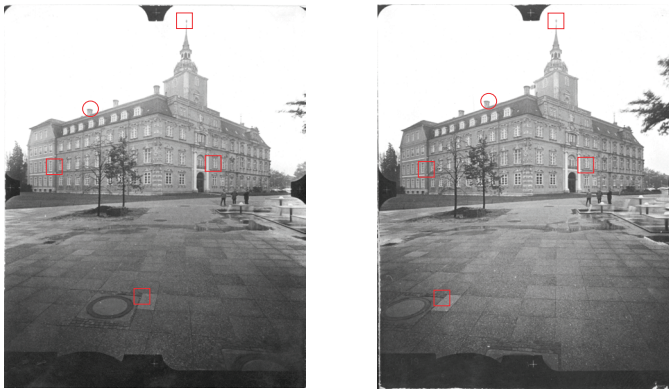
Fig. 4.33: Tie points in a stereo pair; □: correctly matched points; ○: wrongly matched points.

Homologous points can be identified visually, either by stereoscopic viewing or by monoscopic measurement of single images. Non-targeted object points can be identified more reliably by stereoscopic viewing. Correspondence between homologous points can also be performed by digital stereo image matching (see

section 5.5.4). Here similarities or cost functions in grey level patterns are compared in order to match corresponding points (image correlation).

Normally there is no orientation information available during the tie point measurement stage and so there are few controls to prevent false measurements. In larger photogrammetric projects, gross errors (blunders) are therefore almost always present in the observations due to errors in measurement or identification. Fig. 4.33 shows an example of four correct and one incorrect tie points.

4.3.1.3 Orientation of stereo image pairs

The orientation of a stereo pair provides exterior orientation parameters of both images (Fig. 4.34). In principle, this task can be solved separately for each image by space resection (see section 4.2.3) but three-dimensional reference points (full reference points) are then required for each photo (see Fig. 4.9). The reference points can be identical or different for each image. In this procedure the geometric relationship between the two images in a stereo model is not used.
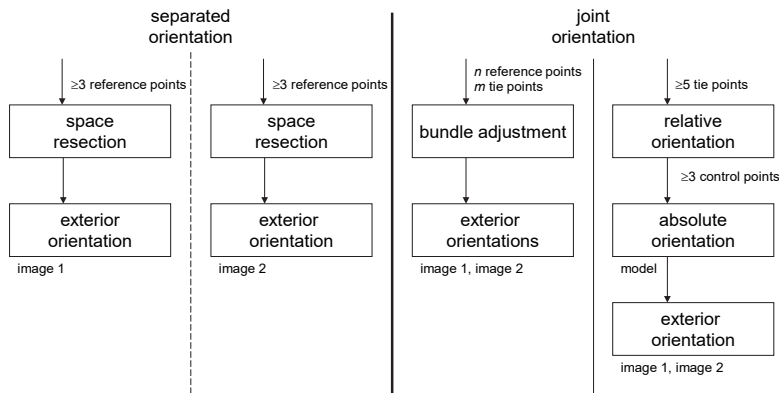
Fig. 4.34: Orientation methods for stereo images.

Fig. 4.33 shows a typical stereo pair where both images cover the object with an overlap of at least 50 % (a standard overlap is 60 %, see also Fig. 3.38a). This feature can be employed in the orientation of both images.

The one-step joint orientation employs the principle of bundle triangulation (section 4.4, Fig. 4.56) for the special case of two images. Here the orientation elements of both images are determined simultaneously in one step using the image coordinates of the reference points and additional tie points.

The traditional two-step solution of this problem works as follows. In the first step the correspondence between the images, and the coordinates of model points, are determined in a local coordinate system (relative orientation, see section 4.3.3). In the

second step the transformation into the global object coordinate system is performed using reference points (absolute orientation, see section 4.3.5).

Once the stereo model has been oriented, three-dimensional object reconstruction is possible using the methods outlined in section 4.3.6.

4.3.1.4 Normal case of stereo photogrammetry

The normal case of stereo photogrammetry is, in fact, the special case in which two identical ideal cameras have parallel axes pointing in the same direction at right angles to the stereo base. With respect to an XYZ coordinate system located in the left perspective centre, object coordinates can be derived from the ratios indicated in Fig. 4.35. Using real cameras, the normal case can only be achieved with low accuracy requirements (example in Fig. 1.32). Its use is mainly in the calculation of approximate coordinates and the estimation of the achievable measurement accuracy.
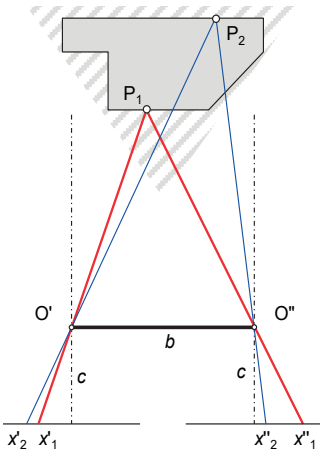
Fig. 4.35: Normal case of stereo photogrammetry.

Stereo images, which approximate to the normal case, can easily be evaluated by visual inspection. Ideally, objects are only observed with a distance-dependent, horizontal shift between the images, and a vertical shift (y-parallax), which is detrimental to stereoscopic viewing, will not exist. This horizontal shift is known as x-parallax or disparity. X-parallax increases with shorter object distances and is zero for objects at infinity. In the example of Fig. 4.35, the x-parallax for point P_1, with $px_1 = x'_1 - x''_1$, is greater than for point P_2 with $px_2 = x'_2 - x''_2$.

Real stereo image pairs can be digitally transformed (rectified) into the distortion-free normal stereo case (normal-case images or epipolar images, section 4.3.3.5) if their interior and relative orientation is known (relative orientation, section 4.3.3).

4.3.2 Epipolar geometry

Fig. 4.36 shows the geometry of a stereo pair imaging any object point P. The base **b**, and the projected rays **r'** and **r"** from each perspective centre to the object point, define an e*pipolar plane*, sometimes called a *basal plane*. This plane intersects the image planes along lines k' and $k"$, which are known as *epipolar lines*. In the case of convergent images, the epipolar lines are convergent. In the special case of normal stereo photogrammetry, the epipolar lines are parallel to the x' direction (Fig. 4.37, see also section 4.3.6.2). In the case of deviations from the central projection, e.g. due to distortion or with panoramic cameras, the epipolar lines become curves.
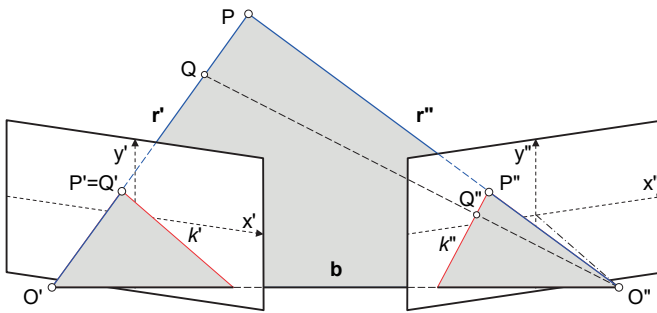
Fig. 4.36: Epipolar plane for convergent images.
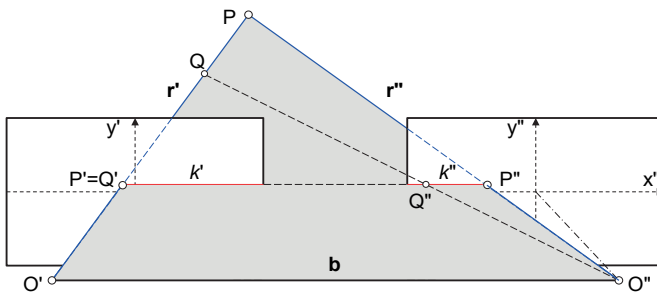
Fig. 4.37: Epipolar plane for normal case of stereo photogrammetry.

The importance of epipolar geometry lies in the fact that, assuming an error-free ray intersection, an image point P'' in the right image, corresponding to P' in the left image, must lie on the epipolar plane and hence on the epipolar line k''. Thus, the search space for matching corresponding points can be significantly reduced. Assuming an additional object point Q lying on ray O'P, it is obvious that the difference in distance (depth) between Q and P results in a parallax along the epipolar line k''. As already noted, in the normal case of stereo photogrammetry the parallax is purely in the x' direction and is known as x-parallax or horizontal parallax.

If the orientation parameters are known, the position of the epipolar line k'' corresponding to P' (or vice versa) can be calculated. Given an arbitrary image point P', and with projection equations 4.7, two points P and Q on the ray $\mathbf{r}'$ can be calculated for two different arbitrary values of scaling factor m. The XYZ coordinates of points P and Q can subsequently be projected into the right image using the collinearity equations (4.10). The epipolar line k'' is then defined by the straight line containing image points P'' and Q''.

Fig. 4.38 shows a convergent stereo image pair, and the images derived from them which correspond to the normal case of stereo photogrammetry (epipolar images, see section 4.3.3.5). For a point measured in the left and right images, the corresponding epipolar lines are shown. In the convergent case they run at an angle through the images and in the normal case they are run parallel on identical y image-coordinate values.
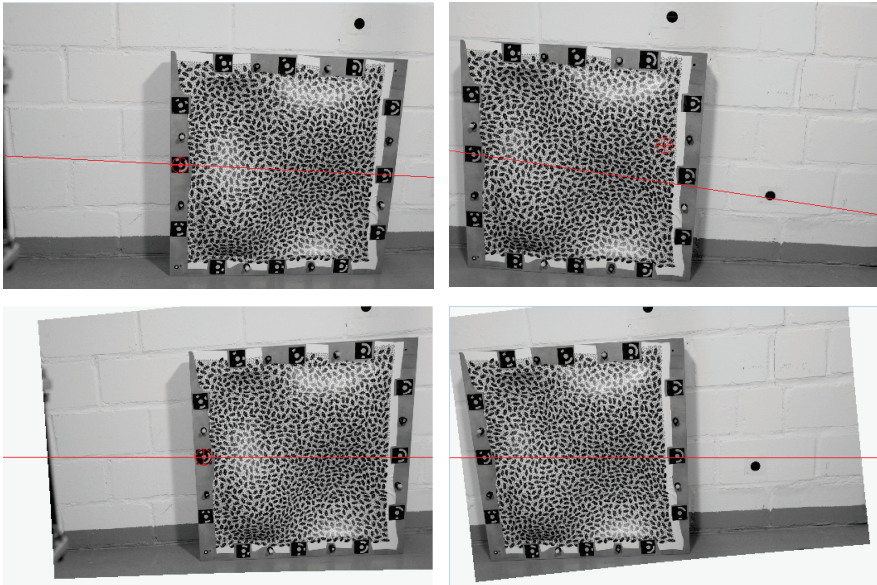
Fig. 4.38: Epipolar lines in stereo image pair; top: original images; bottom: epipolar images.

The epipolar lines can also be calculated from the parameters of relative orientation which are derived below (see section 4.3.3.4).

4.3.3 Relative orientation

Relative orientation describes the translation and rotation of one image with respect to its stereo partner in a common local model coordinate system. It is the first stage in the two-step orientation of a stereo image pair (see Fig. 4.34).
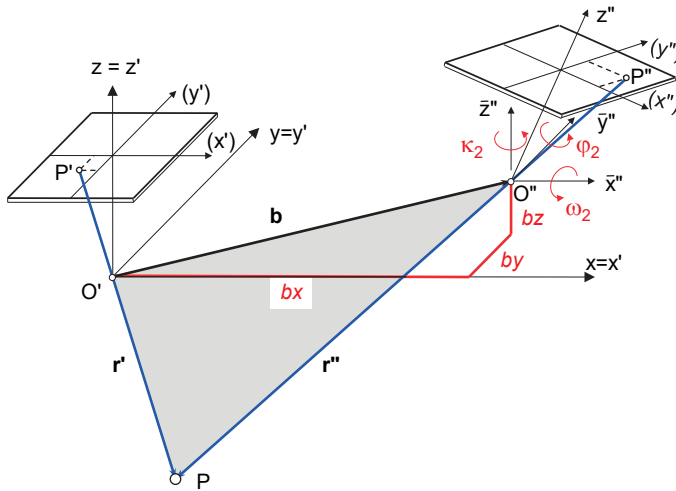
Fig. 4.39: Model coordinate system and relative orientation (left image fixed).

The numerical method of relative orientation can be easily developed for the following case. A local three-dimensional model coordinate system xyz is located in the perspective centre of the first (left) image and oriented parallel to its image coordinate system Fig. 4.39). The parameters of *exterior* orientation of the left image with respect to the model coordinate system are therefore already given:

$$
\begin{aligned}
x_{01} &= 0 & \omega_1 &= 0 \\
y_{01} &= 0 & \varphi_1 &= 0 \\
z_{01} &= 0 & \kappa_1 &= 0
\end{aligned}
\tag{4.55}
$$

Now the second (right) image is oriented in the model coordinate system by 3 translations and 3 rotations:

$$
\begin{aligned}
x_{02} &= bx \\
y_{02} &= by \\
z_{02} &= bz
\end{aligned}
\qquad
\begin{aligned}
\omega_2 \\
\varphi_2 \\
\kappa_2
\end{aligned}
\qquad\qquad (4.56)
$$

The base space vector **b** between the perspective centres O' and O" is defined by the base components bx, by and bz. It is stated in section 4.3.3.1 that the condition for correct relative orientation is that all pairs of homologous rays must be coplanar with the base. Suppose that the right-hand perspective centre is displaced along the base line towards O' and that the image is not rotated. It is clear that the homologous rays of Fig. 4.39 will still be coplanar with the base and that they will intersect in a point lying on the line between O' and P'. Consideration of similar triangles shows that the scale of the model will be directly proportional to the length of the base. That is to say, the model coordinate system can be scaled by an arbitrary factor depending on our choice of base length. One of the base components is therefore set to a constant value, commonly

$$bx = 1$$

Five independent elements by, bz and ω_2, φ_2, κ_2 therefore remain for the definition of the relative orientation.
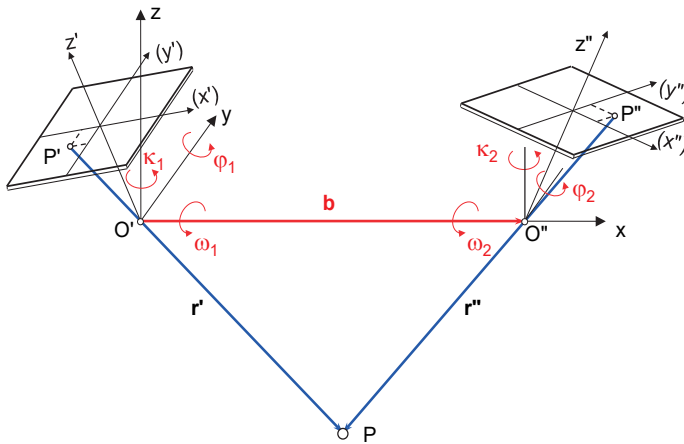
Fig. 4.40: Model coordinate system with base defining the x axis.

For an alternative formulation of the relative orientation, the x axis of the model coordinate system is defined by the stereo base and the origin of the system is located in the left-hand perspective centre (Fig. 4.40). The parameters of exterior orientation in the model coordinate system are then given by:

$$x_{01} = 0 \qquad\qquad \omega_1$$
$$y_{01} = 0 \qquad\qquad \varphi_1$$
$$z_{01} = 0 \qquad\qquad \kappa_1$$
$$x_{02} = bx \qquad\qquad \omega_2 = 0$$
$$y_{02} = 0 \qquad\qquad \varphi_2$$
$$z_{02} = 0 \qquad\qquad \kappa_2$$

The five elements to be determined are here expressed by five independent rotation angles ω_1, φ_1, κ_1 and φ_2, κ_2. Instead of ω_1 (rotation about x axis), ω_2 may be used as an alternative. The scale is again set to an arbitrary value, normally with $bx = 1$.

4.3.3.1 Coplanarity constraint

The computational solution of relative orientation utilizes the condition that an object point P and the two perspective centres O' and O" must lie in a plane (*coplanarity constraint*). This is the epipolar plane defined by vectors **b**, **r'** and **r"**, which also contains the image points P' and P".

The coplanarity constraint is only fulfilled if rays **r'** and **r"** strictly intersect in object point P, i.e. if the positions of image points P' and P", as well as the orientation parameters, are free of error. For each pair of homologous image points, one coplanarity constraint equation can be derived. Consequently, in order to calculate the five unknown orientation parameters, a minimum of five homologous points (tie points) with measured image coordinates is required. The constraint is equivalent to the minimization of y-parallaxes at all observed points P. The term y-parallax is defined by eqn. (4.64).

The coplanarity constraint can be expressed using the scalar triple product of the three vectors. They lie in a plane if the volume of the parallelepiped they define is zero:

$$\left(\mathbf{b} \times \mathbf{r'}\right) \cdot \mathbf{r"} = 0 \tag{4.57}$$

Alternatively, eqn. (4.57) can be expressed by the determinant of the following matrix. The base vector **b** is replaced by the three base components, the image vector **r'** is replaced by the image coordinates in the left image and the image vector **r"** is given by the image coordinates of the right image, transformed by the relative rotation parameters.

$$\Delta = \begin{vmatrix} 1 & x' & \bar{x}" \\ by & y' & \bar{y}" \\ bz & z' & \bar{z}" \end{vmatrix} = 0 \qquad \text{where} \qquad \begin{bmatrix} \bar{x}" \\ \bar{y}" \\ \bar{z}" \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} x" \\ y" \\ z" \end{bmatrix} \tag{4.58}$$
$$\quad\ \mathbf{b} \quad \mathbf{r'} \quad \mathbf{r"} \qquad\qquad\qquad \mathbf{r"} = \mathbf{A}_2 \cdot \mathbf{x"}$$

Here $\mathbf{A}_2$ is the rotation matrix of the right-hand image, so that the coefficients a_{11}, a_{12}...
are functions of the rotation angles $\omega_2, \varphi_2, \kappa_2$. Different principal distances for both
images may be introduced as $z' = -c_1$ and $z'' = -c_2$. For each measured tie point P_i one
observation equation can be established using (4.58).

If the calculation of relative orientation is solved with the aid of the collinearity
equations (4.10), then the 5 homologous point pairs give rise to $5 \times 2 \times 2 = 20$
observation equations. Opposite these are five unknowns of relative orientation as
well as $5 \times 3 = 15$ unknown model coordinates, so that again 5 tie points provide a
minimum solution. The determination of relative orientation using the fundamental
and essential matrices is discussed in section 4.3.4.

4.3.3.2 Calculation

The calculation of the five elements of relative orientation follows the principle of
least-squares adjustment (see section 2.4.2.2). Based on the coplanarity condition, the
following correction equation can be set up for each tie point:

$$v_\Delta = \frac{\partial \Delta}{\partial by} dby + \frac{\partial \Delta}{\partial bz} dbz + \frac{\partial \Delta}{\partial \omega_2} d\omega_2 + \frac{\partial \Delta}{\partial \varphi_2} d\varphi_2 + \frac{\partial \Delta}{\partial \kappa_2} d\kappa_2 + \Delta^0 \tag{4.59}$$

In the case of approximately parallel viewing directions, the initial values required
for linearization are as follows:

$$by^0 = bz^0 = \omega_2{}^0 = \varphi_2{}^0 = \kappa_2{}^0 = 0 \tag{4.60}$$

Δ^0 is the volume of the parallelepiped calculated from the initial values. The
differentials can again easily be computed using the following determinants:

$$\frac{\partial \Delta}{\partial by} = \begin{vmatrix} 0 & x' & \overline{x}'' \\ 1 & y' & \overline{y}'' \\ 0 & z' & \overline{z}'' \end{vmatrix} \qquad \frac{\partial \Delta}{\partial bz} = \begin{vmatrix} 0 & x' & \overline{x}'' \\ 0 & y' & \overline{y}'' \\ 1 & z' & \overline{z}'' \end{vmatrix}$$

$$\frac{\partial \Delta}{\partial \omega_2} = \begin{vmatrix} 1 & x' & 0 \\ by & y' & -\overline{z}'' \\ bz & z' & \overline{y}'' \end{vmatrix} \qquad \frac{\partial \Delta}{\partial \varphi_2} = \begin{vmatrix} 1 & x' & -\overline{y}''\sin\omega_2 + \overline{z}''\cos\omega_2 \\ by & y' & \overline{x}''\sin\omega_2 \\ bz & z' & -\overline{x}''\cos\omega_2 \end{vmatrix} \tag{4.61}$$

$$\frac{\partial \Delta}{\partial \kappa_2} = \begin{vmatrix} 1 & x' & -\overline{y}''\cos\omega_2\cos\varphi_2 - \overline{z}''\sin\omega_2\cos\varphi_2 \\ by & y' & \overline{x}''\cos\omega_2\cos\varphi_2 - \overline{z}''\sin\varphi_2 \\ bz & z' & \overline{x}''\sin\omega_2\cos\varphi_2 + \overline{y}''\sin\varphi_2 \end{vmatrix}$$

The approximate values are iteratively improved by the adjusted corrections until
there is no significant change.

Here the standard deviation of unit weight s_0 provides little information about
achieved accuracy because the volumes of the parallelepipeds are used as

observations instead of the measured image coordinates. Residuals of the estimated orientation elements result in skew intersection of the rays **r'** and **r"**, thus generating y-parallaxes in model space. It is therefore advantageous to analyse the quality of relative orientation using the calculated model coordinates.

Example 4.5:

The following data are available to calculate the relative orientation of the image pair in Fig. 4.38.

	Image 1 (left):		Image 2 (right)	
Interior orientation:	$c = -24.2236$ mm		$c = -24.2236$ mm	

Image coordinates:

P_1	$x'_1 = -0.0395$	$y'_1 = -6.3033$	$x''_1 = -8.1592$	$y''_1 = -6.1394$
P_2	$x'_2 = 6.6590$	$y'_2 = -6.2948$	$x''_2 = -1.0905$	$y''_2 = -6.5887$
P_3	$x'_3 = 9.0086$	$y'_3 = -1.3473$	$x''_3 = 1.1945$	$y''_3 = -1.4564$
P_4	$x'_4 = 7.3672$	$y'_4 = 4.5216$	$x''_4 = -0.8836$	$y''_4 = 4.8255$
P_5	$x'_5 = 0.2936$	$y'_5 = 4.7133$	$x''_5 = -8.3009$	$y''_5 = 5.2626$
P_6	$x'_6 = -2.0348$	$y'_6 = -0.7755$	$x''_6 = -10.4401$	$y''_6 = -0.2882$

Parameters of relative orientation are required (left image fixed).

Base components:	$bx = 1$	$by = -0.0634$	$bz = -0.1280$
Rotation angles:	$\omega = 1.4493°$	$\varphi = 4.1055°$	$\kappa = 2.5182°$

It can be seen that the base components in the y and z directions are significantly smaller than in the x direction and that convergence in the configuration is most apparent in the φ rotation angle about the y axis.

4.3.3.3 Model coordinates

The relationship between image and model coordinates can be expressed by the following ratios:

$$\frac{x}{x'} = \frac{y}{y'} = \frac{z}{z'} = \lambda \qquad \text{: scale factor for a particular point in left image}$$

$$\frac{x - bx}{\overline{x}''} = \frac{y_2 - by}{\overline{y}''} = \frac{z - bz}{\overline{z}''} = \mu \qquad \text{: scale factor for the same point in right image} \qquad (4.62)$$

Elimination of model coordinates gives the scale factors as:

$$\lambda = \frac{bx \cdot \overline{z}'' - bz \cdot \overline{x}''}{x'\overline{z}'' - z'\overline{x}''} \qquad \qquad \mu = \frac{bx \cdot z' - bz \cdot x'}{x' \cdot \overline{z}'' - z' \cdot \overline{x}''} \qquad (4.63)$$

and hence the model coordinates

$$x = \lambda x'$$
$$y_1 = \lambda y'$$
$$y = \frac{y_1 + y_2}{2}$$

$$z = \lambda z'$$
$$y_2 = by + \mu \overline{y}''$$
$$py = y_2 - y_1$$

(4.64)

Due to uncertainties in measurement there are two solutions for the model coordinates in the y direction, i.e. corresponding rays are skew and do not exactly intersect, which results in y-parallax py.

Additional arbitrary homologous image points can be measured in the relatively oriented model, and transformed into model coordinates xyz using equations (4.64). They describe a three-dimensional object surface, correctly shaped, but at an arbitrarily defined scale resulting from our arbitrary choice, $bx = 1$. The transformation of model coordinates into a global object coordinate system at true scale is performed by absolute orientation (see section 4.3.5). The set of equations 4.64 describe a special case of spatial intersection (see also sections 4.3.6.2, 4.4.7.1).

Example 4.6:

The homologous points in example 4.5 have the following model coordinates:

Model coordinates	x	y	z
P_1	−0.0043	−0.6802	−2.6141
P_2	0.7346	−0.6944	−2.6722
P_3	0.9520	−0.1424	−2.5598
P_4	0.7314	0.4489	−2.4050
P_5	0.0285	0.4574	−2.3507
P_6	−0.2066	−0.0787	−2.4592

The z coordinates are approximately 2.5 times the base length 1. This gives rise to an average height-to-base ratio of 2.5:1.

4.3.3.4 Calculation of epipolar lines

The equation of the epipolar line k'' in the right-hand image is given in parametric form as:

$$k'': \quad \mathbf{k}'' = \mathbf{p}'' + t(\mathbf{q}'' - \mathbf{p}'')$$

(4.65)

Here $\mathbf{k}''$ is the locus of points on the straight line through image points $\mathbf{p}''$ and $\mathbf{q}''$, which correspond to the arbitrary model points P and Q lying on the ray $\mathbf{r}'$ (Fig. 4.36). If the parameters of relative orientation (exterior orientation of both images in the model coordinate system) are inserted into the collinearity equations (4.10), the image coordinates in the right-hand image are obtained (with $z'' = -c_2$):

$$x"_i = z" \frac{r_{11}(x_i - bx) + r_{21}(y_i - by) + r_{31}(z_i - bz)}{r_{13}(x_i - bx) + r_{23}(y_i - by) + r_{33}(z_i - bz)}$$

$$y"_i = z" \frac{r_{12}(x_i - bx) + r_{22}(y_i - by) + r_{32}(z_i - bz)}{r_{13}(x_i - bx) + r_{23}(y_i - by) + r_{33}(z_i - bz)}$$

(4.66)

Here r_{ik} are the elements of the rotation matrix of $\omega_2, \varphi_2, \kappa_2$. The perspective centre O' can be used in place of point P:

$$\begin{aligned} x_P &= 0 \\ \text{P:} \quad y_p &= 0 \\ z_p &= 0 \end{aligned}$$

Point Q is given by multiplication of the image vector **x'** by an arbitrary scaling factor λ:

$$\begin{aligned} x_Q &= -\lambda x' \\ \text{Q:} \quad y_Q &= -\lambda y' \\ z_Q &= -\lambda z' \end{aligned}$$

By inserting the model coordinates into (4.66), the image coordinates of points **p"** and **q"** are obtained and hence the straight line equation of the epipolar line. Due to unavoidable measurement errors, the search for point P", the homologous point to P', should not be done along straight line $k"$ but within a narrow band either side of this line.

4.3.3.5 Calculation of normal-case images (epipolar images)

Digitized convergent stereo images can be rectified by epipolar resampling in order to correspond to the normal case of stereo photogrammetry. After rectification they are suitable for ocular stereo viewing. In addition, the epipolar lines are parallel to the x' direction, enabling simplified algorithms for stereo image matching to be applied (see section 5.5.4.1).

Fig. 4.41 illustrates the spatial position of normal-case images with respect to the stereo model. With given exterior orientations for both images, e.g. in model coordinate system xyz, three-dimensional image coordinates $x'_n, y'_n, -c_n$ in the normal-case images can be transformed using (4.10) into the image coordinates x', y' of the original image (and analogously for the second image). Rectification is performed when, for all points in the images, the colour level of the original image $g'(x', y')$ is copied to position x'_n, y'_n in the normal-case image (see section 5.3).
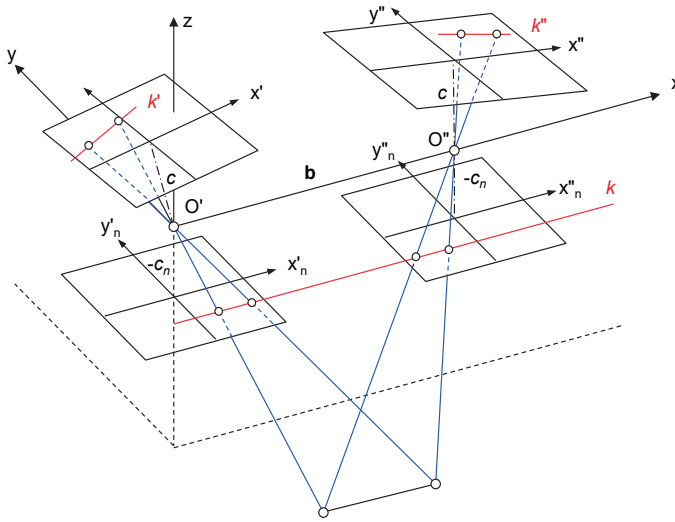
Fig. 4.41: The geometry of normal-case stereo images.

Fig. 4.38 shows a strongly convergent stereo image pair and the distortion-free normal-case stereo images derived from them. Homologous points then lie on the same y coordinate values. The principal points of the normal-case images are exactly at the image centres. The large areas with no image information are the result of the convergence of the original images and could be removed.
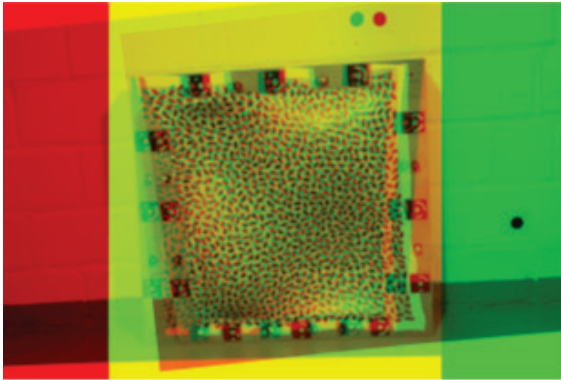
Fig. 4.42: Anaglyph image derived from normal-case stereo images.

Fig. 4.42 shows an anaglyph stereo image which has been created from the normal-case image pair of Fig. 4.38. Note that the x-parallaxes of homologous points vary with object distance.

4.3.3.6 Quality of relative orientation

The existence of y-parallax *py* (defined in 4.64) at a point in the model indicates failure of homologous rays to intersect at that point. The y-parallaxes, considered over the whole model, may be used as a measure of the quality of relative orientation; y-parallax at photo scale gives a normalized figure.

If the y-parallax at a point, *i*, in the model is py_i, then the y-parallax at photo scale may be taken as

$$py'_i = \frac{z'}{z_i} py_i \tag{4.67}$$

Assuming that *by* and *bz* are small compared to *bx*, the following expression, based on a number of tie points, *n*, gives a measure of the quality of the relative orientation:

$$s_{py'} = \frac{1}{n} \sqrt{\sum_{i=1}^{n} py'^2_i} \tag{4.68}$$

The intersection angle α of homologous image rays is the angle between the two spatial vectors **r'** and **r"** where:

$$\cos\alpha = \frac{\mathbf{r'}^T \cdot \mathbf{r''}}{|\mathbf{r'}| \cdot |\mathbf{r''}|} \tag{4.69}$$

Taking all *n* tie points into account, the mean intersection angle can be calculated:

$$\bar{\alpha} = \frac{1}{n} \sum_n \alpha \tag{4.70}$$
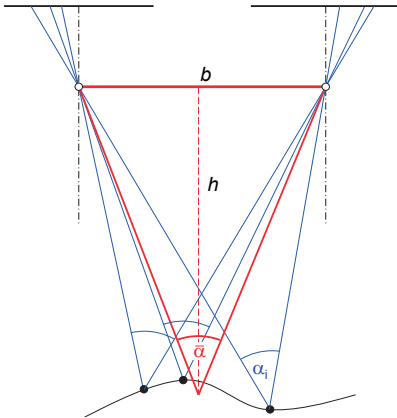

Fig. 4.43: Mean intersection angle.

The mean intersection angle also approximately describes the ratio of the stereo base *b* to the mean object distance *h*, as given by (Fig. 4.43):

$$\tan\frac{\overline{\alpha}}{2} \approx \frac{b}{2h} \tag{4.71}$$

Accuracy of relative orientation and point determination will be optimized when the mean intersection angle is close to a right angle.

The quality of relative orientation depends on the following criteria:

- Accuracy of image coordinates:
  The accuracy of image coordinates depends partly on the measuring accuracy of the instrument and partly on the ability to identify matching points in both images. Image patches with poor structure can be matched less accurately than areas with a significant grey level structure, regardless of the image processing method (visual interpretation or digital image matching).

- Number and distribution of tie points in model space:
  Tie points should be chosen in model space to ensure a robust geometric link between both images. A point distribution as recommended by von Gruber is particularly suitable. This has a tie point in each corner of the model space and one in the middle of each long side. This distribution is strictly possible only in the normal case (Fig. 4.44).

  If the object structure, for example containing large homogeneous areas, does not allow an optimum distribution of homologous points, model errors, which cannot be controlled, may occur in the regions not covered. If all tie points lie on a common straight line, the resulting normal system of equations becomes singular.

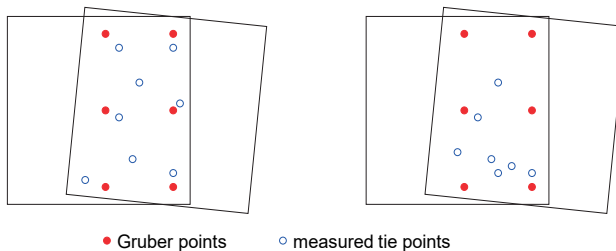● Gruber points   ○ measured tie points

Fig. 4.44: Good and bad distribution of tie points in model space.

To properly control the relative orientation, at least 8–10 well distributed tie points should be measured.

- Height-to-base ratio:
  If the base of the stereo model is small relative to object distance (height) then ray intersection angles are poor. The parameters of relative orientation are then

determined with greater uncertainty. As mentioned above, an optimal configuration is achieved with intersection angles of around 90 degrees.

– Distribution of tie points in object space:

There are a few exceptional cases where singular or weakly conditioned normal equations occur, even though there is a good point distribution in model space. Amongst other cases this applies to the danger cylinder, where the object points used as tie points and the perspective centres of both images lie on a common cylindrical surface (Fig. 4.45). This effect can also occur where object surfaces have small curvatures and the imaging lens has a long focal length. The same problem exists for the space resection, if the image to be oriented is also located on a danger surface.

The result of this configuration is that the cameras do not have unique positions on the cylindrical surface because at different camera positions the tie points subtend the same angles and so have the same image positions (a fundamental property of circles).
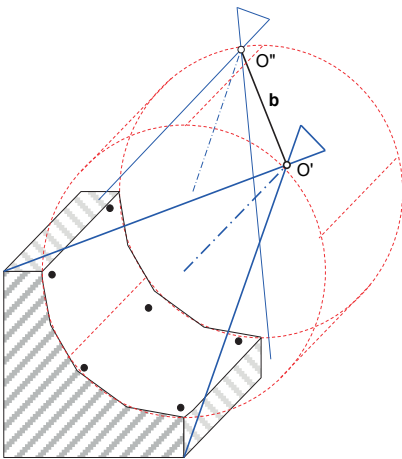
Fig. 4.45: Danger cylinder above a curved surface.

4.3.3.7 Special cases of relative orientation

The method of relative orientation is widely used for traditional stereo image analysis on analytical stereo instruments. These applications typically have parallel imaging directions which permit stereoscopic viewing (as shown in Fig. 4.37). The orientation elements are then relatively small so that initial values for the iterative adjustment can be zero.

Close range photogrammetry, in contrast, often involves arbitrary, convergent, multi-image configurations. In this case relative orientation is not used as the actual orientation method but only as one step in the calculation of approximate values for

the subsequent bundle adjustment (see section 4.4.4.1). Here the image pairs may have orientation values that differ significantly from those of the normal case (examples in Fig. 4.46).
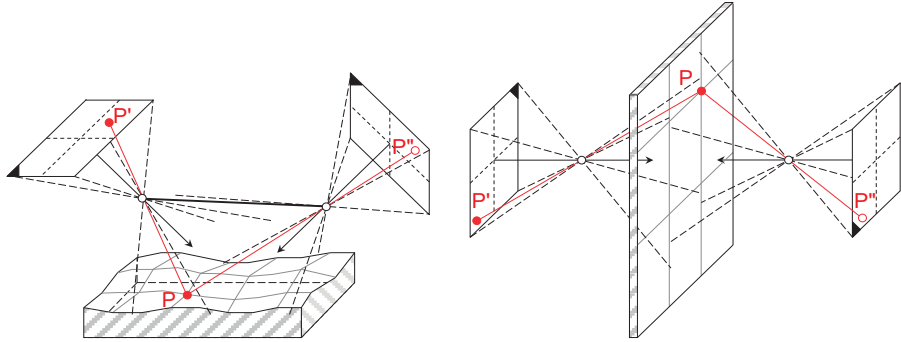
Fig. 4.46: Convergent image pair configurations.

Fig. 4.47 shows the distribution in image space of 12 homologous points in an image pair, in which the right-hand image has significant tilts with respect to the left-hand image.
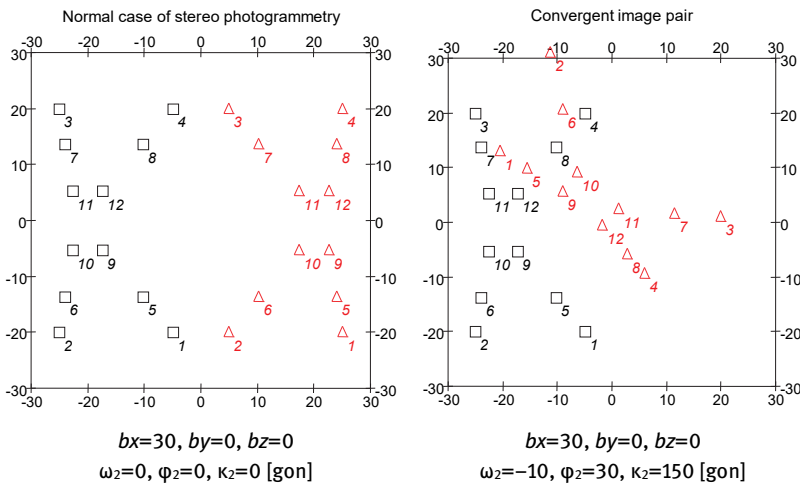
Fig. 4.47: Image point distribution for different orientations; □ left image, △ right image.

It is not possible to define approximate values for the relative orientation of arbitrarily oriented images according to eqn. (4.60). Instead the methods of spatial similarity transformation (see section 2.2.4) or space resection (see section 4.2.3) can be applied.

For these methods, tie points require 3D object coordinates which may be calculated by transforming image coordinates in both images by an approximate scale factor. Alternatively, linear orientation methods allow calculation without approximate values (see section 4.3.4).

In multi-image applications it is possible that two or more images are exposed at the same point but with different orientations. It is also possible that images are located behind one another on the same viewing axis. In these and similar cases, both images cover a common model space which does not provide distinct intersections at object points (Fig. 4.48). The calculation of relative orientation then leads to poor results or fails completely. Such images can, of course, be included with others in a multi-image bundle adjustment.

A further special case occurs for relative orientation using images of strictly planar surfaces. This happens often in close-range applications, for example in the measurement of flat façades or building interiors. In this case only 4 tie points are required because both bundles of rays can be related to each other by a projective transformation with 8 parameters. In order to solve the adjustment problem, the planarity of the object surface can be handled by an additional constraint equation. This constraint can replace one of the required coplanarity conditions so that only four tie points are necessary.
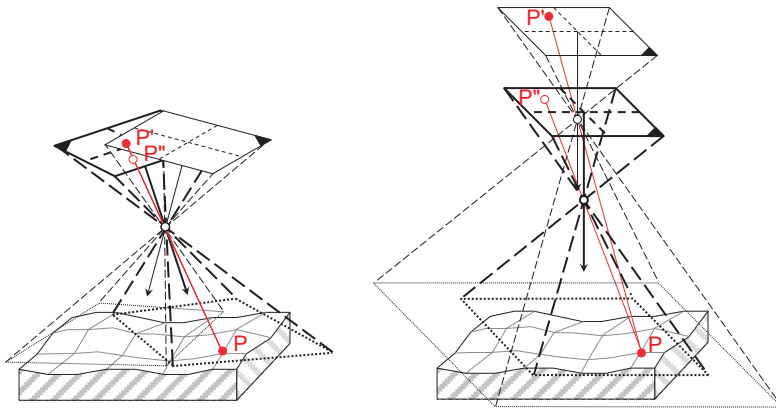
Fig. 4.48: Overlapping image pairs with insufficient spatial ray intersections.

4.3.4 Fundamental matrix and essential matrix

The relationship between two images can also be derived with the aid of projective geometry. According to eqn. (4.32), a projection matrix can be defined for each image:

$$\mathbf{P}_1 = \mathbf{K}_1 \cdot \mathbf{R}_1 \cdot \begin{bmatrix} \mathbf{I} & | & \mathbf{0} \end{bmatrix} \qquad\qquad \mathbf{P}_2 = \mathbf{K}_2 \cdot \mathbf{R}_2 \cdot \begin{bmatrix} \mathbf{I} & | & \textbf{-b} \end{bmatrix} \qquad\qquad (4.72)$$

where

R₁, R₂: rotation matrices
K₁, K₂: calibration matrices
b: base vector

For relative orientation (fixed left-hand image) $\mathbf{R}_1 = \mathbf{I}$. The image coordinates of homologous points in both images can be transformed into a local camera coordinate system which initially has an arbitrary principal distance $c_1 = c_2 = 1$.

$$x_1 = \mathbf{R}_1^{-1} \cdot \mathbf{K}_1^{-1} \cdot x' \qquad\qquad x_2 = \mathbf{R}_2^{-1} \cdot \mathbf{K}_2^{-1} \cdot x'' \tag{4.73}$$

The transformed image coordinates now lie in the same coordinate system as the base vector b and can be used in the following coplanarity condition (compare with eqn. 4.57):

$$x_1 \cdot (b \times x_2) = x_1^T \cdot \mathbf{S}_b \cdot x_2 = 0 \tag{4.74}$$

$$\text{where } \mathbf{S}_b = \begin{bmatrix} 0 & -bz & by \\ bz & 0 & -bx \\ -by & bx & 0 \end{bmatrix}$$

The vector product of vectors b and x_2 can therefore be expressed using the skew symmetric matrix $\mathbf{S}_b$. Insertion of (4.73) in (4.74) leads to the linearized condition:

$$x_1^T \cdot \mathbf{S}_b \cdot x_2 = x'^T \cdot (\mathbf{K}_1^{-1})^T \cdot \mathbf{S}_b \cdot \mathbf{R}_2^{-1} \cdot \mathbf{K}_2^{-1} \cdot x'' = 0 \tag{4.75}$$

Using

$$\mathbf{F} = (\mathbf{K}_1^{-1})^T \cdot \mathbf{S}_b \cdot \mathbf{R}_2^{-1} \cdot \mathbf{K}_2^{-1} \qquad\qquad \text{: fundamental matrix} \tag{4.76}$$

the coplanarity condition for homologous image points can be expressed in the simple form:

$$x'^T \cdot \mathbf{F} \cdot x'' = 0 \qquad\qquad \text{: coplanarity condition} \tag{4.77}$$

The fundamental matrix $\mathbf{F}$ is a homogeneous 3x3 matrix, i.e. multiplication by a scalar does not alter the projection. As a consequence, it can be described by $9-1 = 8$ degrees of freedom. $\mathbf{F}$ contains all necessary relative orientation data, including the parameters of interior orientation. Using the fundamental matrix, the epipolar line in the partner image can be calculated:

$$x'^T \cdot \mathbf{F} = k'' = \begin{bmatrix} A_{k''} \\ B_{k''} \\ C_{k''} \end{bmatrix} \qquad\qquad \text{: epipolar line in right image} \tag{4.78}$$

Here $A_{k''}, B_{k''}, C_{k''}$ correspond to the parameters of a straight line defined according to eqn. (2.90).

If the interior orientation of both images is known (known calibration matrix $\mathbf{K}$), the fundamental matrix reduces to the *essential matrix* $\mathbf{E}$:

$$\boldsymbol{x'}_k^{T} \cdot \mathbf{S}_b \cdot \mathbf{R}_2^{-1} \cdot \boldsymbol{x''}_k^{T} = \boldsymbol{x'}_k^{T} \cdot \mathbf{E} \cdot \boldsymbol{x''}_k^{T} = 0 \qquad (4.79)$$

where

$$\boldsymbol{x'}_k = \mathbf{K}_1^{-1} \cdot \boldsymbol{x'} \qquad\qquad \boldsymbol{x''}_k = \mathbf{K}_2^{-1} \cdot \boldsymbol{x''} \qquad\qquad \mathbf{E} = \mathbf{S}_b \cdot \mathbf{R}_2^{-1}$$

At least eight homologous points are required to calculate the fundamental matrix $\mathbf{F}$. In comparison only five are required to solve for the essential matrix $\mathbf{E}$. The linear system of equations in each case can be solved, for example, using the singular value decomposition.

4.3.5 Absolute orientation

4.3.5.1 Mathematical model

Absolute orientation describes the transformation of the local model coordinate system xyz, resulting from a relative orientation with arbitrary position, rotation and scale, into the object coordinate system XYZ via reference points. Reference points are object points measured in the model coordinate system which have one or more known coordinate components in object space, e.g. XYZ, XY only or Z only. The reference points can be identical to the tie points already used for relative orientation, or they can be measured subsequently as model points in the relatively oriented model.

Absolute orientation consists of a spatial similarity transformation with three translations, three rotations and one scaling factor as described in eqn. 2.65 (see section 2.2.4.1). In order to solve the system of equations, a minimum of seven suitable point elements are required, for example taken from three spatially distributed XYZ reference points.

Fig. 4.49 illustrates the transformation of the model coordinate system, xyz with origin at M, into the object coordinate system, XYZ. The coordinates of M in the XYZ system are $\mathbf{X_M}$. The rotation matrix $\mathbf{R}$ is a function of the three rotation angles ξ, η, ζ about the axes XYZ. The transformation for a model point with coordinates *xyz* (vector $\mathbf{x}$) is given by:

$$\mathbf{X} = f(X_M, Y_M, Z_M, m, \xi, \eta, \zeta, x, y, z)$$
$$= \mathbf{X_M} + m \cdot \mathbf{R} \cdot \mathbf{x}$$

(4.80)

or

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_M \\ Y_M \\ Z_M \end{bmatrix} + m \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

where m is the scale factor between model and object coordinates.
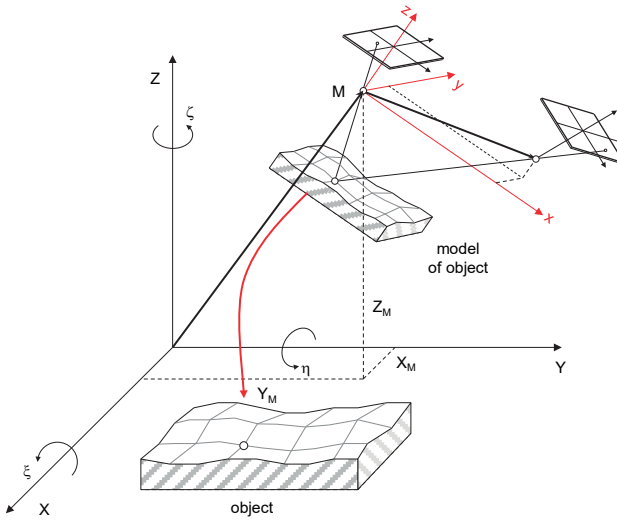
Fig. 4.49: Absolute orientation.

Equations 4.80 are non-linear and are solved in the usual way; if there is redundancy the solution will be based on a least-squares adjustment (see section 2.4.2.2) in which each coordinate component of a reference point provides one linearized correction equation:

$$X + v_X = dX_M + \frac{\partial F}{\partial m} dm + \frac{\partial F}{\partial \xi} d\xi + \frac{\partial F}{\partial \eta} d\eta + \frac{\partial F}{\partial \zeta} d\zeta + X^0$$

$$Y + v_Y = dY_M + \frac{\partial F}{\partial m} dm + \frac{\partial F}{\partial \xi} d\xi + \frac{\partial F}{\partial \eta} d\eta + \frac{\partial F}{\partial \zeta} d\zeta + Y^0$$

$$Z + v_Z = dZ_M + \frac{\partial F}{\partial m} dm + \frac{\partial F}{\partial \xi} d\xi + \frac{\partial F}{\partial \eta} d\eta + \frac{\partial F}{\partial \zeta} d\zeta + Z^0$$

(4.81)

4.3.5.2 Definition of the datum

In many close-range applications, 3D reference points (full reference points) are available. Each reference point therefore provides three correction equations. In aerial photogrammetry it is possible that some reference points only have known plan position (XY) and others only height (Z), resulting in a reduced set of correction equations. In order to solve the absolute orientation, at least 2 X coordinates, 2 Y coordinates and 3 Z coordinates must be available (see also section 4.4.3.2).

The reference points should be well distributed over the object space to be transformed. If all reference points lie on a common straight line, a singular or weak system of equations results. In over-determined configurations of control points, inconsistencies between coordinates (network strain) can distort the transformation parameters and give rise to higher standard deviations in the transformation.

4.3.5.3 Calculation of exterior orientations

From the parameters of relative and absolute orientation for an image pair, the exterior orientation parameters of each image can be calculated.

The position of the perspective centre, $\mathbf{X}_{0_i}$, of an image i is derived from the origin of the model coordinate system $\mathbf{X_M}$ and the transformed components of the base $\mathbf{b}$:

$$\mathbf{X}_{0_i} = \mathbf{X_M} + m \cdot \mathbf{R}(\xi\eta\zeta) \cdot \mathbf{b}_i \tag{4.82}$$

For the left-hand image ($i = 1$) the base components are zero, hence $\mathbf{X}_{0_i} = \mathbf{X_M}$.

In order to calculate the rotation matrix of image i, $\mathbf{R}_i(\omega\varphi\kappa)$, with respect to the object system, the rotation matrix $\mathbf{A}_i(\omega\varphi\kappa)$ of the relative orientation is pre-multiplied by the rotation matrix $\mathbf{R}(\xi\eta\zeta)$ of the absolute orientation:

$$\mathbf{R}_i(\omega\varphi\kappa) = \mathbf{R}(\xi\eta\zeta) \cdot \mathbf{A}_i(\omega\varphi\kappa) \tag{4.83}$$

After absolute orientation, object coordinates are available for the model points. As an alternative, therefore, the parameters of exterior orientation can also be determined by space resection using the transformed model coordinates in object space.

4.3.5.4 Calculation of relative orientation from exterior orientations

If the parameters of exterior orientation, in the object coordinate system, are available for both images of the stereo pair, then the parameters of relative orientation can be derived.

For relative orientation with origin in the left-hand image, rotation matrices of relative orientation $\mathbf{A}_i$ are obtained by multiplication of the exterior orientation matrices $\mathbf{R}_i$ by the inverse rotation of the left image:

$$\mathbf{A}_1 = \mathbf{R}_1^{-1} \cdot \mathbf{R}_1 = \mathbf{I} \qquad\qquad \mathbf{A}_2 = \mathbf{R}_1^{-1} \cdot \mathbf{R}_2 \qquad\qquad (4.84)$$

The required base components in the model coordinate system are obtained from the vector between the two perspective centres, transformed by the inverse rotation of the left-hand image. The resulting vector **B** is then scaled to create a base vector **b** where $bx = 1$.

$$\mathbf{X}_{0_1} = \mathbf{0} \qquad \mathbf{B} = \mathbf{R}_1^{-1} \cdot (\mathbf{X}_2 - \mathbf{X}_1) = \begin{bmatrix} Bx & By & Bz \end{bmatrix}^T$$

$$\mathbf{b} = \frac{1}{Bx} \cdot \mathbf{B} = \begin{bmatrix} 1 & by & bz \end{bmatrix}^T \qquad\qquad (4.85)$$

Example 4.7:

The exterior orientation parameters of the stereo pair in Fig. 4.38 are given as follows:

	Image 1 (left):	Image 2 (right)
Translation:	$X_{01} = -471.890$ mm	$X_{02} = -78.425$ mm
	$Y_{01} = 11.030$ mm	$Y_{02} = -12.630$ mm
	$Z_{01} = 931.070$ mm	$Z_{02} = 916.553$ mm
Rotation:	$\omega_1 = -13.0592°$	$\omega_2 = -11.6501°$
	$\varphi_1 = -4.4402°$	$\varphi_2 = -0.3134°$
	$\kappa_1 = 0.7778°$	$\kappa_2 = 3.4051°$

Solution for relative orientation:

Rotations are calculated using eqn. (4.84):

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{A}_2 = \begin{bmatrix} 0.99647 & -0.04378 & 0.07160 \\ 0.04572 & 0.99862 & -0.02556 \\ -0.07038 & 0.02874 & 0.99710 \end{bmatrix}$$

Rotation angles: $\omega = 1.4513°$ $\varphi = 4.1059°$ $\kappa = 2.5162°$

Base components are calculated as:

$$\mathbf{B} = \begin{bmatrix} 390.471 \\ -25.071 \\ -49.890 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 1 \\ -0.0642 \\ -0.1277 \end{bmatrix}$$

The result is largely identical to the relative orientation calculated in example 4.5. Any small differences are due to the fact that the parameters of exterior orientation (used here) originate in a bundle adjustment composed of 16 images.

4.3.6 Stereoscopic processing

4.3.6.1 Principle of stereo image processing

Stereo processing covers all visual or computational methods for the processing of a stereo image pair. Traditionally, it has greatest application in (interactive) aerial

photogrammetry. However, it is also an elementary method for the automatic calculation of dense surface models.

In close-range work, stereo photogrammetry is used in the following example applications:

– Visual processing of natural features:

The reconstructed object is measured in a stereo plotting instrument using binocular, stereoscopic optical viewing systems. The operator observes an optically or digitally generated "floating mark", the apparent spatial position of which is under his or her control. A measurement is taken when the floating mark appears to lie on the virtual surface of the object; the point measured corresponds to the pair of homologous image points simultaneously viewed stereoscopically. The movement of the floating mark on to the surface is controlled interactively by the operator. Single points may be measured or continuous lines may also be measured as the operator moves the floating mark over the virtual surface.

– Visual or digital reconstruction of free-form surfaces:

Object surfaces of arbitrary shape can be evaluated by stereo photogrammetry if the surface structure permits the identification (matching) of homologous points. Surfaces with insufficient visual pattern or structure must therefore be prepared with a suitable texture, e.g. by pattern projection or other method. Image processing is performed either by the visual method above, or by image processing algorithms which implement stereo image matching of corresponding points (see section 5.5.4). The final goal is the complete 3D reconstruction of the free-form surface, for example as a digital surface model or a dense point cloud.

– Image acquisition with stereometric cameras:

Stereometric cameras (see section 3.5.4) are usually configured to correspond to the normal case of stereo photogrammetry. They provide a simple method of imaging and of photogrammetric object reconstruction which, to a large extent, avoids complicated orientation procedures.

– Point-by-point (tactile) object measurement with online dual camera systems:

Online photogrammetric systems comprising two digital metric cameras can be treated as stereo systems, although they can easily be extended to incorporate more than two cameras. The object is measured by spatial intersection of targeted points (targets, probes) which can be detected and identified automatically. If the exposure of both cameras is synchronized, the object can be measured by hand-held contact probes (see section 6.5.3). Furthermore, it is then possible to capture moving objects in full 3D.

– Control of vision-based machines, e.g. autonomous robots:

There are a number of applications in computer vision (stereo vision, shape from stereo) where a scene is analysed by stereo-based algorithms which reflect the mechanisms of natural human vision. Examples are the control of autonomous robots in unknown environments (avoidance of collisions, see example in Fig. 6.39) and the control of production tools.

In summary, the principle of stereo processing is based on the correspondence of homologous points lying in an epipolar plane. The epipolar plane intersects the image planes in epipolar lines (see section 4.3.2). For the normal case of stereo photogrammetry (Fig. 4.35) the epipolar lines are parallel and depth information can be determined by measuring the x-parallax px'.

4.3.6.2 Point determination using image coordinates

4.3.6.2.1 Coordinate calculation in normal case

The normal case of stereo photogrammetry is, in fact, the special case in which two identical cameras have parallel axes pointing in the same direction at right angles to the stereo base. With respect to an XYZ coordinate system located in the left perspective centre, object coordinates can be derived from the ratios indicated in Fig. 4.50:

Parallel to the image plane:

$$X = \frac{h}{c} \cdot x' = m \cdot x' \qquad\qquad Y = \frac{h}{c} \cdot y' = m \cdot y' \qquad (4.86)$$

In the viewing direction:

$$\frac{h}{c} = \frac{b}{x' - x''} = m$$

and it follows that:

$$Z = h = \frac{b \cdot c}{x' - x''} = \frac{b \cdot c}{px'} \qquad (4.87)$$

(4.86) and (4.87) can also be derived from the collinearity equations (4.10). The rotation angles of both images are zero. The right image is shifted in the X direction by the base length b with respect to the left image. Hence it follows that:

$$x'_{01} = y'_{01} = x'_{02} = y'_{02} = 0$$
$$X_{01} = Y_{01} = Z_{01} = Y_{02} = Z_{02} = 0 \qquad\qquad X_{02} = b$$
$$\omega_1 = \varphi_1 = \kappa_1 = \omega_2 = \varphi_2 = \kappa_2 = 0 \qquad\qquad \mathbf{R}_1 = \mathbf{R}_2 = \mathbf{I}$$

The x-parallax (disparity) px' is measured either by visual examination and coordinate measurement in a stereo plotter (stereo comparator), or by methods of digital image matching. As an example, Fig. 4.51 shows the measurement of two object points. Point P_1 on the manhole cover closest to the cameras has a much larger x-parallax than the more distant point P_2 at the top of the tower. The y' image coordinates are almost equal, i.e. y-parallax does not exist.
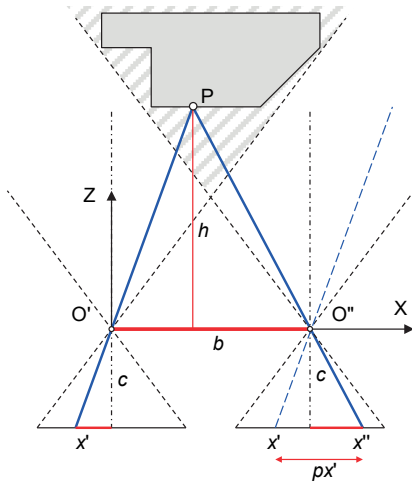
Fig. 4.50: Normal case of stereo photogrammetry.

Example 4.8:

The following image coordinates have been measured for the example image pair above (stereo camera SMK 120, $b = 1.2$ m, $c = -60.2$ mm). The object coordinates of the two points are to be calculated.

1. Point P_1 $x' = -3.924$ mm $x'' = -23.704$ mm
                  $y' = -29.586$ mm $y'' = -29.590$ mm
                  $px' = x'-x'' = 19.780$ mm

2. Point P_2 $x' = 7.955$ mm $x'' = 6.642$ mm
                  $y' = 45.782$ mm $y'' = 45.780$ mm
                  $px' = x'-x'' = 1.313$ mm

3. Z coordinate (distance) $Z = \dfrac{b \cdot c}{px'}$ $Z_1 = -3.65$ m
                                                              $Z_2 = -55.02$ m

4. XY coordinates $X_1 = -0.24$ m $X_2 = 7.27$ m
                     $Y_1 = -1.79$ m $Y_2 = 41.84$ m
                     $m = 61$ $m = 914$

The result shows that point P_1 lies beneath the left-hand camera at a slope distance of 4.08 m. Point P_2 is located to the right and above the left camera at a slope distance of 69.23 m and with a larger image scale number.
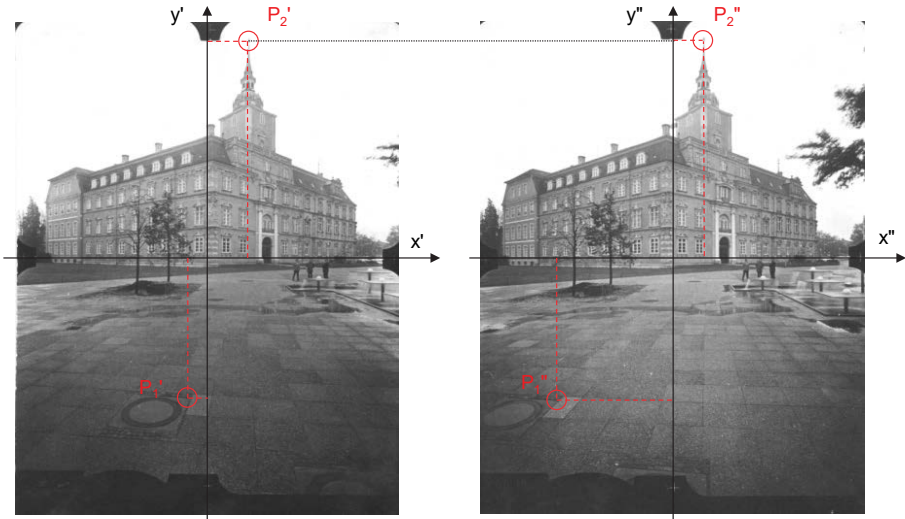
Fig. 4.51: Measurement of two object points in a stereo image pair.

4.3.6.2.2 Accuracy

Differentiation of eqn. (4.87), and application of error propagation, gives the following accuracy estimation of the object coordinate in viewing direction Z (c and b are assumed to be free of error).

$$s_Z = \frac{Z^2}{b \cdot c} s_{px'} = \frac{h}{b} \cdot \frac{h}{c} s_{px'} = q \cdot m \cdot s_{px'} \tag{4.88}$$

The equation shows that the accuracy in the viewing direction is a function of the accuracies of parallax measurement, image scale $m = h/c$ and height-to-base ratio h/b, which corresponds to the design factor q introduced in section 3.3.1.2. The equation also shows that, since b and c are constant for any particular case, the accuracy falls off in proportion to the square of the distance, Z. The height-to-base ratio, or more correctly distance-to-base ratio, describes the intersection geometry. If the base is small compared with the distance, ray intersection is weak and accuracy in the viewing direction is poor.

In general, parallax measurement accuracy can be estimated as

$$s_{px'} = \frac{s_{x'}}{\sqrt{2}} \tag{4.89}$$

i.e. it is slightly more accurate than a single measured image coordinate. In monoscopic measurement, error propagation for $px' = x' - x''$ gives:

$$s_{px'} = s_{x'} \sqrt{2} \tag{4.90}$$

The accuracy of the X and Y coordinates can be similarly derived from eqn. (4.86):

$$s_X = \sqrt{\left(\frac{x'}{c}s_Z\right)^2 + \left(\frac{Z}{c}s_{x'}\right)^2} = \sqrt{\left(\frac{x'}{c}q \cdot m \cdot s_{px'}\right)^2 + (m \cdot s_{x'})^2}$$

$$s_Y = \sqrt{\left(\frac{y'}{c}s_Z\right)^2 + \left(\frac{Z}{c}s_{y'}\right)^2} = \sqrt{\left(\frac{y'}{c}q \cdot m \cdot s_{px'}\right)^2 + (m \cdot s_{y'})^2}$$

(4.91)

In (4.91) the dominant term is the second summand of the square root. Object accuracy can therefore usually be estimated as

$$s_X = s_Y = m \cdot s_{x'y'}$$

(4.92)

The following object accuracies result from the configuration in example 4.8:

Example 4.9:

Assume image coordinates are measured to an accuracy of 10 μm, resulting in a parallax accuracy of 7 μm. Principal distance and base length are assumed free of error.

Object accuracies:

1. Point P_1: $s_X = 0.6$ mm $s_Y = 0.6$ mm $s_Z = 1.4$ mm $q = h/b = 3.2$

2. Point P_2: $s_X = 9.1$ mm $s_Y = 9.1$ mm $s_Z = 315$ mm $q = h/b = 48$

It is clear that the accuracy of point P_2 is significantly decreased in Z and coordinate determination of this point is not practicable.
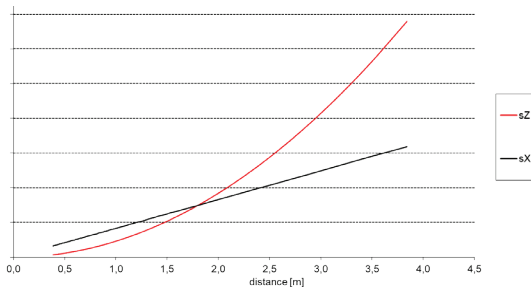

Fig. 4.52: Object accuracy for the normal case of stereo photogrammetry (from example 4.8).

Fig. 4.52 shows the object accuracies for additional point distances taken from example 4.8. For object distances less than about 1.7 m ($q = 1.4 = 1{:}0.7$), the accuracy s_Z in the viewing direction is higher than in the other directions. At longer distances the increase in uncertainty s_Z is quadratic whilst the increase in s_X remains linear.

4.3.6.2.3 Intersection in the general stereo case

A stereo pair which is not configured according to the strictly normal case has orientation parameter values which are not equal to zero. In addition, the images can have arbitrary parameters of interior orientation.
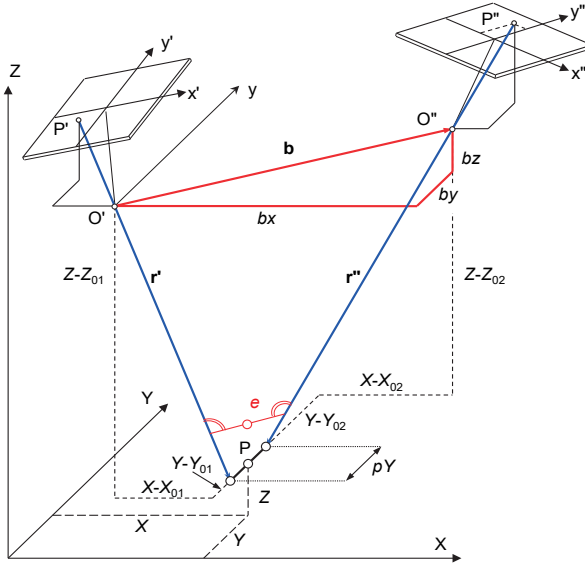

Fig. 4.53: Spatial intersection for the general stereo case.

Object coordinates *XYZ* can be calculated by spatial intersection of the rays **r'** and **r"** if the parameters of interior and exterior orientation of both images are known (Fig. 4.53). Both spatial rays are defined by the measured image coordinates, transformed by the orientation parameters. For the special case of a stereo pair, the spatial intersection can be calculated as follows (see also the calculation of model coordinates, section 4.3.3.3):

1. Transformation of image coordinates:

$$
\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} X_{01} \\ Y_{01} \\ Z_{01} \end{bmatrix} + \mathbf{R}_1 \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \qquad \begin{bmatrix} X'' \\ Y'' \\ Z'' \end{bmatrix} = \begin{bmatrix} X_{02} \\ Y_{02} \\ Z_{02} \end{bmatrix} + \mathbf{R}_2 \begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix} \tag{4.93}
$$

2. Stereo base components:

$$
bx = X_{02} - X_{01} \qquad\qquad by = Y_{02} - Y_{01} \qquad\qquad bz = Z_{02} - Z_{01} \tag{4.94}
$$

For the simple version of the intersection, the skew rays intersect the XY plane at elevation Z of object point P, giving rise to two possible solutions (Fig. 4.53), i.e.:

$$X = X_1 = X_2 \qquad\qquad Z = Z_1 = Z_2 \qquad\qquad Y = \frac{Y_1 + Y_2}{2} \qquad\qquad (4.95)$$

3. Scale factors:

The scale factors for the transformation of image coordinates are:

$$\lambda = \frac{bx(Z''-Z_{02}) - bz(X''-X_{02})}{(X'-X_{01})(Z''-Z_{02}) - (X''-X_{02})(Z'-Z_{01})}$$
$$\mu = \frac{bx(Z'-Z_{01}) - bz(X'-X_{01})}{(X'-X_{01})(Z''-Z_{02}) - (X''-X_{02})(Z'-Z_{01})} \qquad\qquad (4.96)$$

4. Object coordinates:

$$X = X_{01} + \lambda(X'-X_{01}) \qquad\qquad Y_1 = Y_{01} + \lambda(Y'-Y_{01})$$
$$Z = Z_{01} + \lambda(Z'-Z_{01}) \qquad\qquad Y_2 = Y_{02} + \mu(Y''-Y_{02}) \qquad\qquad (4.97)$$
$$Y = (Y_1 + Y_2)/2 \qquad\qquad pY = Y_2 - Y_1$$

Here the Y-parallax in object space pY is a quality measure for coordinate determination. It is zero when the two rays exactly intersect. However, pY may be zero if image measuring errors occur in the direction of epipolar lines.

The solution is not completely rigorous but works in most cases where the base is approximately aligned with the X direction and where bx is large in comparison with by and bz.

Example 4.10:

For the stereo pair in Fig. 4.38 there are image coordinates of homologous points from example 4.5, as well as parameters for interior and exterior orientation in the object coordinate system of example 4.7. Object point coordinates computed from a spatial intersection are required.

Solution according to eqn. (4.97):

Object coordinates:	X [mm]	Y [mm]	Z [mm]	pY [mm]
P1	−390.9432	−477.5426	0.0168	0.0043
P2	−101.5422	−479.1967	0.1027	0.0069
P3	−23.2276	−256.8409	−0.0839	−0.0065
P4	−116.8842	−21.0439	0.0190	0.0138
P5	−392.1723	−21.4735	0.0974	0.0140
P6	−477.5422	−237.6566	−0.1844	−0.0055

Separate evaluation of the stereo pair shows measurement noise in the object space of the order of 10 μm to 20 μm. Taking the object-space y-parallaxes pY as a measure of the quality of the spatial intersection, it can be seen that these lie within the measurement noise.

In the general case of two images with arbitrary orientations, point P is calculated as the midpoint of the shortest distance *e* between both rays (Fig. 4.53; for calculation see section 2.3.2.1). The spatial intersection can also be expressed as an over-determined adjustment problem based on the collinearity equations. In this form it can be extended to more than two images (see section 4.4.7.1).

4.3.6.3 Point determination with floating mark

4.3.6.3.1 Setting a floating mark onto the surface

The term *floating mark* is used here for a digitally generated stereoscopic mark that can be moved through the virtual 3D space of the stereo model. The floating mark is set onto the object surface in order to measure a surface point. Although the floating mark primarily serves the interactive and visual analysis of the stereo image, its principle is also useful for automatic, digital stereo measurement.

The numerical reconstruction of homologous rays is performed in digital stereoplotters. Using separate optical paths, the operator observes two floating marks which fuse into one common mark if set correctly onto the object surface. When the floating mark appears to touch the surface the corresponding XYZ coordinates are recorded.

The XYZ coordinates which correspond to a spatially controlled floating mark can be transformed into image coordinates using equations (4.10). Fig. 4.54 shows how the transformed marks only identify corresponding image patches (homologous points) if the XYZ coordinates represent a point P on the object surface. A measuring position below P at point R, as well a position above P at point Q, result in image points which do not correspond, namely the imaging positions corresponding to the intersections of the non-homologous rays with the object surface. The correct position of the floating mark is controlled either by a visual check or by a digital corre-spondence algorithm applied to the two calculated image positions. The mechanical effort is reduced to a separate real-time shift of both image planes. Starting from an approximate position, the XYZ coordinates of the floating mark are iteratively corrected until correspondence of both images is achieved. It may be noted that this approach enables the measurement of Z above regular XY grid positions.

4.3.6.3.2 Vertical line locus

The image lines *g'* and *g''* correspond to the projection of the vertical line *g* which passes through P (Fig. 4.54). These straight lines are epipolar lines only if *g* is located in the epipolar plane of P. The image lines *g'* and *g''* are known as vertical line loci (VLL), in allusion to aerial photogrammetry (Z ≈ viewing direction). With given orientation parameters they can be easily calculated by a variation of the *Z* coordinate of P.

In order to measure a surface point, it is possible to calculate all points P_i at small intervals ΔZ between two points, e.g. Q and R, re-project them into the images and search for the best correspondence on the straight lines g' and g''.

For arbitrarily oriented object surfaces, g should lie in the direction of the normal vector to the surface at the target point. The method is not restricted to the stereo case but can be extended to an unlimited number of images per point.
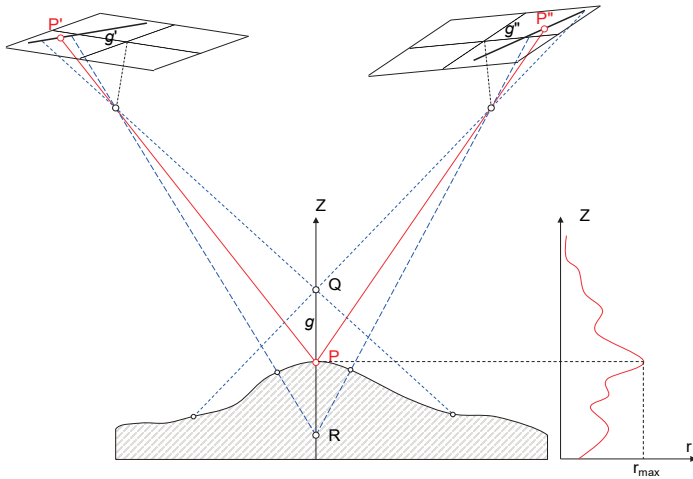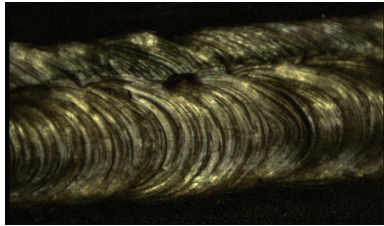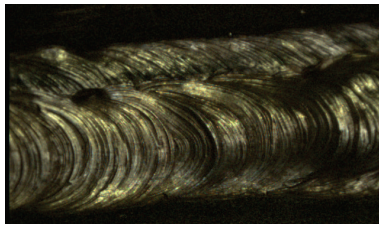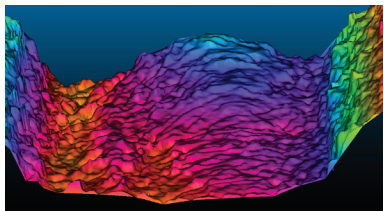

Fig. 4.54: Imaging positions and correlation coefficient r of a vertically shifted floating mark.
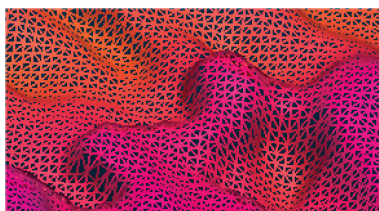

Left image

Right image

Colour-coded surface model

Magnified detail

Fig. 4.55: Reconstruction of a welding seam surface by stereo photogrammetry.

As an example, Fig. 4.55 shows the reconstruction of a welding seam using a stereo image. Object resolution is about 0.04 mm, the accuracy of Z coordinates is estimated to be around 0.2 mm (imaging distance 70 mm, height-to-base ratio 10:1). Imaging matching was performed by semi-global matching (section 5.5.4.2).

4.4 Multi-image processing and bundle adjustment

4.4.1 General remarks

4.4.1.1 Objectives

Bundle adjustment (bundle triangulation, bundle block adjustment, multi-image triangulation, multi-image orientation) is a method for the simultaneous numerical fit of an unlimited number of spatially distributed images (bundles of rays). It makes use of photogrammetric observations (measured image points), survey observations and an object coordinate system (Fig. 4.56). Using tie points, single images are merged into a global model in which the object surface can be reconstructed in three dimensions. The connection to a global object coordinate system can be provided by a minimum number of reference points so that larger areas without reference points can be bridged by multi-image sub-sets. The most important geometric constraint is that all corresponding (homologous) image rays should intersect in their corresponding object point with minimum inconsistency. This overall process is increasingly also referred to as *alignment*, which can include the process of image matching (feature detection and matching).
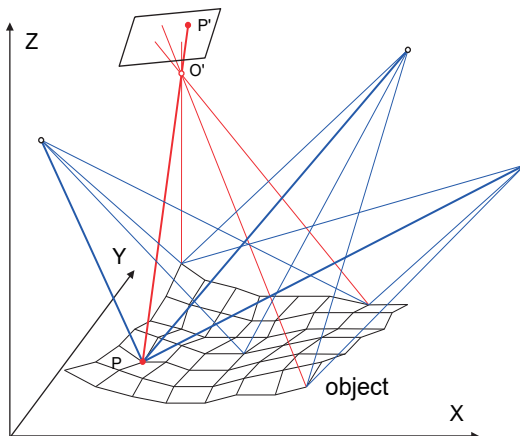

Fig. 4.56: Multi-image triangulation.

In an over-determined system of equations, an adjustment technique estimates 3D object coordinates, image orientation parameters and any additional model parameters, together with related statistical information about accuracy and reliability. All observed (measured) values, and all unknown parameters of a photogrammetric project, are taken into account within one simultaneous calculation which ensures that homologous rays optimally intersect. In this way, the ray bundles provide strong geometry for a dense, high-accuracy measurement network (example in Fig. 4.57). The bundle triangulation therefore represents the most powerful and accurate method of image orientation and point determination in photogrammetry.
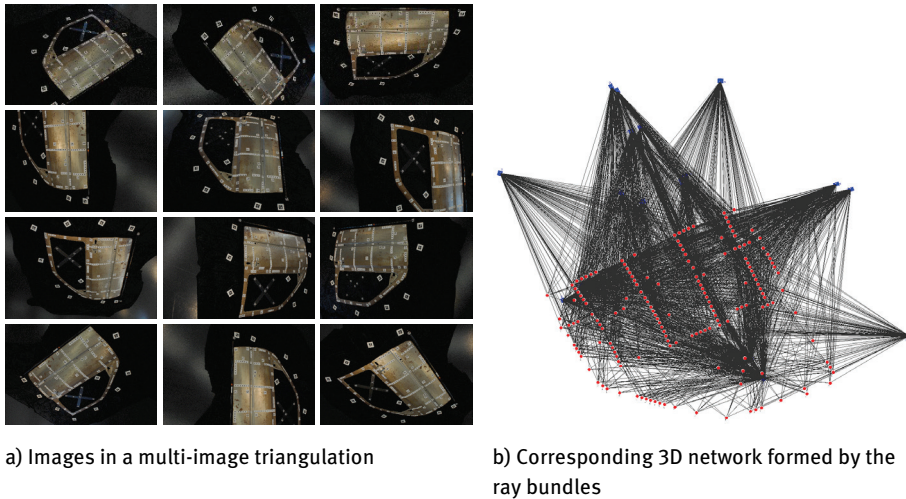

a) Images in a multi-image triangulation

b) Corresponding 3D network formed by the ray bundles

Fig. 4.57: Multi-image network for measuring a car door.

4.4.1.2 Development

The bundle triangulation method has been known since the late 1950s, i.e. from the very beginning of analytical photogrammetry. Bundle adjustment is a very general technique which combines elements of geodetic and photogrammetric triangulation, space resection and camera calibration. These are individually well understood and so the practical problems in the implementation of bundle adjustment do not lie in the mathematical formulations but in the following areas:

–   solution of large systems of normal equations (up to a few thousand unknowns);
–   generation of approximate values for the unknowns;
–   detection and elimination of gross data errors.

The development of practical bundle adjustments is closely related to increases in computing power. In this respect it is worth noting that similar programs for aerial

photogrammetry have largely been developed independently of those for close-range applications.

The triangulation of aerial images is characterized mainly by:

– predominant use of regular strip arrangements of images (Fig. 4.58) and hence
– advantageous structure of normal system of equations and
– easier generation of approximate values, e.g. rotations approximately zero;
– large numbers of images and object points and
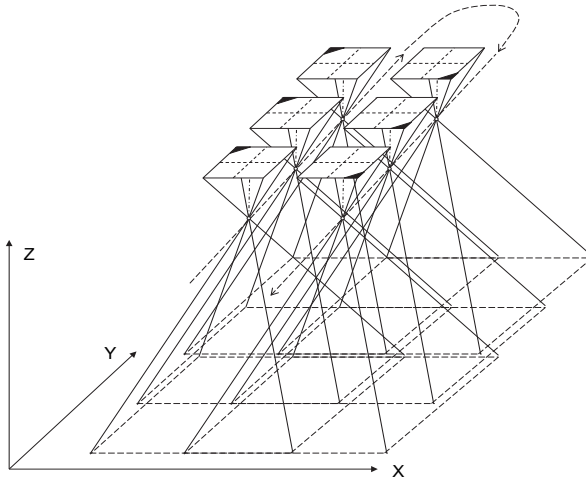– use of only *one* calibrated metric camera.

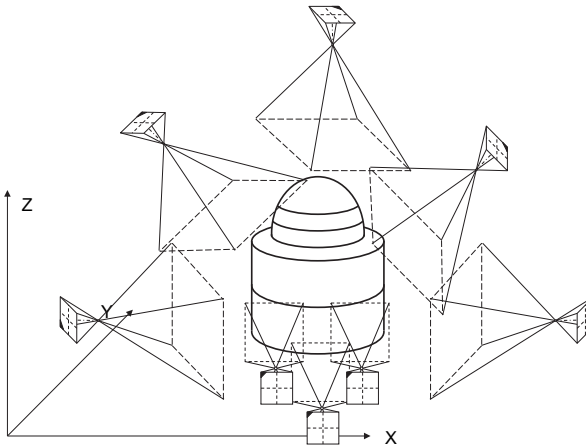**Fig. 4.58:** Image configuration for aerial photogrammetry.

Fig. 4.59: Arbitrary close-range image configuration.

In contrast, the standard case in close-range photogrammetry is characterized by:
– irregularly arranged, arbitrary and often unfavourable image configurations (Fig. 4.59);
– more complex structure of normal system of equations;
– arbitrarily oriented object coordinate systems;
– more demanding solutions for generating approximate values;
– combined adjustment of survey observations and conditions;
– (several) imaging systems to be calibrated simultaneously.

Since the early 1980s the bundle adjustment has been accepted in all areas of photogrammetry. As a result of diverse requirements and applications there are many different bundle adjustment packages on the market.

Since its introduction for close-range use, the method of bundle adjustment has considerably widened the application spectrum as a result of its ability to handle almost arbitrary image configurations with few restrictions on the image acquisition systems. The general concept of bundle adjustment allows for the inclusion of other sensor information and can be performed without any image data at all.

4.4.1.3 Data flow

Fig. 4.60 shows the principle data flow for a bundle adjustment process. Input data for bundle adjustments are typically photogrammetric image coordinates generated by manual or automatic (digital) image measuring systems. Each measured image point is stored together with a unique point identifier and the corresponding image number, hence prior successful matching of image points is assumed (Fig. 4.61). This is sufficient to reconstruct the three-dimensional shape of the object surface, as represented by the measured object points.

Additional information in the object space, such as measured distances, angles, points, straight lines and planes, can also be taken into account. They provide the definition of an absolute scale and the position and orientation of the object coordinate system (datum definition). This information is entered into the system as, for example, reference point files or additional observations, e.g. constraints between object points (see section 4.4.2.3).

In order to linearize the functional model, approximate values must be generated. For simpler image configurations they can be extracted from planning data or project sketches. The generation of approximate values for more complex configurations (larger number of images, arbitrary orientations) is performed by iterative calculation methods (see section 4.4.4). Sensors for direct georeferencing, e.g. GNSS, IMU, also provide approximate values for the orientation unknowns.
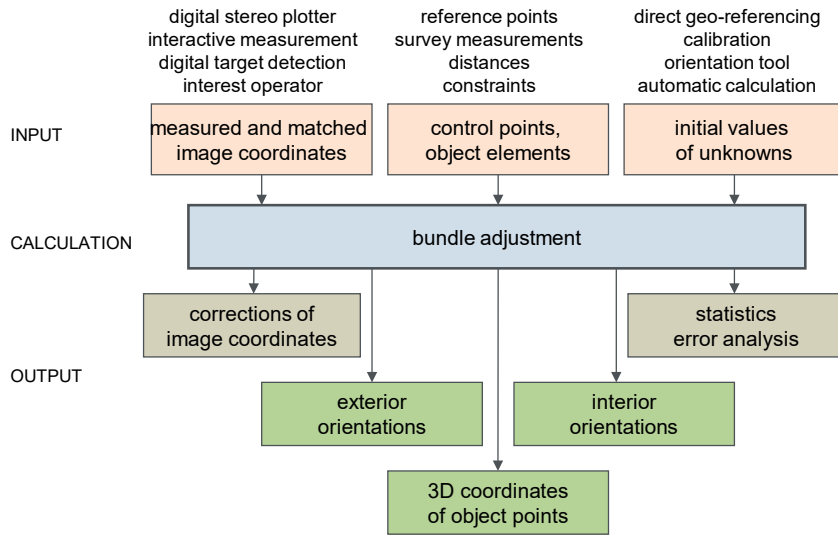
Fig. 4.60: Data flow for bundle adjustment process.

```
17       2    -1.077410    2.913010    0.000200   -0.000100    1
17       3    -1.884320    3.632870    0.000000   -0.000000    1
17       4    -2.904780    4.434720    0.000400    0.000300    1
17       5    -6.194840    0.078450   -0.000000    0.000300    1
17       7    -8.940230   -5.323440    0.000300   -0.000200    1
17       8    -8.559290   -6.360560    0.000500    0.000300    1
18       2    -7.412580    3.507980   -0.000100    0.000100    1
18       3    -8.247020    2.808140   -0.000100    0.000200    1
18       4    -9.185490    1.896390    0.000400   -0.000200    1
18       5    -5.220940   -1.985600   -0.000300    0.000200    1
18       7    -0.466490   -5.242080   -0.000000    0.000100    1
18       8     0.529900   -5.010630    0.000300   -0.000300    1
18       9     2.719160   -4.119760    0.000300    0.000000    1
19       3     7.378300    5.144030    0.000400   -0.000100    1
19       8    -5.978770    4.462030   -0.000400   -0.000100    1
19       9    -6.481080    2.217100    0.000000   -0.000100    1
19      10    -6.920640    0.723660   -0.000300   -0.000100    1
```

Fig. 4.61: Example of an image coordinate file (values in mm) with
image number, point number, x', y', sx', sy', code.

The principal results of bundle adjustment are the estimated 3D coordinates of the object points. They are given in an object coordinate system defined by reference points or free net adjustment (see section 4.4.3).

In addition, the exterior orientation parameters of all images are estimated. These can be further used, for example, in analytical plotters or for subsequent spatial intersections computed outside the bundle adjustment. The interior orientation parameters are estimated if the cameras are calibrated simultaneously within the adjustment.

In order to analyse the quality of the bundle adjustment, it is possible to calculate image coordinate residuals (corrections), standard deviations of object points and orientation data, correlations between parameters and reliability numbers for the detection of gross errors.

4.4.2 Mathematical model

4.4.2.1 Adjustment model

The mathematical model of the bundle adjustment is based on the collinearity equations (see section 4.2.2).

$$x'_P = x'_0 + z' \frac{r_{11}(X - X_0) + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} + \Delta x'$$

$$y'_P = y'_0 + z' \frac{r_{12}(X - X_0) + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} + \Delta y'$$

(4.98)

The structure of these equations allows the direct formulation of primary observed values (image coordinates) as functions of all unknown parameters in the photogrammetric imaging process. The collinearity equations, linearized at approximate values, can therefore be used directly as observation equations for a least-squares adjustment according to the Gauss-Markov model (see section 2.4.2).

It is principally the image coordinates of homologous points which are used as observations[2]. The following unknowns are iteratively determined as functions of these observations:

- three-dimensional object coordinates for each new point i (total u_P, 3 unknowns each);
- exterior orientation of each image j (total u_I, 6 unknowns each);
- interior orientation of each camera k (total u_C, 0 or ≥ 3 unknowns each).

The bundle adjustment is completely general and can, for example, represent an extended form of the space resection (see section 4.2.3, eqn. 4.16):

$$x'_i + vx'_i = F(X_{0j}, Y_{0j}, Z_{0j}, \omega_j, \varphi_j, \kappa_j, x'_{0k}, z'_k, \Delta x'_k, X_i, Y_i, Z_i)$$

$$y'_i + vy'_i = F(X_{0j}, Y_{0j}, Z_{0j}, \omega_j, \varphi_j, \kappa_j, y'_{0k}, z'_k, \Delta x'_k, X_i, Y_i, Z_i)$$

(4.99)

where i: point index
         j: image index
        k: camera index

---

2 Additional observations such as object point coordinates, distances or directions are introduced in section 4.4.2.3.

The non-linear equations (4.98) are linearized using a Taylor series expansion with approximate values for all unknowns (in red) inside the brackets in (4.99). Here the derivatives, already determined in equations (4.19), are extended by the derivatives with respect to object coordinates:

$$\frac{\partial x'}{\partial X} = -\frac{z'}{N^2}(r_{13}k_X - r_{11}N) \qquad\qquad \frac{\partial y'}{\partial X} = -\frac{z'}{N^2}(r_{13}k_Y - r_{12}N)$$

$$\frac{\partial x'}{\partial Y} = -\frac{z'}{N^2}(r_{23}k_X - r_{21}N) \qquad\qquad \frac{\partial y'}{\partial Y} = -\frac{z'}{N^2}(r_{23}k_Y - r_{22}N) \qquad\qquad (4.100)$$

$$\frac{\partial x'}{\partial Z} = -\frac{z'}{N^2}(r_{33}k_X - r_{31}N) \qquad\qquad \frac{\partial y'}{\partial Z} = -\frac{z'}{N^2}(r_{33}k_Y - r_{32}N)$$

If the interior orientation parameters are introduced as unknowns, the following derivatives are added ($c = -z'$):

$$\frac{\partial x'}{\partial x'_0} = 1 \qquad\qquad \frac{\partial y'}{\partial y'_0} = 1 \qquad\qquad (4.101)$$

$$\frac{\partial x'}{\partial c} = -\frac{k_X}{N} \qquad\qquad \frac{\partial y'}{\partial c} = -\frac{k_Y}{N}$$

Derivatives with respect to additional parameters of distortion are introduced in a similar way (see section 3.3.3). If linearization is done numerically (see section 2.4.5.1), the projection equations and selected distortion model can be programmed directly into the source code and a rigorous differentiation is not required.

In standard form, the linearized model is given by

$$\underset{n,1}{\mathbf{l}} + \underset{n,1}{\mathbf{v}} = \underset{n,u}{\mathbf{A}} \cdot \underset{u,1}{\hat{\mathbf{x}}} \qquad\qquad (4.102)$$

and the corresponding system of normal equations is

$$\underset{u,u}{\mathbf{N}} \cdot \underset{u,1}{\hat{\mathbf{x}}} + \underset{u,1}{\mathbf{n}} = \underset{u,1}{\mathbf{0}} \qquad\qquad (4.103)$$

where

$$\underset{u,u}{\mathbf{N}} = \underset{u,n}{\mathbf{A}^T} \cdot \underset{n,n}{\mathbf{P}} \cdot \underset{n,u}{\mathbf{A}} \qquad\qquad \underset{u,1}{\mathbf{n}} = \underset{u,n}{\mathbf{A}^T} \cdot \underset{n,n}{\mathbf{P}} \cdot \underset{n,1}{\mathbf{l}}$$

The solution vector and its covariance matrix are estimated in an iterative adjustment:

$$\underset{u,1}{\hat{\mathbf{x}}} = \underset{u,u}{\mathbf{Q}} \cdot \underset{u,1}{\mathbf{n}} = (\underset{u,n}{\mathbf{A}^T} \cdot \underset{n,n}{\mathbf{P}} \cdot \underset{n,u}{\mathbf{A}})^{-1} \cdot \underset{u,n}{\mathbf{A}^T} \cdot \underset{n,n}{\mathbf{P}} \cdot \underset{n,1}{\mathbf{l}} \qquad\qquad (4.104)$$

where

$$\mathbf{Q}_{u,u} = \mathbf{N}_{u,u}^{-1} \qquad : \text{cofactor matrix}$$

$$\mathbf{C}_{u,u} = \hat{s}_0^2 \cdot \mathbf{Q}_{u,u} \qquad : \text{variance-covariance matrix}$$

4.4.2.2 Normal equations

The number of unknowns in the adjustment system can be calculated as follows:

$$u = u_I \cdot n_{images} + u_P \cdot n_{points} + u_C \cdot n_{cameras} (+u_{datum}) \tag{4.105}$$

where

$u_I = 6$: parameters of exterior orientation per image

$u_P = 3$: XYZ coordinates of new points

$u_C = 0\ldots{\geq}3$: parameters of interior orientation per camera

In addition to the unknown orientation parameters and point coordinates, up to seven parameters are still required for datum definition. However, these can be eliminated by use of reference points or appropriate condition equations (datum defect, see section 4.4.3.1). Table 4.1 gives examples of the number of observations and unknowns for different image configurations.

Table 4.1: Number of unknowns and observations for different image configurations.

	Example 1	u	u_{total}	Example 2	u	u_{total}	Example 3	u	u_{total}
	aerial set-up (Fig. 4.62)			closed loop set-up (Fig. 4.64)			test field calibration (Fig. 7.33d)		
n_{images}	8	6	48	16	6	96	8	6	48
n_{points}	14	3	42	25	3	75	13+6	3	57
$n_{ref.\ pts.}$	6	0	0	0	0	0	3 distances	0	0
$n_{cameras}$	1	0	0	1	5	5	1	7	7
u_{datum}	0		0	7		(7)	6		(6)
u_{total}			90			176			112
n_{obs}			112			384			304
$r=n-u$			22			208			192

Example 1 represents a regular arrangement for aerial photogrammetry or the measurement of plane façades (Fig. 4.62). The number of unknowns is $u = 90$ if 8

images and 14 new points are assumed. The datum is defined by 6 reference points. For image acquisition, a metric camera with known interior orientation is used.[3]

The connection matrix in Fig. 4.62 shows which point is measured in which image. The numbering of object points by measurement strip produces the typical diagonal structure of connections, which is also seen in the structure of the Jacobian matrix **A** and the resulting normal equations (Fig. 4.63). In order to solve the normal equations, the order of observations can be further optimized by suitable sorting algorithms (see section 2.4.5.3).
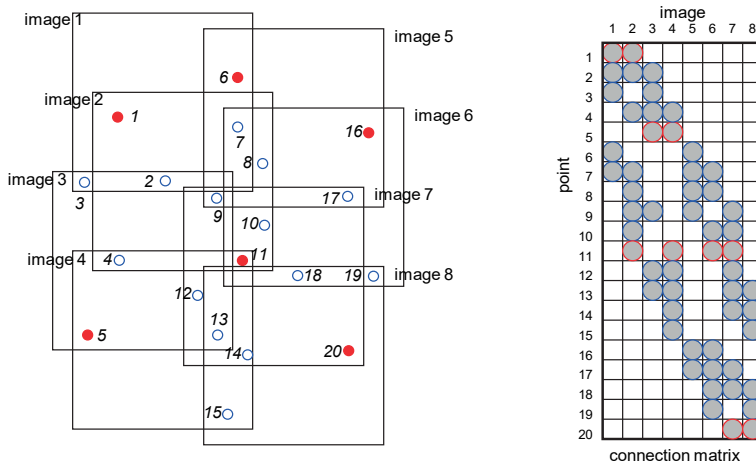
Fig. 4.62: Example 1: aerial or façade arrangement.

On average each object point is measured in 2.8 images. Each measured image point provides 2 observations. Hence, with a total number of $n = 112$ observations and a total redundancy number of $r = n\text{-}u = 112 - 90 = 22$, redundancy in the adjustment system is relatively weak.

---

3 For a clearer illustration, a reduced number of object points is used in this example.
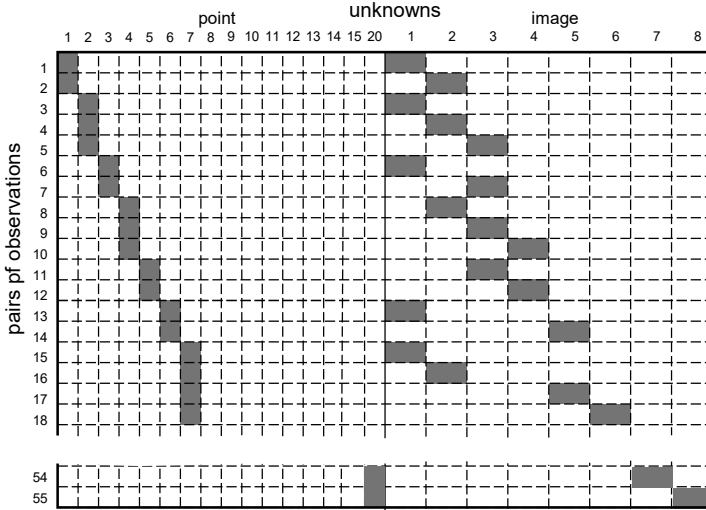
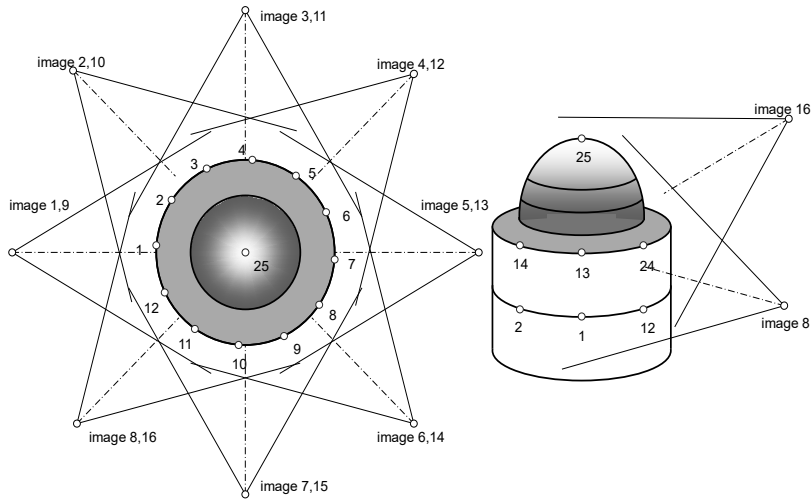**Fig. 4.63:** Structure of design matrix A according to example 1.

Fig. 4.64: Example 2: closed-loop arrangement.

Example 2 shows a typical closed-loop image configuration for an object formed from a cylinder and hemisphere (Fig. 4.64). The object is recorded in two image acquisition sets which have a relative vertical shift (bottom-set images 1–8, top-set images 9–16). The object point on top of the dome appears in all images. As there are no reference points available, the datum defect of 7 is eliminated by a free net adjustment (see section 4.4.3.3). A non-metric camera is used so 5 parameters of interior orientation

must be simultaneously calibrated. A total of 176 unknowns must therefore be estimated.

The corresponding connection matrix (Fig. 4.65) shows an average of 7.7 images for each object point. The redundancy is therefore much higher than in example 1 (aerial or façade arrangement). With 384 measured image coordinates, the total redundancy is $r = 384 - 176 = 208$.

However, the extent to which the adjustment system can be calculated, and the quality of the results, are less a question of total redundancy than the geometric configuration of the system. Consider the arrangement of example 1 (Fig. 4.62) which allows for the determination of plane coordinates (parallel to the image plane) to an acceptable accuracy whilst a different point accuracy, which varies as a function of the height-to-base ratio, applies along the viewing direction. It is not practicable to perform a camera calibration with this set-up. In contrast, the arrangement of example 2 (Fig. 4.64) represents a very stable geometry which can provide not only self-calibration but also a high and homogenous point accuracy in all three coordinate axes.
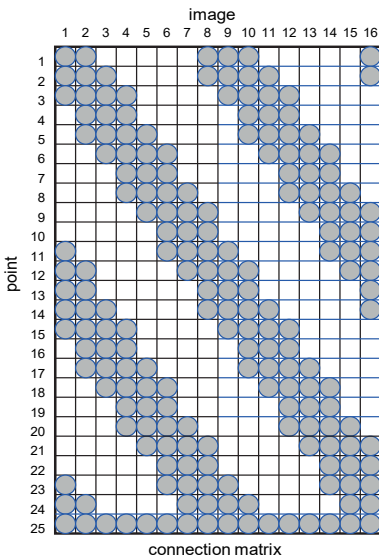

Fig. 4.65: Connection matrix of example 2 (Fig. 4.64).

Finally, example 3 in Table 4.1 refers to an arrangement for test field calibration according to Fig. 7.29 (see section 7.3.1.1). If the test field, in this case with 13 points and 3 distances (6 points), is completely covered by each of the 8 images, a total number of $u = 112$ unknowns and $n = 304$ observations results. Even with this simple set-up a high redundancy and stable geometric configuration are both achieved.

4.4.2.3 Combined adjustment of photogrammetric and survey observations

The system of equations in (4.98) describes the original model of bundle triangulation by defining the image coordinates x',y' (observations) as a function of the unknowns, specifically of the object coordinates X,Y,Z. Additional information about the object or additional non-photogrammetric measurements are not considered in (4.98).

An extended model for the bundle adjustment takes additional observations into account as, for example, measured distances, directions or angles from survey instruments or laser trackers. Other constraints on the object can also be integrated, such as known points, coordinate differences, straight lines, planes or surfaces having rotational symmetry (Fig. 4.66).

All additional observations can be weighted according to their accuracy or importance and therefore have a rigorous stochastic treatment in the adjustment process.
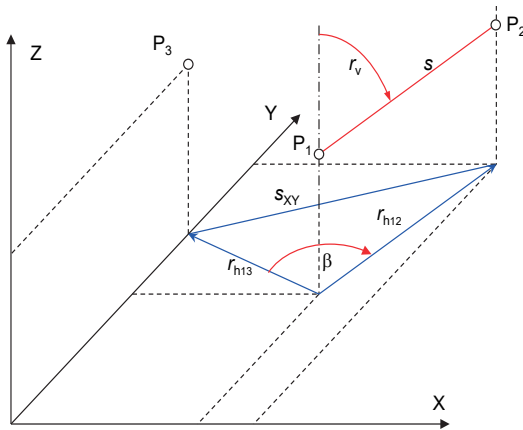
Fig. 4.66: Survey observations.

4.4.2.3.1 Coordinates, coordinate differences and distances

It is particularly easy to introduce observed coordinates, coordinate differences or distances. The following observation equations result from the introduction of

object coordinates:
$$X = X \qquad\qquad Y = Y \qquad\qquad Z = Z \qquad\qquad (4.106)$$

coordinate differences:
$$\Delta X = X_2 - X_1 \qquad\qquad \Delta Y = Y_2 - Y_1 \qquad\qquad \Delta Z = Z_2 - Z_1 \qquad\qquad (4.107)$$

slope distances:
$$s = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2 + (Z_2 - Z_1)^2} \qquad\qquad (4.108)$$

distances in XY plane:

$$s_{XY} = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \tag{4.109}$$

distances in XZ plane:

$$s_{XZ} = \sqrt{(X_2 - X_1)^2 + (Z_2 - Z_1)^2} \tag{4.110}$$

distances in YZ plane:

$$s_{YZ} = \sqrt{(Y_2 - Y_1)^2 + (Z_2 - Z_1)^2} \tag{4.111}$$

4.4.2.3.2 Exterior orientations

Known position or orientation data of a camera can be taken into account by the following additional observation equations:

exterior orientation:

$$
\begin{aligned}
X_0 &= X_0 & \omega &= \omega \\
Y_0 &= Y_0 & \varphi &= \varphi \\
Z_0 &= Z_0 & \kappa &= \kappa
\end{aligned}
\tag{4.112}
$$

Normally the exterior orientation cannot be measured directly. Older metric cameras may be combined with surveying instruments; in this case surveyed angles (azimuth, elevation) can be introduced as rotation angles of exterior orientation; the rotation matrix must be properly defined (see section 4.2.1.2). In addition, exterior orientation parameters may be known from previous calculations and so can be processed with a corresponding weight.

4.4.2.3.3 Relative orientations

For the orientation and calibration of stereo cameras which are stably mounted relative to each other (example in Fig. 3.127), it is useful to introduce constraint equations which fix the relative orientation at all imaging locations:

Base constraint for fixed relative orientation between images i and j:

$$\sqrt{(X_{0i} - X_{0j})^2 + (Y_{0i} - Y_{0j})^2 + (Z_{0i} - Z_{0j})^2} = b_{ij} \tag{4.113}$$

Rotation constraint for fixed relative orientation:

$$
\begin{aligned}
a_{11} \cdot b_{11} + a_{12} \cdot b_{12} + a_{13} \cdot b_{13} &= \alpha \\
a_{21} \cdot b_{21} + a_{22} \cdot b_{22} + a_{23} \cdot b_{23} &= \beta \\
a_{31} \cdot b_{31} + a_{32} \cdot b_{32} + a_{33} \cdot b_{33} &= \gamma
\end{aligned}
\tag{4.114}
$$

Here a_{ij} and b_{ij} are the elements of the rotation matrices **A** and **B** of the images i and j. Each of the constraint equations forms a scalar product, hence α, β, γ define the rotation angles between the axes of the two cameras.

4.4.2.3.4 Directions and angles

Survey directions and angles, for example observed by theodolite, can also be introduced. In a conventional levelled use where the XY plane is horizontal and the theodolite's vertical axis corresponds to the Z axis (Fig. 4.66), the equations are:

horizontal direction:

$$r_h = \arctan\left(\frac{X_2 - X_1}{Y_2 - Y_1}\right) \tag{4.115}$$

horizontal angle:

$$\beta = \arctan\left(\frac{X_3 - X_1}{Y_3 - Y_1}\right) - \arctan\left(\frac{X_2 - X_1}{Y_2 - Y_1}\right) \tag{4.116}$$

vertical angle:

$$r_v = \arcsin\left(\frac{Z_2 - Z_1}{\sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2 + (Z_2 - Z_1)^2}}\right) \tag{4.117}$$

This can be extended to complete sets of surveyed directions (several measured horizontal and vertical directions and distances from one station). With modification, the equations can also apply to non-levelled instruments such as laser trackers. With the above observation types, pure 3D survey nets can also be adjusted.

4.4.2.3.5 Auxiliary coordinate systems

The introduction of auxiliary coordinate systems is a very elegant way of formulating additional information in object space. An auxiliary coordinate system $\bar{X}\,\bar{Y}\,\bar{Z}$ is a 3D coordinate system arbitrarily oriented in space and used to define additional observations or constraints. For example, this can be local reference point configurations with a defined relation to each other or local geometric elements, e.g. rotationally symmetric shapes.

The auxiliary coordinate system can be transformed into the object coordinate system X,Y,Z using a spatial similarity transformation (see section 2.2.4):

$$\bar{\mathbf{X}} = \bar{\mathbf{X}}_0 + \bar{m} \cdot \mathbf{R}_{\bar{\omega}\bar{\varphi}\bar{\kappa}} \cdot (\mathbf{X} - \mathbf{X}_0)$$ (4.118)

or

$$\begin{bmatrix} \bar{X} \\ \bar{Y} \\ \bar{Z} \end{bmatrix} = \begin{bmatrix} \bar{X}_0 \\ \bar{Y}_0 \\ \bar{Z}_0 \end{bmatrix} + \bar{m} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{bmatrix}$$

where

X,Y,Z: spatial coordinates in global system

$\bar{X},\bar{Y},\bar{Z}$:   spatial coordinates in local system

X_0,Y_0,Z_0:   centroid of point cloud in XYZ system

                           (constant coordinate values in the adjustment process)

$\bar{X}_0,\bar{Y}_0,\bar{Z}_0$:   origin of system XYZ with respect to system $\bar{X}\,\bar{Y}\,\bar{Z}$

$\mathbf{R}_{\bar{\omega}\bar{\varphi}\bar{\kappa}}$:   rotation matrix transforming XYZ parallel to $\bar{X}\,\bar{Y}\,\bar{Z}$

$\bar{m}$:   scaling factor between both systems

The seven transformation parameters must be introduced as unknowns in the adjustment; hence approximate values must be provided. It is now possible to define functional relationships between object points within the auxiliary coordinate system. They can be expressed by parameters B_i, which can be transformed into the object coordinate system using (4.118). Parameters B_i can describe simple geometric conditions and also object surfaces of higher order.

$$f(X,Y,Z,\bar{\omega},\bar{\varphi},\bar{\kappa},\bar{X}_0,\bar{Y}_0,\bar{Z}_0,\bar{m},B_i) = g(\bar{X},\bar{Y},\bar{Z},B_i) = C$$ (4.119)

C is a constant in the constraint equation, expressed as function f in the object coordinate system or as function g in the auxiliary coordinate system. An arbitrary object plane is therefore defined by

$$\bar{X} = C$$ (4.120)

A straight line in the auxiliary coordinate system is defined by

$$\bar{X} = C \qquad\qquad\qquad \bar{Y} = C$$ (4.121)

A rotational solid is given by

$$\bar{X}^2 + \bar{Y}^2 + h(\bar{Z},B_i) = C$$ (4.122)

The function h defines the shape of the solid in the direction of the rotational axis, e.g. for a circular cylinder with radius r

$$h(\bar{Z},B_i) = 0 \qquad \text{and} \qquad C = r^2$$ (4.123)

or for a sphere with radius r:

$$h(\bar{Z},B_i) = \bar{Z}^2 \qquad \text{and} \qquad C = r^2$$ (4.124)

4.4.2.3.6 Applications of additional observations

With the aid of the additional observations above, the adjustment system can be significantly influenced:

– Weak image configurations can be stabilized by introducing additional object information.
– Surveyed or other measured object data, and photogrammetric observations, can be adjusted in one step, e.g. in order to minimize net strains (inconsistencies between object points) or to handle measurement data from different sources in a balanced way.
– Unlimited numbers of known distances (scales) between new points can be observed and processed according to their accuracy.
– Single object points can be forced onto an object surface by geometric conditions, e.g. points on a cylindrical surface (pipe line). [4]
– Information about exterior orientation provided by instruments such as inertial navigation units (INU), GNSS location or gyroscopes, can support the adjustment process. Mechanically defined conditions, such as the movement of a camera along a straight line or circle, can also be added as additional observations. This applies equally to geometric constraints between the orientation parameters of different cameras, e.g. for stereo cameras or camera/projector arrangements.
– Additional information which is introduced with an unrealistic weight can negatively affect the adjustment result. This point is particularly important in practice and demands a careful choice of weights and analysis of results.

4.4.2.4 Adjustment of additional parameters
4.4.2.4.1 Self-calibration

Functions for the correction of imaging errors are referred to as *additional parameter functions*. Functional models which describe real optical characteristics of image acquisition systems (parameters of interior orientation) have been summarized in section 3.3. When these parameters are determined within the bundle adjustment, the procedure is known as *self-calibration*.

The linearized model of the adjustment is extended by derivatives for principal distance and principal point, as given in (4.101). Again, approximate values of the additional unknowns must be provided, although the following initial values are usually sufficient:

---

4 Conditions include both weak constraints defined by additional observations with standard deviations, as well as fixed constraints which force an exact condition.

principal distance: $c \approx -f$ (focal length)

principal point: $x'_0 \approx y'_0 \approx 0$

radial distortion: $A_1 \approx A_2 \approx 0$

decentring distortion: $B_1 \approx B_2 \approx 0$

affinity and shear: $C_1 \approx C_2 \approx 0$

The ability to determine individual parameters depends, on one hand, on the modelling of the physical imaging process. Parameter sets based on faulty physical assumptions can lead to weakly conditioned systems of equations, over-parametrized equations or high correlations between parameters. On the other hand, interior orientation parameters can only be reliably calculated if image configuration and distribution of object points are well chosen. Section 7.3.2 summarizes suitable image configurations for self-calibration.

The parameters of interior orientation can additionally be handled as observed values with a corresponding weight. The observation equations in this case are given by:

interior orientation:

$$
\begin{aligned}
c &= c & A_1 &= A_1 \\
x'_0 &= x'_0 \;, & A_2 &= A_2 & \text{(4.125)} \\
y'_0 &= y'_0 & &\text{etc.}
\end{aligned}
$$
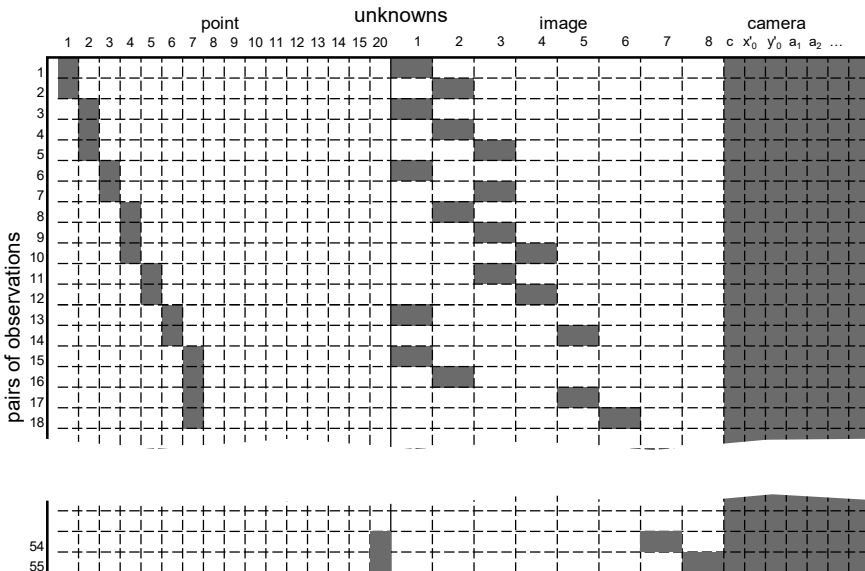
Fig. 4.67: Structure of design matrix **A**, extended by additional unknowns for the interior orientation of a camera (compare Fig. 4.63).

If additional parameters are introduced, the structure of the normal system of equations changes significantly. While the design matrix illustrated in Fig. 4.63 shows a distinctive diagonal structure with a large number of zero elements, the normal equation matrix of adjustment systems with additional parameters contains larger areas with non-zero elements. Each measured image coordinate is functionally connected with the unknowns of interior orientation, hence the right-hand side of the matrix is completely filled (Fig. 4.67).

In order to invert the resulting normal equation matrices, much more storage and computing power is required. Sparse matrix techniques have been successfully applied in practice for the efficient use of computing resources (see section 2.4.5.3).

4.4.2.4.2 Calibration with variable interior orientation

The models normally used for interior orientation assume constant camera parameters during the period of image acquisition. If the camera geometry changes over this period, for example due to refocusing or a change of lens, a "new" camera with its own parameters must be assigned to the corresponding images. The simultaneous determination of more than one group of parameters requires an image configuration appropriate for the calibration of each camera (see section 7.3.2).

If stable camera parameters cannot be guaranteed for longer periods, the interior orientation must be calibrated individually for each image. This approach to image-variant calibration provides corrections for a shift of the perspective centre (camera constant and principal point coordinates) but otherwise assumes stable distortion parameters for all images. For this purpose, the image-variant parameters are introduced as observed unknowns with approximate values and a priori standard deviations which correspond to the expected shift of the perspective centre. The numerical stability of the adjustment is maintained provided there is a suitable number of object points each with an appropriate number of image rays. Depending on image configuration and current state of the camera, applications using DSLR cameras report an accuracy increase by a factor of 2 to 4 compared with cases in which image-variant parameters are not used (see section 3.3.4.4). The approach can be extended to the simultaneous calculation of a correction grid to allow for sensor deformation (see section 3.3.4.5).

4.4.3 Object coordinate system (definition of datum)

4.4.3.1 Rank and datum defect

A network composed of purely photogrammetric observations leads to a singular system of normal equations because, although the shape of the network can be determined, its absolute position and orientation in space cannot be determined. The resulting system of equations has a *rank defect*

$$d = u - r \tag{4.126}$$

where

u: number of unknowns

r: rank($\mathbf{A}$)

The rank defect is caused by a datum defect in the observed network which, for a three-dimensional network, can be removed by defining 7 additional elements:

3 translations
3 rotations
1 scaling factor

If at least one known distance is observed, then the datum defect is reduced by 1. The information is introduced as an additional distance observation according to equations 4.108ff.

Translational datum defects can be eliminated if control or reference points with known object coordinates are observed. Options for this are reviewed in the next section.

Rotational datum defects can also be eliminated by reference points, as well as by directly measured directions or angles.
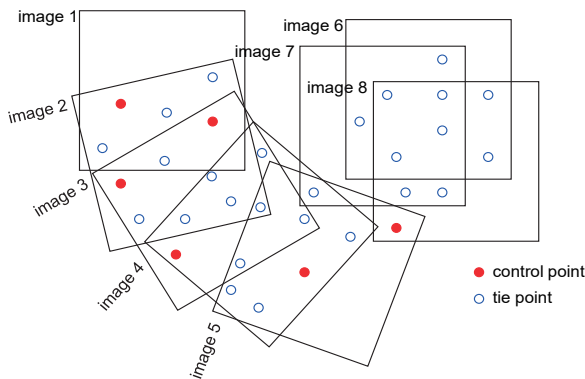
Fig. 4.68: Example of a configuration defect in an arrangement of images; images 6,7,8 cannot be oriented with respect to the rest of the images.

In addition to datum defects, observed networks can contain a configuration defect if, due to missing observations, some portions of the network cannot be determined unambiguously. In photogrammetric networks this problem seldom arises. If some images, for example, contain an insufficient number of tie points, they cannot be oriented with respect to the other images (see example in Fig. 4.68).

4.4.3.2 Reference points
4.4.3.2.1 Error-free reference points

Reference points are used for the definition of a global object coordinate system (datum definition). They can be introduced as error-free reference points into the bundle adjustment if their nominal coordinates are known to a high accuracy. Such coordinates could be introduced into the bundle adjustment as constants. Some bundle adjustment programs allow their input as measurements with zero standard deviations. Logically this approach leads to an error-free definition of the datum.

It is assumed, however, that the accuracy of the reference point coordinates is, for example, better by a factor 5–10 than the photogrammetric point determination, and that there are no inconsistencies in reference point coordinates. Errors in reference point coordinates are interpreted as errors in observations and are therefore difficult to detect. The definition of the object coordinate system using error-free reference points gives rise to a fixed datum.

In principle the spatial distribution of reference points should follow the recommendations made for absolute orientation (see section 4.3.5). The stereo model discussed there must here be considered a model defined by all images (bundles), i.e. a minimum of three reference points is required for the definition of the object coordinate system. As a minimum, the following coordinate components must be given by these three reference points (the example components in brackets relate to an image plane parallel to XY and viewing direction parallel to Z):

– a minimum of 2x2 coordinates parallel to the primary object plane (X_1, Y_1, X_2, Y_2)
– a minimum of 3 coordinates perpendicular to the primary object plane (Z_1, Z_2, Z_3)

The minimum configuration can be established, for example, by 2 full reference points (2x *XYZ*) and one additional reference height point (1x *Z*) or 2 plane reference points (2x *XY*) and 3 reference height points (3x *Z*). However, in many applications of close-range photogrammetry there are more than 2 full reference points available.

Reference points should be widely and uniformly distributed over the area covered by the images. Fig. 4.69 (left) shows an image set-up with a suitable distribution of 4 reference points, leading to a stable datum definition and homogeneous accuracies. In Fig. 4.69 (right) the 4 reference points are distributed inefficiently in one corner of the image configuration. As a consequence, the whole system can rotate around this point cloud, which results in discrepancies at more distant points and correspondingly higher standard deviations.

Reference points used for datum definition must not lie on a common straight line, as the normal system of equations then becomes singular. Unfavourable distributions of reference points which come close to this restriction will result in numerically weak systems of equations.
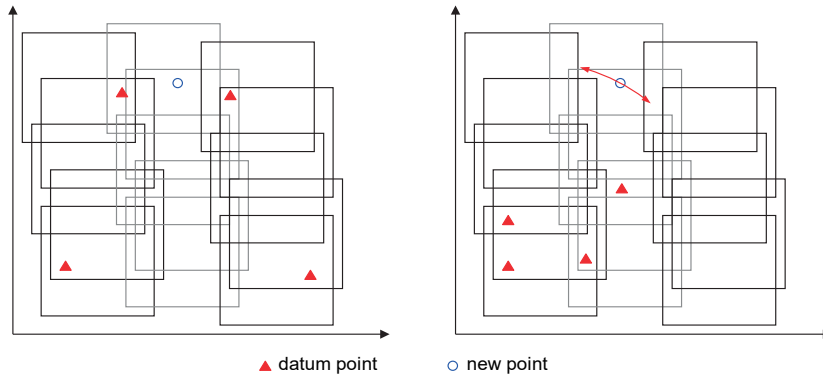
▲ datum point   ○ new point

Fig. 4.69: Good (left) and bad (right) distribution of reference points in a multi-image configuration.

4.4.3.2.2 Coordinates of reference points as observed quantities

Coordinates of reference points can also be introduced as observed quantities with a weight corresponding to their real point accuracy, e.g. depending on the measuring systems used for coordinate determination (see section 4.4.2.3). Within the adjustment system, the coordinates of the reference points are treated as unknowns and receive corrections and accuracy values in the same way as other observations. Standard deviations of weighted reference points can be interpreted as quality measures for the reference points themselves.

Partial compensation for inconsistencies between reference point coordinates can be done by an appropriate variation of weights, provided that these inconsistencies are not directly transferred to the photogrammetric observations. A coordinate system defined in this way is known as a *weighted datum*. Using coordinates of reference points in this way also compensates completely for rank defects in the adjustment system. If the weighted datum also results in a weak definition of the coordinate system, then higher standard deviations for new points are usually obtained.

4.4.3.2.3 Unconstrained datum definition using reference points (3-2-1 method)

Using a minimum amount of object information, it is possible to define the object coordinate system in order to avoid any possible influence of inconsistencies in the reference points. For this purpose, scale is given by a known distance *S*. The coordinate axes can be then defined according to the following scheme known as the *3-2-1 method* (Fig. 4.70):

1. point	X	Y	Z
P_1	7.0	2.0	2.0
P_2	-	4.5	3.5
P_3	-	2.0	-

real coordinates

point	X	Y	Z
P_1	0	0	0
P_2	-	0	0
P_3	-	0	-

fictitious coordinates
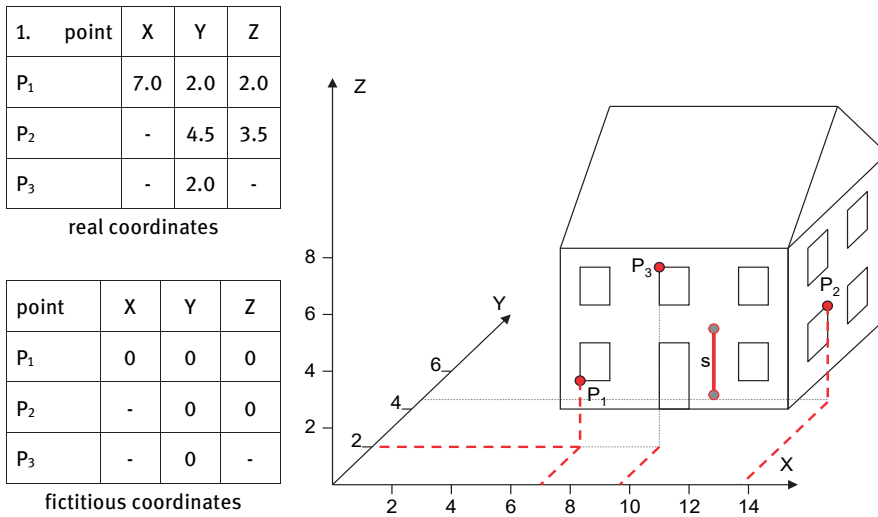
Fig. 4.70: Unconstrained coordinate system definition using reference points (3-2-1 method).

1. Fixing the X,Y,Z coordinates of point 1 defines an arbitrary 3D reference point in the object coordinate system which can, for example, represent the origin ($X = Y = Z = 0$).
2. Fixing the Y,Z coordinates of point 2 defines the X axis. At this stage the system can still be rotated about a line joining point 1 and point 2.
3. Fixing the Y coordinate of point 3 defines the XZ plane (alternatively define the XY plane with fixed Z or YZ plane with fixed X). Hence the coordinate system is uniquely defined without any constraints.

In some configurations, ambiguities in transformation are possible, e.g. a point could be transformed to the mirror image of itself in a plane. In these cases, some very approximate additional data can be used to choose between alternative positions.

Fig. 4.70 shows an example. If real reference point coordinates are not available, the system can also be defined by fictitious coordinates.

Scaling information can, of course, also be derived from the known distances between reference points or simply by the inclusion of a scale bar, e.g. element "s" in Fig. 4.70.

The unconstrained datum definition by reference points does not affect the shape of the photogrammetric network. It is true that the absolute coordinates of object points are related to the arbitrarily selected system origin but the distances between points are independent of the datum. In contrast, the estimated accuracies derived from the covariance matrix are influenced by the datum definition. If datum points are distributed according to Fig. 4.69 (right) then object point accuracy based on

datum point accuracy is interpreted too optimistically when they are close to the datum points and too pessimistically in more distant parts of the network.

4.4.3.3 Direct georeferencing

Direct georeferencing is a term originating in geospatial applications. It refers to the datum definition of a set of images by direct measurement of their exterior orientations without using reference points. Instead, exterior orientation is determined using sensor systems such as GNSS receivers (translations) or inertial measuring units (rotations). However, inertial, gyroscopic and wireless positioning devices are often applied in manufacturing situations, as are mechanical positioning devices such as robot arms (translations and rotations). The term will therefore be used generally in this book. In all cases there is a spatial offset in six degrees of freedom between camera system and the additional sensing devices which must be determined by system calibration. For aerial systems this process is also called boresight calibration.

Practical examples of direct georeferencing can be found in applications such as mobile mapping, or the use of drones for aerial imaging, where no reference points are available (see section 6.12). Further examples are found in the use of mechanical devices (robot, articulated arm, coordinate measurement machine) which can provide camera orientation parameters to a sufficient accuracy for many applications, again without photogrammetric reference points.

However, externally measured exterior orientation data are not usually as accurate as photogrammetrically determined orientation data using reference points. For aerial applications, directly measured orientation data may be introduced as appropriately weighted observations in order to bridge areas where there is a weak distribution of reference or tie points. The combination of direct sensor data with reference points is often known as integrated georeferencing. A similar situation can arise with mechanically guided camera systems where the exterior orientation is additionally supported by measured reference points (see section 6.9.5).

4.4.3.4 Free net adjustment

If no reference points or equivalent datum definitions are available, the problem of datum definition can be avoided by means of a *free net adjustment* which fits the network onto the initial coordinates of the unknown points. Initial values for new points are required in any case for the linearization of the correction equations. They can be generated by the procedures described in section 4.4.4.

The initial values of all unknown points (new object points and perspective centres) form a spatial point cloud. This point cloud can be transformed by three translations, three rotations and one scaling factor onto the photogrammetrically determined model of object points, without affecting the shape of the point cloud, i.e.

without any geometrical constraint. The photogrammetric observations are not influenced by this transformation.

A rank defect of 7 in the normal equation matrix is avoided if exactly 7 observations can be found which are linearly independent with respect to each other and to the other observations. This requirement is exactly fulfilled if the normal system of equations is extended by a matrix **B** with $d = 7$ rows and u columns, where u is the number of unknowns (see section 2.4.2.4):

$$
\underset{d,u}{\mathbf{B}} = \left[
\begin{array}{cccccc:ccccccc}
1 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 1 \\
0 & -Z_{01} & Y_{01} & r_{11} & r_{12} & r_{13} & \cdots & 0 & -Z_1^0 & Y_1^0 & \cdots & 0 & -Z_p^0 & Y_p^0 \\
Z_{01} & 0 & -X_{01} & r_{21} & r_{22} & r_{23} & \cdots & Z_1^0 & 0 & -X_1^0 & \cdots & Z_p^0 & 0 & -X_p^0 \\
-Y_{01} & X_{01} & 0 & r_{31} & r_{32} & r_{33} & \cdots & -Y_1^0 & X_1^0 & 0 & \cdots & -Y_p^0 & X_p^0 & 0 \\
X_{01} & Y_{01} & Z_{01} & 0 & 0 & 0 & \cdots & X_1^0 & Y_1^0 & Z_1^0 & \cdots & X_p^0 & Y_p^0 & Z_p^0
\end{array}
\right]
$$

$$\underbrace{}_{\text{exterior orientations}} \qquad\qquad \underbrace{}_{\text{new points } 1..p} \qquad\qquad (4.127)$$

If the scale of the photogrammetric network is known, the potential rank defect decreases to 6 and the last row of matrix **B** can be eliminated. Practically, scaling can be introduced as additional conditions from observed coordinate differences (e.g. eqn. 4.108) or other given scale information. Additional aspects about definition of reference scales are discussed in section 6.3.1.

The first six columns of matrix **B** are related to the unknown parameters of exterior orientation. They are not further discussed as the perspective centres are normally not of interest in the definition of the datum.

The next columns correspond to the unknown object points. Three condition equations are used for the translation of the system, included in the first three rows of matrix **B**:

$$
\begin{aligned}
dX_1 && +dX_2 && + \cdots &= 0 \\
dY_1 && +dY_2 && + \cdots &= 0 \\
dZ_1 && +dZ_2 & + \cdots &= 0
\end{aligned}
\qquad (4.128)
$$

If the sum of coordinate corrections at all object points becomes zero, the centroid of the initial points is identical to the centroid of the adjusted object points.

The next three rows of matrix **B** contain differential rotations of the coordinate system:

$$
\begin{aligned}
&& -Z_1^0 dY_1 & +Y_1^0 dZ_1 && -Z_2^0 dY_2 & +Y_2^0 dZ_2 & + \cdots &= 0 \\
Z_1^0 dX_1 && & -X_1^0 dZ_1 & +Z_2^0 dX_2 && -X_2^0 dZ_2 & + \cdots &= 0 \\
-Y_1^0 dX_1 && +X_1^0 dY_1 && -Y_2^0 dX_2 & +X_2^0 dY_2 & & + \cdots &= 0
\end{aligned}
\qquad (4.129)
$$

The final row 7 of matrix **B** defines scale:

$$X_1^0 dX_1 \quad +Y_1^0 dY_1 \quad +Z_1^0 dZ_1 \quad +X_2^0 dX_2 \quad +Y_2^0 dY_2 \quad +Z_2^0 dZ_2 \quad + \quad \cdots \quad = \quad 0 \qquad (4.130)$$

The extended system of normal equations then becomes:

$$\begin{bmatrix} \mathbf{A}^T\mathbf{A} & \mathbf{B}^T \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} \hat{\mathbf{x}} \\ \mathbf{k} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T\mathbf{l} \\ \mathbf{0} \end{bmatrix} \qquad (4.131)$$

Here **k** consists of seven Lagrange multipliers. The solution of the extended system of normal equations can be obtained from the pseudo inverse (Moore-Penrose inverse) **Q⁺**:

$$\begin{bmatrix} \hat{\mathbf{x}} \\ \mathbf{k} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T\mathbf{A} & \mathbf{B}^T \\ \mathbf{B} & \mathbf{0} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \mathbf{A}^T\mathbf{l} \\ \mathbf{0} \end{bmatrix} = \bar{\mathbf{N}}^{-1} \cdot \bar{\mathbf{n}} = \mathbf{Q}^+ \cdot \bar{\mathbf{n}} \qquad (4.132)$$

The pseudo inverse has the property that the resulting covariance matrix has a minimum trace:

$$\text{trace}\{\mathbf{Q}^+\} = \min. \qquad (4.133)$$

Hence, the standard deviations of the unknowns (object coordinates) are estimated with minimum quantities. The centroid of object points becomes the origin for the datum which is a fixed point with zero standard deviation. A datum is therefore defined which does not affect the total accuracy of the system.

4.4.3.4.1 Full trace minimization

If all unknown new points are used for datum definition, the full trace of the covariance matrix **Q⁺** is minimized, as explained above. All points therefore contribute to the datum definition but they are not considered free of error and do not have a priori standard deviations.

Fig. 4.71 shows an example of a network for which the datum is defined alternatively by three reference points and one distance (left, 3-2-1 method) and by free net adjustment (right, all points used as datum points). When reference points are used, error ellipses illustrate clearly that standard deviations are smaller for object points close to reference points than for those at the edges of the net. In contrast, error ellipses in the free net adjustment are significantly smaller and more homogenous. Their large semi-axes point towards the centroid of the points. The vertical lines in the ellipses indicate the error in Z, which behaves similarly to those in X and Y.
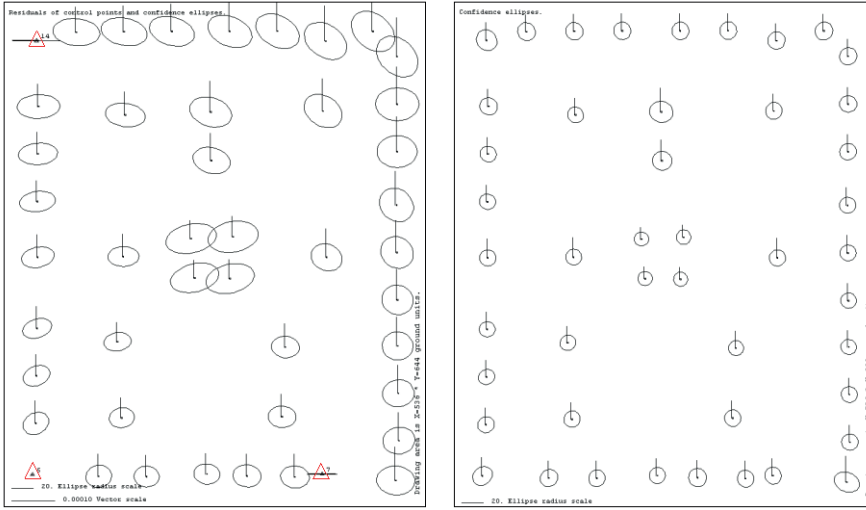
Fig. 4.71: Example of variation in 2D error ellipses estimated from a network based on reference points (left, red points) and by free net adjustment (right).

4.4.3.4.2 Partial trace minimization

There are some applications where not all the object points should be used for datum definition in a free net adjustment. This can occur, for example, when a subset of new points represents an existing network into which the remaining object points should be optimally fitted. In this case, those columns of matrix **B**, which relate to points not used in datum definition, must be set to zero. However, the rank of **B** must not be smaller than $u-d$.

$$
\mathbf{B}_{d,u} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 1 \\
0 & -Z_1^0 & Y_1^0 & 0 & 0 & 0 & \cdots & 0 & -Z_i^0 & Y_i^0 & \cdots & 0 & -Z_p^0 & Y_p^0 \\
Z_1^0 & 0 & -X_1^0 & 0 & 0 & 0 & \cdots & Z_i^0 & 0 & -X_i^0 & \cdots & Z_p^0 & 0 & -X_p^0 \\
-Y_1^0 & X_1^0 & 0 & 0 & 0 & 0 & \cdots & -Y_i^0 & X_i^0 & 0 & \cdots & -Y_p^0 & X_p^0 & 0 \\
X_1^0 & Y_1^0 & Z_1^0 & 0 & 0 & 0 & \cdots & X_i^0 & Y_i^0 & Z_i^0 & \cdots & X_p^0 & Y_p^0 & Z_p^0
\end{bmatrix}
$$

$$
\underbrace{\hspace{3cm}}
$$

new points   non-datum   new points   (4.134)
points

Eqn. (4.134) shows a modified matrix **B** where all those elements are eliminated which are related to the perspective centres.

4.4.3.4.3 Practical aspects of free net adjustment

With free net adjustment, as with the unconstrained datum definition using reference points, the photogrammetric network is not influenced by possible inconsistencies between reference points. The object coordinate residuals are only affected by the photogrammetric observations and the quality of the model. A free net adjustment therefore provides an optimal precision that can be better analysed than standard deviations of unconstrained, or even over-determined, datum definitions.

The free net adjustment is therefore a very flexible tool if

– no reference points are available,
– only the relative positions of object points are of interest, or
– only the quality of the model is to be analysed, for example the model of interior orientation used in self-calibration.

However, the standard deviations of object points are not suitable for a direct assessment of accuracy. They only provide information about the internal quality of a photogrammetric network, i.e. they express how well the observations fit the selected model. Accuracy can only be properly assessed using comparisons with data of higher accuracy (see also section 4.4.5.4). The display of error ellipse or ellipsoids as in Fig. 4.71 is only useful when the required confidence intervals are based on eigenvalues and eigenvectors (section 2.4.3.5).

4.4.4 Generation of approximate values

The generation of approximate values, to be used as starting values or initial values in the iterative solution of a photogrammetric problem, is often a complex task. Approximate values are required for all unknowns to be estimated, i.e. all orientation parameters and all new points or tie points. Since arbitrary image configurations in arbitrarily oriented object coordinate systems may well occur in close-range applications, the manual calculation of approximate values is virtually impossible.

Fig. 4.72 depicts (within solid lines) that information which is necessary for the generation of approximate values required by the bundle adjustment, and (within dotted lines) information which is useful but optional. The key component is a module for the automated calculation of approximate values based on measured image coordinates, camera parameters and, if available, coordinates of reference points. Directly measured values for exterior orientation may also exist, e.g. from GNSS data for UAV or mobile mapping systems. This process is also known as *multi-image orientation*, whereby bundle triangulation is expressly not implied.
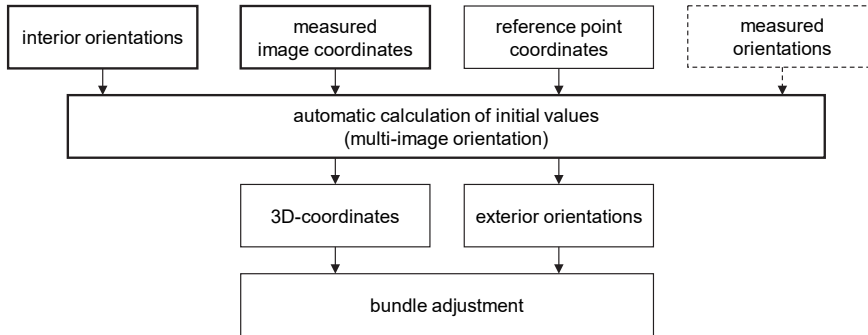
Fig. 4.72: Methods and procedures for the calculation of approximate values.

In many cases, additional information about the image configuration is available, such as surveyed orientation data, parameters derived from free-hand sketches or CAD models. Manual intervention in the procedures for calculating approximate values can sometimes be necessary and such additional information can support this manual process as well as helping to define the coordinate system. Approximate values can, however, also be generated fully automatically.

The following principal methods for the generation of approximate values can be identified. They can also be applied in combination:

– Automatic calculation of approximate values:

   In order to generate approximate values automatically, three strategies are feasible for complex photogrammetric configurations:

   – combined intersection and resection
   – successive forming of models
   – transformation of independent models

   All three strategies are based on a step by step determination of the parameters of exterior orientation, as well as the object coordinates, in a process where all images in the project are added sequentially to a chosen initial model.

– Generation of approximate values by automatic point measurement:

   Digital photogrammetric systems allow for the automatic identification and measurement of coded targets. Reference and tie points can be matched by this method and, by further applying the above orientation strategies, a fully automatic generation of approximate values is possible.

– Direct measurement of approximate values:

   Approximate values of object points, especially of imaging stations and viewing directions, can be measured directly, for example by survey methods. It is also possible to use separate measuring equipment for the location of camera stations, such as GNSS or inertial navigation units, as these are now increasingly incorporated into UAV or mobile mapping systems.

4.4.4.1 Strategies for the automatic calculation of approximate values

In order to calculate approximate values automatically, the following information must be provided:

- a file containing camera data (interior orientation data which can be approximate);
- a file containing measured image coordinates (see Fig. 4.61);
- a file containing reference point coordinates (optional, if available) or other information for the definition of the object coordinate system;
- other known object information (optional, if available), e.g. approximately known exterior orientations.

On the basis of this information, an iterative process is started which attempts to connect all images via approximate orientations. At the same time all measured tie points can be approximately calculated in object space. The bundle adjustment can then be executed.

The following procedures show sample strategies which employ suitable combinations of various methods for orientation and transformation (resection, relative and absolute orientation, similarity transformation, intersection). In all cases, a reasonable starting model formed by two images is defined.

4.4.4.1.1 Starting model and order of calculation

From a multi-image configuration, a *starting model* is provided by one image pair for which relative orientation may be computed. The resulting model coordinate system provides an arbitrary 3D coordinate system for including all subsequent images or models. The choice of starting model, and processing sequence of subsequent images, is not arbitrary and is critical to the success of the automatic calculation of approximate values, especially for complex and irregular image configurations.

Theoretically, n images of a multi-image configuration lead to $n(n-1)/2$ possible models. Image pairs with fewer than 5 homologous points would, of course, be eliminated. Quality criteria can be calculated for each possible model and the selected starting model should have the following properties:

- Number of tie points:
  A large number of tie points leads to a more stable relative orientation where possible gross errors (outliers) can eliminated more easily.

  starting model: maximum number of tie points
- Accuracy of relative orientation:
  The standard deviation of unit weight of the relative orientation (s0) should represent the expected image measuring accuracy.

  starting model: minimum s_0
- Mean intersection angle at model points:
  The mean intersection angle of homologous image rays provides information about the quality of a computed relative orientation. Models with small mean

intersection angles, e.g. less than 10°, have an unfavourable height-to-base ratio and should not be used as starting models.

　　starting model: mean intersection angle close to 90°

– Mean residuals of model coordinates:

The mean intersection offset of homologous image rays which are skew is a quality measure of model coordinate determination.

　　starting model: minimum mean intersection offset

– Number of gross errors in relative orientation:

Models with few or no blunders are preferable for orientation.

　　starting model: no blunders

– Image area covered by tie points:

　　starting model: maximum image area

By giving appropriate weight to these criteria an optimal starting model, as well as a sorted list of further models in order of calculation, can be selected. In general, the model with maximum tie points and best mean intersection angle is a suitable starting model. An unfavourable starting model, chosen without regard to these criteria, can cause the iterative orientation procedure to diverge.

4.4.4.1.2 Combination of space intersection and resection

Fig. 4.73 shows the principal steps in generating approximate values by combined space intersections and resections. A starting model is first selected according to the criteria outlined in the preceding section.

This starting model is used to calculate a relative orientation. Subsequent images are oriented to this model by space resection, provided they have at least 3 spatially distributed tie points with model coordinates known from a previous relative orientation. Model coordinates of new unknown object points are then calculated by intersection. When additional points are calculated, relevant images can be oriented again by resection in order to improve their exterior orientation in the model system. The image configuration is iteratively stabilized in this process since the number of intersected model points continually increases, thus improving the orientation parameters calculated by resection.

Once all images have been oriented with respect to the model coordinate system, a final absolute orientation (similarity transformation) can be performed. For this purpose, reference points with known coordinates in both model and object coordinate systems are used. The final result delivers approximate exterior orientation values of all images and approximate coordinates for all 3D object points.
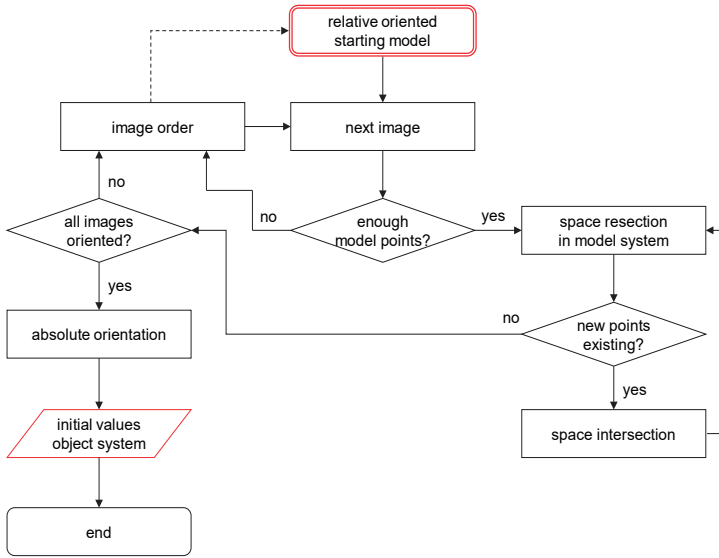
Fig. 4.73: Generation of approximate values with combined space intersection and resection.
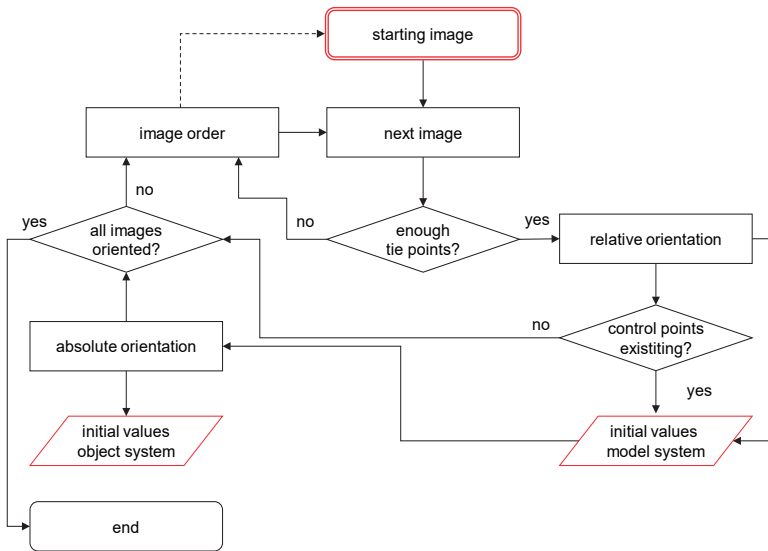
Fig. 4.74: Initial value generation with successive model creation.

4.4.4.1.3 Successive creation of models

Fig. 4.74 illustrates the process of initial value generation by successive creation of models. A relative orientation is first calculated for a suitable starting model. If this

model contains enough reference points, an absolute orientation can immediately be performed. If not, overlapping photos can be successively oriented to the initial model if they contain a sufficient number of homologous points. Again, if a set of models contains enough reference points it can be absolutely oriented. In this way approximate values of unknown object points in a model or object coordinate system can be computed. At the end of the process the parameters of exterior orientation can be derived from the parameters of relative and absolute orientation.

4.4.4.1.4 Transformation of independent models

The process of initial value generation by transformation of independent models is illustrated in Fig. 4.75. A connection matrix of all image pairs (models) is first established. For each possible model a relative orientation is calculated and stored together with corresponding model coordinates and quality estimators.

The starting model is selected according to the following criteria:
a) a maximum number of tie points,
b) appropriate intersection angles and
c) minimum intersection offsets.
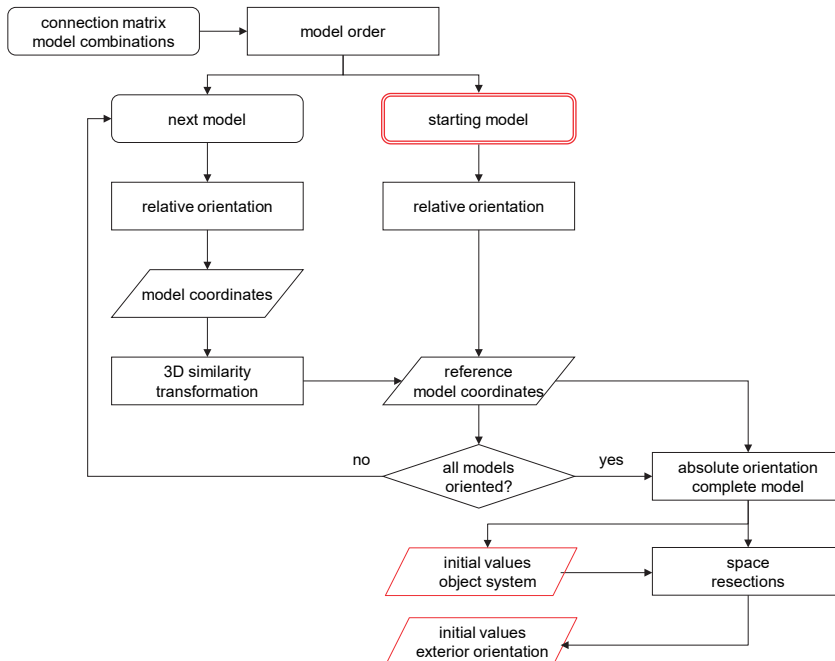
Fig. 4.75: Initial value generation by transformation of independent models.

The model coordinates of the relatively oriented starting model are used to define a local coordinate system. All other relatively oriented models are subsequently transformed into this local system using their independent model coordinates as input to a 3D similarity transformation.

When all points in all models are calculated with respect to the local system, a final absolute transformation of the complete model into the object coordinate system is calculated using reference points with known object coordinates. As a result, approximate values of object points in the object coordinate system are generated.

Finally, the exterior orientations of all images are calculated by space resection (single image orientation) using the object coordinates computed above. If any object points remain without initial values, these can be provided by spatial intersection.

4.4.4.2 Initial value generation by automatic point measurement

Fig. 4.76 shows one photo of a multi-image configuration which has been taken to measure a set of targeted points. Targets with coded point numbers have been placed on several points. These can be automatically identified and decoded by the image processing system (see section 6.2.1.4). The majority of points are marked by standard, non-coded targets. In addition, a local reference tool (front right) is placed in the object space. It consists of a number of coded targets with calibrated local 3D object coordinates. The reference tool need not be imaged in all photos. The example above also shows a reference scale bar which provides absolute scale but is not relevant to initial value generation.
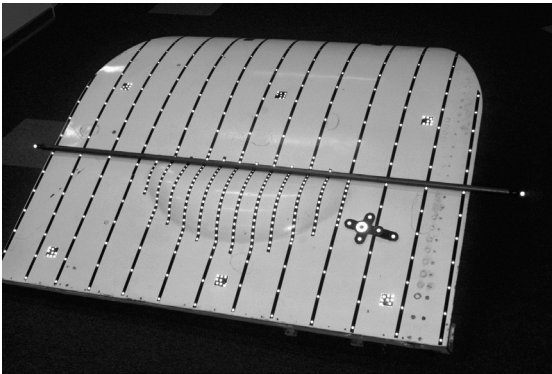
Fig. 4.76: Image with artificial targets and local reference system (GSI).

Approximate values for the freely configured set of images can be generated as follows (Fig. 4.77): A pattern recognition process detects all coded targets and other potential object points. Those photos in which the reference tool is imaged can be individually oriented by space resection into the tool's coordinate system. Remaining

images are oriented by relative orientation using the coded targets and the strategies described above for calculation of approximate values. Object coordinates for coded targets are calculated by intersection. A first bundle adjustment generates improved object coordinates and orientation parameters. Remaining photos, which are not at this stage oriented, are iteratively integrated by resection and bundle adjustment until all images are oriented.

A subsequent processing stage identifies and consecutively numbers all non-coded targets using a matching process based on epipolar geometry (see section 5.5.4). A final bundle adjustment provides the coordinates of all object points. This and similar methods form part of digital online and offline measuring systems (see section 6.6). Initial value generation and precise coordinate determination are integrated in one common procedure.
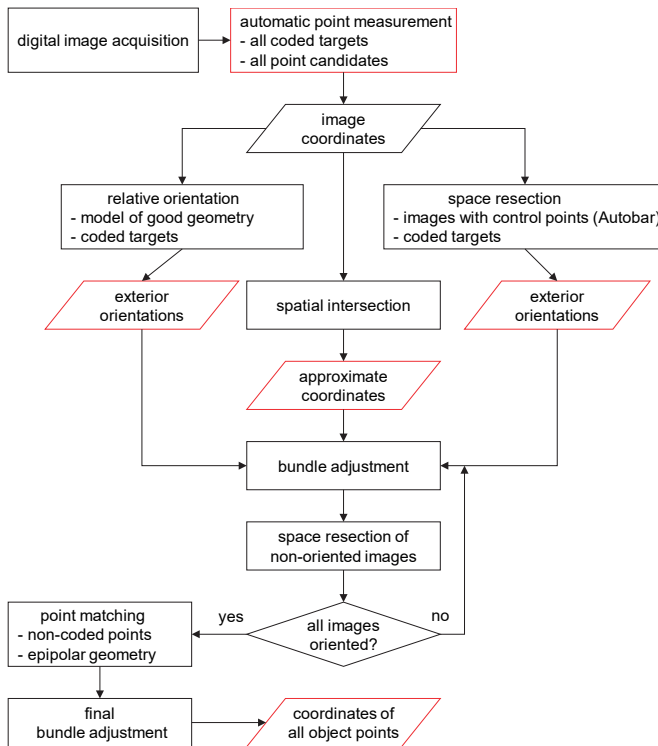
Fig. 4.77: Fully automated generation of initial values and orientation.

4.4.4.3 Practical aspects of the generation of approximate values

The automatic calculation of approximate values (multi-image orientation) is often a time-consuming process for the complex image configurations associated with close-

range photogrammetry. The stability of a multi-image project depends mainly on the distribution of object points and on the configuration of bundles of rays. If there are no, or only coarse, approximate values of orientation parameters, then even a very few gross data errors can lead to divergent solutions for initial values. Since the imaging configuration is then poorly defined, the detection of gross errors is more difficult.

As a result, effective systems for generating initial values should incorporate the following features which increase the level of automation and the scope for error detection.

– Use of algebraic rotation matrices:

Modules for relative orientation, spatial similarity transformation (absolute orientation) and space resection should use rotation matrices based on algebraic functions (quaternions) instead of trigonometric functions. The use of trigonometric functions in the rotation matrices can lead to singularities or ambiguities. In addition, algebraic definitions improve the convergence of solutions (see section 2.2.2.2).

– Robust blunder detection:

The detection of gross data errors should be sufficiently robust that more than one blunder can be detected and eliminated more or less automatically. Estimation methods based on RANSAC are particularly suitable for this purpose (see section 2.4.4.6).

– Automatic and interactive definition of the order of images:

With a suitable starting model, the order of image orientation can be determined automatically. However, situations occur where the suggested order does not lead to convergence, for example due to images which cannot be controlled by other images. In these cases, an interactive definition of image order should be possible.

– Manual activation and deactivation of points and images:

During the process of initial value generation, it can be necessary to deactivate faulty points or images with weak geometry in order to provide an initial image set which can be oriented. Once a sufficient number of images have been oriented, the deactivated images can then be successively added.

– Manual input of approximate values:

It is efficient to use any additional information about the object or the image configuration which may be available, e.g. from additional or previous measurements. In particular, the principal distance should be known in advance, for example approximated by the focal length.

– Approximate values in SfM approaches:

Approaches based on structure-from-motion (SfM, section 5.5.2.2) are particularly robust when there are significant image overlaps and the object surface has unique textures. For calculation of initial values usually a RANSAC-based method is applied that selects five random points from the high number of

candidate tie points. These usually contain a large number of outliers. Through successive adding of additional overlapping images exterior orientations and 3D coordinates of tie points in the model coordinate system are calculated.

In principle, photogrammetric measuring systems capable of automatic measurement and identification of image points, generate significantly fewer gross errors than interactive systems. With appropriate image configurations these systems therefore provide a fully automatic procedure, from the calculation of initial values through to the final result of the bundle adjustment.

4.4.5 Quality measures and analysis of results

4.4.5.1 Output report

Typical bundle adjustment programs report on the current status of processing including parameters such as number of iterations, corrections to unknowns and error messages. When the program is complete, an output report is generated which summarizes all results. It should contain the following information:
– list of input files and control parameters, date, project description;
– number of iterations and standard deviation of unit weight s_0;
– list of observations (image measurements) including corrections, reliability numbers and test values, sorted by images;
– mean standard deviations of image coordinates, sorted by image and divided into x' and y' values;
– list of blunders detected and eliminated;
– list of reference points;
– list of adjusted object points (new points) with standard deviations;
– mean standard deviations of new points, divided into X, Y and Z values;
– maximum corrections with (numeric) identifiers of the corresponding points;
– parameters of interior orientation with standard deviations;
– correlations between the parameters of interior orientation;
– parameters of exterior orientation with standard deviations;
– correlations between the parameters of interior and exterior orientation;
– list of additional (survey) observations with standard deviations.

4.4.5.2 Sigma 0 and reprojection error

A first quality check is given by the a posteriori standard deviation of unit weight (sigma 0, $\hat{s}_0$). For a bundle adjustment based on a functional model without systematic errors and observations without outliers, $\hat{s}_0$ should have a value similar to the expected image measuring precision (see below). Equations (2.209) and (2.226) show clearly that $\hat{s}_0$ describes the mean error after back-projection of the adjusted unknowns (points) into the image plane (reprojection error). It represents an

approximate estimation of accuracy only under the conditions stated and does not replace a final quality check using independent references (scale bars, artefacts) which have not previously been included in the bundle adjustment. The division by n in eqn. (2.226) suggests that $\hat{s}_0$ can be "arbitrarily" reduced by increasing the number of observations. However, the quality of object reconstruction is not necessarily improved in this way.

4.4.5.3 Precision of image coordinates

The precision of image coordinates is calculated from the cofactor matrix:

$$\hat{s}_i = \hat{s}_0 \sqrt{q_{ii}} \qquad (4.135)$$

with q_{ii} the principal diagonal elements of matrix $\mathbf{Q}_{\hat{l}\hat{l}}$ (see equations 2.225 ff).

If the bundle adjustment is calculated using only equally weighted photogrammetric image coordinates as observations, $\hat{s}_0$ should represent the accuracy of the instrumental combination of camera and image measuring device.

a) With systematic effects   b) Without systematic effects
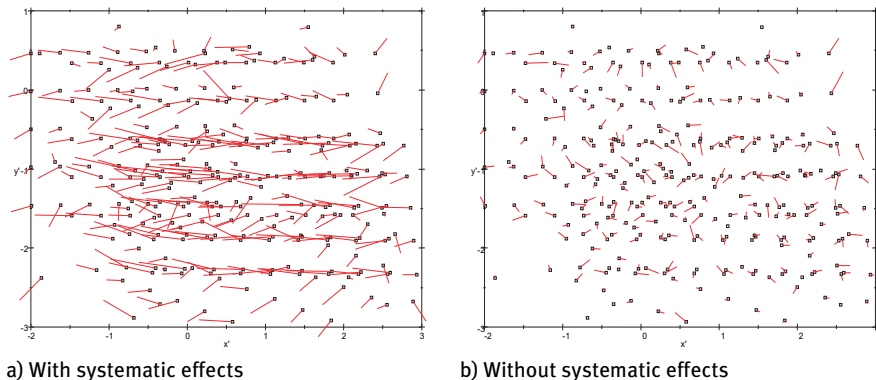
Fig. 4.78: Residuals of image coordinates.

The standard deviations of the measured image coordinates should be similar in both x and y directions. Different values imply that the measuring device or the camera generates systematic errors. This may occur, for example, with digital images generated by an image scanner whose mechanical construction is different in x and y. Bundle adjustment programs with integrated estimation of variance components (see section 2.4.4.3) allow different weighting for separate groups of observations in order to process observations according to their importance or precision.

A graphical analysis of the distribution of residuals should, in the end, show no systematic errors Fig. 4.78). However, a rigorous analysis of point accuracy in object space is not possible with this information.

4.4.5.4 Precision of object coordinates

Usually, the precision of adjusted points is of major importance for the quality analysis of a bundle adjustment. The analysis should consider two criteria (example in Fig. 4.79):

– Root mean square error (RMSE):
  The root mean square error is a measure of the general precision level of the adjustment. Taking the mean image measuring accuracy sx'y' and the mean image scale m into account, the equation for accuracy estimation can be checked (see sections 3.3.1 and 7.2.1):

$$s_{XYZ} = q \cdot m \cdot s_{x'y'} \tag{4.136}$$

  The design factor q reflects the accuracy of the imaging configuration with typical values between 0.7 and 1.5.

– Maximum residuals of single points:
  In contrast to the mean standard deviation, maximum residuals indicate the loss of precision which can be expected at problem points or unfavourable areas within the image configuration. If object coordinates are to be used for further calculation or analysis, then the maximum residuals should stay within specified precision limits.

Both the above quality criteria can be used only to analyse the statistical precision of the photogrammetric procedure. It is necessary here to take into account whether the object coordinate datum was defined without constraints (by minimum number of reference points or by free net adjustment), or if inconsistencies in reference points could be influencing the accuracy of object points.

If image observation accuracy is homogeneous, accuracy differences in object points are mainly caused by:

– different image scales or camera/object distances;
– different numbers of image rays per object point;
– different intersection angles of image rays;
– different image quality of object points;
– variations across the image format, e.g. lower quality imaging near the edges where distortion is often less well determined;
– inconsistencies in reference points.

The true accuracy of a photogrammetric project can be estimated only by comparing photogrammetrically determined points or distances with reference values measured independently to a higher accuracy. However, a rigorous evaluation is possible only with independent reference points which have not already been used as reference points or for reference distances in the bundle adjustment. Only through independent control will all properties of a photogrammetric system become visible and a rigorous accuracy assessment become possible. Suggestions for the verification of photogrammetric systems are further discussed in section 7.2.3.

```
Point Nr. |    x     |     y     |     z    |  sx   |  sy   |   sz    |  Rays
      2     1027.6098     56.8830    34.1252   0.0063   0.0054   0.0075       29
      3     1093.4013    143.4041    30.7408   0.0065   0.0052   0.0077       29
      4     1179.2012    278.0186     0.8845   0.0070   0.0054   0.0090       24
      5      787.0394    635.6436     8.3199   0.0056   0.0053   0.0085       27
      6      367.3523   1010.0567    11.0867   0.0055   0.0066   0.0089       23
      7      240.6758    932.3660    48.2087   0.0049   0.0059   0.0073       32
      8      151.1113    856.2551    60.1397   0.0050   0.0056   0.0068       33
      9      -26.2640    664.4885    65.5948   0.0058   0.0052   0.0068       29
     10     -133.1238    567.6979    57.3798   0.0062   0.0048   0.0065       31
     11     -272.4712    447.0657    48.6376   0.0069   0.0047   0.0065       31
     12     -429.7587    258.2250     7.6041   0.0079   0.0056   0.0080       22
     13     -214.8574     73.6836    11.2978   0.0065   0.0051   0.0064       34
    ...         ...         ...        ...       ...      ...      ...        ..
   1148      624.8693    506.9187    34.8052   0.1658   0.2092   0.4223        4
   1150     -347.0995    557.8746    -5.0205   0.0137   0.0114   0.0250        4

                  Minimum standard deviation:   0.0044   0.0043   0.0060
                  Maximum standard deviation:   0.1658   0.2092   0.4223
                  RMS of standard deviation:    0.0144   0.0165   0.0325
```

Fig. 4.79: Report file showing adjusted object coordinates (Ax.Ori).

4.4.5.5 Quality of self-calibration

The adjusted interior orientation parameters, and their correlations, should be carefully examined if the image acquisition system has been calibrated simultaneously in the bundle adjustment.

Fig. 4.80 is a part of a calculation report showing the adjusted interior orientation data, associated standard deviations and a matrix of correlations between the parameters. The following comments can be made about individual parameters:

– Principal distance c:

The value of c normally corresponds approximately to the focal length. When plane surfaces are imaged without oblique views, the principal distance cannot be uniquely determined and either the value of c is unreasonable or its corresponding standard deviation is higher than expected. In this case, the principal distance is often highly correlated with the exterior orientation in viewing direction. However, this does not mean that the object points are determined to lower accuracy. If the results of the bundle adjustment are intended to calibrate a camera which is subsequently used on other projects, then the principal distance must be calculated accurately and with a standard deviation of order of the image measuring accuracy.

– Principal point x'_0, y'_0:

The principal point normally lies very close to the foot of the perpendicular from the projection centre to the focal plane. When plane surfaces are imaged without oblique views, the position of the principal point cannot be uniquely determined; in this case, however, this does not mean that the object points are determined to lower accuracy. The importance attached to an accurate knowledge of the principal point will depend on the configuration of the network. If the results of the self-calibration are intended to be subsequently used on other projects, then

the position of the principal point must be calculated accurately and with a standard deviation of order of the image measuring accuracy.

It should be remembered that the definition of radial distortion depends on the location of the principal point. A large shift of the principal point may result in irregular parameters for radial distortion. An iterative pre-correction of measured image coordinates is recommended in those cases.

– Radial (symmetric) distortion A_1, A_2, A_3:

The parameters of radial distortion are normally the most effective additional parameters. Their related standard deviations should be much smaller than the parameters themselves. Parameter A_2 is often significantly correlated with A_1 (as shown in Fig. 4.80). However, this does not necessarily affect the overall result, especially for the object coordinates. Parameter A_3 can normally be determined to a significant value only in special cases, for example when fisheye lenses are employed.

– Tangential (decentring) distortion B_1, B_2, affinity and shear C_1, C_2:

The statements concerning A_1 and A_2 can also be applied to the optional additional parameters B_1, B_2, C_1, C_2. For many digital cameras, departures from orthogonality and equality of scale between the axes of the image coordinate system, e.g. due to imaging elements which are not square, are barely detectable. However, cameras utilizing analogue data transfer and a frame grabber should definitely use these parameters.

```
Camera/R0:                  1          8.79576
C       :         -24.1707    0.00029584
x0      :        -0.0662334   0.000368011
y0      :         -0.103883   0.000332042
A1      :      -0.000162606   1.75349e-007
A2      :       2.5697e-007   2.64915e-009
A3      :      4.65242e-011   1.17937e-011
B1      :     -1.64762e-006   1.94141e-007
B2      :      4.10302e-006   1.81828e-007
C1      :      -0.000117619   3.92306e-006
C2      :     -1.28944e-005   3.28251e-006

Correlations:
C    1.000
X0   -0.009   1.000
Y0   -0.054  -0.002   1.000
A1    0.030  -0.006   0.005   1.000
A2    0.029   0.001  -0.019  -0.947   1.000
A3   -0.025   0.004   0.023   0.885  -0.980   1.000
B1   -0.004   0.894   0.015  -0.006   0.010  -0.008   1.000
B2   -0.022  -0.023   0.798  -0.004   0.011  -0.012  -0.009   1.000
C1    0.406  -0.000  -0.056   0.025   0.094  -0.077   0.015  -0.005   1.000
C2    0.011  -0.006  -0.010   0.037  -0.010  -0.003  -0.002  -0.002   0.031   1.000
        C       x0      y0      A1      A2      A3      B1      B2      C1      C2
```

Fig. 4.80: Result of a self-calibration with correlation between the parameters.

Correlations between the parameters of interior orientation and exterior orientation should be avoided by an appropriate imaging configuration (see section 7.3.2). They

do not need to be considered for point determination if the calculation is performed in a single stage mathematical process. However, if single parameters are extracted for use in further external calculations, e.g. orientation parameters are applied in separate spatial intersections, correlations, or more exactly the variance-covariance matrix, cannot be further taken into account, which leads to mistakes in the functional model.

4.4.6 Strategies for bundle adjustment

In many practical applications, either the generation of initial values or the complete bundle adjustment may not run successfully at the first attempt, or the final result might not meet the specification. To avoid these problems, a number of practical tips and suggestions for strategies and procedures in bundle adjustments is given below.

4.4.6.1 Simulation

Simulation of the imaging configuration is one method of project planning (see section 7.1.4). Simulation provides a priori accuracy estimation for optimizing an imaging configuration without real measurements. For this purpose, the measuring object must be represented by simulated object points similar to the actual measurement in terms of number and distribution. The required 3D coordinates can be provided by manual input if the object is not too complex. Alternatively, the coordinates can be obtained from a CAD model or previous measurements (Fig. 4.81).
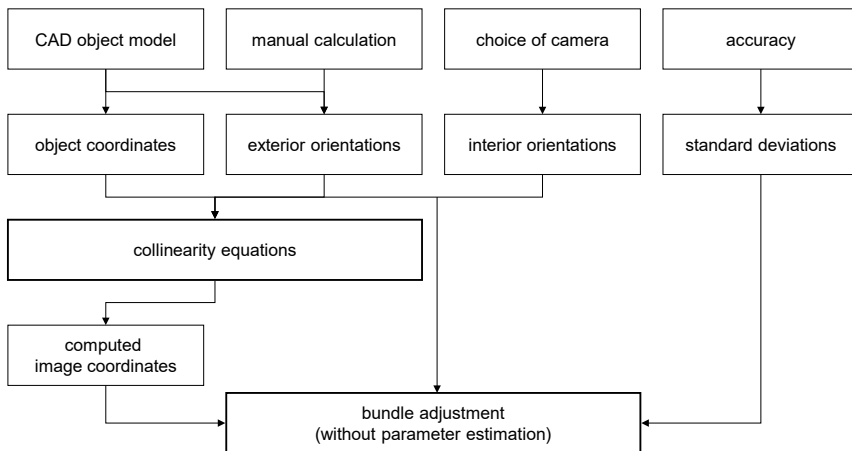
Fig. 4.81: Simulation by bundle adjustment.

The a priori definition of camera stations and viewing directions (exterior orientations) is much more complicated. While orientation data can easily be generated for regular image configuration patterns (example in Fig. 4.58), the measurement of complex structures can often be configured only on site. On the one hand, selection of image configurations is more flexible but, on the other hand, unforeseen problems can often occur in the form of occlusions or restricted camera stations.

Camera data and a priori accuracies must also be defined. Image coordinates can then be calculated using collinearity equations and the simulated object coordinates and orientation data. Using these simulated data as input, a bundle adjustment can be calculated. Here it is only necessary to compute the covariance matrix of the unknowns and the standard deviations of interest, instead of a complete parameter estimation. Now object points, imaging configurations, selection of cameras and accuracies can be varied until a satisfactory adjustment is achieved. By applying the Monte-Carlo method (section 7.1.4.2), the input data can be altered within specific noise ranges based on normal distribution. The computed output values will vary as a function of the noisy input data.

4.4.6.2 Divergence

Bundle adjustments which do not converge are often a serious problem in practical applications. A major reason for this is that standard statistical methods of error detection only work well if at least one iteration has been successfully calculated. Divergence in adjustments can be caused by:

– Faulty input data:
  Error in data formats, units of measurement, typing errors etc. should be detected by the program but often are not.
– Poor initial values:
  Poor initial values of the unknowns lead to an inadequate linearization of the functional model.
– Gross errors in the input data:
  Errors in identifying point or image numbers and measuring errors larger than the significance level of the total measurement. They can be detected in part by statistical tests if robust estimation methods are applied (see section 2.4.4).
– Weak imaging geometry:
  Small intersection angles, a small number of rays per object point and/or poor interior orientation data lead to poorly conditioned normal equations and a lower reliability in the adjustment.

The following steps should therefore be applied to handle divergent bundle adjustments:

1) Check of input data.
2) Controlled program abortion after the first iteration, with checking of differences between initial values and "adjusted" unknowns (see section 2.4.4.1) – high deviations indicate problematic input data.
3) Pre-correction of image coordinates by known distortion values.
4) Adjustment without camera calibration in the first run; subsequent adding of additional parameters (distortion) when blunders have been eliminated from the input data.
5) Check of the geometric configuration of images where gross errors have been detected as the smearing effect of least-squares solutions can lead to misidentification of blunders.

4.4.6.3 Elimination of gross errors

Generally speaking, the bundle adjustment is very sensitive to gross errors (blunders) in the measured data. In complex imaging configurations, gross errors arise easily due to false identification of object points or mistakes in image or point numbering. In contrast, pure measuring errors occur relatively infrequently. The detection of gross errors can fail, especially in geometrically weak configurations where statistical tests based on redundancy numbers are not significant (see section 2.4.4). This is of major importance if gross errors in observations occur at *leverage points* which have a strong influence on overall geometry but which cannot be controlled by other (adjacent) observations.

Most bundle adjustment programs permit both manual and automatic elimination of blunders. In both cases only one blunder should be eliminated per program run, usually the one with the largest normalized correction. The adjustment program should allow the corresponding observation to be set as inactive, not deleted, in order that it can be reactivated later in case it is discovered to be correct.

If observations are eliminated, the corresponding object area and image should be analysed. Uncontrolled elimination of observations can lead to weak imaging geometries if the object is recorded by a small number of images, or represented by only a few object points. A system that is based on the few remaining observations cannot be controlled by a rigorous blunder test and may produce a plausible result even if gross errors are still present.

4.4.7 Multi-image calculation of points and geometric elements

This section deals with analytical methods for object reconstruction based on measured image coordinates from an unlimited number of oriented photos. Digital, multi-image methods which additionally process grey values at image points are discussed in Chapter 5.

The following methods require known parameters of interior and exterior orientation which are normally calculated by a bundle triangulation (see section 4.4). In addition, image coordinates must be reduced to the principal point and corrected for distortion (see section 3.3.2). On this basis, object points, surfaces and basic geometric elements can be determined.

4.4.7.1 General spatial intersection

The general spatial or forward intersection takes measured image coordinates from multiple images, together with their known orientation parameters, and calculates the spatial point coordinates X,Y,Z. The calculation is based on the collinearity equations (4.10) used as observation equations in a least-squares adjustment:

$$x'_i + vx'_i = F(X_{0j}, Y_{0j}, Z_{0j}, \omega_j, \varphi_j, \kappa_j, x'_{0k}, z'_k, \Delta x'_k, X_i, Y_i, Z_i)$$
$$y'_i + vy'_i = F(X_{0j}, Y_{0j}, Z_{0j}, \omega_j, \varphi_j, \kappa_j, y'_{0k}, z'_k, \Delta y'_k, X_i, Y_i, Z_i)$$

(4.137)

where

i: point index
j: image index
k: camera index

In order to calculate the three unknowns X_i, Y_i, Z_i at least three observations (image coordinates) are required. Two images already provide a redundancy of 1 and with each additional observation the redundancy increases by 2.

To set up the normal equations, the derivatives of the unknown object coordinates are calculated according to (4.100).

A global measure of the quality of point determination is given by the shortest distance between the two skew rays (see section 2.3.2.1). If the point accuracy is to be analysed separately for each axis, the covariance matrix must be evaluated. Given the cofactor matrix of unknowns

$$\mathbf{Q}_{\hat{x}\hat{x}} = \begin{bmatrix} q_{XX} & q_{XY} & q_{XZ} \\ q_{YX} & q_{YY} & q_{YZ} \\ q_{ZX} & q_{ZY} & q_{ZZ} \end{bmatrix}$$

(4.138)

the standard deviations of adjusted point coordinates are as follows:

$$\hat{s}_X = \hat{s}_0 \cdot \sqrt{q_{XX}} \qquad\qquad \hat{s}_Y = \hat{s}_0 \cdot \sqrt{q_{YY}} \qquad\qquad \hat{s}_Z = \hat{s}_0 \cdot \sqrt{q_{ZZ}}$$

(4.139)

Hence the mean point error is given by:

$$\hat{s}_P = \sqrt{\hat{s}_X^2 + \hat{s}_Y^2 + \hat{s}_Z^2}$$

(4.140)

It should be noted that these standard deviations depend on the datum definition (origin of the object coordinate system). Spectral decomposition of the corresponding

variance-covariance matrix (see eqn. 2.251) leads to eigenvalues λ_i and eigenvectors $\mathbf{s}_i$. It is then possible to calculate the error or confidence ellipsoid which contains the "true" point with a confidence level of $1-\alpha$ (see section 2.4.3.5). The size and orientation of the confidence ellipsoid is independent of the coordinate system.
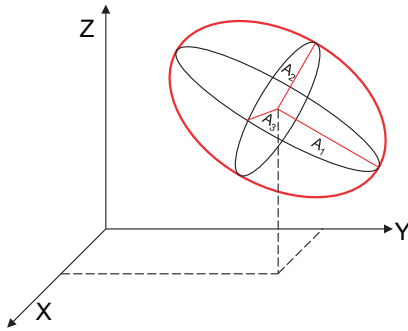
Fig. 4.82: Confidence ellipsoid.

The directions of the semi-axes of the ellipsoid are defined by the eigenvectors and their length is given by the eigenvalues (Fig. 4.82).

good intersection geometry   poor intersection geometry
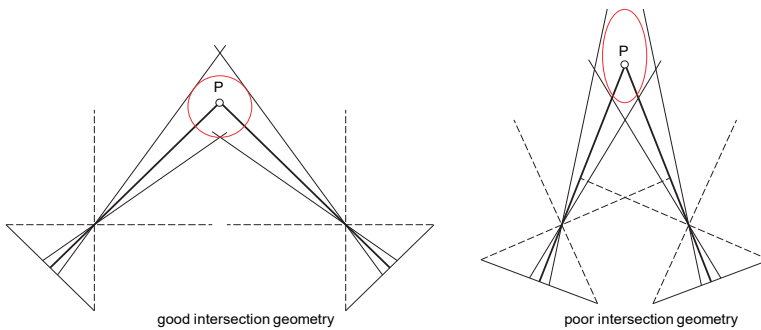
Fig. 4.83: On the geometry of spatial intersection.

In contrast to the standard bundle adjustment, possible correlations between adjusted point coordinates and orientation parameters are not taken into account in the spatial intersection. Assuming image coordinates of equal accuracy in all images, a stretched error ellipsoid indicates a weak intersection of homologous rays (Fig. 4.83).

Fig. 4.84 shows the confidence ellipsoids (in XY view) resulting from the spatial intersection of object points in two cases. Based on the image set shown in Fig. 4.57,

the first case (a) shows the points observed by two adjacent images and the second case (b) shows the points observed by five images surrounding the points. It is obvious that in case (a) the ellipses grow in size as their distance to the cameras increases and that the measuring precision in the viewing direction is lower than in a direction parallel to the images. When multiple images surround the area of interest as in (b) the ellipses are all generally smaller and rounder.

a) From two adjacent images   b) From five surrounding images
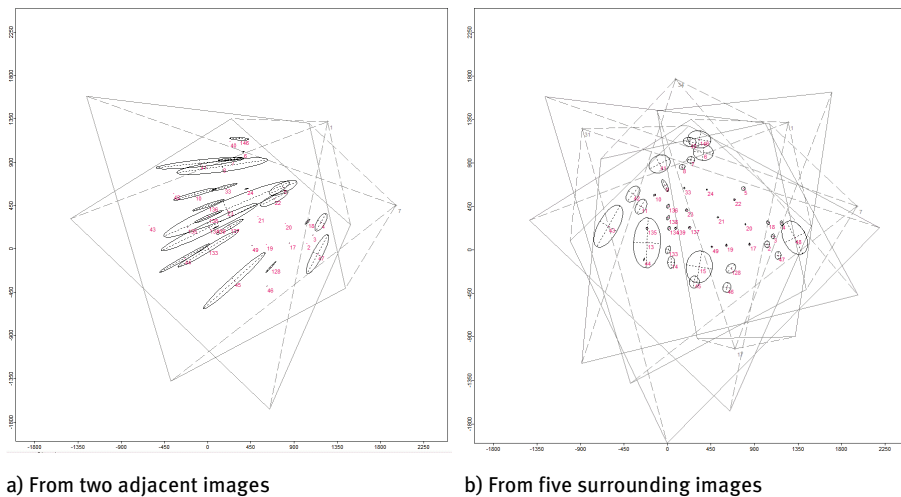
Fig. 4.84: Confidence ellipses for spatial intersection.

4.4.7.2 Direct determination of geometric elements
The direct determination of geometric elements can also be done photogrammetrically. Geometric 3D elements (straight line, circle, sphere, cylinder etc.) can be determined in two ways:
- calculation of best-fit elements from measured 3D coordinates;
- best-fit adjustment of elements to measured 2D contours in multiple images.

The first method is preferred in conventional coordinate metrology. The required 3D coordinates are delivered by coordinate measuring machines (CMM) from direct probing of the object surface. The calculation of best-fit elements, and the intersection of these elements, is of major importance because most industrial objects offered for inspection contain regular geometric elements. The most important algorithms are discussed in section 2.3.

The second method[5], contour measurement, is based on the idea that the imaged edges of geometric elements generate unique grey level contours which can be

---

5 The following derivations are due to Andresen (1991).

extracted by suitable edge operators or manual measurements (see section 5.2.4). Normally, there are no discrete or identifiable points along the imaged edge which would enable individual 3D points on the object surface to be measured, for example by intersection. However, if sufficient edge image points belonging to a common object surface can be detected in a well-configured set of images, an adjustment can be formulated for estimating the parameters of the unknown element. Assuming known parameters of interior and exterior orientation, each image point defines a light ray in space which, in principle, touches or is tangent to the surface of the element. For this purpose, a distance offset is defined between the light ray and the spatial element which is minimized by adjustment of the element's parameters.

The method can be used to calculate, for instance, straight lines, 3D circles, cylinders or other elements which generate an appropriate contour in the image. The number of images is unlimited. However, for a reliable determination, more than two images are usually required. The image coordinates used must be corrected in advance for principal point position and distortion. A detailed description of the individual methods can be found in earlier editions of this textbook.

4.5 Panoramic photogrammetry

Panoramic photogrammetry is a special branch of close-range photogrammetry that uses panoramic images instead of conventional perspective imagery. Panoramic imagery is created either by digitally stitching together multiple images from the same position (left/right, up/down) or by rotating a camera with conventional optics and either an area or line sensor, in a specially designed mounting fixture (see section 3.5.6). This section gives an overview of the basic panoramic imaging model, orientation methods and algorithms for 3D reconstruction.

4.5.1 Cylindrical panoramic imaging model

The most common method of panoramic photogrammetry is based on a cylindrical imaging model, as generated by numerous analogue and digital panoramic cameras or by a computational fusion of individual central perspective images. Assuming the camera rotation corresponds to a horizontal scan, the resulting panoramic image has central perspective imaging properties in the vertical direction only.

An image point P' can be defined either by the cylindrical coordinates r,ξ,η or by the Cartesian panoramic coordinates x,y,z (Fig. 4.85, see also section 2.2.5). The panorama is assumed to be created by a clockwise rotation when viewed from above. The metric image coordinates x',y' and the pixel coordinates u,v are defined within the cylindrical surface of the panorama, which is a plane when the cylinder is unrolled.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} r\cos\xi \\ -r\sin\xi \\ \eta \end{bmatrix} \tag{4.141}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} r\cdot\xi \\ z \end{bmatrix} \tag{4.142}$$
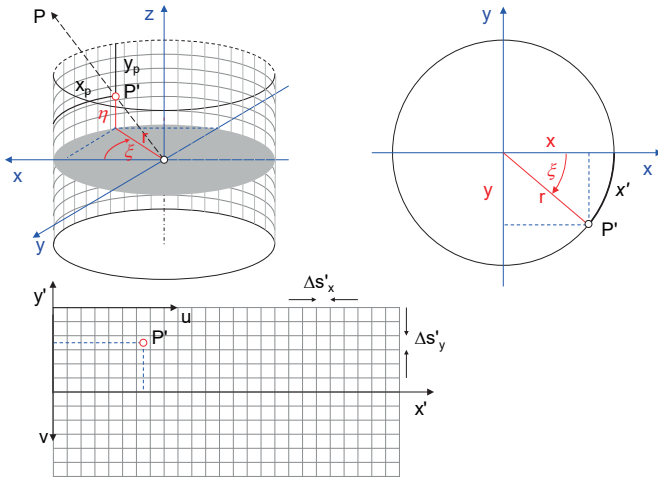
Fig. 4.85: Coordinate systems defining a cylindrical panorama.
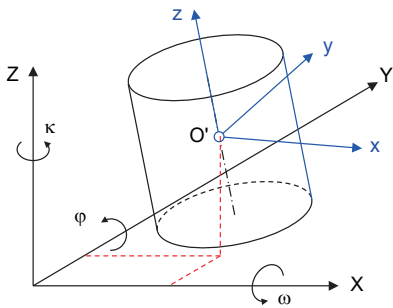
Fig. 4.86: Exterior orientation of cylindrical panorama.

The digital plane panoramic image has n_C columns and n_R rows with a pixel resolution $\Delta s'_x$ and $\Delta s'_y$. The circumference of the panorama is therefore $n_C \cdot \Delta s'_x$ and its height is $n_R \cdot \Delta s'_y$. The radius of the digital image is, in the ideal case, equal to the principal

distance of the camera. It can be calculated from the circumference or the horizontal angular resolution $\Delta\zeta$:

$$r = \frac{n_C \cdot \Delta s'_x}{2\pi} = \frac{\Delta s'_x}{\Delta\xi} \tag{4.143}$$

By introducing the parameters of exterior orientation, the transformation between object coordinate system XYZ and panoramic coordinate system xyz (Fig. 4.86) is given by:

$$\mathbf{X} = \mathbf{X}_0 + \frac{1}{\lambda}\mathbf{R}\cdot\mathbf{x} \tag{4.144}$$

Rearrangement gives

$$\begin{aligned}
\mathbf{x} &= \lambda\cdot\mathbf{R}^{-1}\cdot(\mathbf{X}-\mathbf{X}_0) \\
x &= \lambda\cdot[r_{11}(X-X_0)+r_{21}(Y-Y_0)+r_{31}(Z-Z_0)] = \lambda\cdot\overline{X} \\
y &= \lambda\cdot[r_{12}(X-X_0)+r_{22}(Y-Y_0)+r_{32}(Z-Z_0)] = \lambda\cdot\overline{Y} \\
z &= \lambda\cdot[r_{13}(X-X_0)+r_{23}(Y-Y_0)+r_{33}(Z-Z_0)] = \lambda\cdot\overline{Z}
\end{aligned} \tag{4.145}$$

where $\overline{X},\overline{Y},\overline{Z}$ define a temporary coordinate system that is parallel to the panorama system. With the scale factor

$$\lambda = \frac{r}{\sqrt{\overline{X}^2+\overline{Y}^2}} \tag{4.146}$$

the image coordinates in the unrolled (plane) panorama are given by

$$\begin{aligned}
x' &= r\cdot\arctan\left(\frac{\overline{X}}{\overline{Y}}\right) - \Delta x' \\
y' &= y'_0 + \lambda\cdot\overline{Z} - \Delta y'
\end{aligned} \tag{4.147}$$

or directly as collinearity equations between object coordinates and image coordinates:

$$\begin{aligned}
x' &= r\cdot\arctan\left(\frac{r_{12}(X-X_0)+r_{22}(Y-Y_0)+r_{32}(Z-Z_0)}{r_{11}(X-X_0)+r_{21}(Y-Y_0)+r_{31}(Z-Z_0)}\right) - \Delta x' \\
y' &= y'_0 + \lambda\cdot\left[r_{13}(X-X_0)+r_{23}(Y-Y_0)+r_{33}(Z-Z_0)\right] - \Delta y'
\end{aligned} \tag{4.148}$$

Here y'_0 denotes a shift of the principal point in the y direction and $\Delta x',\Delta y'$ are correction parameters for potential imaging errors in the camera.

The pixel coordinates u,v of a digital panoramic image can be derived as:

$$u = \frac{x'}{\Delta s'_x} \qquad\qquad v = \frac{n_R}{2} - \frac{y'}{\Delta s'_y} \tag{4.149}$$

4.5.2 Orientation of panoramic imagery

The method of determining the exterior orientations of one or more panoramic images is analogous to that for central perspective imagery; that is, by application of the panoramic collinearity equations to space resection and/or bundle adjustment. Both methods require suitable approximate values. Also, due to different system designs, the form of the rotation matrix must take account of the mechanical rotation orders and directions.

4.5.2.1 Approximate values

For the usual case of panoramas with an approximate vertical rotation axis, initial values for exterior orientation can easily be derived from reference points. The centre coordinates of the panorama X_0, Y_0 can be calculated by plane resection using three control points P_1, P_2, P_3 and the subtended angles $\Delta\xi_{12}$, $\Delta\xi_{23}$, which can be derived from the corresponding image points P'_1, P'_2 and P'_3 (Fig. 4.87 left).

Fig. 4.87 centre shows the general principal of plane resection. It is simple to show that a panoramic image subtending an angle α at two reference points, P_1, P_2, must lie on the arc of the circle shown (example points I_1, I_2). Fig. 4.87 right then shows the completed solution using a second circle computed using subtended angle β at reference points P_2, P_3. Hence the intersection of the circles locates the image centre at I.
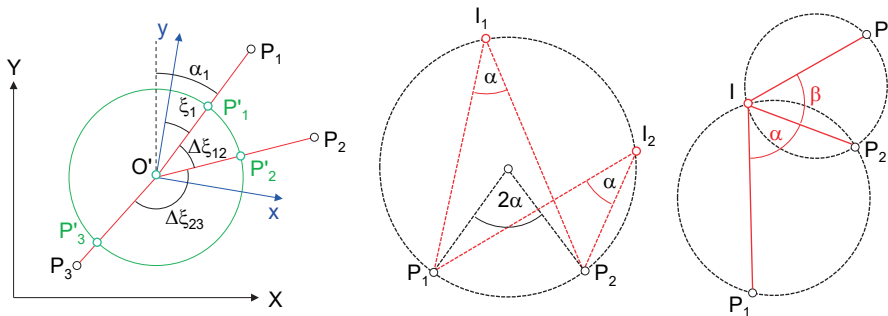
Fig. 4.87: Plane resection.

The Z_0 coordinate can, for example, be calculated from reference point P_1:

$$Z_0 = Z_1 - d_1 \frac{z_1}{r} \tag{4.150}$$

where $d_1 = \sqrt{(X_1 - X_0)^2 + (Y_1 - Y_0)^2}$

The approximate rotation angles of exterior orientation can be given to sufficient accuracy by:

$$\omega = 0 \qquad\qquad \varphi = 0 \qquad\qquad \kappa = \xi_1 - \alpha_1 \qquad\qquad (4.151)$$

4.5.2.2 Space resection

Space resection for a panoramic image can be formulated as an over-determined adjustment problem. For this purpose, the panoramic imaging equations (4.148) can be linearized at the approximate values given above. However, the use of Cartesian panoramic coordinates from (4.148) is much more convenient since they can be directly used as virtual three-dimensional image coordinates for the observation equations of a standard space resection (see section 4.2.3). Each image point provides three individual coordinates x,y,z, while for central perspective images the principal distance $z' = -c$ is constant for all image points.

As for the standard space resection, a minimum of three reference points is required, although this can generate up to four possible solutions. Alternatively, using four reference points a unique solution for the six parameters of exterior orientation is always obtained. If the reference points are distributed over the whole horizon, then even with a small number of points (greater than 4) a very reliable solution is achieved.

4.5.2.3 Bundle adjustment

Bundle adjustment of panoramic images is based on the same general approach as a bundle adjustment of central perspective images, including unknown object points and self-calibration of the camera if required. Some programs permit the simultaneous processing of both panoramic and central perspective images.
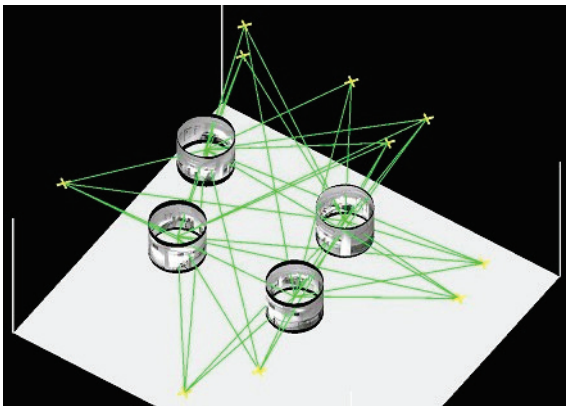
Fig. 4.88: Distribution of 4 panoramas in 3D space.

One advantage of panoramic images is that a stably oriented set of images can be obtained using a comparatively small number of object points. The example in Fig. 4.88 shows a set of four panoramas taken for the interior survey of a room. In this example, and using only four of the unknown object points, the number of unknowns according to Table 4.2 must be calculated.

Table 4.2: Example of unknowns and observations for panoramic bundle adjustment.

Parameter group	Unknowns	Number	Total
Exterior orientations	6	4	24
Object points	3	4	12
Datum definition	−7	1	−7
Sum			29

The 29 unknowns can be determined by measuring the object points in all four images, so providing 4 x 4 x 2 = 32 observations. Since this minimum solution is very sensitive to noise and outliers, the number of object points should be increased to at least 8 in this case. Nevertheless, compared to standard image blocks the number of required object points is much smaller. This is mainly due to the stable geometry of a cylindrical image which can be oriented in 3D space with very little object information.

The bundle adjustment can be extended with parameters for the correction of imaging errors. When a panoramic image results from stitching together single images calibrated using the original camera parameters, no additional corrections are necessary. In contrast, if panoramic images from a rotating line scanner are to be adjusted, parameters specific to the scanner must be introduced (see section 3.5.6.1).

4.5.3 Epipolar geometry

Given two oriented panoramas, epipolar plane and epipolar lines can be constructed for an imaged object point analogously to a stereo image pair (Fig. 4.89). The epipolar plane K is defined by the object point P and either both image points P' and P" or the projection centres O' and O". The epipolar plane intersects the two arbitrarily oriented panoramic cylinders in the elliptical epipolar lines k' and $k"$.

On the unrolled panoramic plane, the epipolar lines are sinusoidal. For an image point P' measured in the left-hand image, an arbitrary object point P can be defined. The corresponding image point P" in the right-hand panorama can be calculated according to (4.148). If the epipolar plane intersects the panoramic cylinder at angle β, corresponding to slope $m = \tan \beta$, the intersection straight line

$$z = m \cdot x \tag{4.152}$$

corresponding to Fig. 4.90 is obtained. With $y' = z$ and $x = r \cdot \cos \xi$ the equation for the epipolar line is given by

$$y' = m \cdot r \cdot \cos \xi \tag{4.153}$$
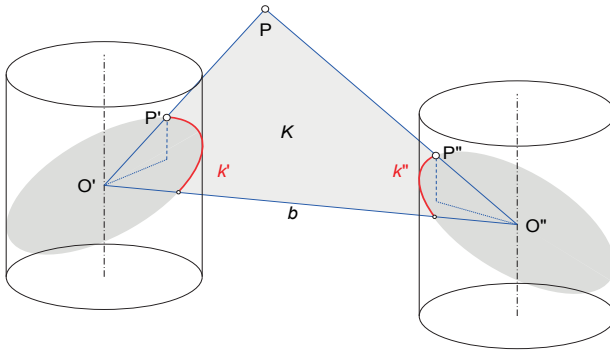
Fig. 4.89: Epipolar geometry for panoramic images.
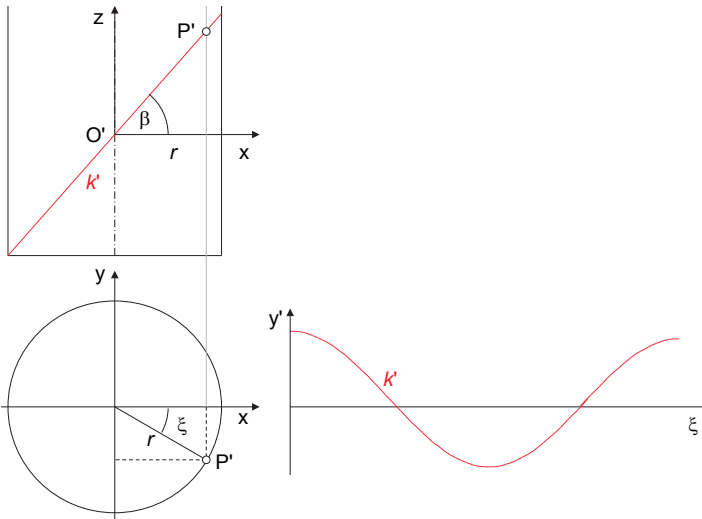
Fig. 4.90: Sine form of panoramic epipolar lines.

Fig. 4.91 shows an example of a panoramic image with the epipolar line superimposed.
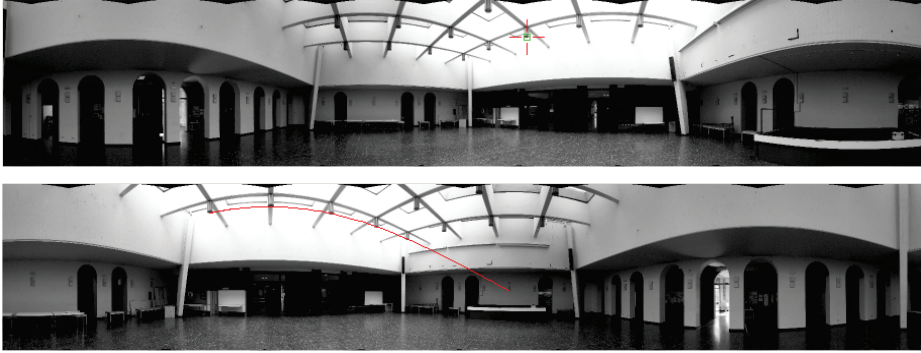
Fig. 4.91: Measured point (top) and sine form of corresponding panoramic epipolar line (bottom).

4.5.4 Spatial intersection

As for space resection and bundle adjustment for panoramic images, a general spatial intersection using three-dimensional panoramic coordinates can also be calculated. Spatial intersection fails if the object point lies on the baseline b (see Fig. 4.89).

If intersection is formulated on the basis of equations (4.148), the derivatives of the observation equations with respect to the unknown object coordinates are required. Using (4.145) they are given by

$$\frac{\partial x'}{\partial X} = \frac{r}{x^2+y^2}(r_{12}x - r_{11}y) \qquad \frac{\partial y'}{\partial X} = \frac{\lambda}{z}\left(\frac{r_{11}x - r_{12}y}{x^2+y^2} + \frac{r_{13}}{z}\right)$$

$$\frac{\partial x'}{\partial Y} = \frac{r}{x^2+y^2}(r_{22}x - r_{21}y) \qquad \frac{\partial y'}{\partial Y} = \frac{\lambda}{z}\left(\frac{r_{21}x - r_{22}y}{x^2+y^2} + \frac{r_{23}}{z}\right) \qquad (4.154)$$

$$\frac{\partial x'}{\partial Z} = \frac{r}{x^2+y^2}(r_{32}x - r_{31}y) \qquad \frac{\partial y'}{\partial Z} = \frac{\lambda}{z}\left(\frac{r_{31}x - r_{32}y}{x^2+y^2} + \frac{r_{33}}{z}\right)$$

As usual, the unknown object coordinates X,Y,Z are calculated by iterative adjustment until corrections to the unknowns become insignificant.

4.5.5 Rectification of panoramic images

4.5.5.1 Orthogonal rectification

As with central perspective images, panoramic images can also be rectified if the geometric relation between object coordinate system and image coordinate system is known. The panoramic images can then be rectified onto any chosen reference plane.

Fig. 4.92 shows the rectification of the panoramic image from Fig. 4.91 onto an interior side wall (XZ plane, upper image) as well as onto the floor and ceiling (XY planes with different Z values, lower images). The resulting images are true to scale with respect to the chosen reference plane but objects outside this plane are distorted. The black circles visible in the XY rectifications represent the areas in the vertical direction which are not covered by the original images.

Orthophoto production by combining panoramic images with 3D point clouds from a terrestrial laser scanner is discussed in section 6.8.1.

4.5.5.2 Tangential images

Central perspective images can be generated from panoramic images by defining a new central projection image plane as a tangential plane to the panoramic cylinder. Then according to Fig. 4.93, every image point P' in that part of the panoramic image which faces the tangential plane, can be projected onto the tangential plane. The result is a central perspective image in which object lines again appear as image lines. The perspective centre is a point on the axis of the panoramic cylinder and the principal distance corresponds to the cylinder radius. The exterior orientation of the tangential image can be directly derived from the exterior orientation of the panorama. The derived central perspective images can be further processed like other conventional photogrammetric images.

a) Rectification onto XZ plane
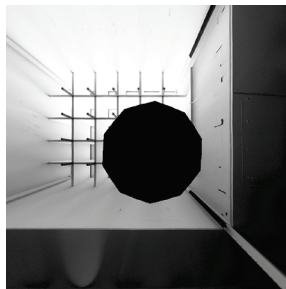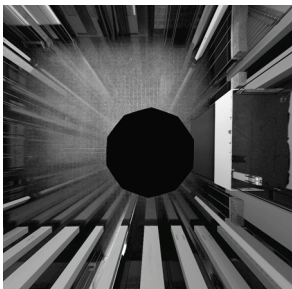
b) Rectification onto XY plane (floor and ceiling)

Fig. 4.92: Rectification of the panoramic images from Fig. 4.91.
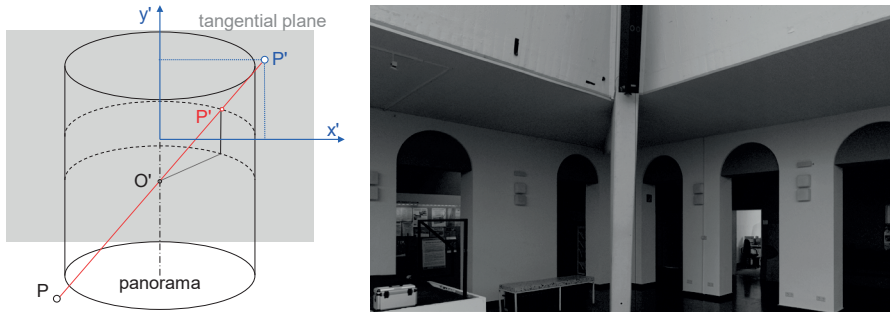
Fig. 4.93: Generation of central perspective tangential images from panoramic images.

4.6 Multi-media photogrammetry

4.6.1 Light refraction at media interfaces

4.6.1.1 Media interfaces

The standard photogrammetric imaging model (see sections 3.1.2, 4.2.1) assumes collinearity of object point, perspective centre and image point. Deviations caused by the lens or sensor are modelled by image-based correction functions. This approach is useful for most image configurations and applications.

If light rays in image or object space pass through optical media with differing refractive indices, they no longer follow a straight line. Using extended functional models for multi-media photogrammetry it is possible to calculate the optical path of the rays through additional media interfaces (MI) and take this into account in object reconstruction. For example, interfaces exist if the optical path intersects the following media:

– walls made of glass in object space (glass container, window panes);
– water (under, through);
– inhomogeneous atmosphere (refraction);
– filter glasses in front of the lens;
– individual lenses within a compound lens;
– glass covers on CCD sensors;
– réseau plates.

For a rigorous model of the optical path, a geometric description of the media interface must be available, for example:

– plane in space;
– second order surface (sphere, ellipsoid);
– wave-shaped surfaces.

Usually the transmission media are assumed to be homogenous and isotropic, i.e. light rays propagate uniformly in all directions inside the media.

Applications of multi-media photogrammetry can be found in fluid flow measurement (recording through glass, water, gas), underwater photogrammetry (underwater archaeology or underwater platform measurement) or hydrology (river bed recording).

4.6.1.2 Plane parallel media interfaces

The simplest multi-media case occurs if there is only one planar interface located parallel to the image plane (Fig. 4.94). An object point P_1 is then projected onto point P'_0, passing through intermediate point P_0 which lies on the interface. As a result, a radial shift $\Delta r'$ occurs with respect to the image point P'_1, which corresponds to a straight-line projection without refraction.
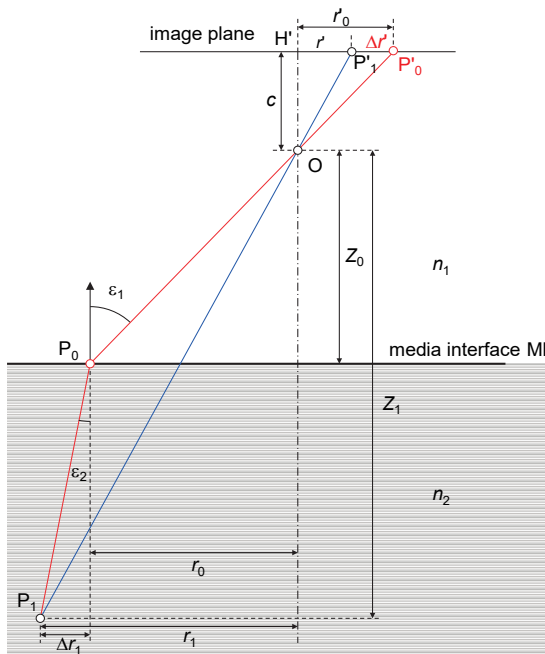
Fig. 4.94: Planar media interface parallel to the image plane.

According to Fig. 4.94, the radial shift $\Delta r'$, taking account of eqn. (3.8), is given by[6]:

---

[6] The following derivations and illustrations are due to Kotowski (1987).

$$\Delta r' = r'_0 \cdot Z_{rel} \left(1 - \frac{1}{\sqrt{n^2 + (n^2 - 1)\tan^2 \varepsilon_1}} \right)$$ (4.155)

where $Z_{rel} = \dfrac{Z_i - Z_0}{Z_i}$ and $r'_0 = c \cdot \tan \varepsilon_1$

Eqn. (4.155) shows that the effect of refraction is a function of distance. The radial shift is zero when:

1. $\varepsilon_1 = 0$: the light ray is perpendicular to the interface and is therefore not refracted;
2. $n = 1$: both media have equal refractive indices;
3. $Z_i = Z_0$: the object point is located on the interface ($Z_{rel} = 0$).

It can be shown that (4.155) can be expressed as the following power series:

$$\Delta r' = Z_{rel} \cdot (A_0 \cdot r'_0 + A_1 \cdot r'^3_0 + A_2 \cdot r'^5_0 + \ldots)$$ (4.156)

This power series expansion is similar to the standard functions for correction of simple radial distortion in (3.55), and also to the correction of distance-dependent distortion (3.73). For applications where Z_{rel} is close to 1, complete compensation for the effect of a plane parallel media interface can be done by the correction function for distance-dependent radial distortion.

The radial shift $\Delta r'$ can become significantly large:

Example 4.11:
A submerged underwater camera with air-filled lens and camera housing, $r'_0 = 21$ mm (full-format), $Z_{rel} = 1$ (media interface located in perspective centre) and $c = 28$mm (wide angle lens) has a maximum radial shift of $\Delta r' = 6.9$ mm.

Example 4.12:
A medium-format camera used for airborne measurement of the seabed, where $r'_0 = 38$ mm, $Z_i = 4$m, $Z_0 = 3$m ($Z_{rel} = 0.25$) (media interface is the water surface) and $c = 80$mm (standard angle lens) leads to a maximum radial shift of $\Delta r' = 2.7$mm.

The model of (4.155) can be extended to an unlimited number of plane parallel media interfaces, for instance for the modelling of plane parallel plates inside the camera (filter, réseau), or glass panes in object space (Fig. 4.95). For p interfaces, the radial shift becomes:

$$\Delta r' = \frac{r'_0}{Z_i} \left[(Z_i - Z_{01}) - \sum_{l=1}^{p} \frac{d_l}{\sqrt{\bar{n}_l^2 + (\bar{n}_l^2 - 1)\tan^2 \varepsilon_1}} \right]$$ (4.157)

where

$$d_l = Z_{l+1} - Z_l \qquad : \text{distance between two adjacent interfaces}$$

$$\overline{n}_l = \frac{n_{l+1}}{n_l} \qquad : \text{relative refractive index}$$

If differential distances between interfaces are used in (4.157), multi-layered media can be modelled. For example, this approach can be used to set up an atmospheric model for the description of atmospheric refraction.

Ray transmission through parallel plane media interfaces, with denser media on the object side, leads to a reduction of the field of view, i.e. the object in the media, for example water, appears larger. Fig. 4.96 illustrates how refraction at the plane interface not only results in a smaller field of view (red) but also, during camera calibration, to an increase in the principal distance (green).
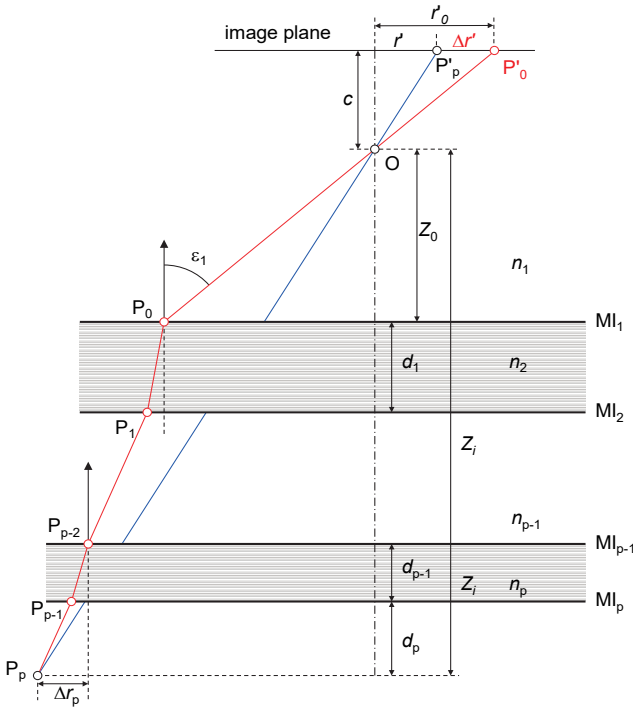
Fig. 4.95: Multiple plane parallel interfaces.
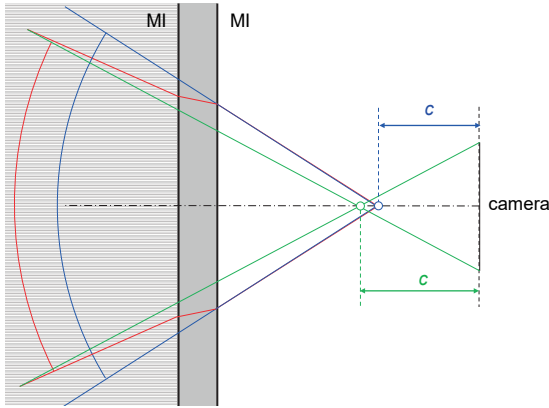
Fig. 4.96: Change in field of view and focal length with plane media interface.

4.6.1.3 Spherical media interfaces

Spherical shells represent a special type of interface. When hemispherical in shape, they are known as hemispherical interfaces or dome ports. If the entrance pupil of the lens (which is approximately the perspective centre O') is positioned exactly at the centre of the corresponding sphere, the photogrammetric imaging model based on collinearity equations (4.10) can also be applied to this multi-media case (Fig. 4.97) The ray from object point P through the perspective centre O' to the image point P' intersects the media interfaces orthogonally and will therefore not be refracted because $\varepsilon_1 = 0$ in eqn. (4.155). Dome ports are, for example, available as cover glasses for underwater camera housings.
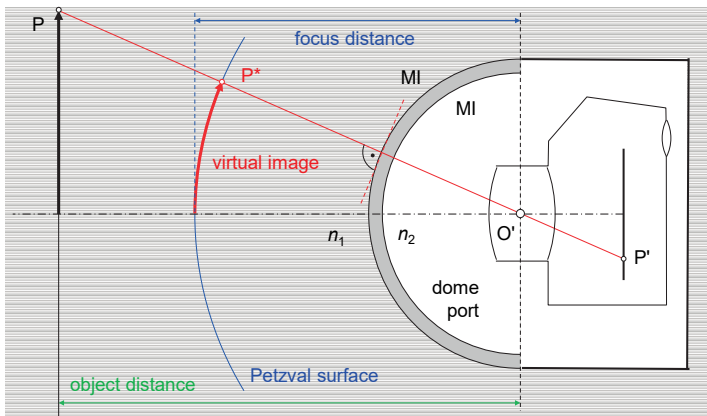
Fig. 4.97: Imaging with hemispherical interfaces and perspective centre at the sphere centre.

The spherical media interface acts as negative lens element which has the effect that an imaged object appears to be closer to the camera than in reality. This virtual image P* is projected onto a spherical surface called the Petzval surface. The camera must therefore be focused on that surface in order to obtain a sharp image of the object.

In summary, hemispherical media interfaces have the following properties in comparison with plane interfaces:

- field of view (viewing angle) and principal distance of the optical system are preserved;
- lens distortion and chromatic aberration are unchanged by the exterior media, e.g. water or air;
- the virtual image appears closer to the media interface and is projected onto a spherical surface;
- the depth of field increases by the ratio of the refractive indices n_1/n_2 (see Fig. 4.97);
- spherical aberration may increase;
- they are more expensive to manufacture;
- they can resist high pressure depending on material (plastic or glass) and thickness.

The differences between plane and spherical interfaces are particular noticeable with wide-angle lenses since the acceptance angle can be very large and therefore refraction at plane interfaces may also be significant towards the edges of the field of view.

Distortions will occur if the entrance pupil is not positioned exactly on the sphere centre. A shift perpendicular to the optical axis can partially be modelled as tangential distortion whilst a shift along the optical axis can be modelled as radial distortion. There are also approaches which rigorously describe the deviations using ray tracing or other methods.

4.6.1.4 Ray tracing through refracting interfaces

If arbitrary interfaces must be taken into account in the imaging model, each light ray must be traced through all contributing media by the successive application of the law of refraction (*ray tracing*). For this purpose, a set of three constraint equations is set up for each refracting point of a media interface. These provide the 3D coordinates of the point of refraction and the path of the ray between the point P_0, the interface point P_1 and the point P_2 (Fig. 4.98):
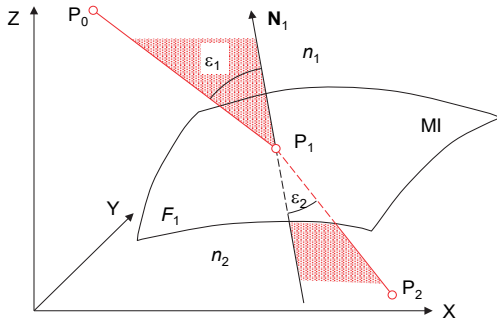
Fig. 4.98: Ray tracing through an optical interface.

1. P_1 is located on surface F_1:
   This condition is fulfilled by

   $$F_1(X_1, Y_1, Z_1) = 0 \tag{4.158}$$

   where for a general second order surface:

   $$\begin{aligned} F_1 &= \mathbf{X}^T \mathbf{A} \mathbf{X} + 2\mathbf{a}^T \mathbf{X} + a = 0 \\ &= f_S(\mathbf{X}, \mathbf{A}, \mathbf{a}, a) \end{aligned} \tag{4.159}$$

   where $\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$, $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ and $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, $a_{ij} = a_{ji}$ for $i \neq j$

   The surface can be parameterized with respect to the temporary coordinate system $\overline{X}, \overline{Y}, \overline{Z}$ according to section 4.4.2.3, e.g. as a rotationally symmetric surface.

2. Compliance with the law of refraction:

   $$F_2 = n_1 \sin \varepsilon_1 = n_2 \sin \varepsilon_2 = 0 \tag{4.160}$$

   The angles of incidence and refraction are introduced as a function of the object coordinates and the normal vector $\mathbf{N}_1$, e.g. for ε_1:

   $$\cos \varepsilon_1 = \frac{\mathbf{N}_1^T \cdot (\mathbf{X}_0 - \mathbf{X}_1)}{\left| \mathbf{N}_1^T \right| \cdot \left| \mathbf{X}_0 - \mathbf{X}_1 \right|} \tag{4.161}$$

3. Surface normal at P1, and the projection rays, lie in a common plane:
   This is implemented by applying the coplanarity constraint to the three vectors:

$$F_3 = \begin{vmatrix} X_0 - X_1 & Y_0 - Y_1 & Z_0 - Z_1 \\ X_1 - X_2 & Y_1 - Y_2 & Z_1 - Z_2 \\ N_X & N_Y & N_Z \end{vmatrix} = 0 \tag{4.162}$$

The three constraint equations are linearized at approximate values. Solving the system of equations results in the object coordinates X_1, Y_1, Z_1 of the refraction point.

For an imaging system of p media interfaces, the system of equations is set up for each refraction point P_l, $l = 1...p$. This principle is used in optics for the calculation of lens systems. This could be used in photogrammetry for a rigorous determination of distortion which took account of all optical elements.

4.6.2 Extended model of bundle triangulation

Using the fundamentals of optical interfaces and ray tracing discussed above, the functional model for bundle adjustment can be extended. For this purpose, the ray tracing algorithm with arbitrary interfaces is integrated into the imaging model of the collinearity equations. It is natural to distinguish two major imaging configurations:
- constant (object invariant) position of interfaces relative to the measured object;
- constant (bundle invariant) location of interfaces relative to the camera system.

4.6.2.1 Object-invariant interfaces
Object-invariant interfaces exist, for example, when an object is observed through glass panes or water, such as when taking pictures of aquariums or pressure tanks from the outside (Fig. 4.99). Object-invariant interfaces are always static with respect to the superior coordinate system.

The extended observation equations can be derived in three steps:
1. Ray tracing according to eqn. (4.159) with p interfaces for point i in image j:

$$\overline{\mathbf{X}}_i^l = f_S(\mathbf{X}_i, \mathbf{X}_{0j}, \mathbf{A}^l, \mathbf{a}^l, a^l, n^l) \tag{4.163}$$

where
$\mathbf{A}^l$, $\mathbf{a}^l$, a^l: parameters of interface
n^l: relative refractive indices
$l = 1...p$: index of interface

2. Spatial rotation and translation:

$$\mathbf{X}_{ij}^{\star} = \mathbf{R}_j \cdot (\bar{\mathbf{X}}_i^1 - \mathbf{X}_{0j})$$

$$\begin{bmatrix} X_{ij}^{\star} \\ Y_{ij}^{\star} \\ Z_{ij}^{\star} \end{bmatrix} = \mathbf{R}_j \cdot \begin{bmatrix} \bar{X}_i^1 - X_{0j} \\ \bar{Y}_i^1 - Y_{0j} \\ \bar{Z}_i^1 - Z_{0j} \end{bmatrix} \tag{4.164}$$

3. Extended collinearity equations ($z' = -c$):

$$\begin{bmatrix} x'_{ij} \\ y'_{ij} \end{bmatrix} = \frac{z'}{Z_{ij}^{\star}} \begin{bmatrix} X_{ij}^{\star} \\ Y_{ij}^{\star} \end{bmatrix} + \begin{bmatrix} x'_{0k} \\ y'_{0k} \end{bmatrix} + \begin{bmatrix} \Delta x'_k \\ \Delta y'_k \end{bmatrix} \tag{4.165}$$
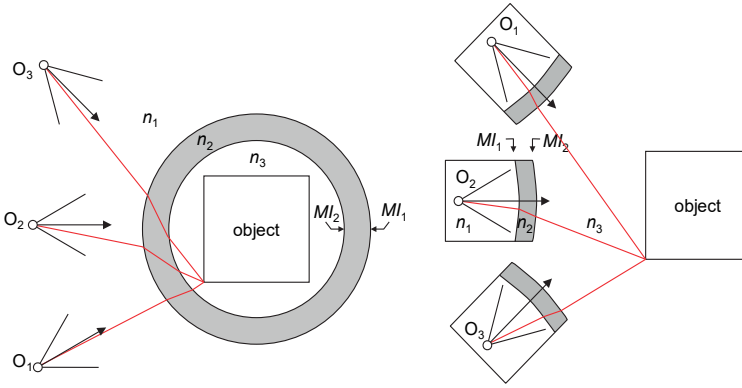
Fig. 4.99: Object-invariant interfaces.　　　**Fig. 4.100:** Bundle-invariant interfaces.

4.6.2.2 Bundle-invariant interfaces

Bundle-invariant interfaces are given by additional optical refracting interfaces which are part of the image acquisition system, e.g. underwater housing, add-on front filter, conversion lenses, parallel plate cover on sensor or réseau plate (Fig. 4.100). Bundle-invariant interfaces are always fixed with respect to the imaging system and therefore move with them in space.

In contrast to the object-invariant approach, the calculation is performed in reverse order:

1. Spatial translation and rotation:

$$\mathbf{X}_{ij}^{\star} = \mathbf{R}_j \cdot (\mathbf{X}_i^1 - \mathbf{X}_{0j}) \tag{4.166}$$

2. Ray tracing:

$$\bar{\mathbf{X}}_{ij}^l = f_S(\mathbf{X}_{ij}^{\star}, \mathbf{A}^l, \mathbf{a}^l, a^l, n^l) \tag{4.167}$$

3. Extended collinearity equations:

$$
\begin{bmatrix} x'_{ij} \\ y'_{ij} \end{bmatrix} = \frac{z'}{\overline{Z}^1_{ij}} \begin{bmatrix} \overline{X}^1_{ij} \\ \overline{Y}^1_{ij} \end{bmatrix} + \begin{bmatrix} x'_{0k} \\ y'_{0k} \end{bmatrix} + \begin{bmatrix} \Delta x'_k \\ \Delta y'_k \end{bmatrix} \tag{4.168}
$$

In relation to the object distance, bundle-invariant interfaces are often in very close proximity and plane-parallel. This mainly causes radially symmetric pixel shifts, whose influence is less strongly distance-dependent ($Z_{rel} \approx 1$, see example 4.8). In practical applications, therefore, models of single-media photogrammetry are often used, which are included in conventional software packages, although this does not reflect the complete imaging geometry. For rotationally symmetric elements (lenses) or plane-parallel plates set parallel to the image plane, these shifts are radially symmetric. With suitable image configurations they can be corrected by camera calibration which employs standard functions for distortion.

4.6.3 Special aspects of underwater photogrammetry

Underwater applications represent the largest area of multi-media photogrammetry. Modern sensor technology, single-board computers and mobile platforms allow increasing application possibilities for photogrammetric measurement tasks, e.g. in the inspection of buildings and ships, the recording of underwater topography or the measurement of plants and animals. In addition to the imaging models already explained above, some special conditions occur under water. Example underwater applications are presented in section 8.6.

Due to the refractive index of water of approx. n_{Water} = 1.33, there is a change in scale compared to air (n_{Air} = 1) of approx. 1.3, i.e. objects under water appear enlarged by this factor (Fig. 4.101). The refractive index of water is influenced by temperature, pressure and conductivity (salinity).

The penetration depth of light in water depends on the wavelength. Light of shorter wavelengths, e.g. blue, is more energetic (see section 3.1.1.2) than light of longer wavelengths, e.g. red, and therefore penetrates deeper into the medium (Fig. 4.102). This results in considerable colour shifts under water, e.g. strong blue or green components. Furthermore, turbidity caused by particles in the water reduces the image quality. These effects can be partially eliminated with digital image processing methods (Fig. 4.103).
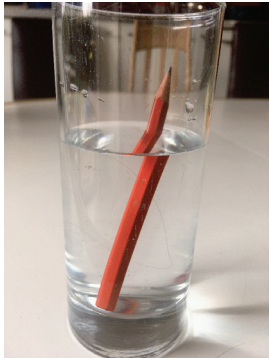
Fig. 4.101: Refraction and magnification through the medium of water.
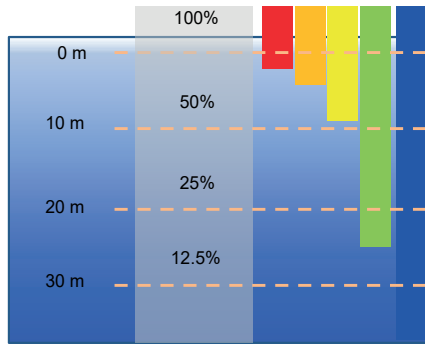
Fig. 4.102: Penetration depth of light in water.
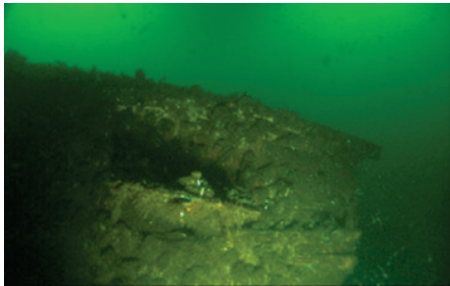
a) Original image
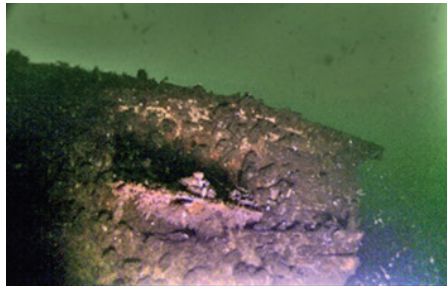
b) Digital image enhancement

Fig. 4.103: Image enhancement of a turbid underwater image.

5 Digital image analysis

5.1 Fundamentals

5.1.1 Image processing procedure

Photogrammetric image processing methods are primarily developed and applied in the fields of image acquisition (sensor technology, calibration), pre-processing and segmentation (image measuring, line following, image matching). Major considerations for these methods are the automation of relatively simple measuring tasks and the achievement of a suitable accuracy. Everyday methods which apply standard tasks such as digital point measurement, e.g. to targets, or stereo photogrammetry (automated orientation, surface reconstruction, driver assistance systems) are well developed. Machine learning methods including, for example, deep learning, are increasingly used to achieve goals in image understanding (object recognition, semantics) which is closely related to artificial intelligence (AI). However, it must be remembered that an adequate amount of training data is required for these purposes.
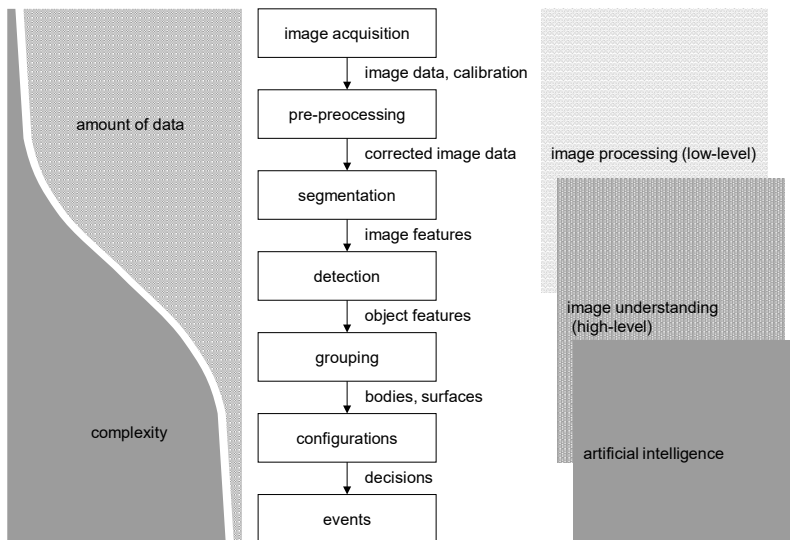
Fig. 5.1: Image processing sequence.

Fig. 5.1 illustrates a typical image-processing workflow, starting with image acquisition and ending with some intelligent initiation of events. As the sequence

proceeds from top to bottom, the volume of data is reduced whilst the complexity of processing increases.

Some characteristic features of current close-range photogrammetry illustrate the possibilities and limitations of automatic image processing:

– The consistent use of object targeting composed of retro-reflective marks, some of which are coded, combined with suitable illumination and exposure techniques, results in quasi-binary images that can be processed fully automatically (Fig. 5.2a).

– Arbitrary image configurations which can result in large variations in image scale, occlusions, incomplete object imaging etc. (Fig. 5.2b). In contrast to simple stereo configurations such as those found in aerial photogrammetry, close-range applications are often characterized by complex object surfaces and convergent multi-image network configurations which require a large amount of interactive processing.

– The application of image analysis and orientation methods from computer vision which are often suitable for fully automated orientation of overlapping image sequences from arbitrary (uncalibrated) cameras. These techniques, such as structure-from-motion, open up interesting opportunities, e.g. for the automatic determination of initial orientation values or for fully automated 3D reconstruction.
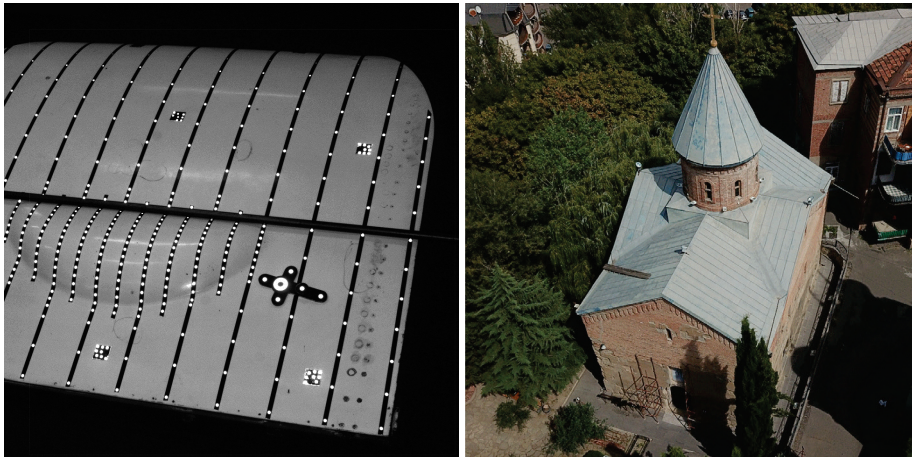
a) Targeted   b) Non-targeted

Fig. 5.2: Targeted and non-targeted object scenes.

This chapter concentrates on those image processing methods which have been successfully used in practical digital close-range photogrammetry. The emphasis is on methods for geometric image processing, image measurement and 3D

reconstruction which aim at high quality outputs. Extensions to basic principles and specialized algorithms can be found in the literature on digital image processing and computer vision, see section 9.5.

5.1.2 Pixel coordinate system

The definition of the pixel coordinate system is fundamental to image processing methods used for image measurement (see section 2.1.1). For an image

$$S = s(x,y) \tag{5.1}$$

conventional processing usually adopts a left-handed xy system of rows and columns which is related to the display of the image on the computer monitor and where x denotes the row direction and y the column direction.
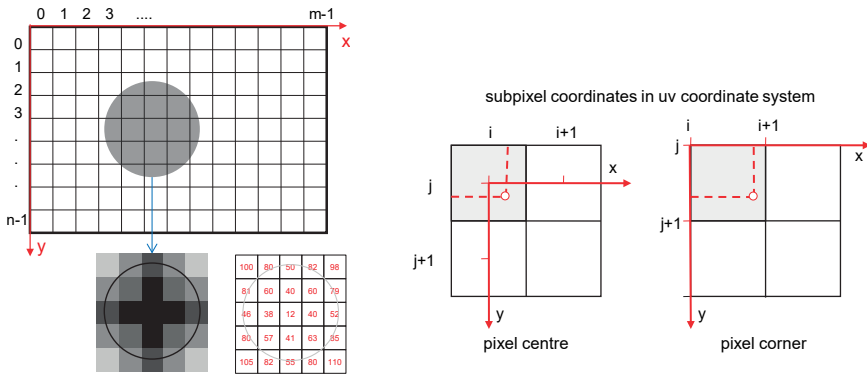
Fig. 5.3: Definition of the pixel coordinate system.

The origin of this system is located in the upper left-hand corner and the first image element has the row and column numbers (0,0). m is the number of columns, n the number of rows. The last image element has the coordinates ($m-1,n-1$). Width and height of a pixel are equal to 1 (Fig. 5.3).

In this discrete grid, each pixel has integer coordinate values. When an object is imaged by a sensor characterized by such a grid, each pixel acquires a grey value corresponding to the local image brightness across its area. Grey values are usually quantized with an 8 bit depth to provide 256 grey levels ranging from 0 (black) to 255 (white). Since human vision can only distinguish about 60 shades of grey, this grey level depth is sufficient for a visual representation of images. However, machine systems can handle a much higher information content and quantizations of 10 bits

(1024 grey levels), 12 bits or 16 bits can be used. True colour images are usually stored with 24 bits per pixel, 8 bits for each red, green and blue (RGB) colour channel.

Due to object size, image scale, pixel dimensions and optical transfer characteristics (MTF, PSF) of the image acquisition system (see sections 3.1.5 and 3.1.6), an imaged object can cover more than one pixel. This leads to a possible sub-pixel position for the coordinates of the imaged object. Measurement to sub-pixel level is only possible if the position of an imaged object can be interpolated over several pixels. Here it is assumed that a small shift of an object edge leads to a corresponding sub-pixel change in the imaged grey values. When sub-pixel coordinates are employed, it is conventional to consider the integer xy-coordinate (i,j) as applying to either the upper left corner or, seldom, the centre of a pixel (Fig. 5.3).

Imaged objects must usually cover more than one pixel in order to be detected or processed. Adjacent pixels belonging to one object are characterized by grey values that have uniform properties within a limited region (connectivity). Within a discrete image raster, each pixel possesses a fixed number of neighbours, with the exception of the image border. In defining connectivity, neighbouring pixels are classified according to the N4 or the N8 scheme (Fig. 5.4). In the following example, three objects A, B and C are imaged. If N4 connectivity is assumed, then object B decomposes into individual pixels and objects A and C are separated. In contrast, N8 connectivity leads to a single integral object B. However, A and C merge together due to their corner connection. Extended algorithms for connectivity must therefore consider the distribution of grey values within certain regions by using, for example, appropriate filters (see section 5.2.3).
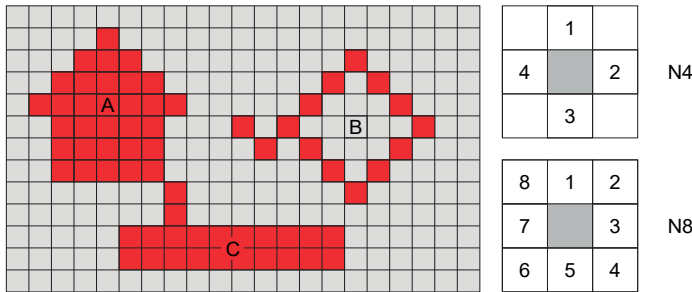
Fig. 5.4: Objects and connectivity.

5.1.3 Handling image data

5.1.3.1 Image pyramids
Image or resolution pyramids describe sequences where successive images are reductions of the previous image, usually by a factor of 2 (Fig. 5.5). Prior to image

reduction the image can be smoothed, for example using Gaussian filters (see section 5.2.3.2). As resolution is reduced, smaller mage structures disappear, i.e. the information content decreases (Fig. 5.6). The total amount of data required to store the pyramid is approximately only 30 % more than for the original image. Image pyramids are typically used in hierarchical pattern recognition or image matching problems which start with a search of coarse features in the lowest resolution image (pyramid top). The search is refined with increasing resolution, working progressively down through the pyramid layers, each time making use of the results of the previous resolution stage.

Original
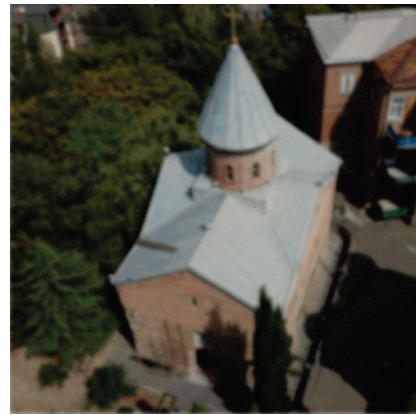
Gaussian low-pass filter
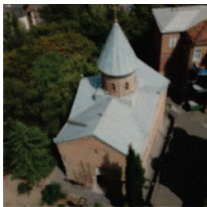
Factor 2   4   8   16

Fig. 5.5: Image pyramids with 5 steps.

5.1.3.2 Data formats

There are many ways to organize digital image data. Numerous data formats have been developed for digital image processing and raster graphics that, in addition to the actual image data, allow the storage of additional information such as image descriptions, colour tables, overview images etc. For photogrammetric purposes, the unique reproducibility of the original image is of major importance. Loss of

information can occur not only in certain data compression methods but also by using an insufficient depth of grey values (bits per pixel).
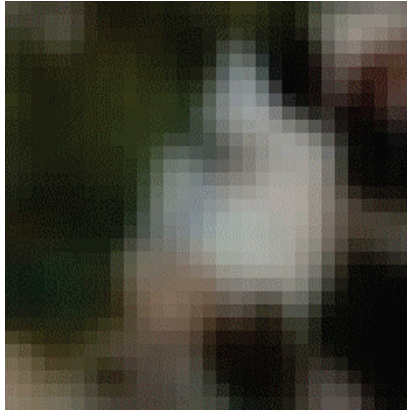
Factor 4   Factor 16

Fig. 5.6: Information content at reduction factors 4 and 16.

From the large number of different image formats the following are in common use:

– Direct storage of raw data:

  Here the original grey values of the image are stored in a binary file without compression. Using one byte per pixel, the resulting file size in bytes is exactly equal to the total number of pixels in the image. The original image format in rows and columns must be separately recorded since the raw image data cannot otherwise be read correctly. Multi-channel images can usually be stored in the order of their spectral bands, either pixel by pixel (pixel interleaved), line by line (line interleaved) or channel by channel (band interleaved). For RGB images based on Bayer mask, some raw formats store only the original intensity values, hence colour reconstruction has to be done separately. Some camera manufacturers offer their own formats for storing raw images.

– TIFF - Tagged Image File Format:

  The TIFF format is widely used due to its universal applicability. It is based on a directory of pointers to critical image information such as image size, colour depth, palettes, resolution etc. which must be interpreted by the import program. Numerous variants of TIFF exist and this can sometimes lead to problems with image transfer. The format permits different methods of image compression (LZW, Huffman, JPEG). GeoTIFF extends this format to store transformation parameters for geo-referencing.

–   BMP - Windows Bitmap:
    The BMP format is the standard format for images within the Microsoft Windows environment. It enables the storage of arbitrary halftone and colour images (up to 24 bit) with varying numbers of grey levels or colours.
–   GIF - Graphics Interchange Format:
    Images stored in GIF format are compressed without loss of information. Grey level and colour images are limited to 8 bits per pixel.
–   PNG - Portable Network Graphics:
    A non-patented replacement for GIF, which further provides an alpha channel (variable transparency), RGB storage and a higher compression level without loss of information.
–   JPEG - Joint Photographic Expert Group:
    JPEG is a format which allows compression levels up to a factor of 100 and in which a certain loss of information is accepted (see below). An updated version (JPEG2000) utilizes wavelet compression (see section 5.1.3.3).

Fig. 5.7 shows the percentage compression levels of different data formats for two images of very different structure. The first image 'mosaic' (Fig. 5.9, detail of the image shown in Fig. 5.25, blue bars) with inhomogeneous structures can be compressed to only 78 % of the original size without loss of information (PNG), while the more homogeneous image 'targets' (Fig. 5.10, red bars) can be reduced to less than 30 % of its original size.
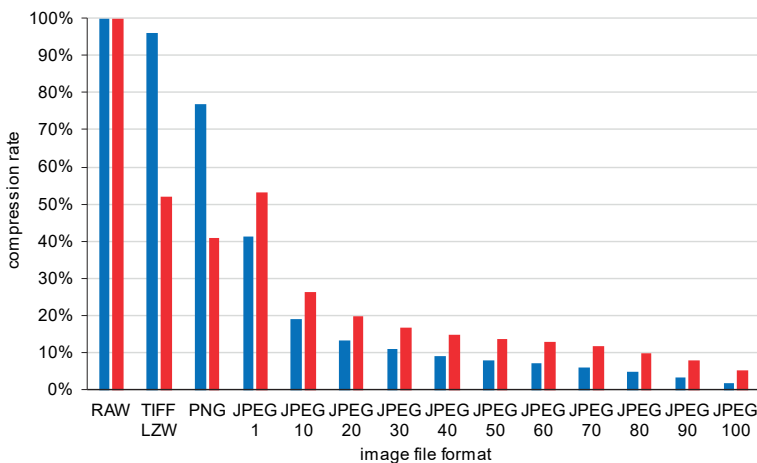
Fig. 5.7: Compression ratios for different image formats
(blue: image "mosaic", Fig. 5.9; red: image "targets", Fig. 5.10).

5.1.3.3 Image compression

Image compression is of major practical importance to digital photogrammetry, due to the large amounts of data which are handled. For example, a monochromatic set of 50 images each of 3000 x 2000 pixels represents some 300 MB of raw data whilst a UAV flight with 400 colour images which are each 4000 x 3000 pixels in size, requires around 14.4 GB of storage for the raw image data.

Run-length encoded compression methods count the number of identical grey values within a line or a region and code the corresponding image area by its grey value and a repetition factor. This method is useful for images with extensive homogeneous regions, but for images of natural scenes this often leads to an increased amount of data.

Frequency-based compression methods apply a spectral analysis to the image (Fourier, cosine or wavelet transformations, see section 5.2.3.1) and store the coefficients of the related functions. Eliminating coefficients of low significance compresses data with loss of information, often called lossy compression.

The basic JPEG image format is also based on a compression method with loss of information. The goal is to preserve the essential image content without a significant loss of visual quality, even at high compression levels. Due to the high compression performance, JPEG is widely used in many graphical and technical fields and the procedure is standardized to ensure consistent and appropriate image quality for the given application.

The compression algorithm is based on the 3-stage *Baseline Sequential* method (Fig. 5.8). A discrete cosine transformation (DCT) is calculated in disjoint 8x8 pixel patches. The resulting coefficients are weighted using a selectable quantization table and stored using run-length encoding. Data decompression is performed in the reverse order. For colour images, an additional IHS (intensity, hue, saturation) colour transformation is performed (section 5.1.1.1) whose channels are compressed differently, analogous to human visual perception.
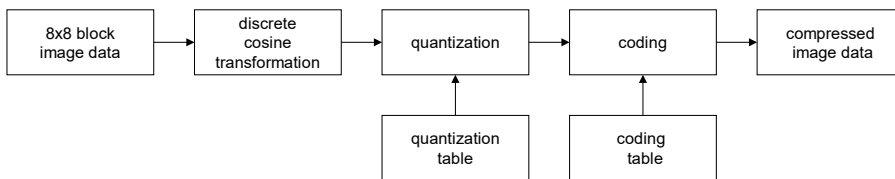
Fig. 5.8: JPEG image compression.

The actual loss of information is controlled by the choice of intervals in the quantization table. Usually the table is designed such that no significant loss of image quality can be visually observed (Fig. 5.9 and Fig. 5.10).

Extensions to the JPEG format are available with JPEG2000. Here the cosine transformation is replaced by a wavelet transformation. In addition, different areas of the image can be compressed to different levels. JPEG2000 also supports transparent images.

a) Original subset   b) JPEG, factor 20   c) JPEG, factor 50   d) JPEG, factor 99
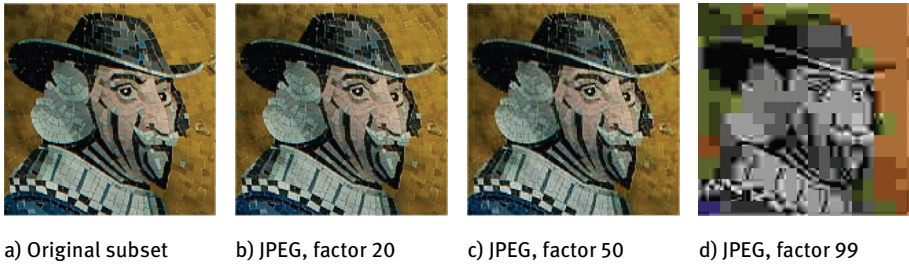
Fig. 5.9: Effect on quality of compression losses in image 'mosaic' (compare with Fig. 5.7).

a) Original subset   b) JPEG, factor 20   c) JPEG, factor 50   d) JPEG, factor 99
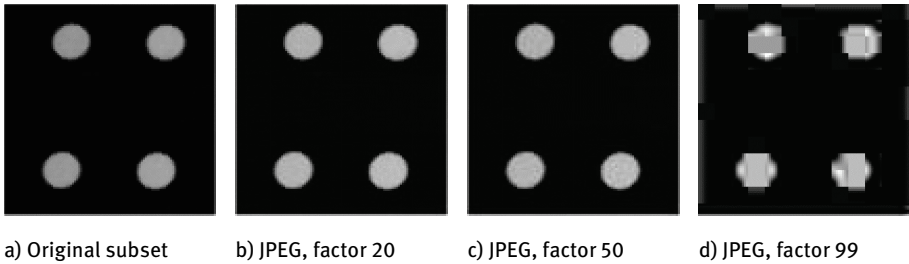
Fig. 5.10: Effect on quality of compression losses in image 'targets' (compare with Fig. 5.7).

The effect of JPEG compression on photogrammetric measurement mainly depends on the image content. In general, JPEG compression can give rise to localized image displacements of the order of 0.1 to 1 pixel. This significantly exceeds the accuracy potential of automatic point measurement which lies around 0.02 – 0.05 pixel (see section 5.2.4.6). In addition, the use of 8x8 pixel tiles within the JPEG process can cause undesirable edge effects.

5.1.3.4 Video formats

The use of digital videos for photogrammetric tasks is steadily increasing, as many camera systems offer an additional function for recording high-resolution videos. Videos also play an important role in the analysis of image sequences (see section 5.5.7). Due to the large amount of data involved, videos are often stored in compressed form, which can lead to a loss of information.

Video compression uses both spatial compression (analogous to JPEG) and temporal compression (from frame to frame). The data compression is called

encoding, the decompression is *decoding*. The implementation of encoding and decoding is known as a *Codec*. Codecs have different profiles depending on the functions used. A commonly used standard is H.264, also known as MPEG-4 Part 10 and MPEG-4 AVC, to which the principles outlined below refer. As the pixel count of videos has steadily increased (5K, 8K) H.264 now has a successor format H.265 or HEVC (high efficient video coding) which delivers higher compression rates.
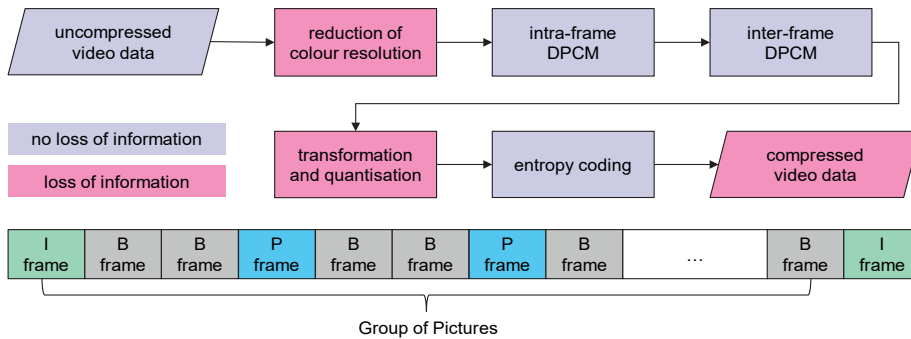
Fig. 5.11: Principle of video compression.

Similar to JPEG compression, each frame is divided into small macro blocks. In the temporal sequence, certain frames are used as a reference against which only the differences in subsequent frames are processed. These compression steps are called DPCM (Differential Pulse Code Modulation). Frames in which only image compression takes place are called I-frames (intra-frame DPCM). Between the I-frames are P-frames (predicted frames) and B-frames (bidirectional frames), in which, in principle, differences, predicted motion vectors and reference frames are stored instead of colour information. With increasing temporal distance from the I-frame, the loss of information increases. As a result, an I-frame is set, for example, every 10 frames. The frames stored between two I-frames are summarised as a Group of Pictures (GOP).

Fig. 5.11 shows the scheme of video compression. The lossy steps largely correspond to JPEG compression.

The (compressed) video data is stored in various container formats, e.g. AVI, MOV or MP4 (MPEG-4). Depending on the format, further data such as audio or metadata can be saved. Container formats can be compatible with several codecs.

5.1.3.5 Integral images
Integral images are used to optimize computation times for the calculation of grey-value sums or square sums of images or image windows. The integral image $I(x,y)$ has

the same size as the original input image $S = s(x,y)$ but at each x,y position it stores the grey-value sum of a rectangular window of S defined from $s(0,0)$ to $s(x,y)$.

$$I(x,y) = \sum_{i=0}^{x}\sum_{j=0}^{y} s(i,j)$$
$$= I(x-1,y-1) + I(x-1,y) + I(x,y-1) + s(x,y)$$

(5.2)

Using $I(x,y)$ it is possible to calculate the grey-value sum of an arbitrary rectangle of S from only four readings. Fig. 5.12 illustrates the principle for the red area-of-interest (AOI). The grey-value sum for this AOI can be derived from the stored sums for the four areas A, B, C, D as follows:

$$\Sigma_{AOI} = \Sigma_A - \Sigma_B - \Sigma_C + \Sigma_D$$
$$= I(x_1-1,y_1-1) - I(x_2,y_1-1) - I(x_1-1,y_2) + I(x_2,y_2)$$

(5.3)
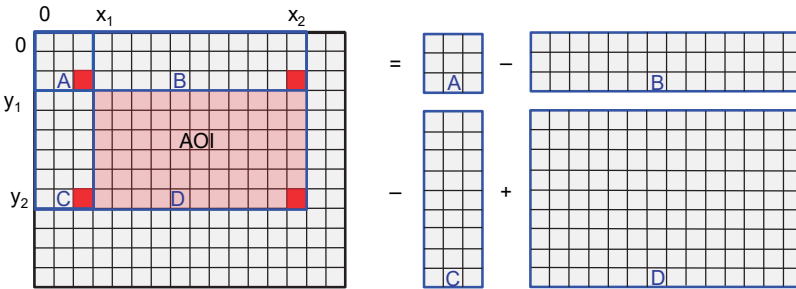
Fig. 5.12: Calculation of grey-value sum using an integral image.

a) Synthetic image   b) Subset of a test-field image
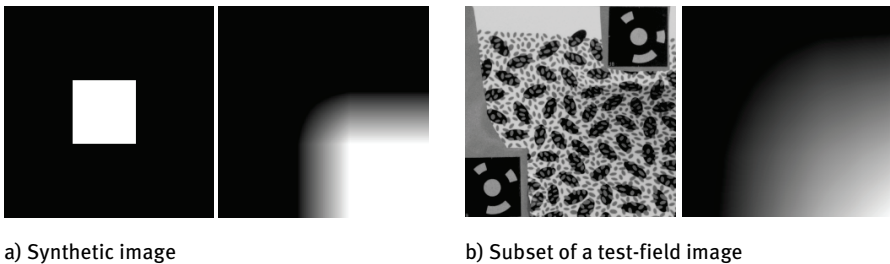
Fig. 5.13: Original images and corresponding integral images (contrast enhanced).

Fig. 5.13 shows two original images and the corresponding integral images (contrast enhanced for display). Integral images can be useful where a large number of grey value sums must be calculated for variable sub-windows, e.g. for convolution filters (section 5.2.3.2), normalized cross-correlation (section 5.4.2.3) or interest operators (section 5.4.3).

5.2 Image pre-processing

5.2.1 Point operations

5.2.1.1 Histogram

The histogram provides the frequency distribution of the grey values in the image. It displays the absolute or relative frequency of each grey value either in tabular or graphical form (Fig. 5.14).
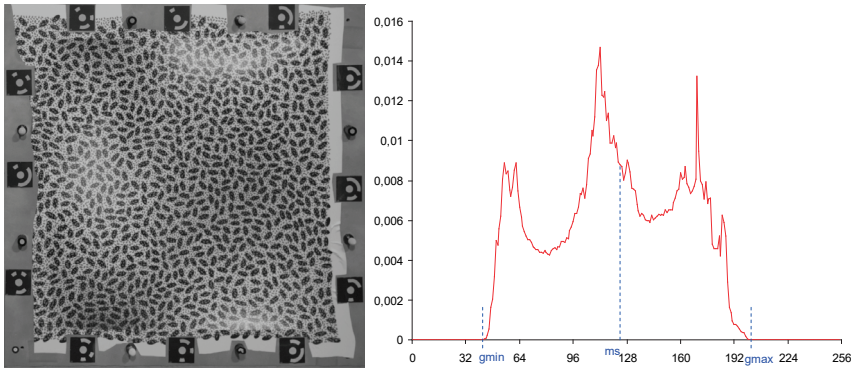
Fig. 5.14: Grey level image with corresponding histogram and parameters.

The most important parameters of a histogram are:

$$p_S(g) = \frac{a_S(g)}{M}$$

: relative frequency $p_S(g)$ and absolute frequency $a_S(g)$ where $M = m \cdot n$ and $0 \leq p_S(g) \leq 1$ (5.4)

g_{min}, g_{max}

: minimum and maximum grey value of the image

$$K = \frac{g_{max} - g_{min}}{g_{max} + g_{min}}$$

: contrast   (5.5)

$$m_S = \frac{1}{M} \sum_{u=0}^{m-1} \sum_{v=0}^{n-1} s(u,v)$$

: mean of grey values   (5.6)

$$= \sum_{g=0}^{255} g \cdot p_S(g)$$

$$q_S = \frac{1}{M} \sum_{u=0}^{m-1} \sum_{v=0}^{n-1} \left[s(u,v) - m_S \right]^2$$

: variance of grey values   (5.7)

$$= \sum_{g=0}^{255} (g - m_S)^2 \cdot p_S(g)$$

$$H = \sum_{g=0}^{255} \left[p_S(g) \cdot \log_2 p_S(g) \right] \qquad \text{: entropy}$$

(5.8)

$$\alpha = \frac{-\sum_{g=0}^{k} \left[p_S(g) \cdot \log_2 p_S(g) \right]}{H} \qquad \text{: symmetry (anisotropic coefficient)}$$

k: minimum grey value where

(5.9)

$$\sum_{g=0}^{k} p_S(g) \geq 0.5$$

Whilst minimum and maximum grey values define the image contrast, the mean is a measure of the average intensity (brightness) of the image. For statistical image processing, variance or standard deviation is also calculated but both are of minor interest in metrology applications.

The information content in an image can be measured by its entropy. It corresponds to the average number of bits necessary to quantize the grey values. Entropy can also be used to calculate an appropriate factor for image compression (see section 5.1.3.3).

The degree of symmetry of a histogram is determined by the anisotropic coefficient. Symmetrical histograms have a coefficient $\alpha = 0.5$. This coefficient can also be used to determine a threshold for bimodal histograms (see section 5.2.1.4).

Example 5.1:

The histogram of the image in Fig. 5.14 has the following parameters:

Minimum grey value:	$g_{min} = 42$
Maximum grey value:	$g_{max} = 204$
Contrast:	$K = 0.658$
Mean value:	$m_s = 120.4$
Variance:	$q = 1506.7$
Standard deviation:	$q^{1/2} = 38.8$
Entropy:	$H = -7.1$
Symmetry:	$\alpha = 0.49$

5.2.1.2 Lookup tables

Lookup tables (LUT, colour tables, palettes) are simple tools for the global manipulation of grey values. Each grey value of an input image is assigned a unique grey value in an output image. This method of processing grey values is simple to implement and can be found in almost all graphics or image processing programs. LUTs are easily displayed in diagrammatic form, see Fig. 5.15. LUT operations are unique and not usually reversible.

Given the lookup table

$$LUT(g) \qquad g=0,1,2,\dots,255$$

(5.10)

the grey values of the output image are calculated:

$$g' = LUT(g) \tag{5.11}$$
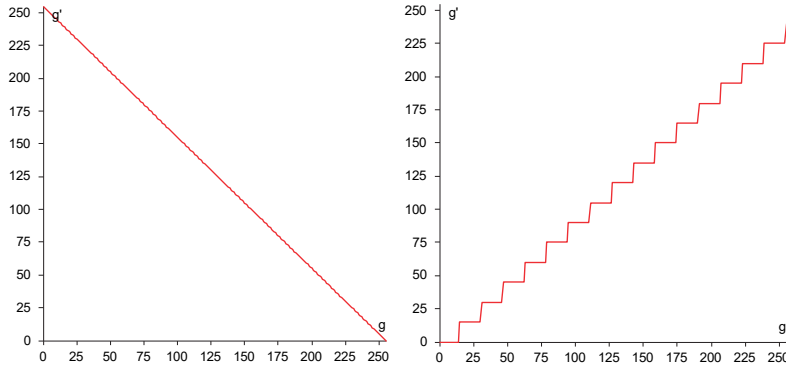
Fig. 5.15: Examples of lookup tables.

5.2.1.3 Contrast enhancement

Manipulating the brightness and contrast of an original image results in a change of the grey value distribution within certain regions, for example along an image edge. Variations of grey values are not only a function of object intensity but are also influenced by the relative position of camera and object. In general, a non-linear manipulation of grey values can affect the geometry of an object's image and should therefore be avoided if possible. However, contrast enhancement can provide a better visual interpretation for interactive image processing.

Image contrast changes are easily applied by a lookup table. Table values can be defined interactively, pre-calculated or derived from the image content itself. Common methods for adjusting brightness and contrast are:

– Linear contrast stretching:

The LUT is a linear interpolation between gmin and gmax (Fig. 5.16). Minimum and maximum grey values can be derived from the histogram, or defined interactively. The calculation of the LUT is derived from a shift (offset) r0 and a scale factor (gain) r1 as follows:

$$LUT(g) = r_0 + r_1 g \tag{5.12}$$

where

$$r_0 = -\frac{255 \cdot g_{min}}{g_{max} - g_{min}} \qquad\qquad r_1 = \frac{255}{g_{max} - g_{min}}$$
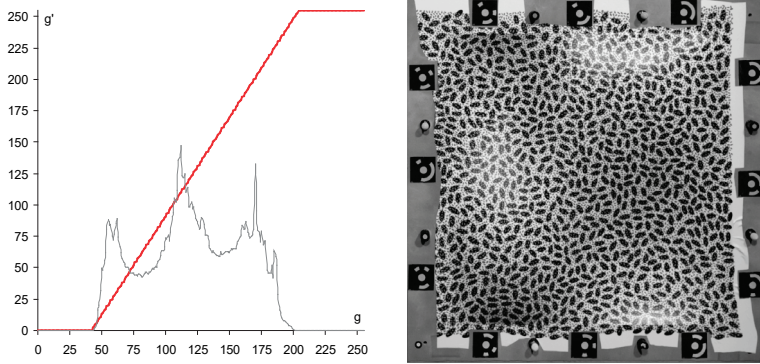
Fig. 5.16: Linear contrast enhancement (original image in Fig. 5.14).

– Histogram equalization:

The cumulative frequency function is calculated from the histogram of the original image:

$$h_S(g) = \sum_{k=0}^{g} p_S(k) \qquad g=0,1,2,\ldots,255 \tag{5.13}$$

where $0 \leq h_S(g) \leq 1$

The LUT values are given by:

$$LUT(g) = 255 \cdot h_S(g) \tag{5.14}$$
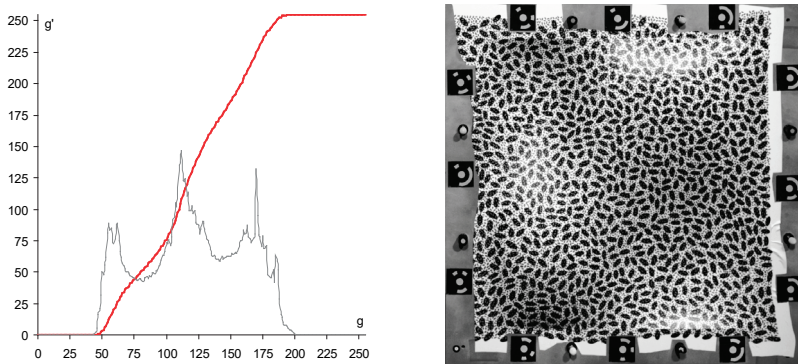
Fig. 5.17: Contrast stretching by histogram equalization (original image in Fig. 5.14).

The function is dependent on the histogram since the slope of the LUT is proportional to the frequency of the corresponding grey value (Fig. 5.17). Image contrast is consequently strongly enhanced in areas where grey values have high

frequencies. The output image S' therefore has a histogram pS' (g') with relative cumulative frequencies which are constant for each grey value according to the definition $h_{S'}(g) = 1/255 \cdot g'$.

– Gamma correction:

Many interactive image processing programs permit a Gamma correction where, in analogy with the gamma characteristic in the photographic process, the slope of the LUT is adjusted logarithmically. This essentially results in an increase or decrease of the mean grey value (Fig. 5.18).

The LUT of a gamma correction is given by:

$$LUT(g) = 255 \cdot \left(\frac{g}{255} \right)^{\frac{1}{\gamma}} \tag{5.15}$$
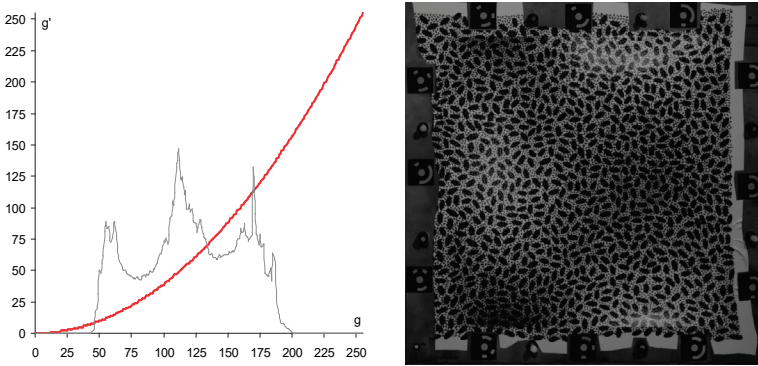
Fig. 5.18: Gamma correction with γ = 0.5 (original image in Fig. 5.14).

– Local contrast adjustment:

In principle, all the above image processing methods can be applied to any image detail (window, area of interest). Using the Wallis filter (section 5.2.3.4) the image is modified so that the local contrast in a filter window is optimized.

5.2.1.4 Thresholding

In general, thresholding is used to clearly differentiate grey values which belong to different object classes, e.g. to separate objects and background. Thresholding is often a pre-processing stage prior to segmentation.

Consider a simple case where the image consists only of two classes, i.e.:

– class K_1: background, e.g. dark
– class K_2: objects, e.g. bright

The corresponding histogram can be expected to be bimodal having two significant groupings of data, each represented by a maximum (peak), which are separated by a minimum (valley). Clearly both classes can be separated by a single threshold t (bimodal thresholding) located within the minimum region between the class maxima, e.g. by defining:

$$t = (m_2 - m_1)/2 \tag{5.16}$$

where m_1, m_2 are the mean grey values of classes K_1, K_2

Applying the lookup table

$$LUT(g) = \begin{cases} g_1 & \text{for } g \leq t \\ g_2 & \text{for } g > t \end{cases} \tag{5.17}$$

where g_1, g_2 are the new grey values for classes K_1, K_2

results in a binary image (two-level image) consisting only of grey values g_1, e.g. 0, and g_2, e.g. 1 or 255. Fig. 5.19 shows the histogram of the image in Fig. 5.20 which has two significant primary maxima ($m_1 \approx 18$, $m_2 \approx 163$), representing background and object. The secondary maximum ($m_3 \approx 254$) is caused by the imaged retro-reflective targets. The binary image of Fig. 5.21 is the result of thresholding with $t = 90$ (between the primary maxima).
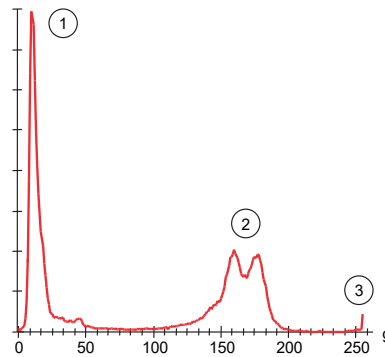
Fig. 5.19: Histogram of Fig. 5.20 with two primary maxima (1, 2) and one secondary maximum (3).

Fig. 5.22 shows the result with threshold value $t = 192$, located near maximum 2. This preserves some image information in addition to the targets. With $t = 230$ (Fig. 5.23) almost all targets are separated or segmented out from the background (but see Fig. 5.40).

For more complex images, the problem of thresholding lies in both the calculation of representative class averages and in the subsequent definition of the threshold value itself. Natural image scenes usually have more than two grey-value classes and this requires a much more complex thresholding procedure (multi-modal or dynamic thresholding).
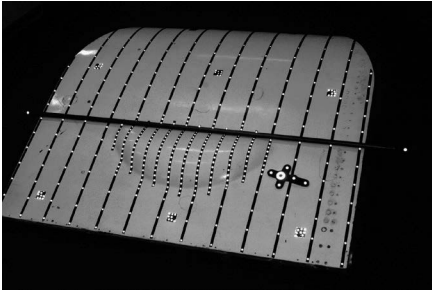
Fig. 5.20: Metric image with retro-targets.
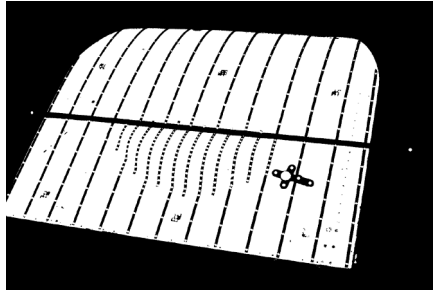
Fig. 5.21: Result after thresholding with $t = 90$.
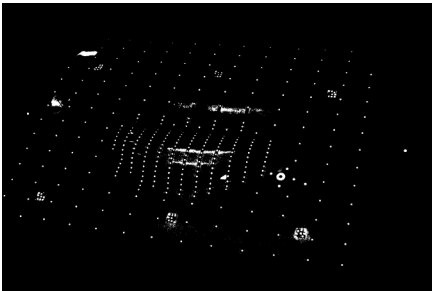
Fig. 5.22: Result after thresholding with $t = 192$.
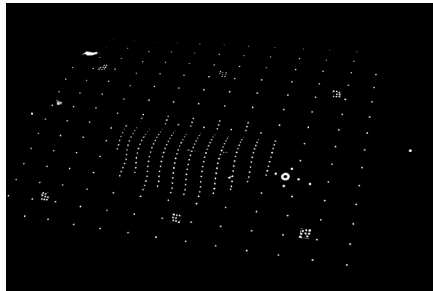
Fig. 5.23: Result after thresholding with $t = 230$.

5.2.1.5 Image arithmetic

Two or more images or image subsets can be combined numerically:

– arithmetic: addition, subtraction, division, multiplication

The grey values of both images are combined arithmetically, e.g. subtracted:

$$s'(x,y) = s_2(x,y) - s_1(x,y) \tag{5.18}$$
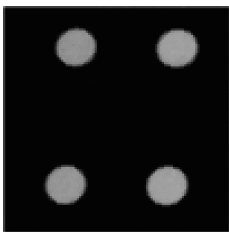
a) Original
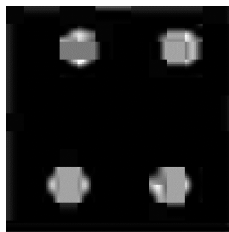
b) Compressed
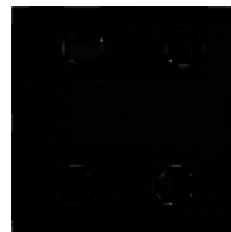
c) Difference

Fig. 5.24: Difference image to illustrate information loss after JPEG image compression.

The results of this image operation (difference image) show the differences between both input images (example in Fig. 5.24). If necessary the grey values of the output image, where negative values are possible, must be transformed into the positive range [0...255], or alternatively stored as 16-bit signed integers.

– logical: =, <, >, ≤, ≥, ≠

The grey values of both images are compared logically, resulting in Boolean values 1 (true) or 0 (false).

– bit-wise: AND, OR, NOT, XOR

The grey values of both input images are combined bit by bit. The XOR operation has practical use for the temporary superimposition of a moving cursor on the image. The original grey value can be recovered without the use of temporary storage by executing two sequential XOR operations with the value 255. Example: 38 XOR 255 = 217; 217 XOR 255 = 38.

5.2.2 Colour operations

5.2.2.1 Colour spaces

True-colour images are composed of three spectral (colour) image channels which store their respective intensity or colour distributions. Storage normally requires 8 bits per channel so that RGB images have a colour depth of 24 bits (Fig. 5.25). An optional alpha channel can be stored to control the transparency of an image (the proportion of the image displayed with respect to the background). RGB images with an alpha channel have a storage depth of 32 bits per pixel.

An image with n colour or spectral channels can be defined as an image vector **S:**

$$\mathbf{S} = \begin{bmatrix} s_0(x,y) \\ s_1(x,y) \\ \vdots \\ s_{n-1}(x,y) \end{bmatrix} \tag{5.19}$$

For an RGB image ($n = 3$), this corresponds to:

$$\mathbf{S}_{RGB} = \begin{bmatrix} s_R(x,y) \\ s_G(x,y) \\ s_B(x,y) \end{bmatrix} = \begin{bmatrix} \mathbf{R} \\ \mathbf{G} \\ \mathbf{B} \end{bmatrix} \tag{5.20}$$

A default conversion of an RGB image to a grey-value image is done for every pixel as follows:

$$s'(x,y) = \left[s_R(x,y) + s_G(x,y) + s_B(x,y) \right] / 3 \tag{5.21}$$

or more simply written as:

$$I = (R + G + B)/3 \tag{5.22}$$

a) Original true-colour image

b) Red channel

c) Green channel

d) Blue channel

Fig. 5.25: True-colour image with separate RGB channels.

The most important colour models in photogrammetry are the RGB colour space and the IHS colour space (also called the HSL colour space), see below and Fig. 5.26. In the RGB colour space, colour is defined by three-dimensional Cartesian coordinates R,G,B. The origin of the coordinate system ($R = 0$, $G = 0$, $B = 0$) defines black and the maximum position ($R = 1$, $G = 1$, $B = 1$) defines white. Values on the principal diagonal, with equal RGB components, define grey values. The interval [0...1] results from the normalisation of the usual grey value range of [0...255]. The additive primary colours, red, green and blue, as well as the subtractive colours (complementary colours) yellow, magenta and cyan (Y,M,C) lie at the corners of the RGB cube. It is simple to convert between additive and subtractive primary colours:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1-C \\ 1-M \\ 1-Y \end{bmatrix} \qquad \begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1-R \\ 1-G \\ 1-B \end{bmatrix} \tag{5.23}$$

Alternatively, a weighted form of intensity calculation can be performed which is closer to human perception, e.g. by

$$I = 0.2989 \cdot R + 0.5870 \cdot G + 0.1140 \cdot B \tag{5.24}$$

In the IHS colour space, colours have components of intensity (I), hue (H) and saturation (S). This is also called the HSL colour space with components hue (H), saturation (S) and luminance (L) where luminance is equivalent to intensity. The IHS values can be interpreted as cylindrical coordinates as shown in Fig. 5.26. Grey values between black and white lie on the cylinder axis which also represents the intensity axis. Primary colours are represented at angle H on the colour circle and their saturation is given by the distance from the cylinder axis.
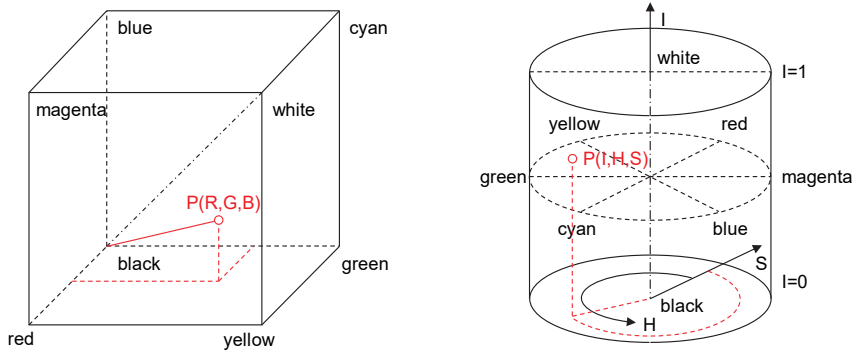
Fig. 5.26: RGB and IHS colour spaces.

The CIE XYZ colour model, named after the Commission Internationale de l'Elcairage, is a standardized representation of colour which relates physical colours to how the average human eye perceives them. Colour representation is in an XYZ system where each colour coordinate is defined in the interval [0,1]. An extension to the CIE model is the CIE *lab* model, where *l*, *a* and *b* represent colour axes. Both models are based on standardized illumination and reflection conditions.

5.2.2.2 Colour transformations

5.2.2.2.1 RGB ↔ IHS

A unique transformation is possible between RGB and IHS colour spaces. However, the formulas used in practice are not consistent. The following is a widely used transformation:

With $I_{min} = \min(R,G,B)$ and $I_{max} = \max(R,G,B)$ then:

$$I = \frac{I_{max} + I_{min}}{2}$$

$$H = \begin{cases} 0° & \text{for } I_{max} = I_{min} \\ 60° \cdot \left(0 + \dfrac{G - B}{I_{max} - I_{min}} \right) & \text{for } I_{max} = R \\ 60° \cdot \left(2 + \dfrac{B - R}{I_{max} - I_{min}} \right) & \text{for } I_{max} = G \\ 60° \cdot \left(4 + \dfrac{R - G}{I_{max} - I_{min}} \right) & \text{for } I_{max} = B \end{cases} \tag{5.25}$$

$$S = \begin{cases} 0 & \text{for } I_{max} = I_{min} \\ \dfrac{I_{max} - I_{min}}{I_{max} + I_{min}} & \text{for } I < 0.5 \\ \dfrac{I_{max} - I_{min}}{2 - I_{max} - I_{min}} & \text{for } I \geq 0.5 \end{cases}$$

The calculated values lie in the intervals $I[0...1]$, $H[0...360]$ und $S[0...1]$ and can subsequently be transformed into the grey-scale range $[0...255]$.

Using the substitutions

$$q = \begin{cases} I(1+S) & \text{for } I < 0.5 \\ I + S - (I \cdot S) & \text{for } I \geq 0.5 \end{cases} \qquad p = 2I - q \qquad H' = H/360$$

$$t_R = H' + 1/3 \qquad\qquad t_G = H' \qquad\qquad t_B = H' - 1/3$$

then for every colour $C \in (R,G,B)$:

$$C = \begin{cases} p + 6(q - p)t_C & \text{for } t_C < 1/6 \\ q & \text{for } 1/6 \leq t_C < 1/2 \\ p + 6(q - p)(2/3 - t_C) & \text{for } 1/2 \leq t_C < 2/3 \\ p & \text{else} \end{cases} \qquad \text{where } t_C = t_C \bmod 1 \tag{5.26}$$

Fig. 5.27 shows the IHS channels which result from the transformation of a true-colour image.

a) Original true-colour image

b) Intensity channel

c) Hue channel

d) Saturation channel

Fig. 5.27: IHS transformation.

5.2.2.2.2 RGB ↔ XYZ

The transformation of RGB values into the CIE XYZ model is achieved with the following operation:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{M} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad \qquad \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \mathbf{M}^{-1} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{5.27}$$

where

$$\mathbf{M} = \begin{bmatrix} 0.4887180 & 0.3106803 & 0.2006017 \\ 0.1762044 & 0.8129847 & 0.0108109 \\ 0 & 0.0102048 & 0.9897952 \end{bmatrix}$$

The transformation matrix $\mathbf{M}$ is defined by the CIE standard. There are deviations from this matrix which are optimized for printing and computer graphics.

5.2.2.2.3 CIE Lab

The CIE Lab system is a colour space better adapted to human vision, in which visually equal perceived colour distances are also mathematically equal. The parameters $L^\star$, $a^\star$ and $b^\star$ are defined with the colours from the XYZ system as follows:

$$L^\star = 116 \cdot f\left(\frac{Y}{Y_n}\right) - 16$$

$$a^\star = 500 \cdot f\left[\left(\frac{X}{X_n}\right) - \left(\frac{Y}{Y_n}\right)\right] \qquad (5.28)$$

$$b^\star = 200 \cdot f\left[\left(\frac{Y}{Y_n}\right) - \left(\frac{Z}{Z_n}\right)\right]$$

with X_n, Y_n, Z_n: colour values of a standardized white reference

and

$$f\left(\frac{X}{X_n}\right) = \begin{cases} \left(\dfrac{X}{X_n}\right)^{\frac{1}{3}} & \text{if } \left(\dfrac{X}{X_n}\right) > \left(\dfrac{6}{29}\right)^3 \\ \left(\dfrac{841}{108}\right)\left(\dfrac{X}{X_n}\right) + \left(\dfrac{4}{29}\right) & \text{if } \left(\dfrac{X}{X_n}\right) \geq \left(\dfrac{6}{29}\right)^3 \end{cases}$$

similarly for Y and Z

$L^\star$ describes the brightness of the image in the range $[0,100]$. The colourimetric values $a^\star$ and $b^\star$ describe the colour gradient along the red-green and yellow-blue axes respectively.

With this, colour distances between colour values can be calculated as follows:

$$\Delta E = \sqrt{\Delta L^{\star 2} + \Delta a^{\star 2} + \Delta b^{\star 2}} \qquad (5.29)$$

There are significant deviations between human colour perception and the Euclidean colour distance in colour areas with high saturation, hence the Euclidean value is only used for small colour differences. Taking environmental conditions into account, the colour distance equation according to CIEDE 2000 can be extended so that the CIE Lab colour space can be seen equidistantly.

The calculation of colour distances is relevant, e.g. for the adjustment of printing profiles, as well as for colour calibration of images (section 5.2.2.2).

5.2.2.3 Colour calibration

Colour calibration methods are used to match the captured colours of an RGB image against a reference chromaticity diagram, so that calibrated colours are produced regardless of local lighting conditions. For this purpose grey or colour charts (examples in Fig. 5.28) are imaged together with the actual scene in at least one image of the block.

The simplest and most common method is white balance, in which a (standardised) grey chart (Fig. 5.28 left) is recorded. The adjustment of the RGB colours is then carried out, for example, in such a way that a lookup table for each colour channel is determined via the colour components of a white, grey and black surface, with which a colour adjustment of the input image is subsequently carried out.

Fig. 5.28: Left: grey charts; right: colour calibration chart (X-Rite ColorChecker).

a) Original image

b) Calibrated image

Fig. 5.29: Result of colour calibration.

Better results are obtained with calibrated colour charts. Fig. 5.28 on the right shows a colour chart with 140 colour patches from which a colour correction matrix (CCM) of the input colours can be calculated in the captured image. Fig. 5.29 illustrates the result of colour calibration for a photogrammetric image.

5.2.2.4 Colour combinations

The following procedures are commonly used to create new colour assignments in grey-scale and colour images:

5.2.2.4.1 Pseudo colour

Pseudo colours are obtained when RGB colour values are assigned to the values in a single-channel, grey-scale image using a colour palette. Fig. 5.30 shows two possible look-up tables which do this. Pseudo colours are used, for example, in the colouring of thermal images (see section 3.5.8) so that the intensity values (representing temperature) can be visualized using a colour scale (see example in Fig. 5.31).

a) Rainbow palette   b) Temperature scale
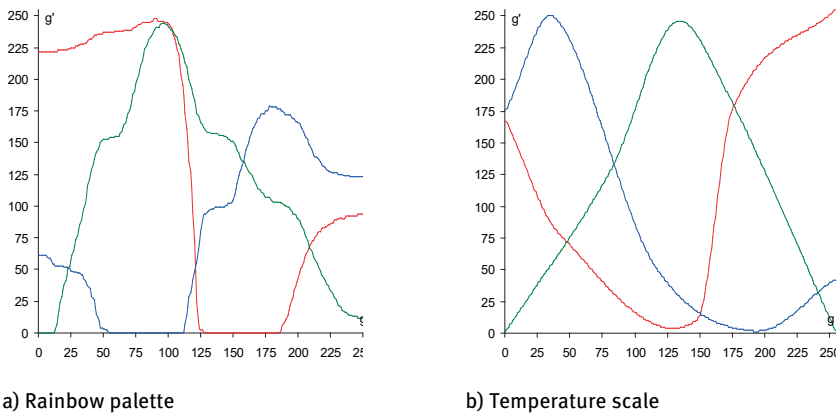
Fig. 5.30: Lookup tables for creating pseudo colour images.

5.2.2.4.2 False colour

False colour images are obtained from multi-channel images, in which every input channel is freely assigned to the channel of an RGB output image. For example, using the following assignments:

$$
\text{a)} \quad \begin{bmatrix} s'_R(x,y) \\ s'_G(x,y) \\ s'_B(x,y) \end{bmatrix} = \begin{bmatrix} s_B(x,y) \\ s_R(x,y) \\ s_G(x,y) \end{bmatrix} \qquad \text{b)} \quad \begin{bmatrix} s'_R(x,y) \\ s'_G(x,y) \\ s'_B(x,y) \end{bmatrix} = \begin{bmatrix} s_{IR}(x,y) \\ s_R(x,y) \\ s_G(x,y) \end{bmatrix} \qquad (5.30)
$$

the following false colour images are obtained. In (a), three standard RGB channels are exchanged and in (b) a false colour infrared image is obtained assigning output red to the input IR channel, output green to input red and output blue to input green channels.
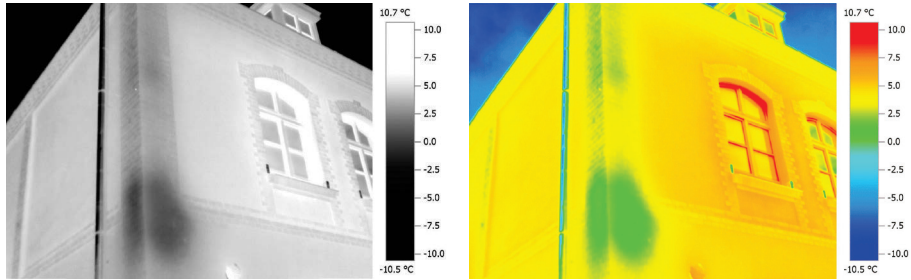
a) Original thermal image   b) Colour-coded thermal image

Fig. 5.31: Pseudo colour for thermal image.

5.2.2.4.3 Pan sharpening

Pan sharpening (also called resolution merging) is the combination of a sensor with high resolution in the panchromatic region (P channel) with the lower resolution channels of a colour sensor (RGB channels). The concept stems from human vision where the retina has a higher number of rods (detecting brightness) and a significantly lower number of cones (detecting colour) which together give the impression of a high-resolution image.
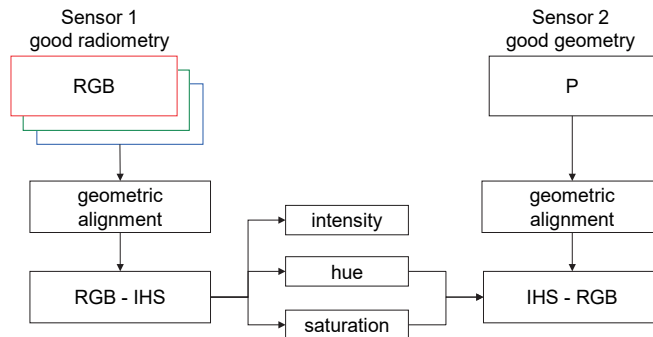
Fig. 5.32: Principle of pan sharpening.

According to Fig. 5.32, the RGB image and the P channel are geometrically aligned so that all input images represent the same object at the same resolution. The RGB image

is subsequently transformed into HIS format. The resulting I channel is then replaced by the P channel and a reverse transformation into RGB space is calculated. The final image then has the geometric resolution of the high-resolution P channel as well as the colour information from the original RGB image.
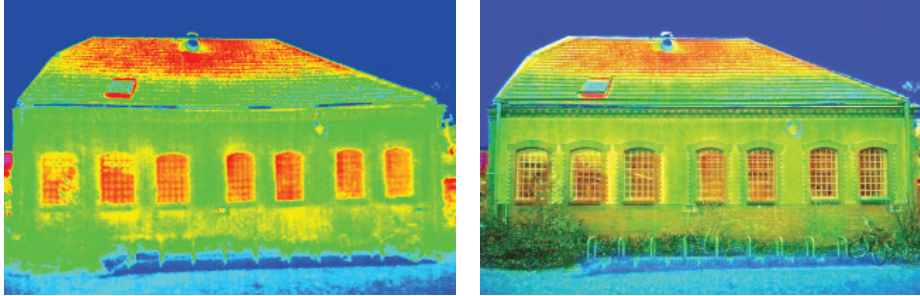
a) Original thermal image   b) Thermal image after pan sharpening

Fig. 5.33: Pan sharpening for a thermal image using a true-colour image.

Fig. 5.33 shows a low-resolution thermal image and the result after combination with the panchromatic channel generated from a high-resolution RGB image. It can be seen that the resulting image shows much better resolution and that numerous object details are visible. However, it is not easy to make a correct measurement of temperature in the pan-sharpened image. For that, a back-transformation into the original thermal image must be calculated.

5.2.3 Filter operations

5.2.3.1 Spatial domain and frequency domain

The theory of digital filtering is based on the fundamentals of digital signal processing (communication engineering, electronics). Its fundamental method is the Fourier transform which represents arbitrary signals (series of discrete values, waves) as linear combinations of trigonometric functions. The discrete one-dimensional Fourier transformation for n samples of an input signal $s(x)$ is given by:

$$F(u) = \frac{1}{n} \sum_{x=0}^{n-1} s(x) \cdot \exp\left(-i\frac{2\pi}{n}ux\right) = \mathrm{Re}\big(F(u)\big) + i \cdot \mathrm{Im}\big(F(u)\big) \tag{5.31}$$

Here u denotes the spatial frequency and $i = \sqrt{-1}$. The inverse Fourier transformation is given by:

$$s(x) = \frac{1}{n}\sum_{u=0}^{n-1} F(u)\exp\left(+i\frac{2\pi}{n}ux\right) \tag{5.32}$$

i.e. the original signal can be exactly reconstructed from its Fourier transform.

The Euler formulas show the connection with the underlying trigonometric functions:

$$e^{-i2\pi ux} = \cos(2\pi ux) - i\sin(2\pi ux)$$
$$e^{i2\pi ux} = \cos(2\pi ux) + i\sin(2\pi ux) \tag{5.33}$$

The power spectrum of $s(x)$ is defined by:

$$P(u) = |F(u)|^2 = \mathrm{Re}^2(F(u)) + \mathrm{Im}^2(F(u)) \tag{5.34}$$

The 1D Fourier transformation can easily be extended to the discrete 2D Fourier transformation.

When applied to an image, the discrete Fourier transformation transforms it from the spatial domain $S = s(x,y)$ into the frequency domain $F(u,v)$. A visual evaluation of the spatial frequencies (wave numbers) in the image can be made through the power spectrum. For example, edges generate high frequencies. A power spectrum example is illustrated in Fig. 5.34, where bright points in the spectrum correspond to large amplitudes. This example shows high amplitudes which are perpendicular to the significant edges in the original image (spatial domain). The two horizontal and vertical lines in the spectrum are caused by the image borders.
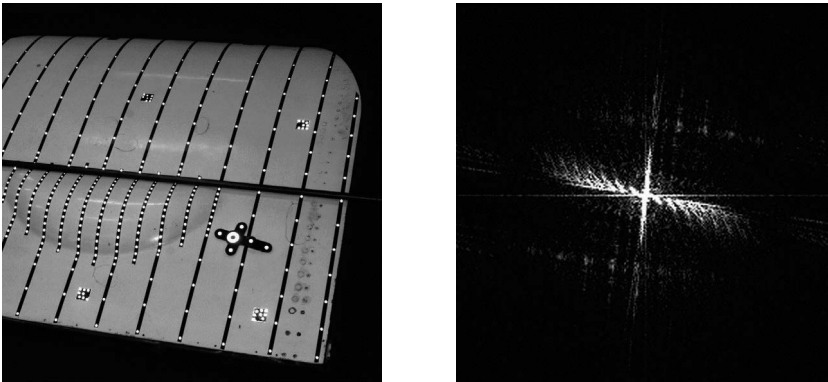
Fig. 5.34: Grey-value image and corresponding power spectrum.

Different basic functions from the Fourier transform result in alternative image transformations. For example, the *discrete cosine transformation* (DCT) uses only cosine terms. The *wavelet transformation* uses various basic functions such as the

Haar function to transform the original signal not only into the frequency domain but also into a scale domain of different resolutions. Wavelets are especially useful for image operations which must simultaneously account for coarse (smoothed) structures and detailed features having high information content. One application of the wavelet transformation is in image compression with information loss (see section 5.1.3.3).
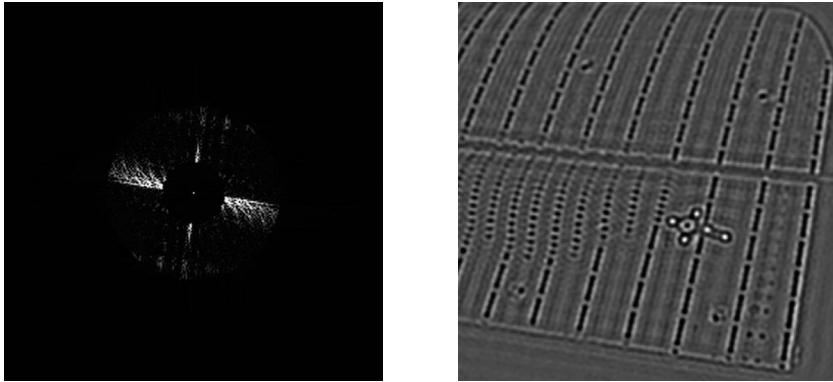
Fig. 5.35: Band-pass filter in power spectrum and resulting image.

Filters can be used to select or suppress certain spatial frequencies in the original image (high-pass filter, band-pass filter, low-pass filter). In the frequency domain, the desired frequencies are multiplied by a filter function which defines the filter characteristics. Fig. 5.35 shows a circular filter in the frequency spectrum. In this example the spectrum within the inner circle and beyond the outer circle is multiplied by 0 and within the ring zone it is multiplied by 1. After transformation back into the spatial domain this band-pass filter produces edge enhancement.
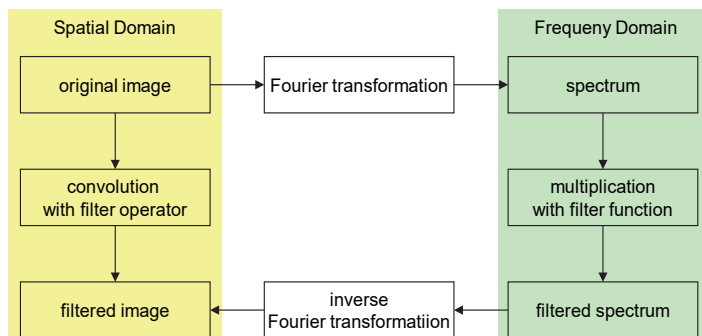
Fig. 5.36: Filtering in the spatial and frequency domains.

In the spatial domain, filters are applied by convolution with a filter operator. It can be shown that both approaches have identical results (Fig. 5.36).

Filter methods based on a one or two-dimensional convolution calculate a weighted sum of grey values in a given pixel region of the input image S. The result is assigned to the output image S' at the position of the region's central pixel.

$$s'(x,y) = \frac{1}{f} \sum_{u=-k}^{+k} \sum_{v=-l}^{+l} s(x-u, y-v) \cdot h(u,v)$$
$$S' = S \otimes H$$
(5.35)

Here $H = h(u,v)$ denotes the filter matrix (filter operator) with $p \times q$ elements, where $k = (p\text{-}1)/2$ and $l = (q\text{-}1)/2$. Usually p and q are odd numbers and often $p = q$. The factor f is used for normalization to the range $[0...255]$. The operator $\otimes$ symbolizes the convolution operation.
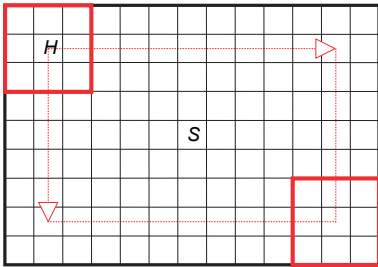
Fig. 5.37: Scheme for image filtering with $p=q=3$.

In order to filter the complete image, the filter operator is shifted over the image in rows and columns as shown in Fig. 5.37. At each x,y position the convolution is calculated and the resulting grey value stored in the output image. The number of computational instructions amounts to $(2k+1)^2$ multiplications and $(2k+1)-1$ additions. For example, an image with 1024 x 1024 pixels, $p = 5$, $k = 2$ requires around $26 \cdot 10^6$ multiplications and $25 \cdot 10^6$ additions. Some filter masks can be split into one-dimensional convolutions which can be separately computed in the x and y directions. In this case only $(4k+2)$ multiplications and $4k$ additions are required. In the example this results in around $10 \cdot 10^6$ multiplications and $8 \cdot 10^6$ additions.

5.2.3.2 Smoothing filters

Smoothing filters (low-pass filters) are mainly used for the suppression of grey-level noise, such as the quantization noise associated with digitization. These types of filter principally divide into linear smoothing filters, based on convolution, and non-linear smoothing filters based on rank orders (median filter). Table 5.1 shows one- and two-

dimensional examples of typical filter operators. Fig. 5.38 illustrates the results of different smoothing filters.

a) Original image

b) Moving average filter (5x5)

c) Binomial filter (5x5)

d) Median filter (5x5)
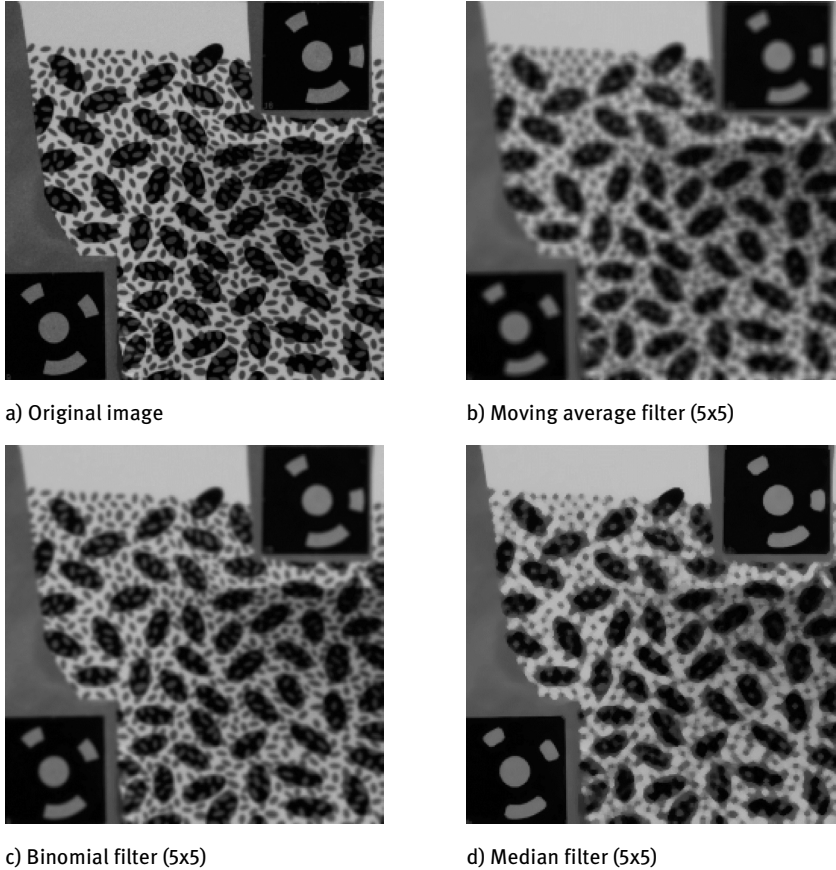
Fig. 5.38: Effect of different smoothing filters.

Linear, low-pass filters smooth the image but, depending on choice of filter coefficients, also smear image edges. The smoothing effect increases with larger filter sizes.

Gaussian filters possess mathematically optimal smoothing properties. The coefficients of the filter are derived from the two-dimensional Gaussian function:

$$f(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{5.36}$$

A suitable filter matrix of size p is usually chosen empirically. As an example, given $\sigma = 1$, a filter size of $p = 7$ is appropriate. The Gaussian filter coefficients can be well approximated by the binomial coefficients of Pascal's triangle (see Table 5.1).

The non-linear median filter performs good smoothing whilst retaining sharp edges (Fig. 5.38d). The median filter is not based on convolution. Instead, it calculates the median value (as opposed to the mean) of a sorted list of grey values in the filter matrix, and uses this as the output grey value. The output image therefore consists only of grey values which exist in the input image. This property is essential for the filtering of images where, instead of intensities, the image elements store attributes or other data. The median filter is a member of the group of rank-order filters.

Table 5.1: Examples of filter operators for image smoothing.

Filter method	1D	2D
Smoothing filter (moving average)	$H_{3,1} = \frac{1}{3} \cdot \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \end{bmatrix}$	$H_{3,3} = \frac{1}{9} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$
Smoothing filter (binomial filter)	$H_{3,1} = \frac{1}{4} \cdot \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1/4 & 1/2 & 1/4 \end{bmatrix}$ $H_{5,1} = \frac{1}{16} \cdot \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \end{bmatrix}$	$H_{3,3} = \frac{1}{4}\begin{pmatrix} 1 & 2 & 1 \end{pmatrix} \cdot \frac{1}{4}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$ $H_{5,5} = \frac{1}{256}\begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$

5.2.3.3 Morphological operations

Morphological operations form their own class of image processing methods. The basic idea is the application of non-linear filters (see median filter, section 5.2.3.2) for the enhancement or suppression of black and white image regions with (known) shape properties, e.g. for the segmentation of point or circular features. Filtering is performed with special *structuring elements* which are tuned to the feature type to be detected and are successively stepped across the whole image. The structuring element is assigned a focus or active point which acts on the corresponding image point and which must be carefully defined in order to avoid offsets in the identified locations of the features being detected.

Two fundamental functions based on Boolean operations for binary images are defined for morphological image processing:

– Erosion:

Erosion leads to the shrinking of regions. The value 1 is set in the output image if all pixels in the filter region, e.g. 3x3 elements, correspond to the structuring element, i.e. the structuring element is in complete agreement with the image region. Otherwise the value 0 is set.

– Dilation:

Dilation leads to the enlargement of connected regions. The number 1 is set if the structuring element includes at least one matching pixel within the image filter region.
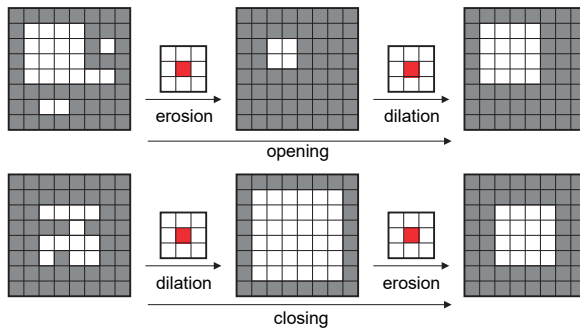
Fig. 5.39: Morphological operations with 3x3 structure element (red: focus point).

Sequential application of dilation and erosion can be used to close gaps or to separate connected regions in the image. The following combinations of basic operations are useful:

– Opening:

Opening is achieved by an erosion followed by dilation (Fig. 5.39 top). Small objects are removed.

– Closing:

The reverse process is referred to as closing. Dilation is followed by erosion in order to close gaps between objects (Fig. 5.39 bottom).

An application of opening is the segmentation of bright targets in a photogrammetric image. Fig. 5.40 shows a problematic situation for an image region with a number of targets lying close together. After simple thresholding, the intermediate result shows several small circular features which do not correspond to target points, as well as connected regions which actually represent separate targets. Segmentation, based on the procedure of section 5.2.1.4 with a minimum point size of 15 pixel, results in nine objects of which two have joined together features from separate adjacent objects.

a) Original image   b) Thresholding   c) Segmentation result
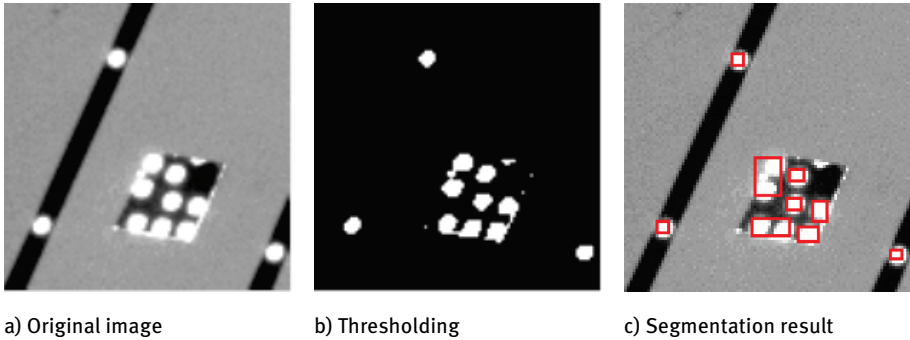
Fig. 5.40: Simple segmentation after thresholding (9 objects).

a) Erosion (3x3)   b) Dilation (3x3)   c) Segmentation result
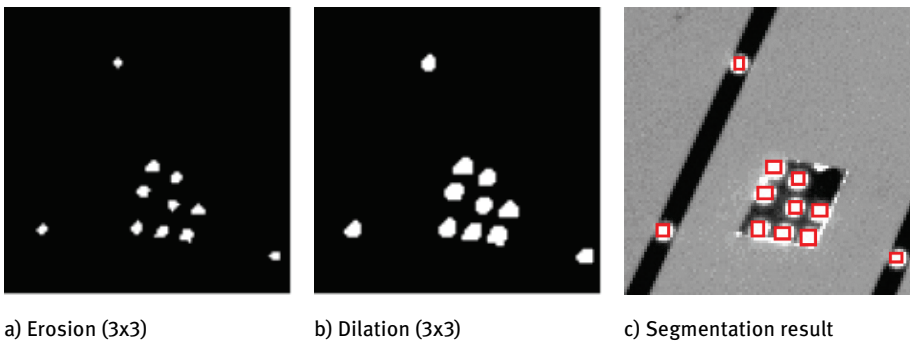
Fig. 5.41: Segmentation after thresholding and opening (11 objects).

Fig. 5.41 shows the same image region processed using the opening operator. Starting with the same binary image, erosion eliminates objects smaller than the applied 3x3 structure element. In addition, the targets shrink. Subsequent application of dilation enlarges the remaining objects to their original size. This type of segmentation correctly extracts all 11 objects.

Morphological operators can be extended to grey-level images. Instead of the Boolean decision for each pixel of the structuring element, the minimum grey-value within the neighbourhood determined by the structuring element (grey-value erosion) or the maximum grey-value (grey-value dilation), is used as the result.

5.2.3.4 Wallis filter

The objective of the Wallis filter is a local optimization of contrast. The grey values in the output image are calculated as follows:

$$s'(x,y) = s(x,y) r_1 + r_0 \qquad (5.37)$$

Here the parameters r_0, r_1 of a linear contrast stretching (eqn. 5.12) are determined in a local filter window of size n x n pixel. The calculation of the filter parameters for each pixel is based on the following values:

$$r_1 = \frac{\sigma' c}{\sigma c + \dfrac{\sigma'}{c}} \tag{5.38}$$

where σ: standard deviation in input window
             σ':   standard deviation in output window
             c: contrast control factor $0 < c < 1.3$

$$r_0 = b m_{S'} + (1 - b - r_1) m_S \tag{5.39}$$

where m_S: average value in input window
             $m_{S'}$:   average value in output window
             b: brightness control factor $0 \leq b \leq 1$
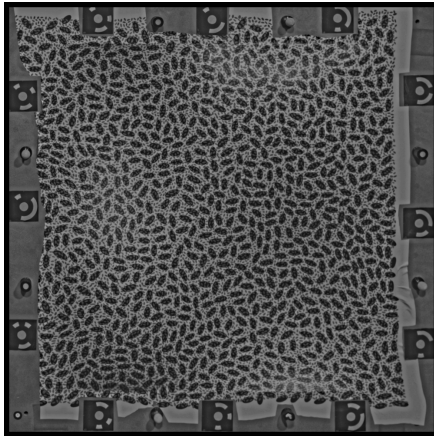
a) Image after application of Wallis filter
(original image in Fig. 5.14)
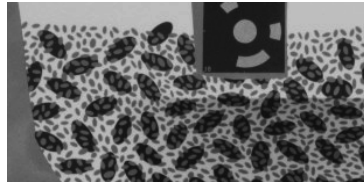
b) Detail from original image
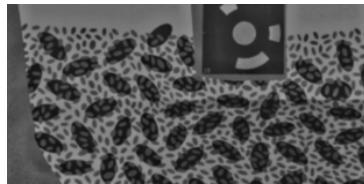
c) Detail from Wallis image

Fig. 5.42: Local contrast improvement with Wallis filter ($n = 25$, $b = 0.6$, $c = 1.6$, $m_{S'} = 120$, $\sigma' = 60$).

By fixing the average value $m_{S'}$, e.g. at 120, and standard deviation σ', e.g. at 60, the brightness and contrast ranges of the output window are defined for the entire image. The control factors b and c are determined empirically, e.g. $b = 0.8$ and $c = 1.3$. The filter size n is set according to the further processing of the image and lies, for example, between $n = 3$ and $n = 25$.

    Amongst other applications, the Wallis filter is used as a preparation for various image-matching techniques (see section 5.4.2.4 and section 5.5) in order to improve the quality of matching independently of image contrast. Fig. 5.42 shows an example

of Wallis filtering. In this case, the selected filter parameters lead to a lower contrast image, but also one with a more consistent contrast across the whole image.

5.2.4 Edge extraction

Grey-value edges are the primary image structures used by the human visual system for object recognition. They are therefore of fundamental interest for digital image measurement and pattern recognition. Every object stands out from the background on the basis of a characteristic change in the relevant image structures (Fig. 5.43).
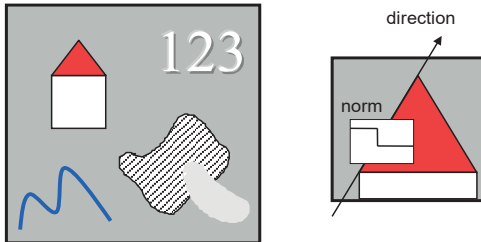
Fig. 5.43: Objects and edges.

This change in image structure can be due to:
– change in grey values (grey-value edge) along the physical object edge
   → edge extraction
– change in colour values (colour edge)
   → multi-spectral edge extraction
– change in surface texture, e.g. hatched vs. point pattern
   → texture analysis

The following discussion concerns only the extraction of grey-level edges which can be characterized by the following properties:
– a significant change in adjacent grey values perpendicular to the edge direction;
– edges have a direction and magnitude;
– edges are formed by small image structures, i.e. the region of grey value change may not be too large;
– small image structures are equivalent to high frequencies in the frequency domain.

If edges are to be extracted by means of filters then, in contrast to smoothing filters, high frequencies must be enhanced and low frequencies suppressed (= high-pass filter or band-pass filter).

An edge, or more precisely ramp, is a significant change of intensity between two grey value areas of a particular size. In contrast, a line is a thin grey value image area which cannot be resolved into two opposite edges.

5.2.5 First order differential filters

The first derivative of a continuous function $s(x)$ is given by

$$s'(x) = \frac{ds}{dx} = \lim_{\Delta x \to 0} \frac{s(x + \Delta x) - s(x)}{\Delta x} \tag{5.40}$$

and for a discrete function by:

$$s'(x) = \frac{s(x+1) - s(x)}{1} = s(x+1) - s(x) \tag{5.41}$$

A filter mask $H_2 = [-1 \ +1]$, known as a *Roberts gradient*, can be derived from (5.41). Larger filter masks, which also offer low-pass filtering, are often used, e.g.

$H_3 = [-1 \ 0 \ +1]$

$H_5 = [-1 \ 0 \ 0 \ 0 \ +1]$ or $[-2 \ -1 \ 0 \ +1 \ +2]$
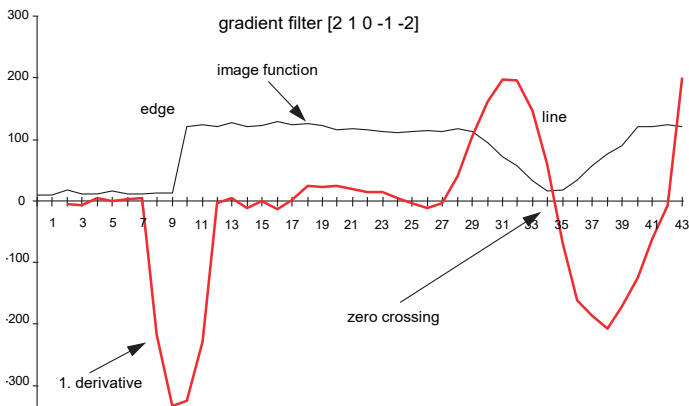
Fig. 5.44: Convolution of an image row with gradient filter $H = [2 \ 1 \ 0 \ -1 \ -2]$.

These filters approximate the discrete first derivative of the image function by (weighted) differences. They evaluate local extremes in grey value distribution from gradients calculated separately in the x and y directions. The zero crossing of the first derivative is assumed to give the position of a grey value line and its maximum or minimum gives the position of an edge (ramp), see Fig. 5.44.

From the gradient

$$\text{grad}(x,y) = g_{xy} = \left(\frac{ds(x,y)}{dx}, \frac{ds(x,y)}{dy} \right)^T \tag{5.42}$$

magnitude and direction of an edge can be derived:

$$\left| g_{xy} \right| = \sqrt{ \left(\frac{ds(x,y)}{dx} \right)^2 + \left(\frac{ds(x,y)}{dy} \right)^2 } \qquad : \text{magnitude of gradient} \tag{5.43}$$

$$\tan \varphi = \frac{ds(x,y)}{dy} \bigg/ \frac{ds(x,y)}{dx} \qquad : \text{direction of gradient} \tag{5.44}$$
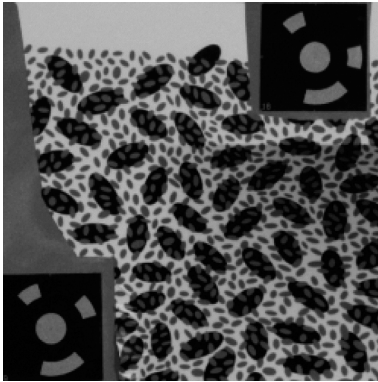
a) Original image
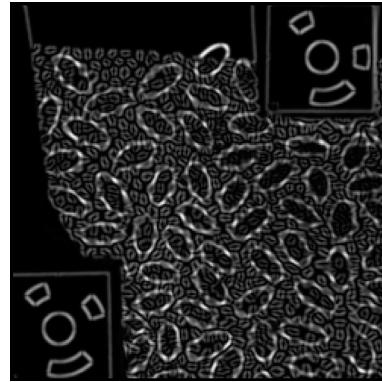
b) Gradient magnitude
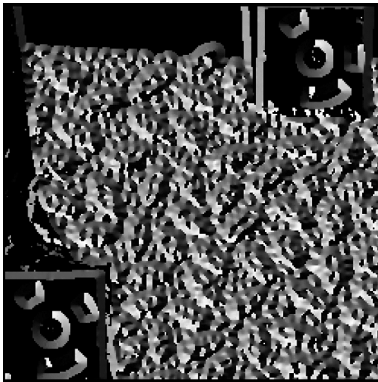
c) Gradient direction
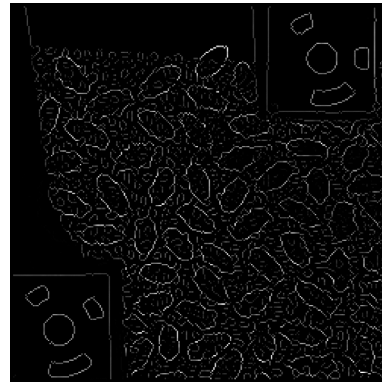
d) Line thinning

Fig. 5.45: Application of a 5x5 gradient operator.

A well-known two-dimensional gradient filter is the Sobel operator. It approximates the first derivatives in x and y by separated convolution with the filter masks

$$H_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \qquad H_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

The operator can be extended to larger filter masks. The magnitude of the gradient is calculated from the intermediate convolution results and stored as a grey value. The direction of the gradient can also be stored as a coded grey value image. Magnitude and direction images can be used for further line and edge extraction, e.g. for line thinning and chaining. Fig. 5.45 shows the application of a 5x5 gradient operator followed by line thinning.

5.2.5.1 Second order differential filters

For a continuous function $s(x)$ the second derivative is given by

$$s''(x) = \frac{d^2 s}{dx^2} = \lim_{\Delta x \to 0} \frac{s(x + \Delta x) - s(x) - [s(x) - s(x - \Delta x)]}{(\Delta x)^2} \tag{5.45}$$

and for a discrete function:

$$s''(x) = \frac{s(x+1) - s(x) - [s(x) - s(x-1)]}{(1)^2} = s(x+1) - 2s(x) + s(x-1) \tag{5.46}$$

Here a filter mask $H_3 = [+1 \ -2 \ +1]$ can be derived. The second derivative can also be generated by double application of the first derivative.
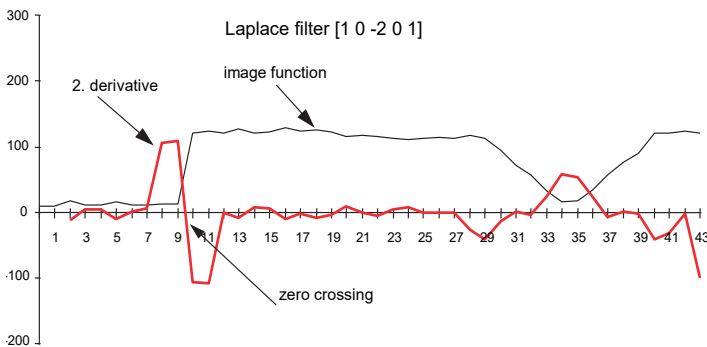
Fig. 5.46: Convolution of an image row by a Laplacian filter $H_5 = [1\ 0\ -2\ 0\ 1]$.

A grey value edge is formed by a ramp change in grey values. The position of the edge is given by the zero crossing of the second derivative (Fig. 5.46).

The second derivative of a two-dimensional function is given by the total second order differential:

$$s''(x,y) = \nabla^2 s = \frac{d^2 s}{dx^2} + \frac{d^2 s}{dy^2} \tag{5.47}$$

For a discrete function $s''(x,y)$, the second derivative can therefore be formed by addition of the partial second derivatives in the x and y directions.

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The resulting convolution mask is called a *Laplace operator* (Fig. 5.47a). Its main properties are:
–   edges are detected in all directions and so it is invariant to rotations;
–   light-dark changes produce an opposite sign to dark-light changes.

a) Laplacian operator   b) Laplacian of Gaussian operator (σ=3.0)
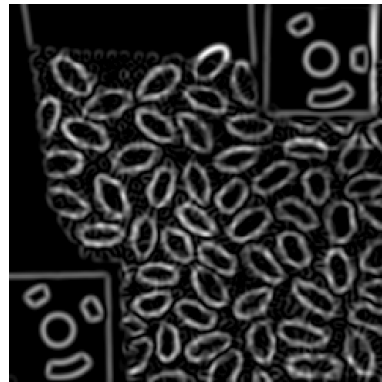
Fig. 5.47: Edge extraction with differential operators.

5.2.5.2 Laplacian of Gaussian filter

Fig. 5.47a illustrates the sensitivity to noise of the Laplace operator, hence minor intensity changes are interpreted as edges. A better result could be expected if the image were smoothed in advance. As the Gaussian filter is an optimal smoothing filter (see section 5.2.3.2), the second derivative of the Gaussian function is regarded as an optimal edge filter which combines smoothing properties with edge extraction capabilities.

The second derivative of the Gaussian function (5.36) is given by:

$$f''(r,\sigma) = \nabla^2_{GAUSS} = -\frac{1}{\pi\sigma^4}\left(1 - \frac{r^2}{2\sigma^2}\right)\cdot\exp\left(-\frac{r^2}{2\sigma^2}\right) \tag{5.48}$$

where $r^2 = x^2 + y^2$

The filter based on this function is known as the *Laplacian of Gaussian* or LoG filter. Due to its shape, it is also called a *Mexican hat* (Fig. 5.48). Empirical analysis shows that a filter size of 11 – 13 corresponds to a ±3σ interval. Fig. 5.47b shows the result of LoG filtering with σ = 2.0.
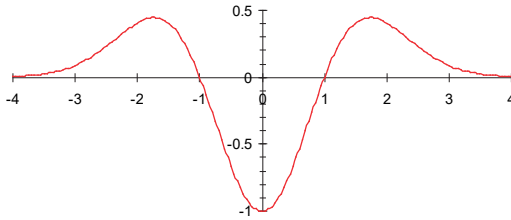
Fig. 5.48: Laplacian of Gaussian (LoG) with σ = 1.

5.2.5.3 Image sharpening

Various image sharpening filters are used to enhance the sharpness of an image. An optically defocused image cannot subsequently be focused without further information, but the visual impression can be improved. Out-of-focus images are characterized by reduced contrast and blurred edges and filters can be defined which reduce these effects. A possible sharpening filter is given by the following definition:

$$\begin{aligned}
s'(x,y) &= 2\cdot s(x,y) - [s(x,y) \otimes h(u,v)]\\
&= s(x,y) + \big(s(x,y) - [s(x,y) \otimes h(u,v)]\big)
\end{aligned} \tag{5.49}$$

Here $h(u,v)$ represents a smoothing filter, e.g. the binomial filter listed in Table 5.1. The subtraction of the smoothed image from the original corresponds to a high-pass filtering. The enhancement in the output image therefore results from adding the high frequencies in the original image. Fig. 5.49 shows an example of the effect of a sharpening filter based on a 5x5 binomial filter.

b) Detail from original image

a) Image after application of binomial
sharpness filter (original image in Fig. 5.14)

c) Detail in sharpened image
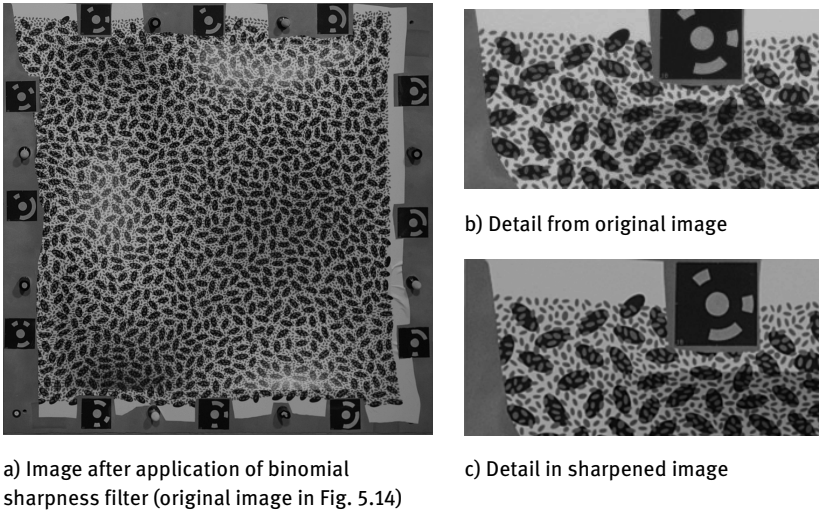
Fig. 5.49: Enhanced sharpness.

5.2.5.4 Hough transform

The Hough transform is based on the condition that all points on an analytical curve can be defined by one common set of parameters. Whilst a Hough transform can be applied to the detection of a wide variety of imaged shapes, one of the simplest solutions is the case of imaged straight lines where all points must fulfil the line equation in Hesse's normal form:

$$r = x \cos\varphi + y \sin\varphi \tag{5.50}$$

In order to determine the parameters r and φ a discrete two-dimensional parameter space (Hough space) is spanned, with elements initialized to zero. For each edge point at position x,y in the image, the direction of the gradient $\varphi+90°$ is known and r can therefore be determined. At position r,φ the corresponding value in Hough space is increased by 1, i.e. each point on the line accumulates at the same position in Hough space (due to rounding errors and noise it is actually a small local area). Hence, straight lines can be detected by searching the Hough accumulator for local maxima.

Fig. 5.50 illustrates the application of the Hough transform to an image with several well-structured edges (a). Edge extraction is performed by a 5x5 gradient filter (in analogy to Fig. 5.45), which delivers a magnitude and direction image (b). Several maxima (clusters) can be recognized in Hough space (c). Clusters which are arranged in one column of the Hough accumulator represent parallel edges in the original image. The value pairs in Hough space can be transformed back into the spatial domain x,y. Analytical lines are determined as a result, although their start and end points cannot be reconstructed (d).

Such a Hough transform is most relevant for the detection of objects predominantly formed by straight lines. The method can be expanded to curves of higher order, e.g. circles of unknown radius, although the dimension of the Hough accumulator is then no longer two-dimensional.

a) Original image

b) Gradient magnitude

d) Detected lines superimposed on the original image   c) Hough accumulator with clusters
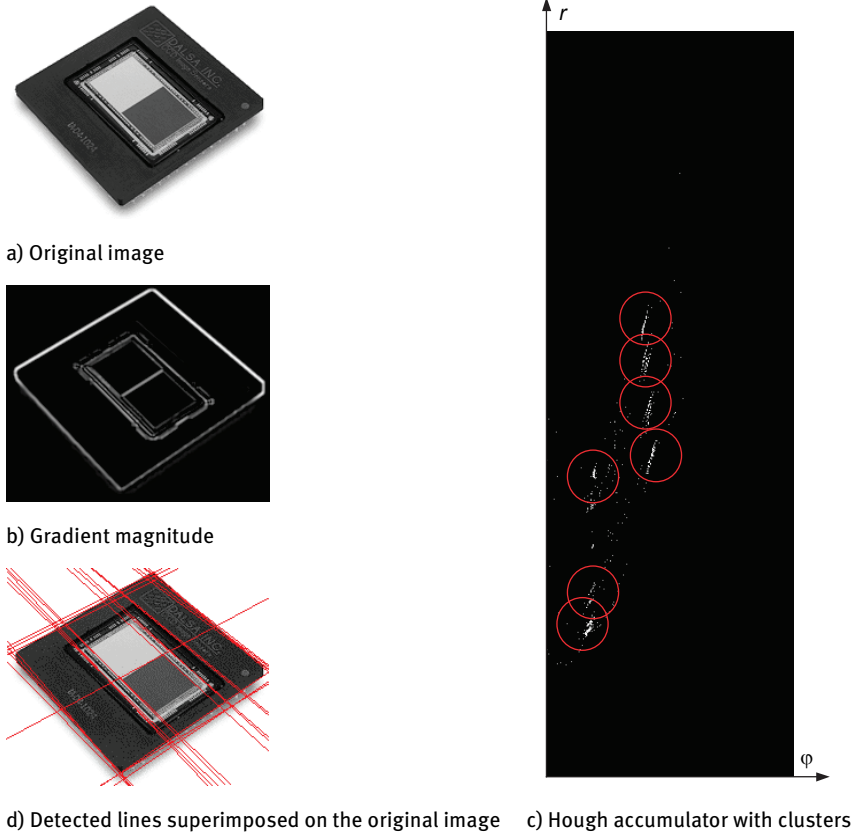
Fig. 5.50: Hough transform.

5.2.5.5 Enhanced edge operators

The simple methods of edge extraction discussed in the previous sections often do not deliver satisfactory results. An edge filter suitable for measurement tasks should have the following properties:

– complete extraction of all relevant edges (robustness);
– simple parametrization (preferably without interactive input);
– high sub-pixel accuracy;
– minimum computational effort.

Numerous methods for edge extraction are given in the literature, such as the following well-established methods. In contrast to simple convolution operators, they are extended by pre and post processing as well as adaptive parameter adjustment. They provide good results even for complex images:

–   Canny operator and Deriche operator:
    The *Canny operator* belongs to the class of operators based on LoG. It optimizes the following quality criteria in edge measurement:
    –   sensitivity to true edges (uniqueness of edge)
    –   robustness to noise (maximum signal-to-noise ratio) accuracy of edge position

    A function is defined for each criterion whose parameters are used to build a non-recursive linear filter. The Canny operator delivers a list of connected contour points with sub-pixel resolution and can be classified as one of the methods of contour following.
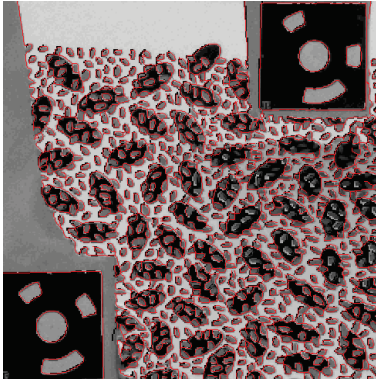
Fig. 5.51: Canny operator with σ = 1.5.   **Fig. 5.52:** Deriche operator with σ = 1.5.
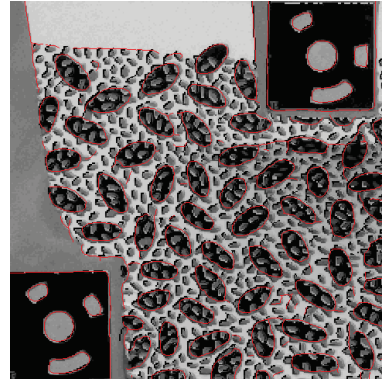
A further development is the *Deriche operator* which achieves the quality criteria above by recursive filtering. Fig. 5.51 and Fig. 5.52 show the application of both filters, where almost identical results are obtained by appropriate parameter settings.

–   Edge extraction in image pyramids:
    Image pyramids (see section 5.1.3.1) represent the image content at different spatial resolutions. Since the ability to detect edges in natural images varies with image resolution (image scale), an approach is sought which determines the optimal scale for each edge pixel. LoG filters, or morphological operators, can be used as edge operators.

– Least-squares edge operators:
A model describing the geometric and radiometric properties of the edge can be determined by least-squares parameter estimation. By adjusting their initial parameter values, a priori edge models (templates) can be fitted to the actual region defining an edge (see section 5.4.2.4, least-squares template matching). An optimal fit could be the least squares estimate in which the differentials have the least entropy.

A global approach to edge extraction is possible if the energy function

$$E = E_{int} + E_{grey} + E_{ext} \tag{5.51}$$

where E_{int}:   curve energy function
          E_{grey}:   grey-value conditions
          E_{ext}:   geometric constraints

is minimized in a least-squares solution. The curve energy function E_{int} describes the behaviour of the curvature along the edge or the sensitivity with respect to possible changes in direction. Grey-value conditions E_{grey} along the edge can be defined by requiring, for example, maximum gradients. Additional geometric constraints such as straight lines or epipolar geometries are specified by E_{ext}.

5.2.5.6 Sub-pixel interpolation
5.2.5.6.1 Sub-pixel resolution
In digital photogrammetry, line and edge filters are used for the measurement of geometric elements (points, lines) which are described by their contours (see section 5.4.2.5). The objective is to locate these patterns to the highest accuracy. As discussed in section 5.1.2, object structures covering several pixels can be measured by interpolation to the sub-pixel level.

The theoretical resolution of the position of a digitized grey-value edge is, in the first instance, a function of the slope of the grey values along the edge and quantization depth (number of bits per grey value). It is defined by a parameter, sometimes called the slack value, d, where the position of the imaged edge can be varied without changing the related grey values (Fig. 5.53). For a step change ($\beta = 0$ in diagram), the slack is maximized and amounts to 1.

It can be shown that, for N quantization steps, the uncertainty of edge positioning is at least $1/(N-1)$. This corresponds to 0.004 pixels for $N = 256$ grey levels. In this case the average deviation of a *single* edge point can be estimated, independently of the slope of the edge, as

$$\bar{\sigma}_d = 0.015 \ \text{pixel} \tag{5.52}$$

This theoretical quantity will be higher in practice due to optical limitations and noise but it defines a lower limit for the estimation of positional accuracy of image measurements. Hence, the accuracy of an edge of length k pixels can be estimated as

$$\overline{\sigma}_K = \frac{\overline{\sigma}_d}{\sqrt{k}} \ \ \text{pixel} \tag{5.53}$$
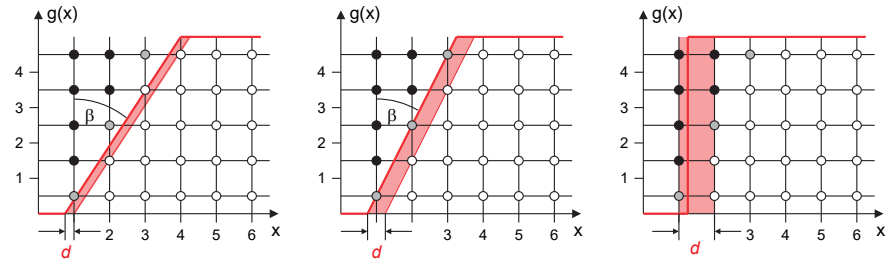
Fig. 5.53: Slack in a digitized edge (according to Förstner 1985).

This is also approximately true for the centre of a circle measured by k edge points. For a circular target of diameter 6 pixels there are approximately $k = 19$ edge points. Hence, in the ideal case, the centre of such a target can be determined to an accuracy of about 0.004 pixel (4/1000 pixel). Further investigations are discussed in section 5.4.2.6.

A selection of edge extraction methods providing sub-pixel interpolation, together with their principal parameters, are summarized in Table 5.2.

a) Zero-crossing interpolation b) Moment preservation
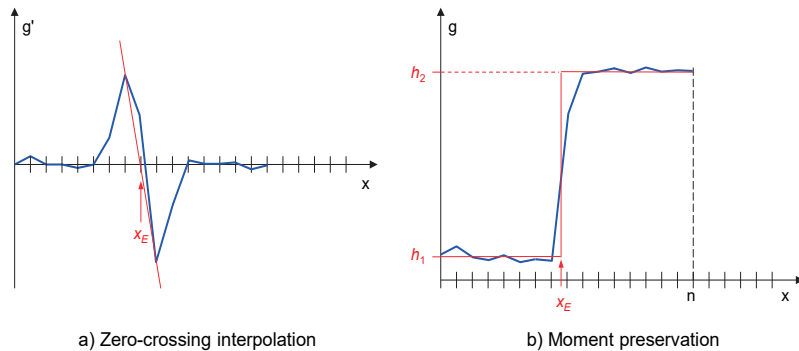
Fig. 5.54: Edge interpolation methods.

5.2.5.6.2 Zero-crossing interpolation

As shown in Fig. 5.44 and Fig. 5.46 the derivative functions do not pass zero at an integer pixel coordinate. The sub-pixel position of an edge can be determined by first or second order interpolation in the neighbourhood of the zero crossing (Fig. 5.54a). This method is used, for example, for the edge measurement of patterns based on

points (see section 5.4.2.5). Interpolation of zero crossings provides edge location with a precision of up to 1/100 pixel.

Table 5.2: Methods for sub-pixel interpolation of edge points.

Method	Model	Intermediate result	Sub-pixel interpolation
Differential filter	deflection point	gradient image	linear interpolation
Moment preservation	grey value plateaux	1st, 2nd and 3rd moments	solution of 3 equations
Feature correlation (template matching)	edge template, cross-correlation	correlation coefficients	2nd order interpolation
Least-squares matching	edge template, geometric and radiometric transformation	up to 8 transformation parameters	shift parameters

5.2.5.6.3 Moment preservation

The basic idea of the moment preservation method assumes that an edge within a one-dimensional image function, e.g. an image row of window size n, can be described by three parameters. These define the left grey-value plateau h_1, the right grey-value plateau h_2 and the coordinate of the grey-value step x_k (Fig. 5.54b). The three required equations are formed by the 1st, 2nd and 3rd moments:

$$m_i = \frac{1}{n}\sum_{j=1}^{n} g_j^i \quad i = 1,2,3 \tag{5.54}$$

The following parameters are determined (without derivation):

$$h_1 = m_1 - \overline{\sigma}\sqrt{\frac{p_2}{p_1}}$$
$$h_2 = m_1 + \overline{\sigma}\sqrt{\frac{p_1}{p_2}}$$

$$p_1 = \frac{1}{2}\left[1 + \overline{s}\sqrt{\frac{1}{4 + \overline{s}^2}}\right] \tag{5.55}$$

where

$$\overline{s} = \frac{m_3 + 2m_1^3 - 3m_1 m_2}{\overline{\sigma}^3}$$
$$\overline{\sigma}^2 = m_2 - m_1^2$$

$$p_2 = 1 - p_1 \tag{5.56}$$

The desired edge position is given by:

$$x_K = n \cdot p_1 \tag{5.57}$$

The moment preservation method is easy to implement and it delivers the sub-pixel edge location without any further interpolation. It is used, for example, with the Zhou operator for the edge extraction of elliptically shaped targets (see section 5.4.2.5).

5.2.5.6.4 Correlation methods

Correlation methods determine the position in a search image which has the highest similarity with a reference pattern (*template*). The reference pattern can be a subset of a natural image, or a synthetically created image. For example, when searching for a vertical edge which switches from dark to light, a template similar to the one in Fig. 5.55 can be used.

The similarity between two patterns can be measured by the normalized cross-correlation coefficient r (see sections 2.4.3.3 and 5.4.2.3). If r is plotted as a function of x, the position of maximum correlation is most likely to be the true position of the reference pattern. If the curve around the maximum is approximated by a quadratic function (parabola), the desired position can be determined to sub-pixel precision (Fig. 5.55).

interpolation by correlation coefficients   template for edge detection
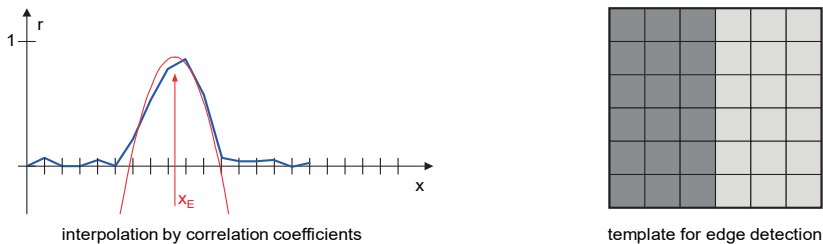
Fig. 5.55: Interpolation by correlation coefficients.

Least-squares matching (see section 5.4.2.4) determines a transformation which describes both the change of contrast and the geometric projection between reference pattern and search image. The approach requires reasonably good initial values of the unknown parameters, in particular the required shift parameters. The adjustment process directly delivers the sub-pixel positions of the pattern (edge).

5.3 Geometric image transformation

The process of modifying the geometric projection of a digital image is here referred to as a geometric image transformation. Related methods are required in photogrammetry for photo rectification and orthophoto production (section 4.2.8), for combining images with CAD models, for calculation of distortion-free images and

normalized stereo images (epipolar images), and for template matching procedures. Rendering and morphing methods also belong to this category. Fig. 5.56 shows an example of the projective rectification of a façade (see also Fig. 4.24).

Fig. 5.56: Geometric rectification of a façade.

The term *rectification* denotes a general modification of pixel coordinates, e.g. for
– translation and rotation;
– change of scale or size (magnification, reduction);
– correction of distortion effects;
– projective rectification (from central perspective to parallel projection);
– orthophoto production (differential rectification);
– rectification of one image with respect to another;
– superimposition of natural structures onto a surface, e.g. a CAD model (*texture mapping*).

Geometric image transformations are generally performed in two stages:
1. Transformation of pixel coordinates (image coordinates) into the target system (rectification) – this transformation is the reverse of the imaging process.
2. Calculation (interpolation) of output grey values.

5.3.1 Fundamentals of rectification

Rectification is founded on the geometric transformation of pixel coordinates from an original image to an output image:

$$s'(x',y') = G\big(s(x,y)\big) = g' \tag{5.58}$$

where
$$x' = f_1(x,y)$$
$$y' = f_2(x,y)$$

Here the grey value *g,* at position (x,y) in the original image, appears in the output image, after grey value interpolation *G,* as grey value *g'* at position (x',y').

The geometric transformations f_1 and f_2 can be almost arbitrary coordinate transformations. The affine transformation (2.12) is often used for standard modifications such as translation, scaling, rotation or shearing. The projective transformation (8 parameters, see eqn. 2.21) is suited to the rectification of images of planar objects. Where a digital surface model is available, arbitrary free-form surfaces can be transformed into orthophotos by use of the collinearity equations (4.10) (see section 4.2.8.2).
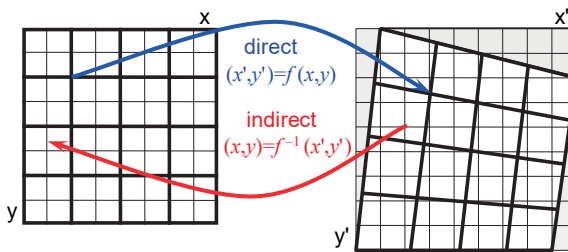
Fig. 5.57: Rectification methods.

In all cases there is a transformation between the pixel coordinates (x,y) of the input image and the pixel coordinates (x',y') of the output image. The grey values of the input image in the region of (x,y) must be stored in the output image at position (x',y'). For this purpose, the indirect rectification method is usually applied in which the output image is processed pixel by pixel. By reversing the geometric transformation, the grey value of the input image is interpolated at the reverse-transformed position (x,y) and then stored in the output image. This algorithm is easy to implement and avoids gaps or overlapping regions in the output image (Fig. 5.57).

5.3.2 Grey value interpolation

The second step in rectification consists of the interpolation of a suitable grey value from the local neighbourhood using an arbitrary non-integer pixel position, and then storing this quantity in the output image (resampling). The following methods are normally used for grey value interpolation:

- zero order interpolation (nearest neighbour);
- first order interpolation (bilinear interpolation);
- second order interpolation (bicubic convolution, Lagrange polynomials).
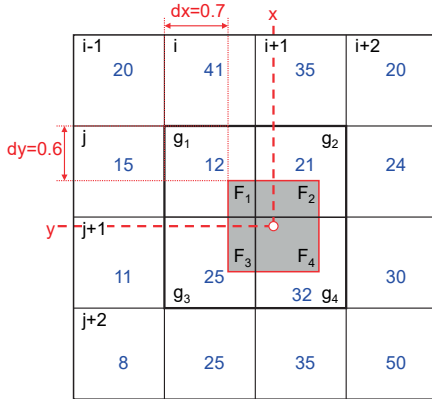
Fig. 5.58: Grey-value interpolation in 2x2 and 4x4 neighbourhood.

For the method of *nearest neighbour*, the grey value at the rounded or truncated real pixel coordinate is used in the output image. The interpolation rule is given by:

$$g' = s(\text{round}(x), \text{round}(y)) \tag{5.59}$$

Nearest-neighbour grey value in the example of Fig. 5.58: $g' = 32 = g_4$

This approach leads to the visually worst rectification result. However, the computational effort is small and the output image consists only of grey values which also exist in the input image.

The *bilinear* or *biquadratic interpolation* takes into account the 2x2 adjacent grey values of the computed pixel position. The interpolated grey value is the result of the weighted average of adjacent grey values in which the weight is given by the relative coverage of the current pixel. The interpolation rule is given by:

$$g' = F_1 \cdot s(i,j) + F_2 \cdot s(i+1,j) + F_3 \cdot s(i,j+1) + F_4 \cdot s(i+1,j+1) \tag{5.60}$$

where $F_1 + F_2 + F_3 + F_4 = 1$

or analogously to eqn. (2.19):

$$\begin{aligned} g' = s(i,j) + dx \cdot [s(i+1,j) - s(i,j)] + dy \cdot [s(i,j+1) - s(i,j)] \\ + dx \cdot dy \cdot [s(i+1,j+1) - s(i+1,j) - s(i,j+1) + s(i,j)] \end{aligned} \tag{5.61}$$

Hence the bilinear grey value in the example of Fig. 5.58: $g' = 25.3 \approx 25$

With modest computational effort, bilinear interpolation generates slightly smoothed rectifications of good quality.

Bicubic convolution and *Lagrange interpolation* are usually applied only in output rectifications where the highest image quality is required. These methods use a 4x4 environment for interpolation which results in computation times up to 10 times higher. The algorithm for bicubic convolution is as follows:

$$df(x) = |x|^3 - 2 \cdot |x|^2 + 1 \qquad \text{for } |x| < 1$$
$$df(x) = -|x|^3 + 5 \cdot |x|^2 - 8 \cdot |x| + 4 \qquad \text{for } 1 \le |x| < 2$$
$$df(x) = 0 \qquad \text{other cases}$$

$$\begin{aligned} a(n) = & \; s(i-1, j+n-2) df(dx+1) \\ & + s(i, j+n-2) \cdot df(dx) \\ & + s(i+1, j+n-2) \cdot df(dx-1) \\ & + s(i+2, j+n-2) \cdot df(dx-2) \end{aligned} \qquad \text{for } n = 1,2,3,4 \qquad (5.62)$$

$$g' = a(1) \cdot df(dy+1) + a(2) \cdot df(dy) + a(3) \cdot df(dy-1) + a(4) \cdot df(dy-2)$$

Bicubic grey value for the example of Fig. 5.58: $g' = 21.7 \approx 22$
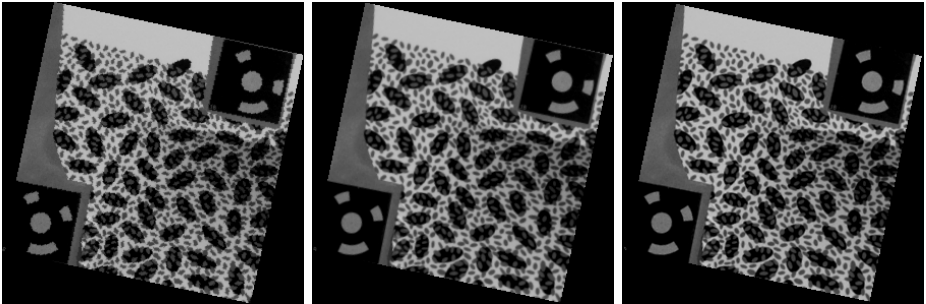
a) Nearest neighbour   b) Bilinear interpolation   c) Bicubic convolution

Fig. 5.59: Rectification results with different methods of grey-value interpolation.

The arbitrarily chosen example of Fig. 5.58 shows clearly that the three different interpolation methods generate quite different values for the interpolated grey level. Fig. 5.59 shows an enlarged image region after rotation, generated by the three methods. The nearest neighbour approach gives rise to clearly visible steps along sloping edges. The other two interpolation methods yield results which are visually very similar.

The problem of grey value interpolation at non-integer positions also occurs in least-squares matching (see section 5.4.2.4). Bilinear interpolation is normally used

as an efficient compromise between computation time and image quality. If edge detection is required in the output image, it is better to apply it prior to rectification. This avoids interpolation bias and less processing power is required to rectify edges than to rectify large pixel arrays.

5.3.3 Textured images

5.3.3.1 Texture mapping

The pattern or structure on a surface is referred to as texture. It results from a variation of small surface elements within a limited neighbourhood. Textures are physically generated by the different reflection properties and geometric characteristics of surface particles which, in the limit, can be as small as molecules. With the aid of textures, visualized objects achieve their realistic appearance.
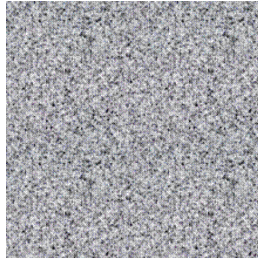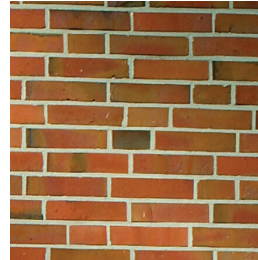
a) Wood   b) Granite   c) Bricks

Fig. 5.60: Artificial and natural textures.

The generation of texture in a visualized scene can be generated most simply by projection of a texture image onto the surface (texture mapping). A texture image is a digital image whose pattern is superimposed on the object in the manner of a slide projection. The texture image can be a real photograph or can consist of artificial patterns (examples in Fig. 5.60).

Texture mapping includes the following principal techniques:
– 2D texture projection onto a planar object surface (2D→2D)
– 3D texture projection onto object surfaces of arbitrary shape (2D→3D→2D)

For 2D texture mapping, a plane transformation between a region in the visualized image and the texture image is calculated. According to the principle of plane rectification (see section 4.2.8.1), corresponding image patches are defined in both images and their corner points are mapped by affine or projective transformation. There is no direct relationship to the 3D object model, i.e. for each new visualization the texture regions must be re-defined, if necessary by manual interaction.

Fig. 5.61 illustrates the schematic process. The relationship between output image (visualization) and texture image is given by a plane coordinate transformation. Its parameters are derived through identical points (control points) on a plane, e.g. triangle, quadrangle or polygon. Then each pixel of the output image is processed according to the blue work flow, i.e. the coordinate transformation is calculated only once per plane image section.
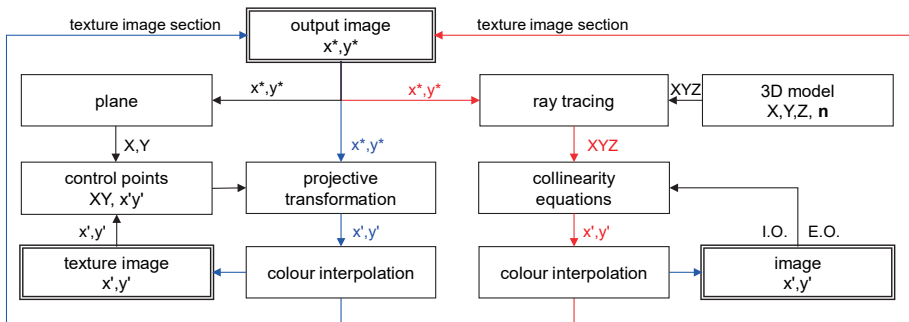
Fig. 5.61: Texture mapping.

Fig. 5.62: Texture mapping with natural patterns.

Texture mapping is more flexible if the known geometric relationship between visualization (projection) and 3D model is utilized. Each output pixel defines a ray in space which is given by selection of viewing direction and geometric projection. This ray is intersected with the 3D model by application of ray tracing methods in order to handle any occlusions. The calculated intersection points are projected back into the original image using the collinearity equations. From the resulting pixel position a suitable colour value is interpolated and transferred into the output image (red path

in Fig. 5.61). If the process has an integrated object-based reflection and lighting model, photo-realistic scenes can be created which include lighting effects, shadows and the natural object structure.

Fig. 5.62 shows an example of texture mapping in which terrestrial photographs are used for the façades, and aerial photographs for the ground. The trees are artificial graphics objects.

5.3.3.2 Synthetic images

For the creation of synthetic images to be used in place of real photogrammetric images, arbitrary interior and exterior orientation data can be defined for the output image. Again, each pixel x^*,y^* defines a spatial ray whose processing follows the red path in Fig. 5.61.

a) Original image   b) Synthetic image

Fig. 5.63: Example of a synthetic image (compare also with Fig. 5.105).

Fig. 5.63b shows an example of a synthetic image which has been generated with a 3D model of a sculpture using textures taken from the original image of Fig. 5.63a. Synthetic images can be used, for example, to test a certain image-measuring method since all input values are known for comparison with the computed results. A well-known tool for rendering synthetic images is the open-source software Blender.

5.4 Digital processing of single images

This section deals with methods for locating objects in single digital images. It distinguishes between algorithms for the determination of single point features (pattern centres) and those for the detection of lines and edges. The common aim of these methods is the accurate and reliable measurement, to sub-pixel resolution, of image coordinates for use in analytical object reconstruction. Three-dimensional image processing methods are discussed in section 5.5.

5.4.1 Approximate values

5.4.1.1 Possibilities
The image processing methods discussed here require initial approximations for the image position which is to be accurately determined. These approximate values can be found in different ways:

– by pre-determined (calculated) image coordinates, e.g. from (approximately) known object coordinates and known image orientation parameters;
– by manually setting an on-screen cursor;
– by pattern recognition (segmentation) during image pre-processing, e.g. by searching for certain grey value patterns in the image;
– using interest operators which detect regions of significant image structure (see section 5.4.3.1).

In close-range photogrammetry, in particular for industrial applications, simple image structures can often be engineered through specific targeting and illumination techniques. These can ensure, for example, that only (bright) object points on a homogeneous (dark) background exist in the image (examples in Fig. 5.2a, Fig. 7.30). In this case, the generation of approximate values reduces to the location of simple image patterns and can, in many cases, be fully automated (see section 5.4.1.2).

The use of coded targets in industrial applications has been particularly successful. These encode a point identification which can be detected and decoded automatically (see section 6.2.1.4).

5.4.1.2 Segmentation of point features
The measurement (segmentation) of bright targets is an important special case in practical photogrammetric image acquisition. If no information about the position of target points is available, the image must be searched for potential candidates. Since measurement tasks in practice involve different image scales, perspective distortion, extraneous lighting, occlusions etc., the following hierarchical process of point segmentation has proven effective:

1. Adaptive binarization by thresholding (see section 5.2.1.4).
2. Detection of connected image regions exceeding a threshold (connected components).
3. Analysis of detected regions with respect to size (number of pixels) and shape.
4. Storage of image positions which meet appropriate conditions of size and shape.

Thresholding is relatively simple for artificial targets. If retro-reflective targets with flash illumination are used, or LED targets, these generate significant peaks in the upper region of the grey value histogram (see Fig. 5.19). For non-reflective targets, e.g. printed on paper, the contrast against the image background is often weaker. In some

circumstances, thresholding cannot then be performed globally but must be adapted to different parts of the image.

Connected components (regions) are detected by a neighbourhood or connectivity analysis. Using the sequential process outlined in Fig. 5.64, the left and the upper three neighbouring pixels are analysed at the current pixel position. If the grey value of one of these neighbours exceeds the threshold, the neighbour belongs to a region already detected. Otherwise a new region is created. Pixels in connected components are marked in lists or registered in corresponding label images. V-shaped objects, which may be assigned to more than one region, are recombined by a contact analysis.
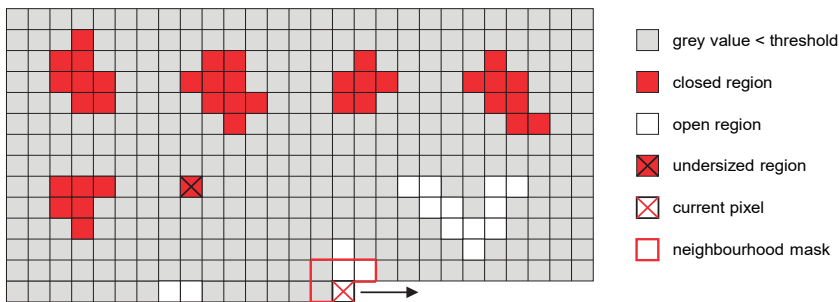
Fig. 5.64: Sequential connectivity analysis (according to Maas 1992).

In a non-sequential connectivity analysis, adjacent pixels are traced recursively in all directions until one or more of the following termination criteria are met:
– the corresponding grey value is less than the threshold;
– the maximum number of pixels permitted in a region is exceeded;
– the pixel already belongs to a known region.

The resulting connected components can now be analysed with respect to their shape and size. Suitable criteria are the number of connected pixels, the standard deviation of a best-fit ellipse from the surrounding contour or the ratio of semi-axes b/a.

The regions which remain after binarization and shape analysis are then finally indexed and stored sequentially.

Fully-automated segmentation is particularly successful if the following criteria are met:
– good contrast in brightness between target and background through suitable targeting and illumination techniques (retro-targets, flash);
– no extraneous reflections from the object surface (avoidance of secondary light sources);
– no occlusions due to other object parts;

–   good separation of individual targets in image space;
–   minimum target size in image space (5-pixel diameter);
–   no excessive differences size between imaged targets.

The problems of target imperfections and false detections are illustrated in the examples of Fig. 5.20ff, Fig. 5.40, Fig. 5.41 and Fig. 5.65a. In some cases, the segmentation steps above must be extended by additional pre-processing, e.g. morphological operations, and further analysis.

5.4.2 Measurement of single point features

Here single point features are taken to mean image patterns where the centre of the pattern is the reference point. Examples are shown in Fig. 5.65.

a) Circular target   b) Retro-reflective target   c) Checkerboard target with cursor in a zoomed window
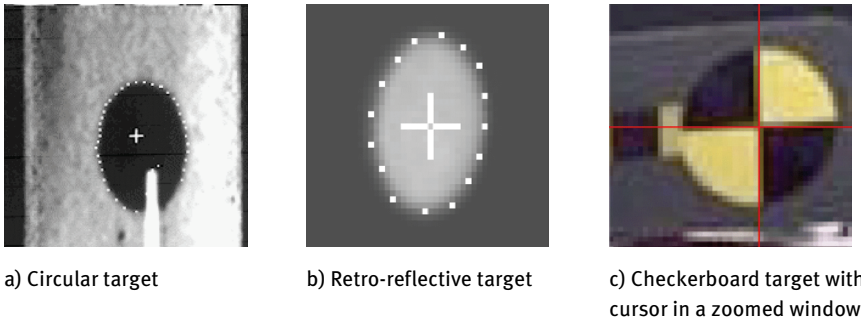
Fig. 5.65: Examples of single point features with detected contour points.

5.4.2.1 On-screen measurement

Arbitrary image features can be measured manually by positioning a digital floating mark (cursor) on the computer screen. For this purpose, the cursor should be displayed as a cross or circle. The minimum movement of the floating mark is 1 pixel in screen coordinates. The average measurement accuracy of non-targeted points, e.g. building corners, is around 0.3-0.5 pixel. If the image is zoomed, the measurement accuracy can be improved to approximately 0.2 pixel.

5.4.2.2 Centroid methods

If the feature to be measured consists of a symmetrical distribution of grey values, the local centroid can be used to determine the centre. The centroid is effectively a weighted mean of the pixel coordinates in the processing window:

$$x_M = \frac{\sum\limits_{i=1}^{n}(x_i \cdot T \cdot g_i)}{\sum\limits_{i=1}^{n}(T \cdot g_i)} \qquad\qquad y_M = \frac{\sum\limits_{i=1}^{n}(y_i \cdot T \cdot g_i)}{\sum\limits_{i=1}^{n}(T \cdot g_i)} \qquad (5.63)$$

Here n is the number of processed pixels in the window, g_i is the grey value at the pixel position (x_i, y_i). The decision function T is used to decide whether a pixel is used for calculation. T can be defined by an (adaptive) grey value threshold t, for example:

$$T = \begin{cases} 0 & \text{for } g < t \\ 1 & \text{for } g \geq t \end{cases}$$

Optionally, the calculation of grey-value sums in the denominator of eqn. (5.63) can be accelerated using the method of integral images (section 5.1.3.4). For features whose structure is defined by grey-value edges, such as the circumference of a circle, it is reasonable to include edge information in the centroid calculation. For this purpose, a weighting function based on gradients is employed:

$$x_M = \frac{\sum\limits_{i=1}^{n}x_i\, grad^2(g_{x,i})}{\sum\limits_{i=1}^{n} grad^2(g_{x,i})} \qquad \text{or} \qquad y_M = \frac{\sum\limits_{i=1}^{n}y_i\, grad^2(g_{y,i})}{\sum\limits_{i=1}^{n} grad^2(g_{y,i})} \qquad (5.64)$$

Centroid operators are computationally fast and easy to implement. In general, they also work for very small features ($\varnothing < 5$ pixel) as well as slightly defocused points. However, the result depends directly on the grey-value distribution of the environment so they are only suitable for symmetrical homogenous patterns as shown in Fig. 5.65b. Defective pixels within the processing window will negatively affect the calculation of the centre.

The theoretical accuracy of the centroid can be estimated by applying error propagation to (5.63):

$$\sigma_{x_M} = \frac{\sigma_g}{\sum g_i}\sqrt{\sum(x_i - x_M)^2} \qquad\qquad \sigma_{y_M} = \frac{\sigma_g}{\sum g_i}\sqrt{\sum(y_i - y_M)^2} \qquad (5.65)$$

The standard deviation of the centroid is clearly a linear function of the grey value noise σ_g, and the distance of a pixel from the centre. It is therefore dependent on the size of the feature.

Example 5.2:

A targeted point with the parameters

point diameter:	6 pixels
operator window size:	13 x 13 pixels
target grey value:	200
background grey value:	20
grey-value noise:	0.5 grey level

results in a theoretical standard deviation of the centroid of $\sigma_{x_M} = \sigma_{y_M} = 0.003$ pixel.

In practice, centroid operators can reach an accuracy 0.03 – 0.05 pixel if circular or elliptical white targets on dark backgrounds are used (see section 5.2.4.6).

5.4.2.3 Correlation methods

In image processing, correlation methods are procedures which calculate a similarity measure between a reference pattern $f(x,y)$ and a target image patch extracted from a larger search area within the acquired image $g(x,y)$. The position of best agreement is assumed to be the location of the reference pattern in the image.

A common similarity value is the normalized cross-correlation coefficient. It is based on the following covariance and standard deviations (see section 2.4.3.3):

$$\rho_{fg} = \frac{\sigma_{fg}}{\sigma_f \sigma_g} \qquad \text{: correlation coefficient} \tag{5.66}$$

where

$$\sigma_{fg} = \frac{\sum[(f_i - m_f)(g_i - m_g)]}{n} \qquad \sigma_f = \sqrt{\frac{\sum(f_i - m_f)^2}{n}} \qquad \sigma_g = \sqrt{\frac{\sum(g_i - m_g)^2}{n}}$$

m_f, m_g : arithmetic mean of grey values

f_i, g_i: : grey value in reference and search image with n pixels, $i = 1...n$

For pattern recognition, the reference pattern is successively shifted across a window of the search image according to Fig. 5.37, with the correlation coefficient calculated at each position.

Fig. 5.66 shows the reference image and correlation result of the search for circular targets in Fig. 5.2. As expected, correlation maxima occur at bright targets, but medium correlation values are also caused by background noise and edges.

To identify pattern positions in the search image, all x,y positions with correlation coefficient ρ greater than a threshold t are stored. The choice of a suitable threshold t depends on the image content. For correlating stereo images of similar appearance, the threshold t can be derived from the auto-correlation function. A suitable threshold lies in the range $t = 0.5$ to 0.7. Where synthetic patterns (templates) are used for image correlation, optimal correlation values may be lower if the corresponding patch in the

search image deviates from the reference image in terms of background or other disturbances.

a) Reference pattern
(enlarged)

b) Correlation result
(black = high correlation)
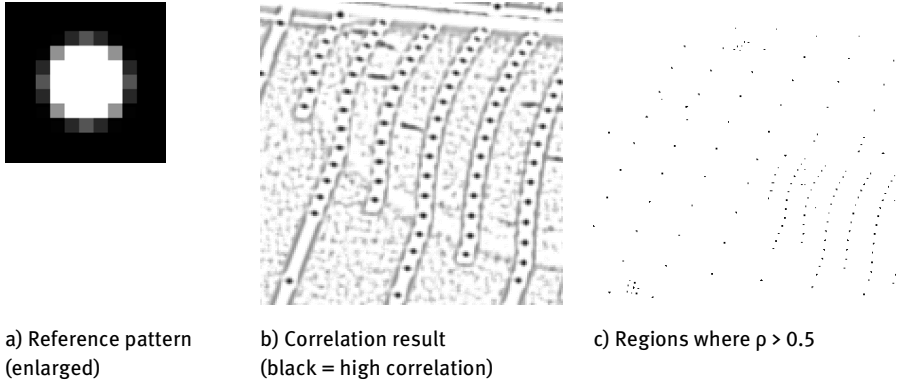
c) Regions where ρ > 0.5

Fig. 5.66: Cross correlation (window of search image of Fig. 5.2a).

In the regions where correlation maxima are detected, a further interpolating function applied to the neighbouring correlation values can determine the feature position to sub-pixel coordinates (see section 5.2.4.6).

The calculation process can be accelerated by prior calculation of the values which are constant in the reference image (σ_f in eqn. 5.66), by using integral images (section 5.1.3.4) for calculation of the grey value sums in eqn. (5.66) or by reducing the image resolution. With reduced resolution the pattern matrix can then be shifted across the image in larger steps but this effectively leads to a loss of information. However, a hierarchical calculation based on image pyramids (see section 5.1.3.1) can be performed in which the search results of one stage are used as prior knowledge for the next higher resolution stage.

Using *auto-correlation* the uniqueness of the correlated search area can be tested. A search window is defined around a pixel and, within that, a smaller central patch is used as a reference to be shifted across the search area. The variation of the correlation coefficient indicates a measure for the detectability of the pattern. Fig. 5.67a shows the auto-correlation of a homogeneous bright area. As always, at the central position the resulting correlation coefficient is 1 while only a poor correlation exists for the rest of the pattern. In Fig. 5.67b the auto-correlation along an edge shows high values while a sharp drop-off is visible perpendicular to the edge. Finally, Fig. 5.67c shows the result for a cross-shaped feature that also appears in the auto-correlation function.

Cross correlation is a robust method, independent of contrast but requiring a high computational effort. Target patterns can have an almost arbitrary structure. However, differences in scale and rotation, or any other distortions between reference image and target image, are not readily modelled and lead directly to a reduction in similarity value. An image measuring accuracy of about 0.1 pixel can be achieved.

a) Homogenous pattern   b) Edge   c) Cross
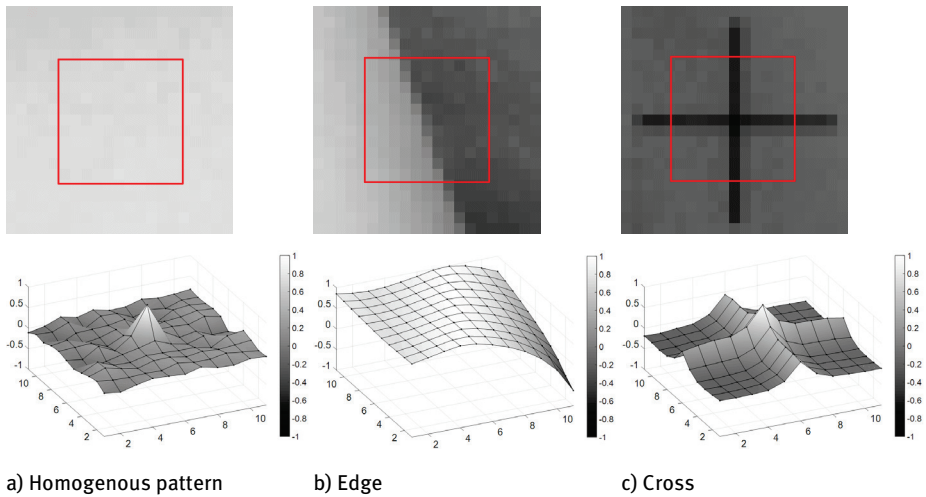
Fig. 5.67: Auto-correlation of different image features (search area 21x21 pixel; reference window 11x11 pixel).

5.4.2.4 Least-squares matching

5.4.2.4.1 Principle

The method of least-squares matching (LSM) employs an iterative geometric and radiometric transformation between reference image and search image in order to minimize the least-squares sum of grey-value differences between both images. The reference image can be a window in a real image which must be matched in a corresponding image, e.g. stereo partner. For a known grey-value structure, the reference image can be generated synthetically and used as a template for all similar points in the search image (*least squares template matching*).

The geometric fit assumes that both image patches correspond to a plane area of the object. The mapping between two central perspective images can then be described by the projective transformation (2.21). For sufficiently small image patches where the 3D surface giving rise to the imaged area can be assumed to be planar, the 8-parameter projective transformation can be replaced by a 6-parameter affine transformation (2.12). The six parameters are estimated by least-squares adjustment using the grey values of both image patches as observations. The radiometric fit is performed by a linear grey-value transformation with two parameters.

template

search image

iteration 1   2   3   4   5   6
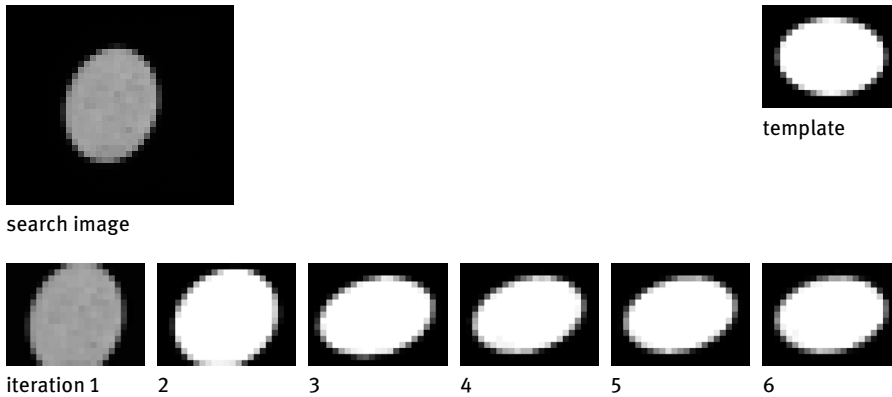
Fig. 5.68: Iterative transformation of an image patch.

Fig. 5.68 shows the iterative geometric and radiometric transformation of a reference pattern to the target pattern, which has been extracted from a larger search area. The formulation as a least-squares problem has the following implications:

– An optimal solution is obtained if the mathematical model is a reasonably good description of the optical imaging process.
– The mathematical model can be extended if additional information or conditions are available, e.g. geometric constraints between images or on the object.
– The approach can be adapted to simultaneous point matching in an unlimited number of images.
– The observation equations are non-linear and must therefore be linearized at given initial values.
– The adjustment is normally highly redundant because all grey values in the image patch are used as observations to solve for only eight unknowns.
– Accuracy estimates of the unknowns can be derived from the covariance matrix. Internal quality measures are therefore available which can be used for blunder (gross error) detection, quality analysis and post-processing.

Least-squares matching was developed in the mid-eighties for digital stereo image analysis and is now established as a universal method of image analysis. In two-dimensional image processing it can, in addition to single point measurement, be applied to edge extraction and line following. For 3D object reconstruction it can be configured as a spatial intersection or bundle triangulation, and also integrated into the determination of object surface models and geometric elements.

5.4.2.4.2 Mathematical model
Given two image patches $f(x,y)$ and $g(x,y)$, identical apart from a noise component $e(x,y)$:

$$f(x,y) - e(x,y) = g(x,y) \tag{5.67}$$

For a radiometric and geometric fit, every grey value at position (x,y) in the reference image f_i is expressed as the corresponding radiometrically and geometrically transformed grey value g_i at position (x',y') in the search image as follows:

$$
\begin{aligned}
f_i(x,y) - e_i(x,y) &= r_0 + r_1 g_i(x',y') &\quad i &= 1,\dots,n \\
x' &= a_0 + a_1 x + a_2 y &\quad n &= p \cdot q \quad \text{(window size)} \\
y' &= b_0 + b_1 x + b_2 y &\quad n &\geq 8
\end{aligned}
\tag{5.68}
$$

Both translation parameters a_0 and b_0 are of major importance as they define the relative shift between reference image and search image. Coordinates x',y' are non-integer values and so the corresponding grey values must be appropriately interpolated, e.g. using bilinear interpolation (see section 5.3.2).

The observation equation (5.68) must be linearized since the image function $g(x',y')$ is non-linear. In summary, the linearized correction equations are as follows (ignoring index i):

$$
\begin{aligned}
f(x,y) - e(x,y) = g(x',y') &+ r_1 g_x da_0 + r_1 g_x x da_1 + r_1 g_x y da_2 + \\
&+ r_1 g_y db_0 + r_1 g_y x db_1 + r_1 g_y y db_2 + dr_0 + dr_1 g(x',y')
\end{aligned}
\tag{5.69}
$$

The partial differentials are given by the grey value gradients (see section 5.2.4):

$$g_{x'} = \frac{\partial g(x',y')}{\partial x'} \qquad\qquad g_{y'} = \frac{\partial g(x',y')}{\partial y'} \tag{5.70}$$

It is convenient and sufficient for most purposes to set initial parameter values as follows:

$$
\begin{aligned}
a_0^o &= a_2^o = b_0^o = b_1^o = r_0^o = 0 \\
a_1^o &= b_2^o = r_1^o = 1
\end{aligned}
\tag{5.71}
$$

If the transformation parameters are written as the vector of unknowns $\hat{\mathbf{x}}$, the partial derivatives as the design matrix $\mathbf{A}$ and the grey-value differences between reference image and search image as the vector of observations $\mathbf{l}$, then the linearized correction equation are given by:

$$\underset{n,1}{\mathbf{l}} + \underset{n,1}{\mathbf{v}} = \underset{n,u}{\mathbf{A}} \cdot \underset{u,1}{\hat{\mathbf{x}}} \tag{5.72}$$

$$\text{where } \underset{u,1}{\hat{\mathbf{x}}^T} = \left[da_0, da_1, da_2, db_0, db_1, db_2, dr_0, dr_1 \right]$$

$$\hat{\mathbf{x}} = (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \cdot (\mathbf{A}^T \mathbf{P} \mathbf{l}) \qquad\qquad \hat{s}_0 = \sqrt{\frac{\mathbf{v}^T \mathbf{P} \mathbf{v}}{n - u}}$$

n: number of observations = number of pixels in window
u: number of unknowns (8)

It is usual to give all observations equal weight by setting $\mathbf{P} = \mathbf{I}$. The adjustment equations must be solved iteratively. In every iteration the unknowns are corrected. This leads to new grey-value differences between search image and transformed (rectified) reference image, until the least-squares sum of the corrections is less than a threshold.

During computation, the estimated parameters should be tested for significance. Depending on image content, i.e. the similarity between reference and search image and the accuracy of initial values, the chosen transformation model (5.68) may have to be simplified or extended in successive iterations. This effect can be demonstrated by the least-squares matching of an elliptical pattern to a circular target in the search image. For this purpose, the affine transformation is over-parametrized because rotation and scaling can either be modelled by a shear angle β and different scales in x and y (Fig. 5.69 left), or equivalently by a global rotation α and a scale factor (Fig. 5.69 right). In this case a 5-parameter transformation without a parameter for shear should be used. In addition, it is useful to compute the geometric parameters first, with radiometric coefficients included in the final iterations.
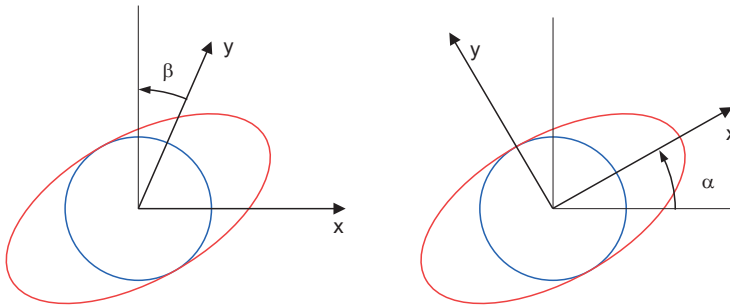
Fig. 5.69: Transformation of a circle into a rotated ellipse.

5.4.2.4.3 Quality of least-squares matching

The adjustment equations for least-squares matching are usually highly redundant. For example, a window size of 21x21 pixels generates $n = 441$ observations for only $u = 8$ unknowns.

Grey-level gradients are used in the linearized correction equations (5.69), and a solution exists only if enough image structures (edges) are available in the matched windows. For homogeneous image patches the normal system of equations is singular.

The approach requires approximate initial values, especially for the shift coefficients a_0 and b_0. As a rule of thumb, the approximate window position should not be displaced more than half the window size from the desired point. Approximate

values can be derived from a previous segmentation process or from known object geometry and orientation data.

After solving the system of adjustment equations (5.72), any residuals describe the remaining grey-value differences between reference image and adjusted search image which are not described by the mathematical model.

$$v_i = f_i(x,y) - \hat{g}_i(x,y) \tag{5.73}$$

They are a measure of the noise level in the image as well as the quality of the mathematical model.

The calculated cofactor matrix $\mathbf{Q}_{ll}$ can be used to judge the quality of parameter estimation. Similar to (2.216), the a posteriori standard deviation of estimated parameter j is given by:

$$\hat{s}_j = \hat{s}_0 \sqrt{q_{jj}} \tag{5.74}$$

where

q_{jj}:   diagonal elements of $\mathbf{Q}_{ll}$

The standard deviations of parameters a_0 and b_0 can reach high accuracies of the order of 0.01–0.04 pixel if there is good similarity between reference and search image. However, the standard deviation is only an analytical error estimate. The example in Fig. 5.70 shows the result of a least-squares matching with 5 geometric and 2 radiometric parameters. Standard deviations of shift parameters for the good-quality search image are less than 0.02 pixel. For low-quality images, standard deviations can still be of the order of 0.08 pixel, although the centre coordinates are displaced by 0.29 pixel and −0.11 pixel with respect to the non-defective optimum point.

a) Reference image (template)   b) High quality search image   c) Defective search image
$a_0 = 17.099 \pm 0.020$ $a_0 = 16.806 \pm 0.088$
$b_0 = 17.829 \pm 0.012$ $b_0 = 17.940 \pm 0.069$

Fig. 5.70: Least-squares matching (5+2 parameters): shift parameters with standard deviations.

Using blunder detection, as explained in section 2.4.4, it is possible to eliminate a limited number of gross errors in the observations. Here blunders refer to pixels whose grey values are caused, for example, by occlusions or other artefacts, and are

not described by the functional model. If the adjustment additionally tests for the significance of parameters, and if non-significant parameters are eliminated automatically, then least-squares matching becomes an adaptive, self-controlled method of point measurement.

5.4.2.4.4 Extensions

The least-squares matching algorithm described above can be extended by the integration of simultaneous processing of multiple images (multi-image matching) and by the introduction of geometric constraints (epipolar geometry), as described in sections 5.5.4 and 5.5.5.

In addition, the geometric transformation of eqn. (5.68) can be replaced by a more rigorous plane-projective transformation (eqn. 2.21) or a polynomial transformation (eqn. 2.17). Both approaches are suitable for larger matching windows where the affine transformation no longer represents a linear approximation, e.g. for highly convergent images or for textures of low resolution. A 2nd order polynomial is also useful if the object surface represented by the matching windows is curved.

5.4.2.5 Structural measuring methods

Structural measuring methods extract edges in the image which are relevant to the object, and reconstruct its geometry with the aid of mathematically defined shapes.

5.4.2.5.1 Circular and elliptical features

In general, circular objects are imaged as ellipses. To a first approximation, the ellipse centre corresponds to the projected circle centre (see section 6.2.1.1). For this purpose, the star operator or the Zhou operator have been proven effective. The centre of the ellipse is determined in several steps:
1. Definition of a search window based on a given approximate position.
2. Extraction of edge points (ellipse boundary).
3. Calculation of ellipse parameters.

The star operator determines points on the ellipse by edge detection (e.g. according to section 5.1.2) along search lines radiating from an approximation to the centre of the ellipse (Fig. 5.71). These search lines intersect the ellipse at favourable angles and grey values must be appropriately interpolated along the lines.

The coordinates of the extracted edge points are subsequently used as observations for calculating the parameters of a best-fit ellipse (see section 2.3.1.3). Individual false edge points can be eliminated by blunder detection. From the 5 parameters of the best-fitting ellipse the centre coordinates can be extracted. Since the calculated edge points depend on the initial value of the centre, it may be necessary to iterate the process in order to improve the accuracy of the ellipse centre.

initial value   edge point
fitted element   centre
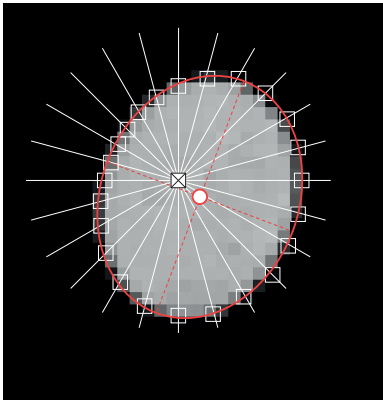
Fig. 5.71: Principle of the star operator.

initial value   edge point   point on
fitted element centre   diameter
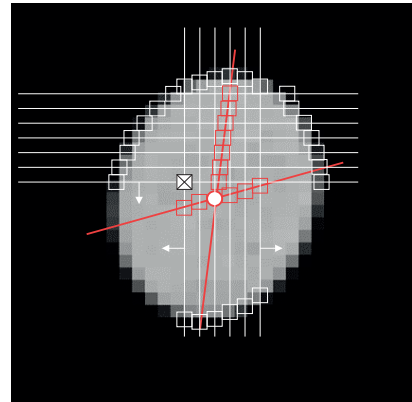
Fig. 5.72: Principle of Zhou ellipse operator.

The *Zhou operator* makes use of conjugate ellipse diameters (see section 2.3.1.3). Ellipse diameters are straight lines connecting mid points of parallel chords. The intersection point of conjugate diameters corresponds to the ellipse centre. In the image, the ellipse edge points are determined within rows and columns and the corresponding middle point is calculated. Two regression lines are determined for the middle points belonging to the diameters whose intersection point then corresponds to the desired ellipse centre (Fig. 5.72).

metric image   search window

target

centre

outlier

edge points   ellipse
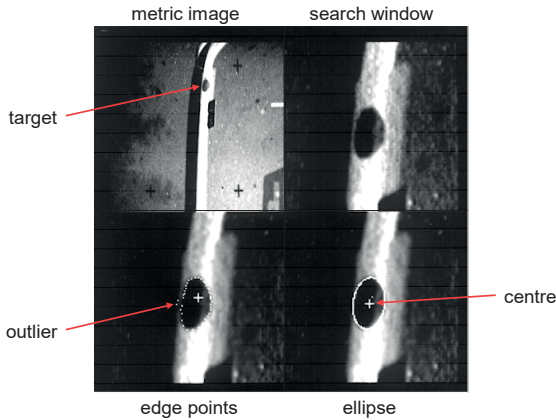
Fig. 5.73: Example of ellipse measurement on curved surface.

This method of ellipse measurement essentially requires good-quality targets. A small number of defective edge points can be handled by robust blunder detection within the ellipse or line adjustment. There are limits to the measurement of damaged or occluded targets (e.g. Fig. 3.152) as the number of false edge points increases. Circular targets cannot be correctly measured by ellipse approximation if they are located on non-planar surfaces (Fig. 5.73).

5.4.2.5.2 Cross-shaped features

Cross-shaped features, e.g. object corners, checkerboard corners and réseau crosses, can be measured in a similar way to ellipse measurement by using edge detection. Here the objective is also to extract the relevant grey value edges which define the cross.

For upright crosses (see example in Fig. 5.65c) the centre points of the bars are extracted along rows and columns, with the central region ignored. As with the Zhou operator, a regression line is fitted to the centre points of each bar. The intersection point of both lines defines the centre of the cross.
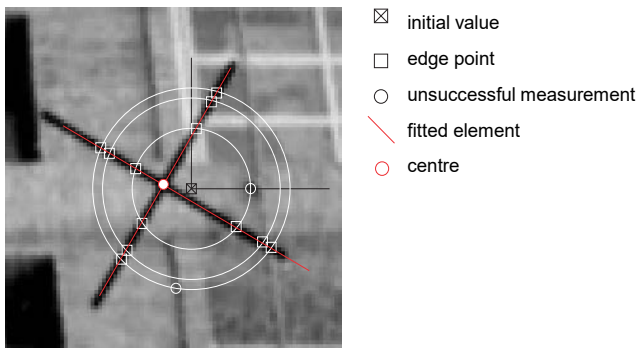
Fig. 5.74: Principle of the ring operator.

Arbitrarily rotated crosses can only be reliably measured by the above algorithm if the extracted line points are analysed in order to correctly assign them to the appropriate cross bar. A rotation-invariant method is provided by the ring operator which extracts edge points within concentric rings around the approximate initial centre point (Fig. 5.74). Extracted edge points within a ring are initially defined by polar coordinates (radius and arc length). These are easily transformed into Cartesian coordinates to which regression lines can be fitted.

Measurements based on edge detection have the advantage that feature and background can be separated relatively easily and only those image points are used which describe the actual shape of the feature. As an example, imaged réseau crosses

are often disturbed by background structures. False edge points can be detected and eliminated a priori by a classification of image gradients (sign and magnitude). Any remaining blunders can be identified as outliers on the regression line.

5.4.2.6 Accuracy issues

The location accuracy of single point features can be assessed as follows:

– Comparison with nominal coordinates of synthetic reference features:
   Single point features with a regular geometric structure, variable centre coordinates, shape parameters and contrast, can be generated synthetically. These can be analysed after known arbitrary sub-pixel displacements are applied, for example by geometric transformations and appropriate grey value interpolations. Accuracy analysis using synthetic features is particularly useful for testing individual effects and parameters of an algorithm.

– Analytical error analysis of adjustment:
   If the centre of a feature is calculated by adjustment, e.g. by least-squares matching or a best-fit ellipse, standard deviations and reliability figures can be computed for the centre coordinates. However, these only indicate the precision to which the chosen mathematical model fits the observations supplied, such as grey values or edge points (see section 5.4.2.4).

– Analysis of bundle adjustment:
   Multi-image bundle adjustment can offer the possibility of a more rigorous accuracy assessment which takes into account all influences in the measurement process (image acquisition system, digitization, point detection operator, mathematical model). This can be achieved, for example, by calibration against a test field of high-accuracy reference points. The image residuals remaining after the adjustment can be interpreted as a quality measure for the target point accuracy, although these are also potentially influenced by systematic effects that are not accounted for in the functional model. Ultimately however, only independently measured reference points, distances or surface shapes provide a strictly rigorous method of analysing point measurement quality (see sections 4.4.5 and 7.2).

Many comparative investigations of different measuring algorithms, applied to synthetic and real test images, have shown the promising potential of digital point measurement.

– Measurement resolution is limited to about 1/1000 – 2/1000 pixel if adjustment methods (least-squares matching, ellipse operators) are applied to the localization of appropriately sized and undamaged synthetic features. This result corresponds to the theoretical positioning accuracy of edge-based operators (see section 5.2.4.6).

– For real imagery with well exposed and bright elliptical targets, accuracies of 2/100 – 5/100 pixel can be achieved if least-squares operators or adaptive

centroid operators are applied to multi-image configurations. Ellipse measurement based on edge detection tends to be slightly more accurate than least-squares matching and centroid methods if there is increased image noise or distinct blunders are present.

– A significant factor in determining point accuracy is the size (diameter) of imaged points. The optimum target size is between about 5 and 15 pixels in diameter. Smaller points do not provide enough object information, which limits the localization accuracy of matching procedures or edge-oriented operators. Larger point diameters result in larger numbers of observations but the number of significant edge points increases only linearly whilst the number of pixels in the window increases quadratically. In addition, disturbances in the image are more likely with the result that the centre of the ellipse would be displaced with respect to the actual centre of the target circle as it increases in size (see section 6.2.1.1).

5.4.3 Feature extraction

5.4.3.1 Interest operators and feature detectors

Interest operators are algorithms for the extraction of distinctive image points which are potentially suitable candidates for image-to-image matching. Suitable candidates for homologous points are grey-value patterns which, as far as possible, are unique within a limited region and are likely to have a similar appearance in the corresponding image. For each pixel, interest operators determine one or more parameters (interest values) which can be used in subsequent feature matching. Interest points, also called key points, are useful for determining approximate points for surface reconstruction. However, they are rarely applied when searching for artificial target points.

Criteria for such distinctive candidate features, and the requirements for an optimal interest operator, can be summarized as follows:
– individuality (locally unique, distinct from background);
– invariance in terms of geometric and radiometric distortions;
– robustness (insensitivity to noise);
– rarity (globally unique, distinct from other candidates);
– applicability (interest values are suitable for further image analysis).

Within a local window, the following is a selection of possible criteria for determining the presence of readily identifiable structures:
– Local variance:
  The grey-value variance in a window can be calculated by eqn. (5.7). Highly structured image patterns have high variances, homogeneous regions have zero variance. The variance does not have any geometric meaning, high numbers can also result from edges and therefore it is not suitable as an interest value.

– Auto-correlation or auto-covariance function:
  The auto-correlation function or auto-covariance function results from calculation of the normalized cross-correlation (section 5.4.2.3) of an image patch with itself. If the function shows a sharp maximum it indicates a distinctive image structure which is not repeated locally (see section 5.4.2.3).
– Self-similarity:
  These operators calculate the covariance matrix of the displacement of an image window. The corresponding error ellipse becomes small and circular for image features which are distinctive in all directions. Examples are given by the Förstner operator and by the Harris operator (section 5.4.3.2).
– Grey-value surface curvature:
  If the grey values in a local region are regarded as a spatial surface (grey-value mountains), distinctive points have a high local curvature in all directions, declining rapidly in the near neighbourhood. Curvature can be calculated by differential operators (section 5.2.4.1) which approximate the second derivative.
– Gradient sums:
  This operator computes the squared gradient sums in the four principal directions of a window. If the smallest of the four sums exceeds a threshold, then a distinctive feature is indicated. The remaining image regions are those which have significant intensity changes in all directions. Distinctive point features are recognizable by the fact that the gradient sums are significant in all directions. Individual edges, which have little structural change along the edge direction, are therefore eliminated. Examples of methods based on gradient sums are the SURF and SIFT operators (section 5.4.3.5).
– Local grey-value comparison:
  If the number of similar grey values in a window is below a threshold value, then this indicates a distinctive point feature (SUSAN and FAST operators).

Feature detectors combine the detection of interest points (or key points) with the extraction of descriptors. Descriptors are compact and distinctive representations of the area surrounding the key point. The descriptors serve as feature vectors for each interest point and can be used directly for correspondence analysis (see below). In general, any interest point detector can be combined with any feature descriptor. However, most successful approaches provide an integrated method of key point detection and descriptor extraction.

5.4.3.2 Förstner operator

The *Förstner operator* is based on the assumption that the region around a point $f(x,y)$ is a shifted and noisy copy of the original image signal $g(x,y)$ (see eqn. 5.67):

$$f(x,y) = g(x + x_0, y + y_0) + e(x,y) \tag{5.75}$$

Linearization at initial values $x_0 = 0$, $y_0 = 0$ gives:

$$dg(x,y) - e(x,y) = \frac{\partial g}{\partial x} x_0 + \frac{\partial g}{\partial y} y_0 = g_x x_0 + g_y y_0 \qquad (5.76)$$

where $dg(x,y) = f(x,y) - g(x,y)$

Using the unknown shift parameters x_0, y_0 and the uncorrelated, equally weighted observations (grey value differences) $dg(x,y)$, the following normal system of equations for a least-squares estimation is obtained:

$$\mathbf{N} \cdot \mathbf{x} = \mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^T \cdot \mathbf{dg}$$

or $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (5.77)$

$$\begin{bmatrix} \sum g_x^2 & \sum g_x g_y \\ \sum g_y g_x & \sum g_y^2 \end{bmatrix} \cdot \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} g_x dg \\ g_y dg \end{bmatrix}$$

The normal equation matrix $\mathbf{N}$ contains the functional model of a displacement of the image window in x and y. Its inverse can be interpreted as a variance-covariance matrix whose eigenvalues λ_1, λ_2 indicate the semi-axes of an error ellipse (section 2.4.3.5). Features forming well-defined points are characterized by small circular error ellipses. In contrast, elongated ellipses are obtained for edge points. Unstructured or noisy image features result in large error ellipses.

Based on the parameters

$$w = \frac{\det(\mathbf{N})}{\text{trace}(\mathbf{N})} = \frac{1}{\lambda_1 + \lambda_2} \qquad \text{: measure of ellipse size; } w > 0 \qquad (5.78)$$

$$q = \frac{4\det(\mathbf{N})}{\text{spur}(\mathbf{N})^2} = 1 - \left(\frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2} \right)^2 \qquad \text{: measure of roundness of ellipse; } 0 \le q \le 1$$

a distinct point is observed if thresholds w_{min} and q_{min} are exceeded. Suitable windows are 5 to 7 pixels in size, appropriate thresholds are in the ranges:

$w_{min} = (0.5 \ldots 1.5) \cdot w_{mean}$, $\qquad\qquad w_{mean} = $ mean of w for the complete image

$q_{min} = 0.5 \ldots 0.75$

Fig. 5.75b shows the result of the Förstner operator applied to a stereo image after a thinning procedure. Distinct corners are detected but a slight displacement can also be observed here. A more precise point location can be obtained if the unknowns x_0, y_0 in (5.77) are calculated. As expected, several points are detected, whilst other similar features fall below the threshold.
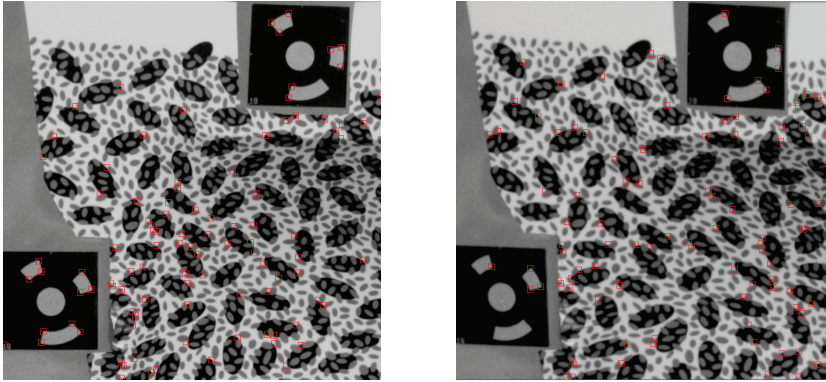
Fig. 5.75: Point detection in a stereo image with Förstner operator (filter size 5x5).

The Förstner operator is in some respects similar to the independently developed *Harris operator*. Today their key idea is sometimes referred to as the Förstner-Harris approach. To avoid the complexity of an eigenvalue decomposition the Harris operator calculates an interest value directly from the matrix **N**:

$$v = \det(\mathbf{N}) - k \cdot \text{trace}(\mathbf{N})^2 \tag{5.79}$$

The parameter k is often set to 0.04 in order to separate point features from edges.

Förstner and Harris operators are only suitable for images with equal scales and minor perspective deviation, e.g. parallel aerial imagery. They are less applicable to images with larger scale differences or projective distortions.

5.4.3.3 SUSAN operator

The *SUSAN operator* (smallest univalue segment assimilating nucleus) compares the intensities of the pixels in a circular window with the grey value of the central pixel, designated as the *nucleus*. A distinctive point has been found when the number of similar grey values in the window lies under a threshold value.

A decision value c is calculated for every pixel in a filter region, u,v:

$$c(u,v) = \begin{cases} 1 & \text{for} & \left| s(x+u, y+v) - s(x,y) \right| \\ 0 & \text{for} & \left| s(x+u, y+v) - s(x,y) \right| > t \end{cases} \tag{5.80}$$

where

t: threshold value for similarity of grey value to nucleus

A more stable decision function is given by:

$$c(u,v) = \exp\left[-\left(\frac{s(x+u,y+v)-s(x,y)}{t}\right)^6\right] \tag{5.81}$$

The sum of all values $c(u,v)$ in the window

$$n(x,y) = \sum c(u,v) \tag{5.82}$$

is compared with the geometry threshold T:

$$R(x,y) = \begin{cases} T-n(x,y) & \text{if } n(x,y)<T \\ 0 & \text{else} \end{cases} \tag{5.83}$$

In order to detect corner features, a threshold value $T = n_{max}/2$ is set (n_{max} = number of pixels in the filter window). When less than half the grey values in the window are similar to one another, a distinctive point with the interest value R is generated at the position x,y.
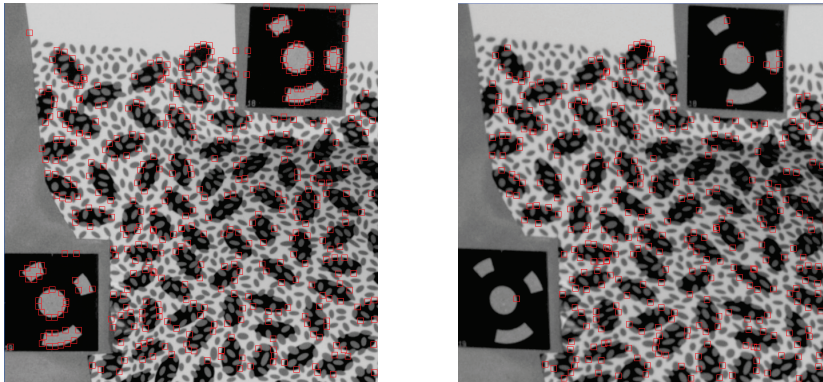
Fig. 5.76: Feature detection in stereo pair using SUSAN operator (search window 5x5).

Fig. 5.76 shows examples of points detected in a stereo pair by the SUSAN operator. Since the operator depends directly on the brightness of the image, many similar points are found in the left and right image in the area of the random pattern, but very different characteristics are recognized in the high-contrast patterns (coded targets).

5.4.3.4 FAST operator

Like the SUSAN operator, the *FAST operator* (features from accelerated segment test) analyses the intensities of the pixels in a circular window. A distinctive feature is assumed to be found when a number of connected pixels with similar grey values is

found in a ring around the central pixel. An analysis based on a ring is largely invariant to changes in scale and rotation.
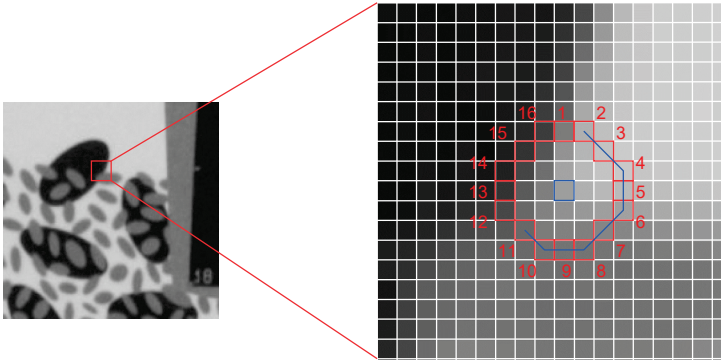
Fig. 5.77: Principle of the FAST operator.

The principle behind the operator is shown in Fig. 5.77. In this example, within a ring which is 16 pixels in length, 10 neighbouring pixels lie below the threshold value. The grey values $s(u)$, $u = 1...16$, are compared with the grey value of the central pixel and assigned a similarity value $c(u)$:

$$c(u) = \begin{cases} d & \text{for} \quad s(u) \leq s(x,y) - t & \text{(darker)} \\ s & \text{for} \quad s(x,y) - t < s(u) < s(x,y) + t & \text{(similar)} \\ b & \text{for} \quad s(x,y) + t \leq s(u) & \text{(brighter)} \end{cases} \tag{5.84}$$

where

t: threshold value for similarity of grey value to central pixel

In the simplest case, a corner point is found when, for 12 neighbouring pixels, $c(u) = d$ or $c(u) = b$. For a more robust analysis of the area, the distribution of the grey values can be evaluated and compared to different types of corner features which are known a priori.

5.4.3.5 SIFT operator

The *SIFT operator* (scale invariant feature transform) detects prominent points (detector) and a corresponding feature vector (descriptor) in a four-step procedure:

1. Extraction of edges and their extrema in image pyramids:
    Firstly, the input image is transformed into image pyramids according to section 5.1.3.1. Subsequently, each pyramid level is smoothed by a Gaussian filter to create a multi-scale space (Fig. 5.78). A subtraction of two images in the multi-scale space is equivalent to a DoG filter (section 5.2.4.2), hence gradients are

magnified. In adjacent DoG images, the local extremum is found by determining the minimum or maximum DoG value within the N8 neighbourhood of the current image and the 3x3 neighbours of the upper and lower level of the pyramid (Fig. 5.79). The extracted points are used as initial key points for the next processing step.

2. Localization of feature points:

The feature candidates from step (1) are further thinned. Based on gradient magnitudes and curvature of the grey-value surface (Hessian matrix, compare Förstner and Harris operator, section 5.4.3.2), features with low contrast or minor edges are eliminated.

3. Calculation of major gradient directions:

In the immediate environment of a feature point, gradient magnitude and direction are calculated and entered into a histogram of 10° classes covering the range 0° to 360°. The gradients are given weights which are inversely proportional to the distance from the central pixel and are summed within the histogram. The histogram class with the highest frequency indicates the major orientation angle of the feature.
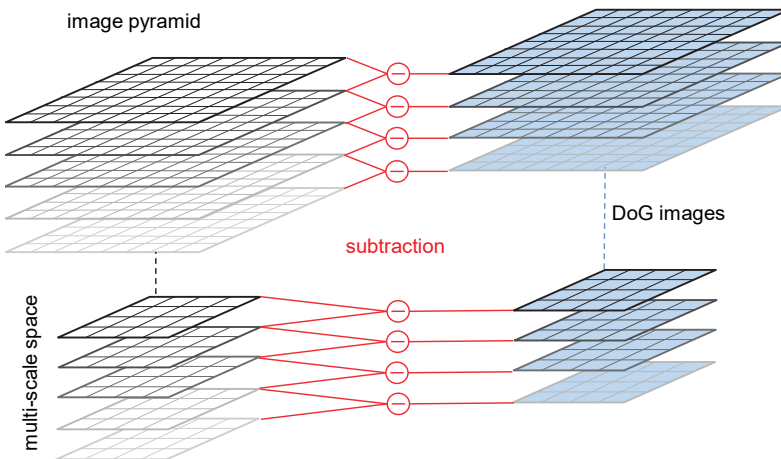
Fig. 5.78: Multi-scale space.

4. Descriptor:

In the final step, an area of 16x16 pixels is divided into 4x4 blocks in which gradient magnitude and direction are again calculated (Fig. 5.80). The gradient direction is related to the main direction extracted in step 3, hence invariance against rotation is achieved. The directions are given weights which depend on the reciprocal of their distance from the feature point and stored in a histogram

with 8 classes (45° steps). From the 16 blocks and the 8 classes, 128 features are derived in the form of a normalized SIFT feature vector.
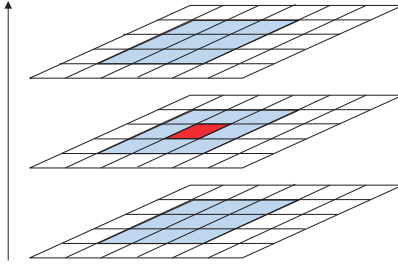
Fig. 5.79: Search for extrema in DoG images.
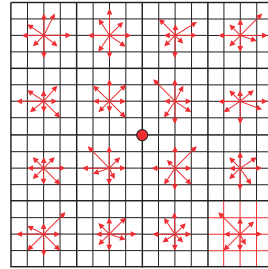
Fig. 5.80: Descriptor from gradient analysis.
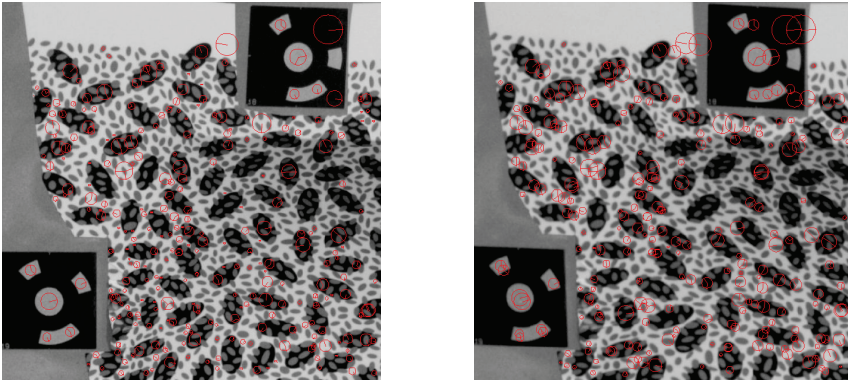
Fig. 5.81: Feature detection in stereo pair using SIFT operator.

Fig. 5.81 illustrated the application of the SIFT operator for feature extraction in a stereo image. The size of the circles indicates the pyramid level in which the feature was detected. The vectors inside the circles indicate the direction of maximum gradient. Typically, the operator detects blobs rather than corner points or other distinctive points. Also typically, the operator does not detect corner points or other points with visual impact, but blob-like structures (compare with Fig. 5.90). The SIFT operator is particularly suitable for assignments between images which have significant differences with respect to scale, rotation and small differences in perspective. (Fig. 5.81). It is in widespread use, for example as a key component of many software solutions based on structure-from-motion (5.5.2.2).

The ASIFT operator is an extension to the SIFT operator which is less sensitive to rotational and perspective distortions. However, it requires much higher computational effort.

5.4.3.6 SURF Operator

The *SURF operator* (speed-up robust features) uses the maxima of the determinant of the Hessian to detect interest points by the second derivations of the grey-value distribution (according to LoG filter, section 5.2.4.2). The computation of the Hessian is accelerated by approximating the underlying Gaussian filter process with simple box filters and using integral images (5.1.3.4) for convolution. The SURF detector (much like the SIFT detector) uses image areas that resemble blobs as interest points.

The neighbourhood of the interest point is characterized by a descriptor using a histogram of Haar wavelet responses. This histogram is computed over a square region oriented along the dominant orientation to achieve invariance to rotation. As with the SIFT, scale space is exploited to gain invariance to scale. Scale space is computed implicitly by varying the size of the filters. Integral images are used to accelerate filtering particularly for large filter sizes. Fig. 5.82 gives an example of the interest points detected using the SURF approach. The size of the circle indicates the scale at which the feature was detected. The ticks indicate the direction of the dominant gradient.
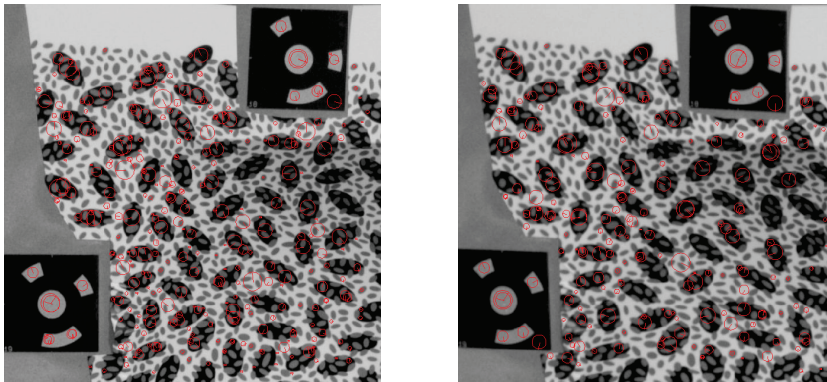
Fig. 5.82: Feature detection in stereo pair using SURF operator.

5.4.3.7 ORB operator

ORB (Oriented FAST and rotated BRIEF) is a feature detector and descriptor formed as an extended version of FAST (section 5.4.3.4) and the binary descriptor BRIEF (see below).

The ORB extension applies a combination of the FAST and Harris operators to image pyramids in order to provide invariance against scale. In addition, an orientation vector of gradients between the centroid (x_c, y_c) and the centre of the patch (x, y) is calculated:

$$\theta = \arctan(\Delta y / \Delta x) \qquad (5.85)$$

with $\Delta x = x - x_c$ and $\Delta y = y - y_c$

The azimuth angles Θ are stored for each pyramid layer to provide angular information to the descriptor, in a similar way to SIFT.

BRIEF computes a binary string where each element is the result of a binary comparison of randomly sampled point pairs (p_j, p_k) with grey values $g(p_j)$, $g(p_k)$ within a patch **p.** This generates high variance and low correlation within a certain region around the feature:

$$t(\mathbf{p}; p_j, p_k) = \begin{cases} 1: & g(p_j) < g(p_k) \\ 0: & g(p_j) \geq g(p_k) \end{cases} \qquad (5.86)$$

Because of a high sensitivity to noise, the image should be low-pass filtered in advance, e.g. by Gaussian filtering. From eqn. (5.86), a feature value is calculated as the sum of n binary results from the randomly selected points:

$$f_n(\mathbf{p}) = \sum_{i=1}^{n} 2^{i-1} t(\mathbf{p}; p_j, p_k) \qquad (5.87)$$

In its original form, BRIEF is not invariant to rotations. By taking the direction information of the extended FAST operator (eqn. 5.85) into account, BRIEF can be steered towards the principal feature direction and so become insensitive to rotations. The main advantage is the constrained selection of sampling pairs to be used in eqn. (5.86) which should be both distinctive (high variation) and unique (low correlation).

Various extensions to the basic ORB concept are available. In comparison to SIFT, it creates similar feature points but does this faster and with much lower computational effort. The operator is very popular in the field of visual odometry (visual SLAM, see section 5.5.7.6) and other real-time applications with limited computing power.

5.5 Image matching and 3D object reconstruction

5.5.1 Overview

Image matching methods are used to identify and uniquely match identical object features (points, patterns, edges) in two or more images of the object. Matching methods are also required for:
– identification of discrete (targeted) image points for 3D point measurement;
– identification of homologous image features for 3D surface reconstruction;
– identification of homologous points for stitching images into mosaics;
– identification and tracking of objects in image sequences.

One of the earliest problems in computer vision, and still one of the most researched topics, was automatic matching of corresponding image features. Correspondence analysis is a fundamental requirement in understanding images of spatial scenes and is closely related to human visual perception. Whilst digital image matching of suitably structured object scenes, e.g. using targets or projected patterns, can surpass human performance in some areas (measuring accuracy, processing speed), the analysis of arbitrary object scenes is still the subject of intensive research.

Correspondence analysis can be classified as an ill-posed problem, i.e. it is uncertain if any solution exists which is unique and robust with respect to variations in the input data. In principle, the following problems may occur when matching arbitrary scenes:

- due to occlusion, an image point P_{ij} (point i in image j) does not have a homologous partner point P_{ik};
- due to ambiguous object structures or transparent surfaces, there are several candidates P_{ik} for image point P_{ij};
- for regions with poor texture, the solution becomes unstable or sensitive with respect to minor disturbances in the image (noise).

Solutions in practice assume the following preconditions for object and image acquisition:

- intensities in all images cover the same spectral regions;
- constant illumination, atmospheric effects and media interfaces for the imaging period;
- stable object surface over the period of image acquisition;
- macroscopically smooth object surface;
- locally unique textures;
- opaque object surface;
- largely diffuse reflection off the surface;
- known approximate values for orientation data (image overlap) and object data (geometric and radiometric parameters).

From the wide range of matching techniques, a selection of successful methods is discussed below. These are directly related to geometric surface reconstruction.

In practical applications, there are essentially two initial situations (Fig. 5.83):

a) The selected images are not yet oriented but the interior orientation may be known to a sufficient accuracy that can assist further processing.
b) The selected images have known interior and exterior orientations.

The first case mainly relates to multi-image configurations which should be automatically oriented. The process usually consists of feature extraction in all images followed by feature-based matching. After robust calculation of relative orientations by RANSAC methods, potentially coupled with repeated bundle

adjustment, a final bundle adjustment is computed for all images. This strategy is now commonly called structure-from-motion (section 5.5.2.2).

In the second case, geometric constraints (epipolar geometry, section 4.3.2) can be set up between images which can also be used to limit the search space for corresponding points. The objective is a precise matching of homologous features and the calculation of spatial intersections in order to create a sparse point cloud (using feature points) or a dense point cloud (for the complete surface).

Mixed forms of these strategies are possible, as illustrated in Fig. 5.83. For example, after orientation of a starting model, epipolar lines can then be used to support matching.
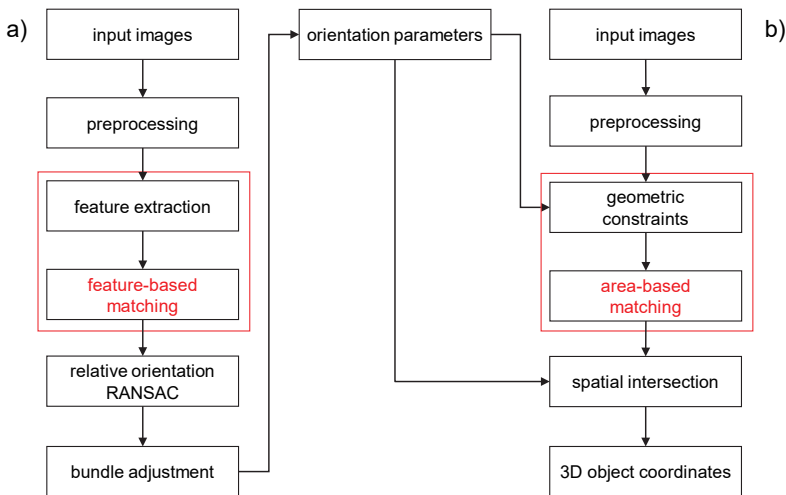
Fig. 5.83: Strategies for image matching.

According to Fig. 5.83 *pre-processing* includes image enhancement (smoothing, noise reduction, contrast adjustment) and the reduction of resolution (image pyramids). For noise reduction, smoothing filters are used (section 5.2.3.2). These are also used in the creation of image pyramids (section 5.1.3.1). Global contrast can be optimized by using filters which adapt to the local contrast, e.g. the Wallis filter (section 5.2.3.4). Additional pre-processing methods include colour reduction (e.g. RGB to grey values, section 5.1.1.1) and global colour transformations.

In the *feature-extraction* step, image features such as distinct points or edges are extracted from the images independently of one another and a large percentage are assumed to be common to all images. For this purpose, it is not only interest operators, feature detectors (section 5.4.3.1) and edge operators (section 5.1.2) which are used, but also segmentation methods for the coarse detection of object points (e.g.

coded targets, section 6.2.1.4). In addition to the geometric parameters, the extracted features are characterized by additional attributes, e.g. topological relationships and point numbers.

The subsequent step of *feature-based matching* attempts to identify as many corresponding features as possible in all images. This is achieved by applying suitable cost or similarity functions (see section 5.5.2). Additional information in the form of knowledge or rules can be used here in order to limit the search space and to minimize mismatches. Starting with a suitable pair of images, a relative orientation based on the measured feature points is calculated. All other images are then successively connected by corresponding points. In the final bundle adjustment, the cameras used for image acquisition are calibrated, precise exterior orientations are calculated for all images and the 3D coordinates of all feature points are generated (sparse point cloud, sparse matching). In general, the matching of features is the most problematical step in correspondence analysis because the type and extent of additional knowledge may vary.

In case (b) of Fig. 5.83 the *area-based precise matching* of original grey values determines corresponding object elements to a high accuracy. Correlation and least-squares methods are classified as area-based matching methods. In this step, additional geometric information such as epipolar geometry (see section 4.3.2, object constraints) can be used to improve accuracy and reliability. 3D object data can be derived from the calculated homologous image elements by intersection, e.g. for each pixel if a dense surface reconstruction is required (*dense matching*).
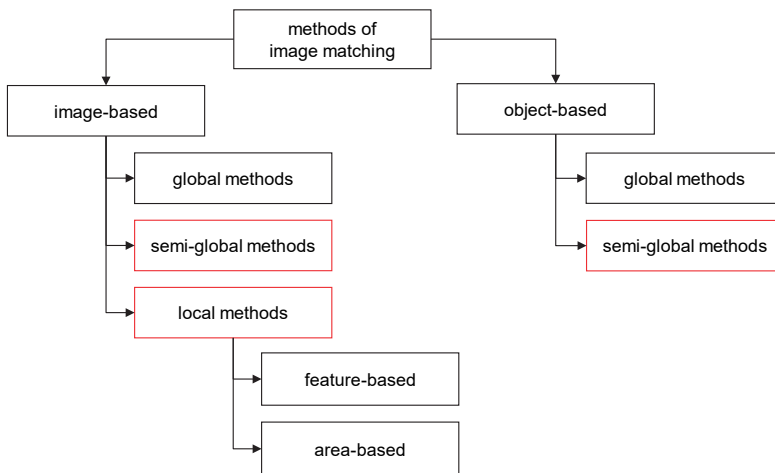
Fig. 5.84: Methods of digital image matching.

Fig. 5.84 shows an organizational structure for different matching methods. It differentiates between image-based and object-based procedures. The *image-based methods* either work globally (similarity between complete images), semi-globally (local similarity with a limited number of global search paths) or locally (feature-based or area-based similarity in local search windows). The methods commonly used in photogrammetry are marked in red. Here image-based methods are more common as they offer more options.

Object-based matching methods solve the matching problem starting from object space. Here there is usually an approximate geometric and radiometric object model that can be refined by back-projection into the corresponding image space and analysis of the related grey-value distributions. As a result, a structured spatial object description in the form of coordinates (point clouds), voxels, geometric elements or vector fields is derived.

The following sections start with strategies for image matching in non-oriented images. They continue with image-based matching methods for oriented stereo and multi-image configurations and conclude with object-based matching methods.

5.5.2 Strategies for matching non-oriented images

5.5.2.1 Coded targets

The initial task at the start of 3D image processing is usually to orient the recorded images and calibrate the camera(s). In applications where the object is marked with suitable targets, the measurement of control and tie points is a robust process which does not require interactive operation. Here the methods outlined in section 5.4.2 are applied, typically calculating the target centres to a high sub-pixel accuracy. Initial values for exterior orientations and 3D object coordinates can be derived by the procedure described in section 4.4.4. The subsequent bundle adjustment then uses comparatively few, but reliable and accurate, image points.

5.5.2.2 Structure-from-motion

For non-targeted image sets, there are a number of established methods which can be classified under the term structure-from-motion (SfM). SfM, with a history dating back to at least 1979, enables an object or structure to be reconstructed three-dimensionally from the motion of a camera. Typically, at its core there is a closed-form solution for the relative orientation of an image pair from corresponding points without the need for prior information (see section 4.3.4). Tried-and-tested workflows use a combination of image processing algorithms, robust orientation methods, bundle adjustment with self-calibration, stereo image analysis and point cloud processing.
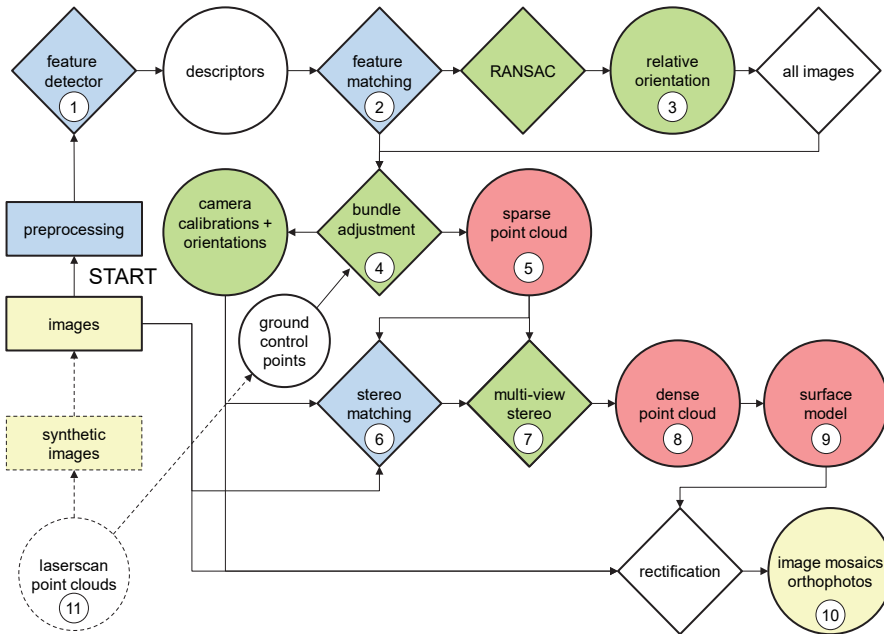
Fig. 5.85: Structure-from-Motion (SfM).

As outlined by Fig. 5.85, all images are pre-processed before applying an interest or feature detector (1) which is often SIFT. After detected features are matched (2) they are used in a RANSAC-based relative orientation of an initial pair of images (3). The selection of this starting stereo model is critical as there is a trade-off between choosing a model with a wide baseline (more accurate but larger image differences) and one with a small baseline (less accurate bus closely similar images). Subsequently, all remaining images are oriented with respect to the starting model before processing by a bundle adjustment (4). Optional datum constraints (reference points) can be added and a simultaneous camera calibration can be calculated. At this stage, the estimated object coordinates from the bundle adjustment already represent a sparse point cloud (5).

If a denser surface model is required, dense stereo matching (6) and subsequent multi-view stereo (7) can be used (see section 5.5.5.1) to generate a dense point cloud (8) and hence a complete surface model (9). From the oriented images and surface models, orthophotos or image mosaics (10) can be derived. Optionally, registered laserscan data (11) can be integrated in some programmes (see below).

SfM is usually successful if there is good surface texture, densely overlapping images and a camera which is either pre-calibrated or is at least mechanically stable during image acquisition. The typically achievable accuracy lies in the range 0.5–2 pixels depending on the method of feature extraction and the measure of similarity

between images. In current software implementations, the computational effort at the orientation stage is essentially defined by the process of feature extraction. However, if dense matching is applied this dominates all other stages. With a large number of high-resolution images, total processing time in this case could be several hours or even days. Well-known software programs which integrate SfM include: Agisoft Metashape, RealityCapture, Pix4Dmapper, Context Capture, MicMac, Meshroom, Bundler, COLMAP, Theia or OpenMVG (see section 6.4.2).

If registered laser scans are available for the recorded object, they can be used to support the orientation of the photogrammetric images. Synthetic images can be calculated for each laser point by projecting the scanner's point cloud onto the side surfaces of a cube centred at the origin of the scanner measurements (Fig. 5.86a). These images are then already calibrated and oriented per se, so that the photogrammetric images can be oriented by matching against similar features in the coordinate system of the laser scans (Fig. 5.86b).
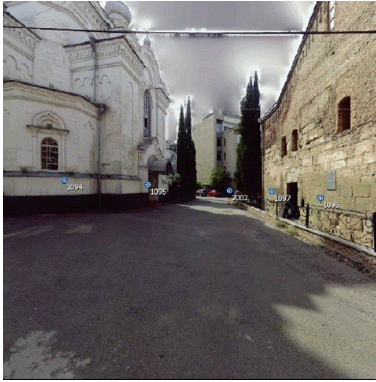
a) Synthetic image generated from a registered laser-scan point cloud
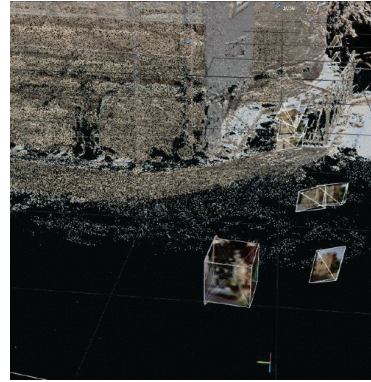
b) 3D point cloud with superimposed original images and synthetic image cubes

Fig. 5.86: Simulated image cubes from terrestrial laser scans.

5.5.2.3 Image masking

If objects are photographed in such a way that their background has few features or features that vary greatly from image to image, it can be helpful for SfM-based methods to mask the images, i.e. to hide the irrelevant background. There are manual and automatic methods for this (see Fig. 5.87 for an example):

– Manual drawing of an image mask:
  A mask image is created for each image with the help of a digital drawing pen, which quickly becomes very time-consuming with a larger number of images.

– Application of background with known colour:
  If it is known that the background in the scene is represented by a known colour, this area can be found and masked automatically. This is known as chroma keying in visual effects.
– Use of a background image without a measurement object:
  For each measurement image, a second image is taken from the same direction without the measurement object; this assumes that identical shooting positions are guaranteed, e.g. by using a tripod (example in Fig. 5.87). Alternatively, one background image for all photos is sufficient if the background is static and identical for all recording positions.
– Learning methods:
  Machine learning methods (deep learning) can be used to calculate an automatic segmentation of foreground and background from a sufficient amount of training data.
– Application of a given 3D model:
  If, after a first reconstruction step, a 3D model of the object and orientation data of the images are known, the circumscribing image contour of the 3D object can be calculated from each view. The reconstruction can then be done again with the masked images, if necessary several times in succession.
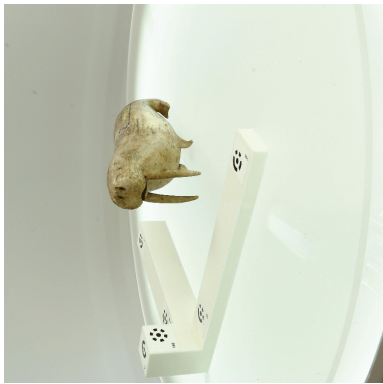
a) Original image
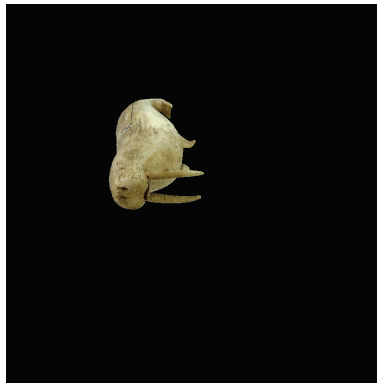
b) Masked image

Fig. 5.87: Example of image masking.

5.5.3 Similarity measures

The similarity between two or more samples of intensity distributions in images can be measured by different mathematical approaches based on similarity or cost criteria. For corresponding image patches, similarity measures should be a maximum, i.e. a similarity value such as a correlation coefficient should be high

between two samples with similar intensity distributions. In contrast, cost (energy) functions measure the effort which leads to a best correspondence, hence these measures should be a minimum for similar distributions. The next sections describe the most common similarity measures for image matching whilst some typical cost functions for semi-global matching are given in sections 5.5.4.2 and 5.5.6.3.

5.5.3.1 Sums of differences

A simple approach for measuring the similarity between two image patches f and g of equal size is given by the sum of absolute grey-value differences (SAD) or the sum of squared grey-value differences (SSD). See example in Fig. 5.88.

$$SAD = \sum |f_i - g_i| \qquad\qquad SSD = \sum (f_i - g_i)^2 \qquad (5.88)$$

The SAD and SSD values depend on contrast and they fail when there are geometric distortions between the two image patches. Consequently, they are only suitable for matching very similar image areas but are easy and fast to compute.

A calculation which is independent of brightness can be obtained by introducing the grey-value mean in both images, m_f and m_g. This gives the measures of zero mean SAD (ZSAD) and zero mean SSD (ZSSD) as follows:

$$ZSAD = \sum |f_i - m_f - (g_i - m_g)| \qquad\qquad ZSSD = \sum [f_i - m_f - (g_i - m_g)]^2 \qquad (5.89)$$
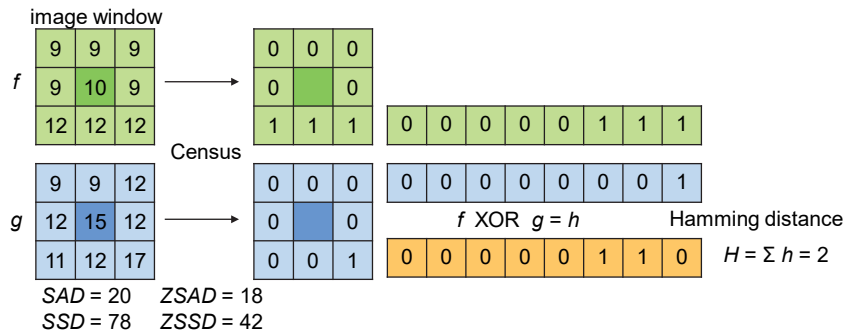
Fig. 5.88: Similarity measures - Census and Hamming distance.

5.5.3.2 Census and Hamming distance

The *Census* similarity measure (Census transform and Hamming distance) compares the grey value of a central pixel in a window with the adjacent pixels (Fig. 5.88). If the neighbouring grey value is smaller, 0 is set at that pixel, otherwise it is set to 1. The chains of binary numbers for both images are then combined bit-wise by XOR. The result is therefore 1 if corresponding bits are different, zero if not. Finally, the

Hamming distance *H* counts the number of 1s, hence both samples are similar for small values of *H*. The method is easy to implement, uses minimum processing power and is invariant to radiometric changes. However, geometric distortions between the image patches cannot be compensated.

5.5.3.3 Hashing

Hashing is a group similarity measure based on dividing the image into *n* x *m* blocks. For each block, a *Block Mean Value* algorithm calculates the mean *S* of colour values and compares it with the corresponding median value *M*. If $S > M$ then 1 is set for the block, otherwise 0. To suppress high frequencies which disturb the calculation, the image size is reduced in advance. Other hashing variants are based on histograms and discrete cosine transforms.

Image size 512 x 512 pixel

Image size 64 x 64 pixel
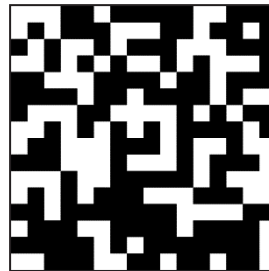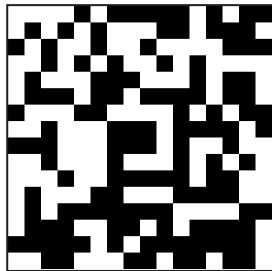
Fig. 5.89: Hashes calculated for an original image (left) and a heavily reduced and lower-contrast image (right) each divided into 16 x 16 blocks (Hamming distance between hashes = 28).

After binarization, the *n* x *m* bits are stored in a bit chain known as a *hash*. To a large extent, the hash is invariant to scaling, image compression and geometric distortions. Differences in contrast and brightness between images only have a minor impact. The Hamming distance can be used to compare two hashes. Fig. 5.89 shows the hash calculated for an original image and one for a much smaller copy with lower contrast. The similarity between the bit patterns is clearly visible. Hashing can also be useful

to search for similar images within large data sets, for example on social media or cloud storage.

5.5.3.4 Normalized cross-correlation

The normalized correlation coefficient (NCC), already introduced in section 5.4.2.3, can be used as similarity measure between two images if there are no large geometric differences between them. Affine or perspective distortions between the images directly lead to a reduced correlation coefficient. In contrast, radiometric differences are widely compensated by normalization.

If a minimum cost value is used instead of a maximum similarity value, the cost value c can be written in terms of the correlation coefficient ρ_{fg} as:

$$c = (1 - \rho_{fg}) / 2 \tag{5.90}$$

As an example calculation based on Fig. 5.88, with the correlation coefficient defined as in eqn. (5.66), the following values are obtained: $m_f = 10.111$, $m_g = 12.111$, $\sigma_f = 1.370$, $\sigma_g = 2.424$, $\sigma_{fg} = 1.543$. These result in a correlation coefficient of $\rho_{fg} = 0.464$ and a cost value of $c = 0.268$.

5.5.3.5 Least-squares matching

Least-squares matching (LSM, see section 5.4.2.4) calculates the geometric and radiometric transformation parameters between two grey value samples, and the related mean standard deviation σ_0 that can serve as a similarity measure. The importance of LSM lies more in the precise calculation of parallaxes or disparities (shift parameters of coordinate transformation) rather than being used as a (robust) similarity value. The methods require starting values for translation, approximately half the window width, to ensure convergence.

5.5.3.6 Euclidian distance between feature vectors

For matching image areas which have been detected by an interest or feature detector with an n-dimensional feature vector, the distance between two vectors serves as a similarity criterion. In the simplest case, the Euclidian distance is used:

$$\begin{aligned} d(\mathbf{a}, \mathbf{b})^2 &= (a_1 - b_1)^2 + (a_2 - b_2)^2 + \ldots + (a_n - b_n)^2 \\ &= (\mathbf{a} - \mathbf{b})^T \cdot (\mathbf{a} - \mathbf{b}) \end{aligned} \tag{5.91}$$

A comparison of two n-dimensional feature vectors $\mathbf{a}$ and $\mathbf{b}$ using the Euclidian distance then indicates similarity when this distance is a minimum. When matching image areas generated by the SIFT operator, this similarity measure is effective if the images contain distinctive textures. Where there are repeated patterns, the minimum distances can easily fail to be effective measures since similar feature vectors occur

in the neighbourhood. In this case, the minimum distance is only accepted if it differs significantly from the next lowest value of the distance. In the example of a normalized stereo image, Fig. 5.90 demonstrates the matching result between two images processed by SIFT and using minimum Euclidian distance as a matching measure. Most of the correspondences appear horizontally, in the direction of the epipolar lines. Correspondences along sloping lines indicate false matchings, and can therefore be excluded if the relative orientation between both images is known.
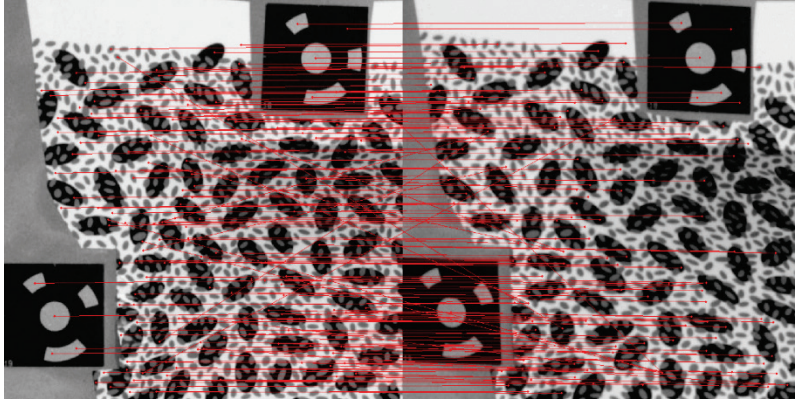
Fig. 5.90: Matching features detected by SIFT using minimum Euclidian distance.

The *Mahalanobis distance* represents an extended measure of distance. This weights individual feature dimensions by the covariance matrix. These can be derived, for instance, from known variances which are known for particular feature types.

$$d(\mathbf{a}, \mathbf{b})^2 = (\mathbf{a} - \mathbf{b})^T \cdot \mathbf{\Sigma}^{-1} \cdot (\mathbf{a} - \mathbf{b}) \qquad : \text{Mahalonobis distance} \qquad (5.92)$$

Geometrically, points of equal Mahalanobis distance form a (hyper) ellipsoid whilst they are lying on a (hyper) sphere for the normal Euclidian distance. This corresponds to the difference between the Helmert error ellipse and a confidence ellipse, see section 2.4.3.5.

5.5.4 Correspondence analysis based on epipolar geometry

5.5.4.1 Matching in image pairs

The correspondence problem can be greatly simplified if the relative orientation of the images is known. This is the case for fixed stereo vision systems and in close-range applications where image orientation is solved in a separate process, e.g. by application of coded targets or by structure-from-motion (section 5.5.2.2). The search

space for a corresponding point in the neighbouring image can then be reduced to an epipolar line (see section 4.3.2 and Fig. 5.91).

In convergent images, the epipolar lines are sloped at an angle to the x-axis (section 4.3.2). In this case, the computational effort may increase since for every image point in the left-hand image an individual epipolar line in the right-hand image must be calculated, possibly taking account of distortion which would result in a curved epipolar line. If convergent stereo images are converted into normalized stereo images (section 4.3.3.5), the epipolar lines are then straight and parallel to the image rows. This, therefore, significantly reduces the computational effort, especially for dense point matching.
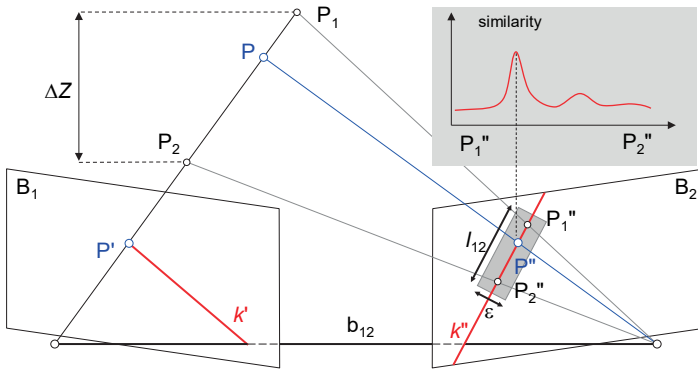
Fig. 5.91: Matching in an image pair based on epipolar lines (according to Maas 1992).

Consider first the matching process in an image pair B_1, B_2 (Fig. 5.91). Point P'', corresponding to P', is located within a small band either side of epipolar line k''. The band width ε depends on the uncertainty of the orientation parameters and the image measurement quality of P'. The search space is strictly reduced to the straight line k'' only for perfect input data. The length of the search space l_{12} is a function of the maximum depth ΔZ in object space.

A correspondence analysis can be performed along the epipolar lines using one of the similarity measures described in section 5.5.2.3. If the grey value pattern around P' is unique, point P'' can be found by a maximum similarity measure or a minimum cost function. Non-textured areas, or repeating patterns, lead to arbitrary results within the search region which may not be solved with two images but require at least one more image from a different view point (section 5.5.5). A consistency check can be done by left-right and right-left matching which must both return the same point P''.

With increasing number of image points n, and area f of the search space, the probability P_a of ambiguous point matches also increases:

$$P_a = 1 - e^{-\frac{n \cdot f}{F}}$$ (5.93)

where

n: number of image points
F: image area
f: area of epipolar search space

The total number of ambiguities N_a for an image pair is given by:

$$N_a = (n^2 - n)\frac{2 \cdot c \cdot \varepsilon \cdot b_{12} \cdot (Z_{max} - Z_{min})}{F \cdot Z_{max} \cdot Z_{min}}$$ (5.94)

where

c: principal distance
b_{12}:   base length between image B_1 and image B_2
ε: width (tolerance) of search space
$Z_{min} < Z < Z_{max}$: depth of object

It therefore increases

– quadratically with the number of image points;
– linearly with the length of the epipolar lines;
– linearly with the base length;
– approximately linearly with object depth;
– linearly with the width of the search space.

Example 5.3:

Consider an image pair (Kodak DCS 420) with parameters $c = 18$ mm, $F = 18 \times 28$ mm², $b_{12} = 1$ m, $Z_{max} = 3$ m, $Z_{min} = 1$ m and $\varepsilon = 0.02$ mm. Depending on the number of image points n the following ambiguities N_a result:

$n = 50$: $N_a = 2$
$n = 100$: $N_a = 10$
$n = 250$: $N_a = 60$
$n = 1000$: $N_a = 950$

The example above shows that epipolar geometry leads to a more reliable match if the number of image points is relatively small. It depends mainly on the application as to whether ambiguities can be reduced by an analysis of interest values, a reduction of object depth using available approximations or by a reduced search width.

5.5.4.2 Semi-global matching

In the ideal case, dense image matching should result in a 3D value for each pixel in order to reconstruct the object surface completely. In particular, object edges (discontinuities) must be modelled precisely and occluded (missing) areas should be

filled using information from additional images. Area-based matching methods often work with a (rectangular) image window which, depending on the size, has a smoothing effect (low-pass filter). With increasing window size, the information available for matching increases, but low-pass filtering also increases. Smaller windows preserve edges more sharply, but cover just a small object area with little texture information. A single pixel (window size 1x1) does not hold sufficient information for matching.

Semi-global matching (SGM) is a method which, for every pixel, determines a homologous point in the corresponding image. For this purpose, the focused search offered by epipolar matching (section 5.5.4.1) is extended by a more complex cost function and the maximum similarity (minimum cost) is calculated along m image paths, e.g. $m=8$ or $m=16$, which is the essence of the semi-global approach, see Fig. 5.92a.

a) Search paths r for SGM ($m = 8$) and cost aggregation along one path

b) Cost calculation for disparities $D(p)$
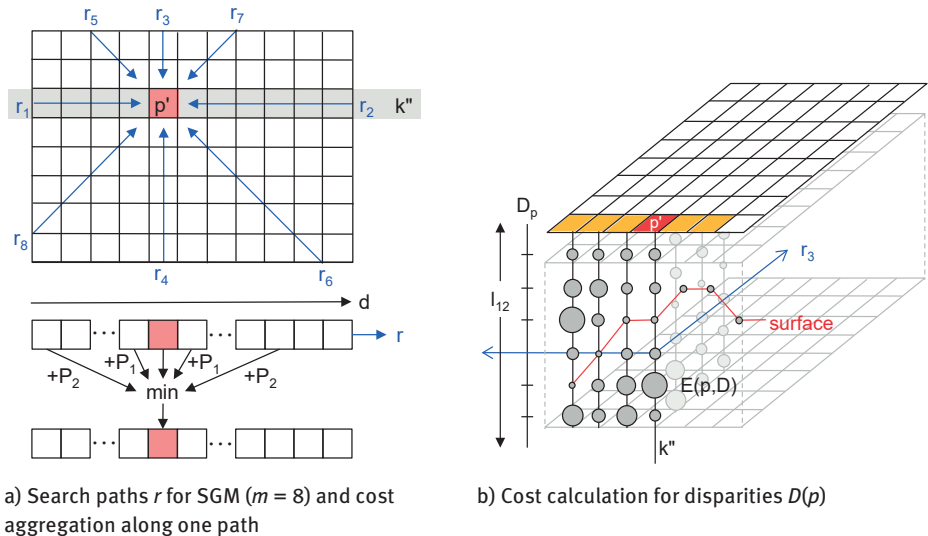
Fig. 5.92: Semi-global matching.

Fig. 5.92b illustrates the cost calculation. For a current pixel p in the left hand image, all cost functions $C(p,D_p)$ within the search region l_{12} along the epipolar line k'' in the right hand image, are calculated and stored in a data structure. For similarity measurement, different cost functions can be used, e.g. the sum of absolute grey-value differences (SAD) or Census (section 5.5.3.2). In Fig. 5.92b the costs are indicated by grey circles whose size represents the cost. The disparity with minimum cost is then compared with the costs of the neighbouring pixels (red line). If the difference of disparities with the neighbouring minimum is larger than zero, the recent cost value is given an extra amount depending on the disparity difference (penalty terms

P_1 and P_2). Eqn. (5.95) shows the resulting energy function $E(D)$. This analysis of the neighbourhood is now applied along m image paths, i.e. the data structure mentioned above must be provided m times. Subsequently, all cost values assigned to a pixel are summed and again stored within the data structure $E(p, D_p)$. Finally, the required disparity $D(p)$ is given at the location of the minimum cost sum.

Mathematically, the semi-global optimization process described above is as follows. The cost or energy function E calculates the costs of a possible disparity (parallax) D of a pixel depending on a cost value C and two penalty terms P_1 and P_2:

$$E(D) = \sum_p C(p, D_p) + \sum_{q \in N_p} P_1 \cdot T[|D_p - D_q| = 1] + \sum_{q \in N_p} P_2 \cdot T[|D_p - D_q| > 1] \qquad (5.95)$$

The first term of (5.95) contains the matching cost C for a pixel p. The second term in eqn. (5.95) adds a (small) penalty P_1 for the current disparity D_p to the cost value C if the difference between D_p and the disparity D_q at a neighbouring pixel q is 1. (The function T returns 1 if $|D_p - D_q| = 1$ and 0 in all other cases.) In an analogous way, the third term adds a larger penalty to the cost value C if the difference exceeds 1. (Here, the function T returns 1 if $|D_p - D_q| > 1$ and 0 in all other cases.) The penalties are defined within the interval of the applied cost function, e.g. -1 to +1 for normalized cross-correlation (NCC) whereas the ratio P_2/P_1 should lie between 2 and 3.

After calculation of the matching costs and their storage in the data structure C, a cost aggregation follows in which the penalties P_1 and P_2 are added to the matching costs (Fig. 5.92a bottom). It can be processed separately for every path r:

$$\begin{aligned}
L_r(p, D) = {} & C(p, D) \\
& + \min[L_r(p-r, D), L_r(p-r, D-1) + P_1, \\
& \qquad L_r(p-r, D+1) + P_1, \min_i L_r(p-r, i) + P_2)] \\
& - \min_k L_r(p-r, k)
\end{aligned} \qquad (5.96)$$

$$\text{where } (p, D) = (x', y', D)$$

The indices i and k are loop variables for the iteration through the aggregated costs at the previous position along one of the paths. For each path, the positions of adjacent pixels are defined as follows, e.g. with $u = 1$, $v = 0$ for a path in the x' direction:

$$L_r(p-r, D) = L_r(x'-u, y'-v, D) \qquad (5.97)$$

The expression in eqn. (5.96) searches for the minimum path costs, including penalties P_1 and P_2 potentially added at the position of the preceding pixel in path direction $(p-r)$. When found, the minimum is added to the matching costs of the current pixel p and the current disparity D. The last term in eqn. (5.96) subtracts the value of minimum path costs for the preceding pixel in order to avoid large numbers in L_r. The results of cost aggregation of all m paths are then accumulated as follows:

$$S(p,D) = \sum_{r=1}^{8,16} L_r(p,D)$$
(5.98)

In $S(p,D)$, the minimum is searched for every pixel p from which the final disparity is derived and stored in the disparity map $D(p)$.

$$D(p) = \arg\min_D S(p,D)$$
(5.99)

SGM creates 3D models with sharp edges and, to a certain extent, is able to bridge regions of poor texture since the semi-global search locates valuable image information at larger distances from the current pixel in support of the matching process. However, the accuracy in areas of reduced textured is lower since a degree of interpolation is involved.
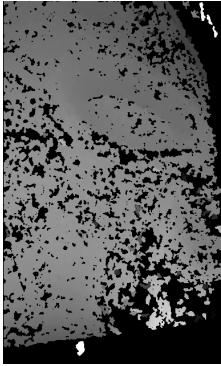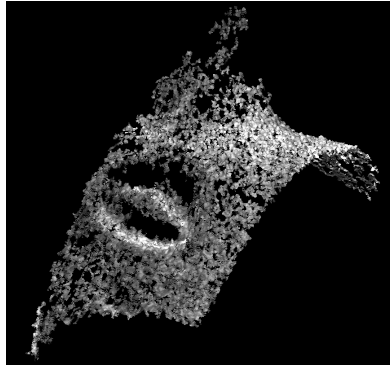
Disparity map   Point cloud (filtered)

Fig. 5.93: Image matching in a stereo pair with correlation-based stereo matching.
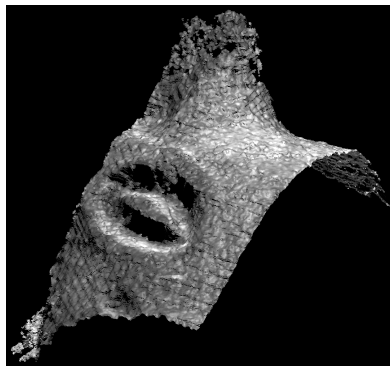
Disparity map   Point cloud (filtered)

Fig. 5.94: Image matching in a stereo pair with SGM.

Fig. 5.93 shows the disparity map and point cloud for the stereo image pair in Fig. 5.96 resulting from correlation-based stereo matching using a window size of 7 pixels. Fig. 5.94 illustrates the result with additional semi-global optimization and matching (compare with Fig. 5.106). In both cases a left/right check and a speckle filter are applied according to the OpenCV semi-global block-matching function. The more complete and smoothed surface from SGM is clearly visible.

5.5.5 Multi-image matching

There are different ways of extending image matching to more than two images. Matched pairs are often combined in order to fuse single point clouds into a global point cloud. Methods using image triplets are also common and can, to a large extent, solve the ambiguities in stereo matching. Finally, matching can be formulated for any number of images.

5.5.5.1 Multi-View Stereo

The term *Multi-View Stereo* (MVS) describes an approach which creates combinations of stereo images from an oriented multi-image configuration. Stereo matching is performed in each pair to give a single disparity map or, after computing associated space intersections, a point cloud. All point clouds are located within one common object coordinate system, hence can easily be fused into a single common point cloud. In MVS a reference image is selected and all overlapping images are applied for image matching. For n neighbouring images, n disparity maps are generated, hence each pixel of the reference image may have up to n disparities. The disparities are normalized in advance by dividing through by the base length (compare with eqn. 4.87). If there are holes in individual disparity maps, these can be filled with disparities from other image pairs. Finally, the normalized disparities are averaged, usually by a median filter weighted by the base length. Disparities due to longer stereo baselines (smaller height-to-base ratio) generate better height values according to eqn. (4.88).

5.5.5.2 Matching in oriented image triples

Ambiguities can be considerably reduced if the number of images is increased. Fig. 5.95 illustrates the matching principle for a configuration of three images. Starting with an image point P' in image B_1, the corresponding epipolar lines k_{12} and k_{13} can be calculated for the other images. For both partner images, ambiguities are represented as candidates P_a'', P_b'', P_c'' in image B_2 and P_d'', P_e'', P_f'' in image B_3. The homologous points of P' cannot therefore be uniquely determined.
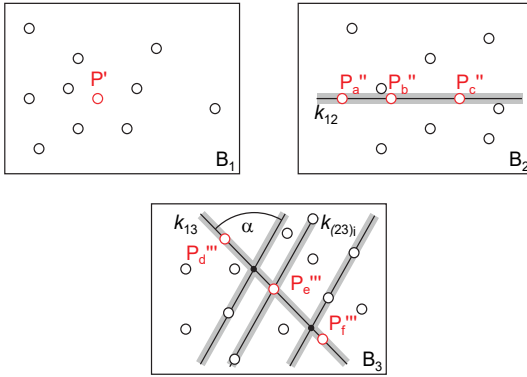
Fig. 5.95: Matching in image triples based on epipolar lines (according to Maas 1992).

If, in addition, the epipolar lines $k_{(23)i}$ are calculated in image B_3 for all candidates P_i in image B_2, it is most likely that only one intersection point with k_{13} is located close to a candidate in image B_3, in this case P_e'''. The search space is therefore restricted to the tolerance region of the intersection points. The number of possible ambiguities for three images is given by:

$$N_a = \frac{4(n^2 - n) \cdot \varepsilon^2}{F \cdot \sin \alpha} \left(1 + \frac{b_{12}}{b_{23}} + \frac{b_{12}}{b_{13}} \right) \tag{5.100}$$

where

n: number of image points
F: image area
ε: width (tolerance) of the search space
α: angle between epipolar lines in image B_3
b_{jk}:   base length between image B_j and image B_k

Ambiguities are minimized if the three cameras are arranged in an equilateral triangle such that $b_{12} = b_{13} = b_{23}$ and $\alpha = 60°$. In the numerical example above ($n = 1000$) the number of ambiguities is then reduced to $N_a = 10$.

Fig. 5.96 shows an image triplet recorded by a three-camera endoscopic system (Fig. 5.97 top left). Point P' is measured in image 1 and the corresponding epipolar lines k_{12} and k_{13} are displayed in images 2 and 3. These contain the corresponding points P" and P"'. If potentially corresponding image points are measured on k_{12}, they result in different epipolar lines k_{23} in image 3. Only one of them contains the correct point and intersects with epipolar line k_{13} in image 3.

image 1 (left)   image 2 (right)   image 3 (top)
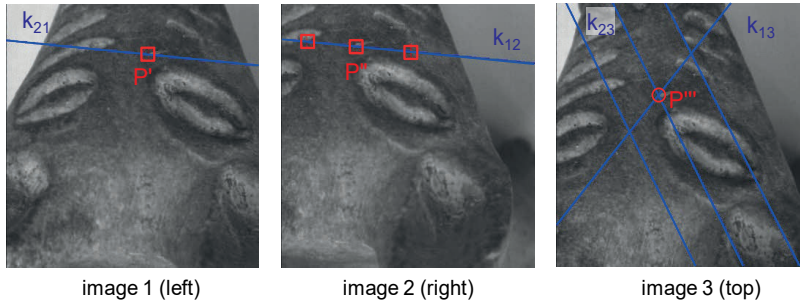
Fig. 5.96: Measured point and corresponding epipolar lines in an image triplet.

a) original

b) trifocal

c) orthogonal

image 1 (left)   image 2 (right)   image 3 (top)
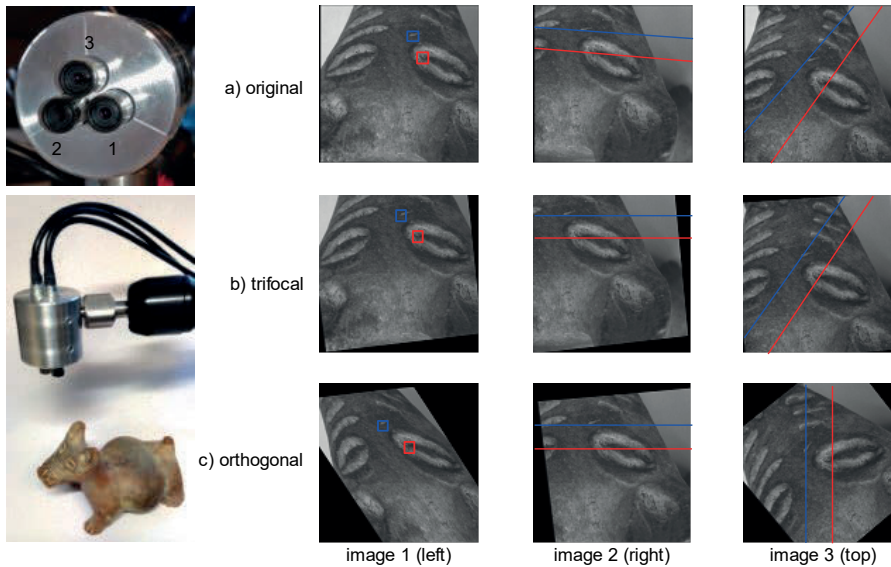
Fig. 5.97: Normalized images and epipolar lines for a three-camera system.

Normalized images can also be calculated for image triplets. Fig. 5.97a shows the three original images recorded by the three-camera system displayed on the left. As an example, two points are measured in image 1 and their corresponding epipolar lines are shown in the neighbouring images. If the three images are rectified onto the plane formed by the three perspective centres (trifocal plane), then images 1 and 2 are rotated parallel to each other by a rotation matrix that is formed by the normal vector of the trifocal plane and the base vector between the perspective centres of image 1 and 2. Then the same rotation matrix is assigned to image 3. The resulting images are displayed in row (b). Whilst the epipolar lines are parallel for images 1 and 2, they are at a slope angle in image 3. However, the content of the rotated images remains similar so that rectangular, non-rotated matching windows can be used. The images

displayed in row (c) are also rectified onto the trifocal plane such that the epipolar lines are horizontal between images 1 and 2 and vertical between 1 and 3. This is achieved by rotating the y-axis of image 3 parallel to the base vector between the projection centres of images 1 and 3 (vertical epipolar lines in image 3), and by applying shear and scale to image 1 using the distortion parameters C_1 and C_2 (see section 3.3.3.3). However, in this case, due to the rotation and distortion in the orthogonal image triplet (c), local matching requires affine transformed matching windows.

5.5.5.3 Matching in an unlimited number of images

The method can be extended to a virtually unlimited number of images. In order to search for homologous points of P' in all other images, a combinatorial approach must be applied which investigates all likely combinations of epipolar line intersections. The following matching strategy represents a practical solution for limiting the required computational effort:

1. Selection of an image point in image B_i.
2. Search for candidates on epipolar lines in images B_{i+j}, until at least 1 candidate is found.
3. Verification of all candidates by calculating their object coordinates using spatial intersection followed by back projection into all images B_{i+j+1} ... B_n.
4. Counting all successful verifications, i.e. the calculated (back projected) image position must contain an image point.
5. Acceptance of the candidate possessing the significantly largest number of successful verifications.

This approach offers some major advantages:
– an arbitrary number of images can be processed;
– an object point need not be visible in every image, e.g. due to occlusions or limited image format;
– interest values can optionally be used in order to reduce the number of candidates;
– approximate values of image points are not required, i.e. there are no pre-conditions for the smoothness of the object surface nor the spatial distribution of object points.

5.5.5.4 Multi-image least-squares matching

The least-squares matching approach introduced in section 5.4.2.4 for two image patches (reference and search image) can be extended by the following features:
– simultaneous matching of one point in multiple images (*multi-image matching*);
– simultaneous matching of multiple points in multiple images (*multi-point matching*);

– introduction of geometric conditions in image space and object space (*multi-image geometrically constrained matching*);
– introduction of object models (object-space matching).

Consider a reference image $f(x,y)$ and m search images $g_i(x,y)$, which are to be matched to the reference image. Eqn. (5.67) can be extended to multiple images:

$$f(x,y)-e_i(x,y)=g_i(x,y) \qquad\qquad i=1,\ldots,m \qquad\qquad (5.101)$$

Here $e_i(x,y)$ indicates random noise in image i. In a similar way to least-squares matching, the following adjustment system results:

$$\underset{n,1}{\mathbf{v}} = \underset{n,u}{\mathbf{A}} \cdot \underset{u,1}{\hat{\mathbf{x}}} - \underset{n,1}{\mathbf{l}} \qquad\qquad (5.102)$$

where

$$\hat{\mathbf{x}}_i^T = \left[da_0, da_1, da_2, db_0, db_1, db_2, r_0, r_1 \right]_i \qquad i=1,\ldots,m$$

m: number of search images
n: total number of observations
   $n = n_1+n_2+\ldots+n_m$; n_i = number of observations (pixel) in image $g_i(x,y)$
u: total number of unknowns, $u = 8 \cdot m$

The m parameter vectors $\hat{\mathbf{x}}_i^T$ can be determined independently within the system of equations (5.102) because they have no cross connections within the design matrix $\mathbf{A}$.

5.5.5.4.1 Functional model

The multi-image approach of (5.101) enables the simultaneous determination of all matches of a point. However, it does not consider the constraint that all homologous points must correspond to one common object point. This constraint can be formulated by the condition that all homologous image rays, taking account of image orientation parameters, must intersect optimally at the object point.

   This constraint is also the basis for the bundle adjustment model (section 4.4.2) where the collinearity equations (4.10) are used. The three-dimensional coordinates of point P in image k are given by[1]:

$$\mathbf{x'}_{Pk} = \frac{1}{m_{Pk}} \cdot \mathbf{R}_k^{-1} \cdot (\mathbf{X}_P - \mathbf{X}_{0k})$$

$$\begin{bmatrix} x'_P \\ y'_P \\ -c \end{bmatrix}_k = \frac{1}{m_{Pk}} \cdot \begin{bmatrix} r_{11} & r_{21} & r_{31} \\ r_{12} & r_{22} & r_{32} \\ r_{13} & r_{23} & r_{33} \end{bmatrix}_k \cdot \begin{bmatrix} X_P - X_{0k} \\ Y_P - Y_{0k} \\ Z_P - Z_{0k} \end{bmatrix} \qquad (5.103)$$

In simplified notation, the image coordinates x'_P, y'_P are given by

---

[1] Image coordinates have their origin at the principal point and are corrected for distortion.

$$x'_{Pk} = -c_k \left(\frac{k_X}{N}\right)_{Pk} = -F^X_{Pk} \qquad\qquad y'_{Pk} = -c_k \left(\frac{k_Y}{N}\right)_{Pk} = -F^Y_{Pk} \qquad (5.104)$$

Based on initial values $(x'_p, y'_p)^0$, the following non-linear observation equations for the required shifts $\Delta x'_p, \Delta y'_p$ result:

$$\Delta x'_{Pk} + F^X_{Pk} + x'^{\,0}_{Pk} = 0 \qquad\qquad \Delta y'_{Pk} + F^Y_{Pk} + y'^{\,0}_{Pk} = 0 \qquad (5.105)$$

Here $\Delta x'_p, \Delta y'_p$ correspond to the shift coefficients da_0, db_0 in (5.102). The equations (5.105) establish the relationship between image space and object space. In contrast to the collinearity equations, which primarily only establishes the functional relationship between observed image coordinates and unknown point and orientation data, this approach uses as observations the original grey values, in combination with least-squares matching. Additional observation equations can be set up using the parameters in terms F_{pk}, e.g. for the simultaneous calculation of object coordinates XYZ or for the formulation of geometric constraints, e.g. $Z = $ const.

5.5.5.4.2 Object restrictions

It is first assumed that the interior and exterior orientation parameters of all images are known, e.g. by a prior bundle triangulation. The system of equations (5.105) must be linearized at initial values of the remaining unknowns X, Y, Z:

$$\begin{aligned}
\Delta x'_{Pk} + \frac{\partial F^X_{Pk}}{\partial X_P} dX_P + \frac{\partial F^X_{Pk}}{\partial Y_P} dY_P + \frac{\partial F^X_{Pk}}{\partial Z_P} dZ_P + F^{X\,0}_{Pk} + x'^{\,0}_{Pk} = 0 \\
\Delta y'_{Pk} + \frac{\partial F^Y_{Pk}}{\partial X_P} dX_P + \frac{\partial F^Y_{Pk}}{\partial Y_P} dY_P + \frac{\partial F^Y_{Pk}}{\partial Z_P} dZ_P + F^{Y\,0}_{Pk} + y'^{\,0}_{Pk} = 0
\end{aligned} \qquad (5.106)$$

In summary the system of additional observation equations is as follows:

$$\underset{m',1}{\mathbf{w}} = \underset{m',3}{\mathbf{B}} \cdot \underset{3,1}{\hat{\mathbf{y}}} - \underset{m',1}{\mathbf{t}} \qquad (5.107)$$

where

$$\hat{\mathbf{y}}^T = \left[dX, dY, dZ \right]$$

m': number of images:

  $m' = m$, if no transformation of the reference image is permitted

  $m' = m+1$, if the reference image is also to be transformed, e.g. with respect to an artificial template

For the extended system of (5.105) and (5.107) the complete vector of unknowns is given by:

$$\overline{\mathbf{x}} = (\mathbf{A}^T \mathbf{P} \mathbf{A} + \mathbf{B}^T \mathbf{P}_t \mathbf{B})^{-1} \cdot (\mathbf{A}^T \mathbf{P} \mathbf{l} + \mathbf{B}^T \mathbf{P}_t \mathbf{t}) \qquad (5.108)$$

The parameter vector $\overline{\mathbf{x}}$ consists of the simultaneously calculated displacements of the image patches in $g_i(x,y)$ as well as the adjusted corrections of the object coordinates. Because of their relationship to the XYZ object coordinates, the image shifts cannot take arbitrary values but are constrained to follow an epipolar line (Fig. 5.98). The influence of this geometric condition can be controlled by the weight matrix $\mathbf{P}_t$.
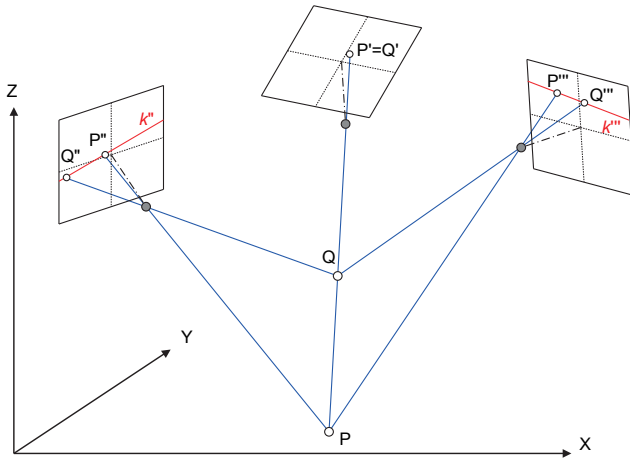
Fig. 5.98: Geometric condition of ray intersection.

The model in (5.106) is appropriate for determining arbitrary 3D object points. In order to measure surface models, some of the coordinate components of the surface points can be fixed:

– Constant X,Y:

If the surface points are located on a particular grid of XY coordinates, for example where $\Delta X = \Delta Y = \text{const.}$, corresponding terms can be eliminated from (5.106) and only Z-value adjustments remain:

$$\Delta x'_{Pk} + \frac{\partial F^X_{Pk}}{\partial Z_P} dZ_P + F^{X\ 0}_{Pk} + x'^{\ 0}_{Pk} = 0$$

$$\Delta y'_{Pk} + \frac{\partial F^Y_{Pk}}{\partial Z_P} dZ_P + F^{Y\ 0}_{Pk} + y'^{\ 0}_{Pk} = 0$$

(5.109)

In consequence, the normal system of equations becomes simpler. The approach corresponds to the manual measurement of terrain models in stereoscopic plotters. In this case the image shifts do not take place within epipolar lines but on the vertical line locus instead (see section 4.3.6.3).

– Constant Z:
  In order to measure contour lines at a given height Z = constant, only coordinate displacements in the X and Y directions are permitted and (5.106) reduces to:

$$\Delta x'_{Pk} + \frac{\partial F^X_{Pk}}{\partial X_P} dX_P + \frac{\partial F^X_{Pk}}{\partial Y_P} dY_P + F^{X\ 0}_{Pk} + x'^{\ 0}_{Pk} = 0$$

$$\Delta y'_{Pk} + \frac{\partial F^Y_{Pk}}{\partial X_P} dX_P + \frac{\partial F^Y_{Pk}}{\partial Y_P} dY_P + F^{Y\ 0}_{Pk} + y'^{\ 0}_{Pk} = 0$$

(5.110)

The possibility of restricting particular coordinate components enables surface models to be recorded with respect to specific reference planes or along specific sections through the surface. The (topological) structure of the surface model is therefore defined directly during measurement and not by a later analysis of an unstructured point cloud which could be the result, for instance, of an active projection method (section 6.7.3 and also see section 5.5.6.1).

5.5.5.4.3 Additional extensions
The concept of additional observation equations, described above, permits the introduction of further constraints on the adjustment system. The influence of additional observations can be controlled by appropriate weighting from $p_i = 0$ where the constraint has no effect through to $p_i = \infty$ where the constraint is strictly enforced.

As examples, additional constraints can be formulated for the following tasks:
– Edge extraction by least-squares matching:
  Single edge points can be determined by a suitable edge template (see example in Fig. 5.55). As edges are linear features, an additional constraint can be introduced which forces the template to move in the direction of the gradient, i.e. perpendicular to the edge.
– Determination of spatial object contours:
  Spatial object contours can be described by geometric elements (straight line, circle etc.) or spline functions. The corresponding analytical parameters can be included as unknowns in a least-squares fit. An example of this approach is the method of LSB snakes which combines edge extraction and determination of a B-spline spatial curve (snake) in a single analysis.
– Measurement of point grids:
  The least-squares matching concept for a single object point can easily be extended to an unlimited number of points. However, if no geometric relationship between the points exists then the result is identical to single point matching. Geometric relations can be formulated either by associating points with a common geometric element, e.g. all point belong to a plane or a cylinder, or by defining neighbourhood relationships.

The latter approach can be compared to the interpolation of digital surface models where points on a grid are connected by additional constraints such as

minimum surface curvature (section 3.3.4.5). The approach of multi-patch matching utilizes this idea by defining constraints between adjacent grid elements (patches).
– Bundle concept:
  If the interior and exterior orientation parameters of images are known only approximately, they can be included as unknowns in the least-squares matching. Eqn. (5.106) is then extended by corresponding differential coefficients of the additional orientation parameters.

5.5.6 Matching methods with object models

The matching methods described in the previous sections are mainly based on geometric relationships between images, or between images and object. Although contrast differences are modelled in the least-squares matching by two radiometric parameters, they are completely independent of the reflection characteristics of the surface.

In order to create a complete object model, it is necessary to combine the geometric object properties (position and shape) with the reflection properties of the surface. This idea is used in various forms of 3D visualization and can also be used for object reconstruction. The features of such a complete reconstruction method are:
– introduction of a surface reflection model (material characteristics, illumination);
– ray tracing through different media (as a minimum through the atmosphere);
– multi-image adjustment based on least-squares matching;
– topological structuring of the surface by surface grid and shape lines.

The global objective is the 3D surface reconstruction by simultaneous calculation of all orientation parameters, object point coordinates and geometric element parameters, as well as illumination and reflection parameters. Grey values in multiple images are available as observations, as are initial values of unknowns. The number of parameters is extremely high and the resulting system of equations is correspondingly complex.

5.5.6.1 Object-based multi-image matching
Object-based multi-image matching is based on a relationship between the intensity (colour) value of a surface element G_i and the grey values g_{ij} of the associated images (Fig. 5.99). The grey values are a function of the orientation parameters $\mathbf{O}_j$ of an image j and the surface parameters $\mathbf{Z}_i$ of an element i:

$$G_i = g_{ij}(\mathbf{Z}_i, \mathbf{O}_j) \qquad i = 1, \ldots, m \qquad j = 1, \ldots, n \tag{5.111}$$

where

m: number of surface elements

n: number of images

In order to solve this system of equations, initial values for the unknown surface and the orientation parameters are required. The difference between the grey values calculated from the initial values can be regarded as a stochastic value which gives:

$$\Delta g_{ij} = G_i^0 - g_{ij}(\mathbf{Z}_i^0, \mathbf{O}_j^0) \tag{5.112}$$
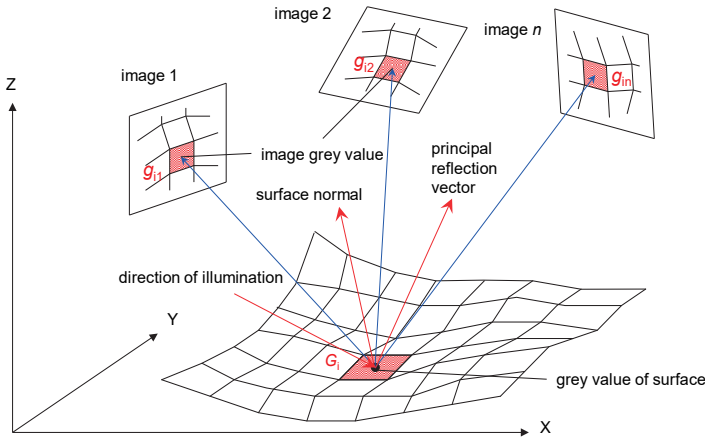
Fig. 5.99: Relationship between intensity value of surface and grey value of image (Schneider 1991).

The imaged light intensity depends on the properties of the surface material and of the geometric orientation of surface element, light source and imaging sensor (Fig. 5.99) as, for example, described by Phong's illumination model. In a similar way to least-squares matching, the grey values in the individual images are matched using two radiometric parameters which are dependent on material:

$$\Delta g_{ij} = G_i^0 - \left(r_{1,j}^0 + r_{2,j}^0 \cdot g'_{ij}(\mathbf{Z}_i^0, \mathbf{O}_j^0) \right) \tag{5.113}$$

where

$r_{1,j}^0, r_{2,j}^0$:   approximate radiometric correction parameters

g'_{ij}:   observed image grey value: $g_{ij} = r_{1,j}^0 + r_{2,j}^0 \cdot g'_{ij}$

Eqn. (5.113) can form the observation equation for each grey value in all images where a specific surface element is visible. Using the substitution

$$r_{2,j} = 1 + r'_{2,j}$$

and re-arranging the linearized correction equations, the following is obtained:

$$v_{ij} = G_i^0 - dG_i - (r_{1,j}^0 + dr_{1,j}) - (r'^0_{2,j} + dr'_{2,j}) \cdot g'_{ij}(\mathbf{Z}_i^0, \mathbf{O}_j^0)$$
$$- \frac{\partial g'}{\partial \mathbf{Z}} d\mathbf{Z} - \frac{\partial g'}{\partial \mathbf{O}} d\mathbf{O} - g'_{ij}(\mathbf{Z}_i^0, \mathbf{O}_j^0) \tag{5.114}$$

The differential coefficients

$$\frac{\partial g'}{\partial \mathbf{Z}} = \frac{\partial g'}{\partial x'} \cdot \frac{\partial x'}{\partial \mathbf{Z}} + \frac{\partial g'}{\partial y'} \cdot \frac{\partial y'}{\partial \mathbf{Z}} \qquad\qquad \frac{\partial g'}{\partial \mathbf{O}} = \frac{\partial g'}{\partial x'} \cdot \frac{\partial x'}{\partial \mathbf{O}} + \frac{\partial g'}{\partial y'} \cdot \frac{\partial y'}{\partial \mathbf{O}} \tag{5.115}$$

contain the grey-value gradients $\partial g'/\partial x'$ and $\partial g'/\partial y'$. This approach can therefore only be applied when an appropriate number of edges or structures exist in the images. The remaining differential coefficients $\partial x'/\partial \mathbf{Z}$ etc. correspond to the derivatives from the space resection and bundle triangulation models (see section 4.4.2).

Now all the relevant parameters are available for an iterative adjustment:

$$\mathbf{v} = \mathbf{A} \cdot \hat{\mathbf{x}} - \mathbf{l}$$

with observations

$$\mathbf{l} = \mathbf{g'}(\mathbf{Z}^0, \mathbf{O}^0)$$

which leads to a vector of unknowns:

$$\hat{\mathbf{x}}^T = (d\mathbf{Z}^T, d\mathbf{O}^T, d\mathbf{r}_1^T, d\mathbf{r}'^T_2, d\mathbf{G}^T) \tag{5.116}$$

The importance of object-based multi-image matching has two aspects. One is the simultaneous determination of all parameters influencing the formation of the image, including the object itself. The other is that the method can be extensively modified by altering object parameters relevant to the application. In this way, the following tasks can be solved by use of appropriate functions:

– Measurement of artificial targets:

Circular targets can be defined as a circle in space (see section 2.3.2.3) whose centre must be determined. An additional transformation converts the coordinates of the plane of the circle into the object coordinate system. The target's grey values have a radiometric model which is a function of the radius and consists, for example, of a white circle on a black background (Fig. 5.100).

The surface element size should approximately correspond to the size of a pixel in image space. The circle centre in object space is determined ignoring any eccentricities (see section 6.2.1.1).
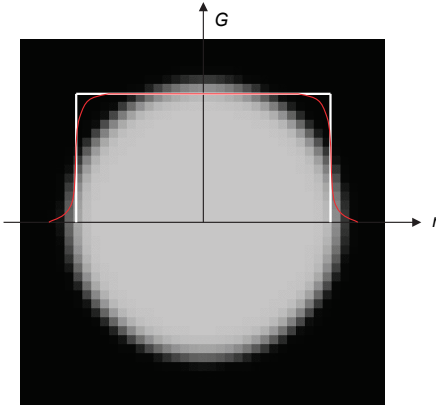
Fig. 5.100: Distribution of smoothed object grey values for a circular target.

– Measurement of surfaces:
Surfaces are recorded as profiles, surface models or free-form surfaces. If the measurement is based on profiles, an arbitrarily oriented section is defined in which the profile is represented by a two-dimensional curve (Fig. 5.101). Within the section, a surface element is defined by a tangential plane whose degrees of freedom are reduced to two rotations and a shift within the section plane.
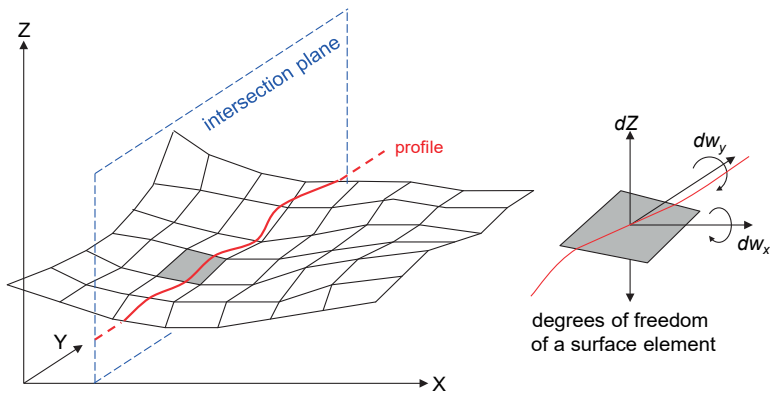
Fig. 5.101: Resulting profile from intersection of section plane and surface.

– Bundle triangulation:
The system of equations (5.113), extended to determine 3D circles and additional parameters of interior orientation (simultaneous calibration), leads logically to

an object-based bundle triangulation for photogrammetric images with circular targets.

– Orthophoto production:
Object-based multi-image matching can also be applied to the production of orthophotos where the calculated grey values of the object are here used for image generation. The required 3D object model can be estimated within the multi-image matching process itself, or it can be generated from other sources. In principle, this approach enables orthophotos to be generated from an arbitrary number of images, with occlusions and radiometric variations no longer affecting the result.

Object-based multi-image matching, briefly described here, has considerable potential for wide application and makes use only of the original image grey values. Feasibility in principle, and some test applications, have been demonstrated but so far practical use has been limited by the high computational effort needed to solve the large systems of normal equations. The generation of appropriate initial values for the unknown parameters is, in addition, a non-trivial problem.

5.5.6.2 Multi-image matching with surface grids

Surface elements calculated by object-based multi-image matching (section 5.5.6.1) are independent of their adjacent elements in terms of geometry and radiometry. However, this assumption is invalid for piecewise smooth surfaces. With the exception of discontinuities (breaklines), adjacent object elements can be connected by radiometric and geometric interpolation functions. The above approach can therefore be extended by coefficients of piecewise linear functions for both the radiometric model and the geometric model (terrain model). For this purpose, the surface is represented by triangular or quadrilateral patches (facets) within which linear or bilinear interpolation can be applied.

Fig. 5.102 illustrates the principle. The required surface $Z(X,Y)$ consists of grey values $G(X,Y)$. Photographic recording of surface points P(X,Y,Z) results in image points P'(x',y') with grey values $g'(x',y')$. Based on an approximately known height model $Z^0(X^0,Y^0)$, surface points are interpolated, e.g. on a fixed grid with separations $\Delta X^0 = \Delta Y^0 =$const. The surface reflection model $G'(X,Y)$ is calculated on a dense grid width $\Delta X_G, \Delta Y_G$ (Fig. 5.103).

The grid width of the reflection model should be adjusted to the maximum spatial frequency in object space. If this information is missing, the grid width should be chosen with respect to the mean pixel size $\Delta x', \Delta y'$ in the images:

$$\Delta X_G \geq m \cdot \Delta x'$$
$$\Delta Y_G \geq m \cdot \Delta y'$$

where m: image scale number (5.117)
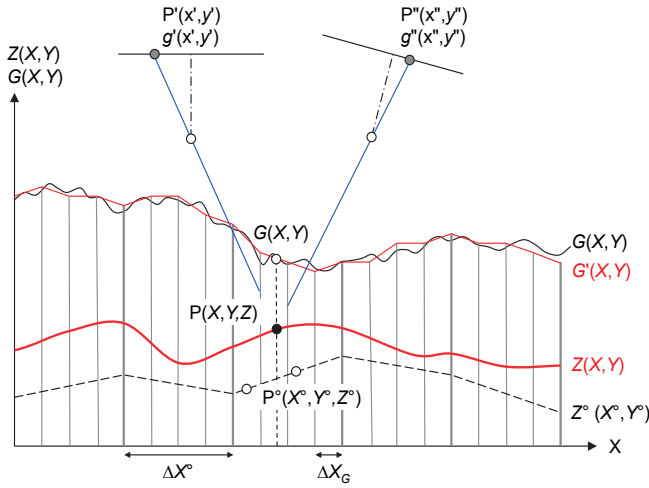
Fig. 5.102: Interpolation and projection of surface and reflection model (after Wrobel 1987).
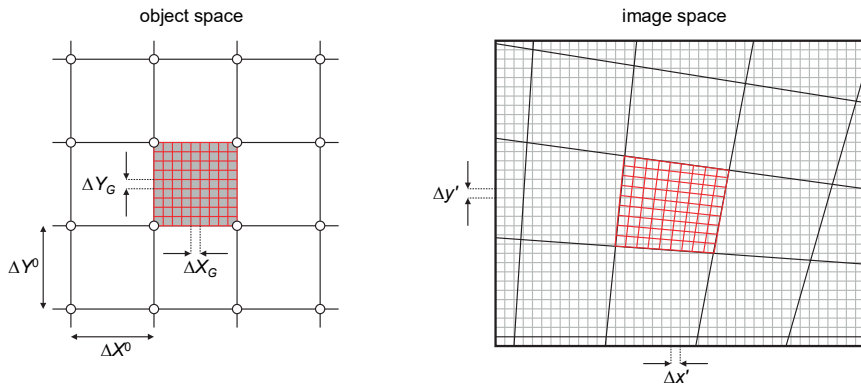
Fig. 5.103: Projection of object facets into image space.

For the height model, the grid width should adapt to the local object shape in order to achieve a reasonable result. For example, grid width can be reduced on object breaklines. Usually a square grid is chosen where

$$\Delta X^0 = \Delta Y^0 > \Delta X_G = \Delta Y_G \tag{5.118}$$

The introduction of radiometric facets with corner-point grey values $G(X_k, Y_l)$ leads to:

$$G(X,Y) = \sum a_{kl}(X,Y) \cdot G(X_k, Y_l) \tag{5.119}$$

$$\text{where } \sum_{kl} a_{kl} = 1$$

Correspondingly, coefficients for the height model's geometric facets $Z(X_r,Y_s)$ are introduced to give:

$$Z(X,Y) = \sum b_{rs}(X,Y) \cdot Z(X_r,Y_s) \tag{5.120}$$

where $\sum_{rs} b_{rs} = 1$

The system of equations (5.113) is extended by (5.119) and (5.120). Coefficients a_{kl} and b_{rs} are determined by the nodal point values of each object grid or triangle facet.

Object-based multi-image matching based on object facets offers additional benefits:

– By choosing suitable weights for the coefficients, the smoothness of a surface can be controlled so that image matching can bridge critical areas containing intensity distributions where texture is not present.
– Occlusions and discontinuities can be handled by correspondingly detailed object surfaces or 3D models.

5.5.6.3 Object-based semi-global multi-image matching (OSGM)

The principle of semi-global matching (SGM, section 5.5.4.2) can be transferred from image to object space. As a first step, the object space is divided into discrete voxels (voxel grid). The voxels are organized as cubes or cuboids with their dimensions (ΔX, ΔY, ΔZ) defining the required resolution for object reconstruction. Voxels should be defined by taking account of the mean pixel size in object space (ground sampling distance, GSD) and the spatial configuration (height-to-base ratios) of the images. The SGM energy function in eqn. (5.95) is then modified such that image coordinates and disparities are replaced by object coordinates:

$$E(Z) = \sum_{X,Y} [C(X,Y,Z) + \sum_{q \in N_p} P_1 \cdot T(|Z - Z_q| = \Delta Z) + \sum_{q \in N_p} P_2 \cdot T(|Z - Z_q| > \Delta Z)] \tag{5.121}$$

The first term in (5.121) contains the matching costs for each voxel whilst the second and third terms introduce penalties for differences between Z value between neighbouring voxels. Hence, the smoothing conditions of the original SGM here cause a smoothing in the direction of a selected axis in space, which in this case is the Z axis.

The calculation of matching costs for each pixel (creation of data structure C in eqn. 5.121) considers grey or colour values from 2 to n images (Fig. 5.104). True multi-image matching is therefore provided. With n images, the conventional calculation of matching costs must be extended. However, since pairwise matching in multi-view approaches (section 5.5.5.1), and the redundant calculation of object coordinates, allow for a robust detection of outliers and occlusions, it makes sense to integrate stereo and multi-image strategies into OSGM as well.

If window-based cost functions are used, e.g. Census or normalized cross-correlation (NCC), the matching windows are defined in object space orthogonal to the reference axis and back-projected into image space. This strategy ensures a high invariance against image rotations and, to some extent, against differences in scale. In OSGM, the matching windows are shifted in object space by ΔZ increments according to the selected resolution. Back-projection of the matching windows into image space leads to subpixel positions, hence cost calculation is done with interpolated grey or colour values. As for SGM, Census or NCC can be used as cost functions.
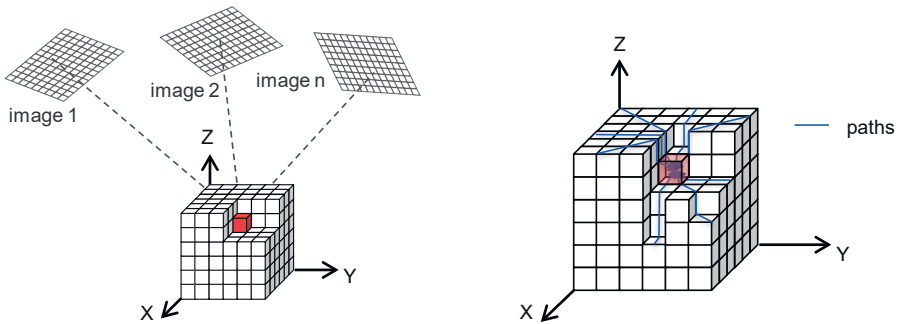
Fig. 5.104: Cost calculation for a voxel (left) and path directions for cost aggregation (right).

After cost calculation, the penalties P_1 and P_2 are added to the costs (cost aggregation). The term of eqn. (5.96) is then modified to

$$
\begin{aligned}
L_r(v,Z) = {} & C(v,Z) \\
& + \min[L_r(v-r,Z), L_r(v-r,Z-\Delta Z) + P_1, L_r(v-r,Z+\Delta Z) \\
& + P_1, \min_i L_r(v-r,i\cdot\Delta Z) + P_2)] - \min_k L_r(v-r,k\cdot\Delta Z)
\end{aligned}
\tag{5.122}
$$

where $(v,Z) = (X,Y,Z)$ and X,Y,Z are the central coordinates of a voxel v. The indices i and k denote loop variables for the iteration through the aggregated costs at a preceding position within the path. The relationship between two voxels, with respect to a reference plane, in this case the XY plane, is defined by:

$$
L_r(v-r,Z) = L_r(X-u\cdot\Delta X, Y-v\cdot\Delta Y, Z)
\tag{5.123}
$$

As an example, $u=1$, $v=0$ for a path in the X-direction. Applying eqn. (5.122) to cost aggregation results in a penalty for differences of Z values compared with neighbouring voxels, hence the smoothing takes effect in the direction of the Z axis. The results of path-wise cost aggregation can be accumulated by:

$$S(v,Z) = \sum_{r=1}^{m} L_r(v,Z) \qquad\qquad (5.124)$$

Finally, the minimum for each voxel v in the Z direction is determined in S. This represents the final matching result. Hence, OSGM directly leads to a 2½D point cloud in object space instead of a disparity map in image space.
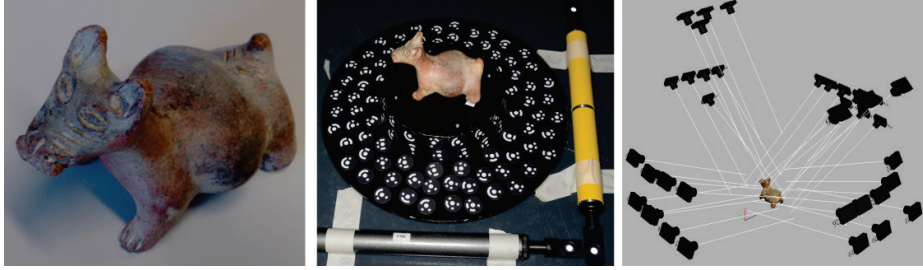
Fig. 5.105: Test object (left), reference points (middle) and imaging situation (right).

Fig. 5.105 shows a ceramic sculpture, imaged in an all-around configuration by 38 images and with all images used in the 3D reconstruction. In a first step, the object space is divided into separate voxel blocks to which those images are assigned that see the block completely. The reconstruction result is displayed in Fig. 5.106. It shows a homogeneous and dense surface with only a few holes due to areas without texture. In this example, the voxel resolution is set to 0.3 mm, corresponding to the mean GSD of the images. The mean standard deviation with respect to a reference model measured by a fringe projection system is 0.16 mm.
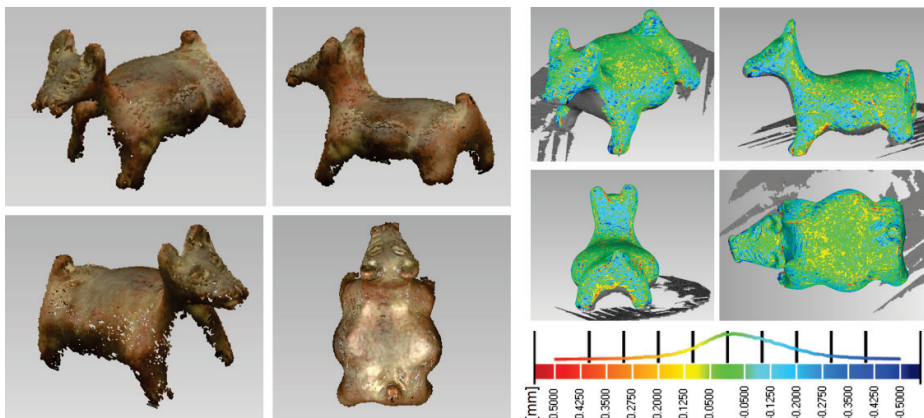
Fig. 5.106: Reconstructed 3D model (left) and comparison with reference data (right).

OSGM is also useful for the matching of nadir and oblique aerial images, i.e. images with very different image scales. Consequently, the voxel structure directly results in a true orthophoto.

5.5.6.4 Additional object-based matching methods

From the variety of additional object-based matching methods known from recent research, a selection is briefly outlined here.

5.5.6.4.1 Plane-sweep methods

The objective of this method is the extraction of object planes from oriented images. In the simplest case, the applied similarity measure is based on the minimum grey-value difference between homologous points which are determined by plane projective transformation from a reference image into participating images. From the known perspective centres of the reference image and the k participating images, the following equation can be formed using the normal vector n_E of the required plane and the distance d_E between plane and camera (see also eqn. 4.76):

$$H(E, \mathbf{P}_k) = \mathbf{K}_k \left(\mathbf{R}_k^T + \frac{\mathbf{R}_k^T X_{0k} \mathbf{n}_E^T}{d_E} \right) \mathbf{K}_{ref}^{-1} \tag{5.125}$$

H is a homogenous projective transformation (homography) which is used to transform the homogenous image vector from the reference image into the neighbouring image:

$$\begin{bmatrix} u_k & v_k & w_k \end{bmatrix}^T = H(E, \mathbf{P}_k) \cdot \begin{bmatrix} x_{ref} & y_{ref} & 1 \end{bmatrix}^T \tag{5.126}$$

$$x_k = u_k / w_k \qquad y_k = v_k / w_k$$

With the cost function

$$C(x, y, E) = g(x_{ref}, y_{ref}) - r_k^{ref} g(x_k, y_k) \tag{5.127}$$

it is assumed that image points lying on the plane have the same colour taking account of a gain factor r_k^{ref}. Through a global optimization, points lying on a plane are extracted in order to calculate the plane parameters n_E and d_E. Extensions of the method allow for the simultaneous estimation of an arbitrary number of planes from images and image sequences, e.g. for autonomous navigation of vehicles. Accurate 3D elements cannot be generated by these methods.

5.5.6.4.2 PatchMatch stereo

This approach is also based on the extraction of local object planes such that for every pixel one plane in space can be assigned whose projection into the stereo image

results in minimal colour differences. In the first step a normal disparity image is calculated using a local image-based similarity measure. Using a left/right check, the result is tested for consistency and possible occlusions are eliminated. For a pixel $p(x,y)$ in the left hand image, random parameters of plane E are estimated using distance information derived from disparity. From the plane parameters (compare with eqn. 2.147) the following cost function is constructed:

$$C(p,E) = \sum_{q \in W} w(p,q) \cdot \rho(q,q') \qquad (5.128)$$

where

W: square window with centre p

$w(p,q)$:   probability that $g(p)$ is equal to $g(q)$

q':   pixel in right hand image with $q - (a_E x_q + b_E y_q + c_E)$

a_E, b_E, c_E: parameters of plane E

ρ: measure of non-similarity between q and q'

The equation shows that the cost function depends on:
– the probability that two pixels on the same plane have similar colours;
– the non-similarity measure ρ which compares q with point q' lying on the epipolar line in the neighbouring image that again depends on the plane.

For each adjacent pixel q, the cost measure is tested to see if it can be improved by an additional pixel, i.e. if it belongs to the same plane. If this is fulfilled, the plane E will be assigned to point q, otherwise a new random selection of plane parameters is chosen.

5.5.6.4.3 Optimization of triangular meshes

The idea here is based on a point cloud derived from a Multi-View Stereo analysis (section 5.5.5.1) used for triangular meshing (Delauney triangulation, see section 2.3.3.1). As a matching strategy, for example, the correspondence of feature points or the plane-sweep method can be applied. In contrast to a voxel representation, the triangular mesh adapts to the density of surface points. After applying a ray tracing test, each triangle is projected into the participating images in order to calculate a similarity measure between them.

Subsequently, the gradient of the similarity measure is determined at the edge points of adjacent triangles. The differences between neighbouring gradients are adjusted by a gradient descent method. This is done by an iterative shift of the 3D points to improve the similarity of projected triangles and to minimize gradient differences. As a result, a surface model is generated that can be densified by dividing the triangles further.

5.5.6.4.4 Global photo consistency

The objective is the reconstruction of a surface S from all views of an object by minimizing an energy function E which consists of three cost terms. The terms indicate if a point x lies on the object or is part of the background under consideration. The evaluation makes use of a photo-consistency function which applies the constraint that an object point has the same colour in all images. For this, Lambertian reflection is assumed.

$$E(S) = \int\limits_{R^S_{Obj}} \rho_{Obj}(x)dx + \int\limits_{R^S_{Back}} \rho_{Back}(x)dx + v\int\limits_{S} \rho_S(x)dS \tag{5.129}$$

where

x:	3D object point
ρ_{Obj}:	binary [0,1] variable for membership of x to object region R_{Obj}
R_{Obj}:	object region in surface S
ρ_{Back}:	binary [0,1] variable for membership of x to background region R_{Back}
R_{Back}:	background region of S
ρ_S:	binary [0,1] variable for photo consistency
v:	weighting factor
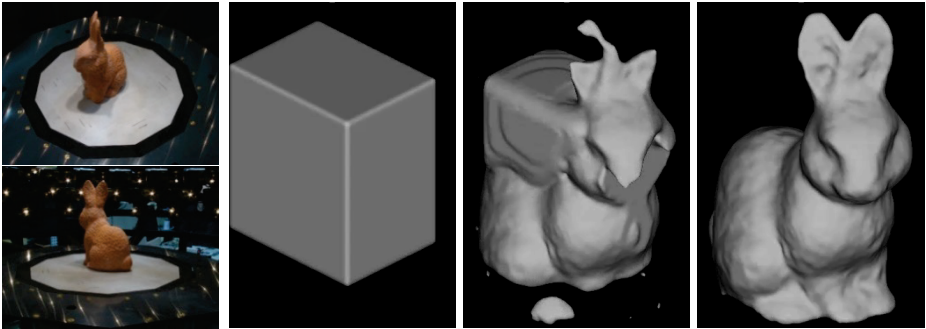
Fig. 5.107: Images from a sequence (left) and iterative 3D modelling by photo consistency (right) (CVG TU Munich).

The surface to be modelled is derived from the minimum of $E(S)$. Fig. 5.107 shows an example of surface reconstruction from an image sequence where a starting model (cuboid) is iteratively adapted to the shape of the object. The object itself is separated from the background.

5.5.7 Matching in image sequences

An image sequence implies a series of images acquired dynamically. Here it is assumed that camera and/or object are moving relative to each other such that the differences from single image (frame) to image are small enough to describe and model the dynamics of the scene. To achieve this, the frame rate (frames per seconds, fps) must be appropriate for the relative velocity between imaging system and object (which may also be changing form). A series of images, or a multi-image set, as used in offline photogrammetry or structure-from-motion, is not regarded here as an image sequence.

As examples, image sequences can be recorded by the following systems:
– videos with frame rates up to 60 fps (video cameras, action cameras, surveillance cameras, digital cameras with video mode);
– high-speed movies with frame rates up to 100000 fps (high-speed cameras, section 3.5.3);
– cameras in autonomous systems (vehicles, robots, drones, section 6.11);
– multi-camera systems for motion capture (section 6.10.3);
– mobile mapping systems (section 6.11).

In principle, image sequences can be acquired by any number of cameras simultaneously. For a precise matching in dynamic stereo or multi-image sequences, the cameras must be synchronized to avoid false 3D reconstructions (see section 6.10.1). All applications must ensure that the measured object coordinates or trajectories are related to a common coordinate system.

The single frames in a sequence can also be processed like any other digital image. If object points are to be tracked within a dynamic image sequence, the problem can be solved, in principle, by the matching procedures described in sections 5.5.4 and 5.5.5.

The following photogrammetric applications can be handled by dynamic image sequences:
– Tracking of targeted object points:
  Targeting an object, e.g. with retro-reflective markers (section 3.6.2), enables the secure identification and matching of, typically, a small number of object points. Example applications are motion capture (section 6.10.3), medical navigation in orthopaedics (section 8.7.2), measurement of trajectories in car crash testing (section 6.10.2.3) or 6DOF measurement of mobile measuring sensors (section 6.9.5).
– Tracking of non-targeted (natural) object points:
  If the object cannot be targeted, natural features must be measured. These could either be randomly identified feature points generated by a feature detector or explicitly defined object points which are physically relevant to the application. Often these specific points are required to be tracked, e.g. for material testing or

mechanical analysis. Example applications are measurements of formed metal sheet (strain analysis, section 6.7.5.3), measurement of deformed free-form surfaces etc.

– Deformation analysis of surfaces:
  Here the object's surface is analysed as a whole, i.e. there is no tracking of specific surface points. Applications include the measurement of changing volumes or the determination of critical object displacements.

– Detection of moving objects:
  Numerous applications of image sequence analysis in computer vision and robotics deal with the detection and tracking of objects, e.g. detection of pedestrians for driver-assistance systems, tracking herds of animals and monitoring large groups of people to detect suspicious patterns of motion. In these cases, it is normally more important to have secure and robust object detection than high measurement accuracy.

– Localization of camera platforms in known or unknown environments:
  Tracking methods utilizing object-space reference data can be used to determine the spatial position and orientation (6DOF, pose) of a camera or its platform. Example applications are the measurement of hand-held probes with integrated camera (section 6.5.2.2), control of UAVs for automatic landing and robot calibration. SLAM algorithms can be used for localization of camera positions with simultaneous mapping of the (unknown) environment (section 5.5.7.6).

5.5.7.1 2D object tracking in single-camera sequences

A sequence of images taken by a single camera can be described mathematically as:

$$\mathbf{S} = s(x,y,t) = \begin{bmatrix} s_0(x,y,t_0) \\ s_1(x,y,t_1) \\ \vdots \\ s_{n-1}(x,y,t_{n-1}) \end{bmatrix} \tag{5.130}$$

where t denotes either the time stamp of a single frame or an identifier of the frame within the sequence. It is assumed that all frames have the same geometric properties, e.g. the same image size.

5.5.7.2 Target tracking

Target tracking can easily be achieved by the methods described in section 5.4.2. Starting with the selection of a target in the initial image S_0 (manually or by automatic segmentation), the measured image coordinates of the target are transferred to image S_1 where they serve as approximations for the actual image position. The process is repeated for all images S_i. If coded targets are used, the identification of targets is robust even in cases where the line of sight is temporarily interrupted. Example

applications are described in section 6.10.2.2 (6DOF tracking) and 8.7.2 (medical navigation).

5.5.7.2.1 Tracking of distinctive image patterns

The following strategies are suitable for tracking distinctive but non-targeted image patterns (feature tracking):

– Interest operator and feature descriptor:

An interest or feature detector provides position and feature parameters for distinctive image regions in the initial frame S0 (e.g. SIFT, section 5.4.3.5). The detection is repeated in the next image. This can be completely independent of the search in the previous image or the search regions in the current image can be restricted to areas which were detected in the previous image. Points are matched by matching feature vectors. As in stereo image matching, this allows corresponding points with large positional changes to be matched.

A disadvantage of feature tracking is that specific object points cannot be tracked. Feature tracking therefore cannot be used to track points representing key mechanical features or locations where deformations are required.

– Area-based image matching:

Firstly, points to be tracked are located in the initial frame S0 (see Fig. 5.108a). Their selection can be done manually, by segmentation of point patterns, or by use of an interest operator. Fig. 5.108 then illustrates the subsequent matching process for a single feature found in S0. In the next frame S1, the search is not repeated independently but starts at the location found in S0. The best match between the patterns in S1 and S0 is determined by a similarity measure, such as normalized cross-correlation or least-squares matching (LSM, section 5.4.2.4). The reference window (template) is the image patch detected in the initial frame whilst the search window covers a predefined search area in the next frame (b). In the following step (c), the rectangular region around the image pattern detected in S1 becomes the reference window and the process is repeated in S2 and so on (d). LSM has the advantage that the matching result can be used in the next template through the estimated geometric and radiometric transformation parameters.
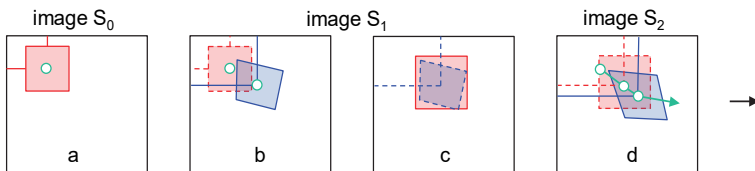
Fig. 5.108: Area-based matching within an image sequence.

With this approach, natural surface points can be tracked through a sequence, e.g. for markerless motion capture or strain analysis of deformed objects (section 6.7.5.3).

5.5.7.2.2 Optical flow

Optical flow is a general term for the apparent motion of imaged objects in a field of view. This motion derives from intensity changes in an image sequence. To a first approximation, it is assumed that the grey values of identical object points do not change in sequential images, i.e. there are no occlusions or illumination changes. The grey or intensity value g can then be expressed as follows:

$$g(x,y,t) = g(x + \Delta x, y + \Delta y, t + \Delta t) \tag{5.131}$$

Applying a Taylor expansion with truncation after the first term:

$$g(x,y,t) = g(x,y,t) + \Delta x \frac{\partial g}{\partial x} + \Delta y \frac{\partial g}{\partial y} + \Delta t \frac{\partial g}{\partial t} \tag{5.132}$$

The intensity changes of an imaged point shifted by Δx and Δy over time interval Δt can therefore be described by the grey-value gradients in the space and time domains. Division by Δt and rearrangement gives:

$$0 = \frac{\Delta x}{\Delta t} \frac{\partial g}{\partial x} + \frac{\Delta y}{\Delta t} \frac{\partial g}{\partial y} + \frac{\partial g}{\partial t} = u \cdot g_x + v \cdot g_y + g_t \tag{5.133}$$

The terms $u = \Delta x/\Delta t$ and $v = \Delta y/\Delta t$ express the velocity of an image point moving through the sequence in the x and y directions. If the grey-value gradients g_x and g_y are zero, no movement exists. Optical flow can therefore only be identified in image regions where gradients occur, i.e. perpendicular to grey value edges (compare with section 5.2.4). Eqn. (5.133) is known as the *optical flow constraint equation*.

In order to generate a homogeneous motion field, i.e. where single image points cannot have arbitrary motion vectors, a constraint equation for regularization is introduced:

$$\left|\nabla u\right|^2 + \left|\nabla v\right|^2 = 0 \tag{5.134}$$

These second derivatives in time (compare with eqn. 5.47) describe the acceleration of image points. As a consequence, the constraint forces the image points to move with constant velocity and in the same direction. An energy function can be derived from equations (5.133) and (5.134):

$$E(u,v) = \sum_x \sum_y [(u \cdot g_x + v \cdot g_y + g_t)^2 + \alpha(\left|\nabla u\right|^2 + \left|\nabla v\right|^2)] \Delta x \Delta y \tag{5.135}$$

Here α is a damping factor which controls the effect of regularization. This energy function is minimized for the whole image, taking into account the neighbourhood around each pixel where the damping term applies.

Fig. 5.109: Image sequence (left) and calculated optical flow (right).

Fig. 5.109 shows a sequence of four images of a rotating rotor blade and the derived optical flow. The motion field is mostly plausible but there are some vectors at stable object points and a few discrepancies at neighbouring points. These occur because of the relatively large time interval between images (low frame rate) and the optical flow approach itself which exclusively measures gradients without tracking discrete features.

5.5.7.3 3D object reconstruction from single-camera image sequences

Where a camera is moving during image acquisition and the exterior orientation of each camera station can be determined, 3D reconstructions can be derived from single-camera image sequences. If the camera is moved parallel to the object, e.g. as in conventional aerial surveying, the base line b between two successive images is determined by the image overlap, and the data can be processed by standard stereo or multi-image analysis (Fig. 5.110a). In contrast, if the camera moves in the direction of its optical axis, as is typical for vehicles with forward looking cameras, corresponding image rays intersect at very narrow angles (Fig. 5.110b). Rectification into epipolar images (normalized stereo images) then requires the method of polar rectification. Here images must be transformed from orthogonal pixel coordinate system into a polar pixel coordinate system.

a) Side-looking image sequence   b) Forward-looking image sequence
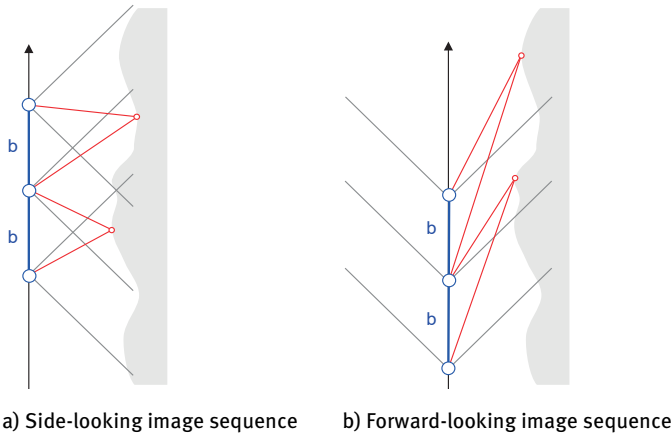
Fig. 5.110: Image matching in a single-camera imaging sequence.

5.5.7.4 Object tracking in multi-camera image sequences

The strategies for processing image sequences from single cameras can also be used for image sequences which use several synchronized cameras. Since multiple cameras are used, stereo and multi-camera methods can also be applied. A further optional restriction is that matching must be consistent from frame to frame in one camera and between cameras. Since matching in single-camera image sequences generates trajectories of 2D image coordinates (static camera, object in motion) or 3D reconstructions (moving camera, static object), multi-camera sequences enable both the measurement of 3D space and of dynamic change.

Fig. 5.111 illustrates the matching scheme for a stereo-image sequence. For an image pair S_k (images S'_k and S''_k) the matching of detected features, optionally with epipolar constraints, results in object points with XYZ coordinates. For the next image pair S_{k+1}, matching can be performed both with image points from the previous image S_k or with new detected features between left and right image. This not only makes matching more robust but points which are missing from one or two images can be reconstructed from other stereo pairs.

At every recording position of a stereo or multi-camera system, a point cloud can be calculated. Using ICP methods (section 6.9.6.2) these single point clouds can be matched in order to derive 6DOF transformation parameters which are equivalent to the exterior orientations of each camera position.

Section 6.7.5.3 describes applications for strain analysis where a key characteristic is the tracking of discrete object points on an object's surface as it deforms.
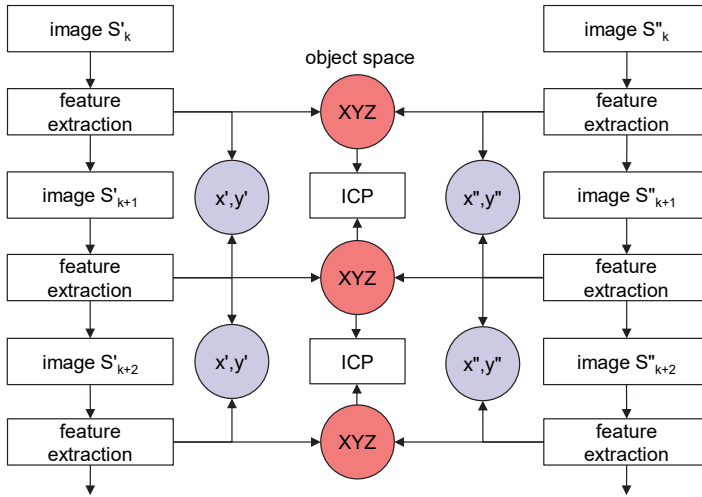
Fig. 5.111: Image matching in a sequence of stereo images.

5.5.7.5 Prediction of subsequent points (Kalman filter)

When object points are lost during tracking, their most probable position in a later frame can be estimated from the preceding trajectory. If the object movement is linear in time and space, the following position is easily extrapolated:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{t_{k+1} - t_k}{t_k - t_{k-1}} (\mathbf{x}_k - \mathbf{x}_{k-1}) \tag{5.136}$$

The point in frame $k+1$ with coordinates $\mathbf{x}_{k+1}$ is derived from the coordinates in preceding frames $\mathbf{x}_{k-1}$ and $\mathbf{x}_k$ and the corresponding time differences between the frames. The coordinate may be two-dimensional (image coordinates) or three-dimensional (object coordinates). It may also be necessary to take into account that linear object movements can result in non-linear movements in the image after central projective transformation.

A more general estimation of movement can be achieved using the Kalman filter. It allows for prediction, filtering, smoothing and correction of the movement based on a prior set of motion parameters. A system of equations can be established to predict the state, in this case point coordinates, at the next instant in time:

$$\mathbf{x}_{k+1} = \mathbf{T} \cdot \hat{\mathbf{x}}_k + \mathbf{C} \cdot \mathbf{w}_k \tag{5.137}$$

T is the transition or dynamic matrix which models the motion of the coordinate vector. In the simplest case, it describes a homogeneous motion at constant velocity (constant velocity model). $\hat{\mathbf{x}}_k$ is the adjusted (corrected) position at time k. **C** is the process noise matrix which models the linear relationship between the state

parameters and the process noise. The process noise vector $\mathbf{w}$ represents the deviations from the assumed motion model. The total system noise is given by $\mathbf{C} \cdot \mathbf{w}_k$.

If measurement noise (observation noise) of the adjusted coordinate vector $\hat{\mathbf{x}}_k$ is given with the covariance matrix $\Sigma_{\hat{x}\hat{x},k}$, then the covariance matrix of the next state can be calculated:

$$\Sigma_{xx,k+1} = \mathbf{T} \cdot \Sigma_{\hat{x}\hat{x},k} \cdot \mathbf{T}^T + \mathbf{C} \cdot \Sigma_{ww} \cdot \mathbf{C}^T \tag{5.138}$$

The *innovation* $\mathbf{d}$ is the difference between measurements $\mathbf{l}$ at state $k+1$ and the predicted value:

$$\mathbf{d}_{k+1} = \mathbf{l}_{k+1} - \mathbf{A} \cdot \mathbf{x}_{k+1} \tag{5.139}$$

where

A: design matrix (see section 2.4.1)

The innovation covariance matrix is given by:

$$\Sigma_{dd,k+1} = \mathbf{A} \cdot \Sigma_{xx,k+1} \cdot \mathbf{A}^T + \Sigma_{ll} \tag{5.140}$$

From this, the gain matrix (Kalman gain) can be derived:

$$\mathbf{K}_{k+1} = \Sigma_{xx,k+1} \cdot \mathbf{A}^T \cdot \Sigma_{dd,k+1}^T \tag{5.141}$$

This is used to smooth and filter the motion, and to calculate the corrected coordinates at state $k+1$:

$$\hat{\mathbf{x}}_{k+1} = \mathbf{x}_{k+1} + \mathbf{K}_{k+1} \cdot \mathbf{d}_{k+1} \tag{5.142}$$

$$\Sigma_{\hat{x}\hat{x},k+1} = \Sigma_{xx,k+1} - \mathbf{K}_{k+1} \cdot \Sigma_{dd,k+1} \cdot \mathbf{K}_{k+1}^T \tag{5.143}$$

The correction to point coordinates is called an *update*.

The Kalman filter method can integrate different sensor data such as images, odometers, GNSS etc. into one common motion model. For example, it can be used to control movements of cars, manoeuvres in flight and robot trajectories. Extensions to Kalman filtering can address issues such as non-linear motion, abrupt disruptions and sensor failure.

Fig. 5.112 shows an example of target tracking which is interrupted by an occluding object. The blue circles show the moving target which is temporarily hidden behind the white object. Using a Kalman filter, the red circles show the predicted position and its uncertainty. The uncertainty increases with increasing distance from the last corrected point. The update generates the points marked in yellow, which brings the tracking back to the actual target.
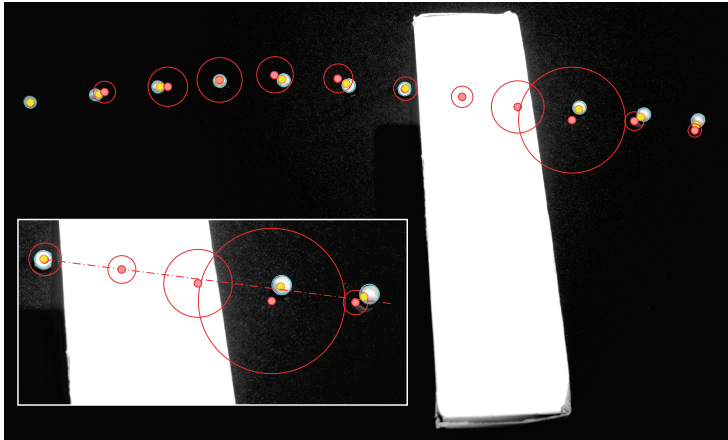
Fig. 5.112: Example of prediction and update by Kalman filtering.

5.5.7.6 Simultaneous Localization and Mapping (SLAM)

A special type of image sequence analysis exists when the camera platform moves along an unknown trajectory within an unknown environment, a task originally developed for mobile robotics. In this case, the exterior orientation of the cameras is not known. If the platform supports two or more synchronized cameras with fixed relative positions, the relative orientation between them can be calculated in advance. For each synchronized set of frames, a set of object coordinates in a local model coordinate system can then be derived according to the process outlined in Fig. 5.111. Depth accuracy will depend on the height-to-base ratio between cameras and object points. In contrast, if only one camera is on board, 3D coordinates can only be calculated using stereo models formed by sequential image frames. In the case of forward-looking cameras, the accuracy in depth (direction of motion) is usually lower since the baseline is small in the X or Y directions (see section 5.5.7.3). For absolute scaling, at least one distance must be known in object space, e.g. from the distance travelled between two frames.

Using the camera motion to derive both the exterior orientations and the object shape in a global coordinate system is known as *Simultaneous Localization and Mapping* (SLAM). Calculating these parameters is essentially also the objective of a sequential bundle adjustment. It should be noted that SLAM is a general technique applicable to different sensor types. When applied to cameras, it may be called *Visual SLAM*. It is also called *visual odometry* when the main emphasis is on the pose estimation of camera or platform, particularly if the trajectory is not closed (loop closure).

A SLAM process can be based on various techniques such as particle filters, graph-based methods and regression-based methods. However, one based on an extended Kalman filter can be structured as follows:

1. Detection and matching of distinctive object features (landmarks):
    The methods described in section 5.5.7.1 lead to position and descriptors of interest points that can be matched across the images.
2. Kalman filter:
    Using the Kalman filter outlined in section 5.5.7.5, the state estimation is predicted by the motion model. The calculation of subsequent landmark positions is done by Kalman update.
3. New landmarks:
    When new landmarks are detected, they are added to the list of detected features and can be used in later epochs until they disappear again. The covariances of calculated 3D points are integrated into the extended Kalman filter.

The error propagation behind the calculation of locations reveals significant drifting, similar to the error effects resulting from an open traverse (polygon) calculation in surveying. If the camera platform passes through a region already mapped, or if it returns to the start position (loop closure), a global bundle adjustment can be calculated which minimizes the discrepancies arising from redundant measurement of object points and the start and end poses of the camera(s). The positioning process can be significantly strengthened by adding additional sensors to the platform. Examples are GNSS navigation sensors, odometers, altimeters, range sensors and laser scanners. A multiplicity of different sensors offers a valuable redundancy of information, particularly when some sensors temporarily do not deliver useful signals.

a) Calculated route   b) 3D point cloud with camera positions   c) Superimposition with aerial image
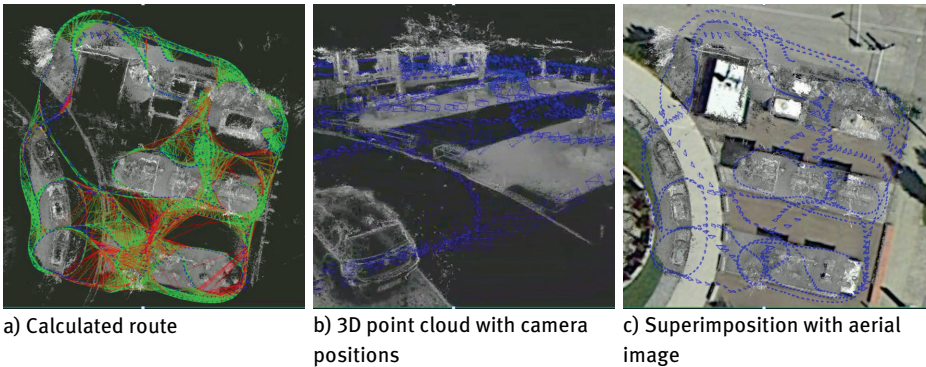
Fig. 5.113: Example of SLAM for environmental mapping with a stereo camera (CVG TU Munich).

Fig. 5.113 is an example of SLAM used for mapping an urban area. A stereo camera is mounted on a car and records stereo-image sequences which use LSD-SLAM (large-scale direct monocular SLAM) to calculate the trajectory (a) and a point cloud of the environment (b).

6 Measuring tasks and systems

6.1 Overview

This chapter presents practical, working systems, often commercially available, which utilize photogrammetric principles to solve a range of measuring tasks. The very wide variety of potential application areas and system solutions in close-range photogrammetry makes it difficult to classify typical system concepts. The chapter starts with an overview about targeting designs and configurations followed by a section on reference systems (scale, control points). Then interactive software programs and tactile probing systems are presented. These are followed by point-by-point and area-based probing systems, including solutions for the measurement of complex free-form surfaces and system for terrestrial laser scanning. The presentation continues with systems for measuring dynamic processes, as well as systems operating from moving platforms such as drones (UAV, RPAS) or robots. The chapter concludes with an overview of visualisation systems (stereo, AR/VR).

Digital processing systems consist of hardware and software components which are subject to rapid technical change. Hence, the solutions and products presented here represent only a snapshot of current technical development. This should, however, also provide clear pointers to the concepts and possibilities of future systems. Note in the following that any stated technical details and measurement accuracies are based on the manufacturer's claims and specifications.

6.2 Targeting

In many applications, locations to be measured on an object need to be marked with artificial targets, e.g.
- to identify natural object feature points which cannot otherwise be identified accurately;
- as uniquely defined points for comparative measurements;
- as control points for geodetic measurement;
- for automatic point identification and measurement;
- for accuracy improvement.

The physical size and type of target to be used depends on the chosen imaging configuration (camera stations and viewing directions, image scale, resolution) and illumination (light source, lighting direction). The manufacture of very small targets and the logistics of very large targets can be prohibitive.

6.2.1 Target designs

6.2.1.1 Circular targets

Because of their radially symmetric form, circular targets are very suitable for representing a target point (the centre) by a surface. Determination of the target centre is invariant to rotation and, over a wide range, also invariant to scale. They are suitable for both manual, interactive image measurement as well as automated digital techniques for point determination.

The centre of circular targets can be found in analogue images by manually centring a measuring mark, which is a circle or dot, over the target image. In a digitized image, the target centre can be computed by a centre-of-gravity calculation, correlation with a reference pattern or numerical calculation of the centre of the target's outline circle or ellipse (section 5.4.2).

In order for accurate measurements to be made, the diameter of the circular target images must be matched to the target detection and measurement process. For digital processing systems, it is generally accepted that target images should be at least 5 pixels in diameter. In addition to practical considerations, such as the maximum target size for attaching to an object, the maximum target diameter is also a function of the maximum allowable eccentricity between the true image position of the circle centre and the centre of the target's elliptical image.
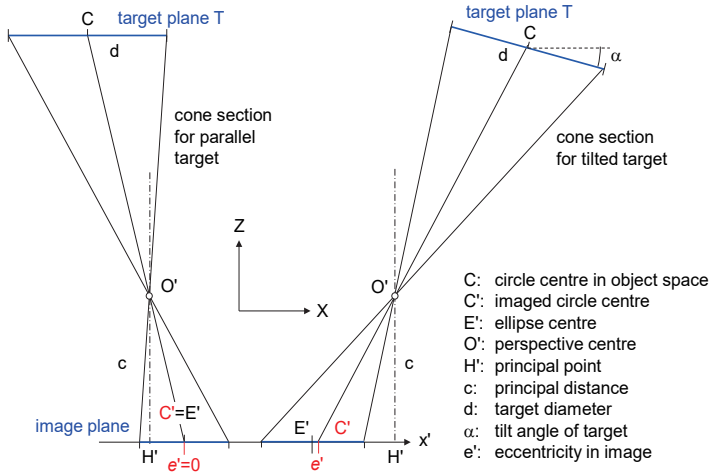
Fig. 6.1: Image eccentricity for a circular target with parallel and inclined target planes.

Fig. 6.1 displays the effect of perspective eccentricity of an imaged circle. The circle centre C is imaged as C' whilst the ellipse centre E' is displaced by the eccentricity e'. Only in the case where circle and image planes are parallel, are both points identical:

C' = E'. The degree of eccentricity depends on the size of the target, viewing direction, lateral offset from the optical axis and image scale. It can be estimated as follows:

$$e' = r_m - \frac{c}{2} \cdot \left(\frac{R_m + \frac{d}{2} \cdot \sin(90 - \alpha)}{Z_m - \frac{d}{2} \cdot \cos(90 - \alpha)} + \frac{R_m - \frac{d}{2} \cdot \sin(90 - \alpha)}{Z_m + \frac{d}{2} \cdot \cos(90 - \alpha)} \right) \tag{6.1}$$

where e':   eccentricity of projection
   d: target diameter in object space
   r_m: radial offset of target in image
   α: viewing direction = angle between image plane and target plane
   R_m: lateral offset of target from optical axis
   Z_m: distance to target
   c: principal distance

Essentially, targets which are smaller and/or placed at greater distances result in eccentricities which are negligible (smaller than 0.3 µm). For greater image scales, larger image formats, bigger targets, strongly convergent camera directions and high accuracy requirements, then eccentricity can be significant and must be taken into account during processing.
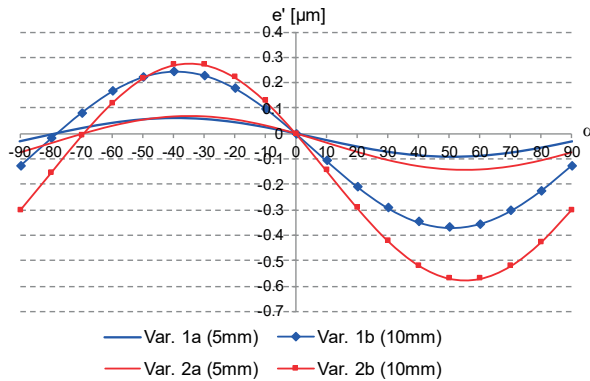
Var. 1: $c = 24$ mm   Var. 2: $c = 8$ mm
       $R_m = 208$ mm $R_m = 187$ mm
       $r_m = 5$ mm $r_m = 3$ mm
       a) target diameter $d = 5$ mm   a) target diameter $d = 5$ mm
       b) target diameter $d = 10$ mm   b) target diameter $d = 10$ mm

Fig. 6.2: Eccentricity of projection as a function of target orientation angle.

The effect of eccentricity is extremely complex for the multi-photo convergent image configurations that characterize most high-accuracy photogrammetric measurement.

It is often assumed that the effect is compensated by the parameters of interior and exterior orientation if they are estimated using self-calibration techniques. However, for high-precision applications (<0.3 µm in image space) it is recommended that small targets are used (see Fig. 6.2 and Example 6.1). In using small targets, the target diameter, or level of ring flash illumination for retro-targets, must be capable of generating target image diameters of the order of 5–10 pixels if demanding measuring accuracies are to be achieved.

Eccentricity does not matter if the target centre is calculated as the centre of a directly defined circle in 3D space, as implemented by the contour algorithm (section 4.4.7.2) or bundle adjustment with geometric 3D elements.

Example 6.1:

Fig. 6.2 illustrates two variant calculations of eccentricity. It shows the resulting eccentricity as a sinusoidal function of the angle α between image plane and target plane. Variants (1) and (2) are each associated with two target diameters (a) 5 mm and (b) 10 mm.

For the first variant, targets are assumed to be imaged at an object distance of 1000 mm by a digital SLR of 24 mm focal length. In object space, the target is laterally offset from the camera axis by 208 mm, resulting in an image position 5 mm from the principal point. For 5 mm targets the maximum eccentricity is always less than 0.1 µm while for 10 mm targets the maximum value is -0.37 µm at α=50°.

For the second variant, targets are assumed to be imaged at an object distance of 500 mm by a typical industrial CCD camera of 8 mm focal length. In object space, the target is laterally offset from the camera axis by 187 mm, resulting in an image position 3 mm from the principal point. For 5 mm targets the maximum eccentricity is always less than 0.15 µm while for 10 mm targets the maximum value is -0.57 µm at α=55°.

6.2.1.2 Spherical targets

Spherical targets have the following properties compared with flat targets:

– spherical targets are always imaged as ellipses with an eccentricity that is radial about the optical axis;
– from Fig. 6.3 it can be seen that the ellipse is the result of a conic projection. Note that the base of the projective conic is smaller than the sphere diameter;
– spherical targets can be viewed consistently over a much wider angular range;
– spheres can be used as touch probes.

However, there are disadvantages with regard to mechanical construction and optical properties:

– the brightness of the image falls off rapidly towards the edges of the sphere;
– the edge of the sphere appears blurred;
– the mounting of the sphere can disrupt the isolation of the target edge in the image;
– sphere diameters for practical use lie between 5 mm and 20 mm;
– accurate retro-reflecting spheres are expensive to manufacture.

The most important advantage of spherical targets is the fact that they can be viewed from almost any direction. Whereas flat, retro-reflective targets can only be viewed over an angle of ±45°, retro-reflecting spheres can be viewed over a range of 240° (Fig. 6.4). They are used, for example, in hand-held probes for photogrammetric navigation systems or as reference points on engineering tools and components (Fig. 6.5).

C: sphere centre in object space
C': imaged sphere centre
E': ellipse centre
O': perspective centre
H': principal point
c: principal distance
R: sphere radius
r: radius of tangential circle C_T
e': eccentricity in image
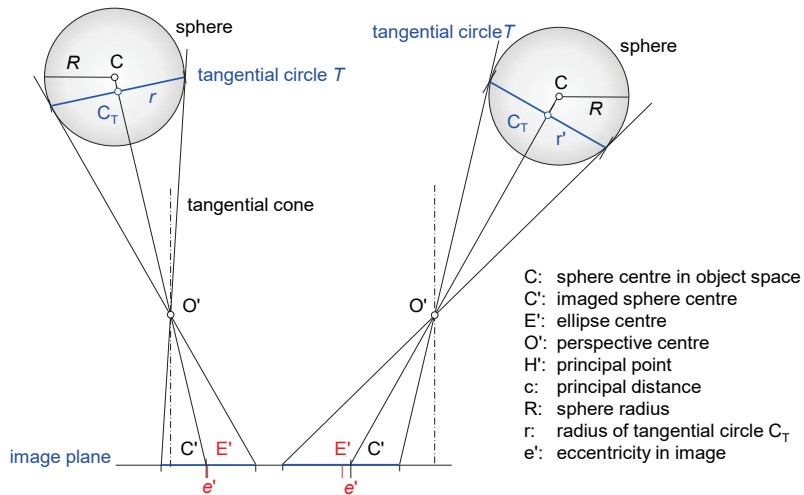
Fig. 6.3: Eccentricity of projection for a spherical target.
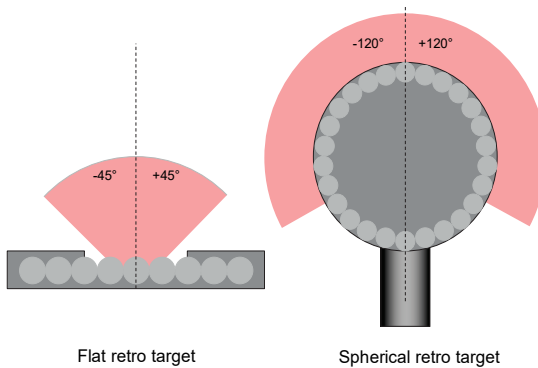
Fig. 6.4: Example viewing angles of flat and spherical retro-targets.

a) Retro-reflecting spheres
(above: IZI, below: Atesos)

b) Hand-held probe for optical navigation
system (AXIOS 3D)

Fig. 6.5: Application of spherical targets.

6.2.1.3 Patterned targets

Some applications make use of targets where the target point is defined by the intersection of two lines. Cross-shaped, checkerboard and sectored targets are examples here. Their advantage is the direct definition of a centre point using well defined edges and their good separation from the background. A disadvantage is the greater effort in locating the centre using digital methods and, in comparison with circular targets, the greater dependency on rotation angle.

Fig. 6.6 shows examples of targets whose centre is defined by the intersection of at least two lines in object space or two curves in image space.
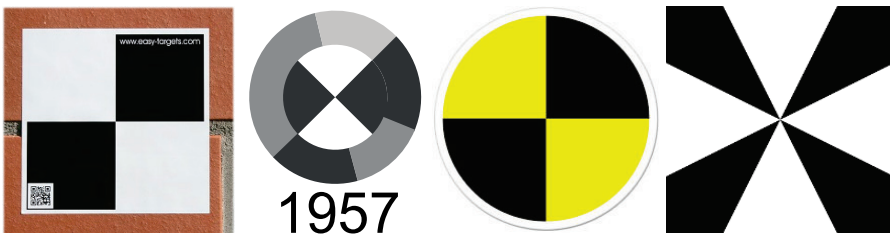
Fig. 6.6: Examples of line-pattern targets.

6.2.1.4 Coded targets

Targets with an additional pattern which encodes an individual point identification number, can be used to automate point identification. The codes, like product bar codes, are arranged in lines, rings or regions around the central target point (Fig. 6.7).

Patterns can be designed which encode many thousands of point-identification numbers.

Coded targets should meet the following requirements:

– invariance with respect to position, rotation and size;
– invariance with respect to perspective or affine distortion;
– robust decoding with error detection (even with partial occlusions);
– precisely defined and identifiable centre;
– sufficient number of different point identification numbers;
– pattern detectable in any image;
– fast processing times for pattern recognition;
– minimum pattern size;
– low production costs.
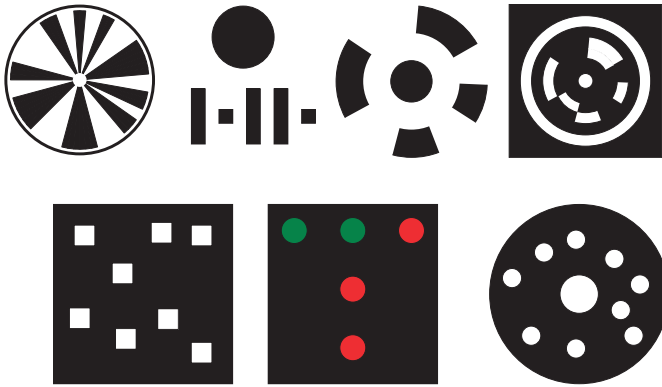
Fig. 6.7: Selection of coded targets;
upper row: barcode patterns, lower row: shape and colour patterns.

The point identification number is decoded by image analysis of the number and arrangement of the elements defining the code. The patterns displayed in Fig. 6.7 (upper row) are based on barcode techniques where the code can be reconstructed from a series of black and white marks (bit series) spread around a circle or along a line. The number of coded characters is limited by the number of bits in the barcode. The patterns displayed in Fig. 6.7 (lower row) consist of clusters of targets grouped in a local coordinate system where the target identification number is a function of the local point distribution. For decoding, image coordinates of all the sub-targets are computed. The planar image pattern can then be perspectively transformed onto reference points in the pattern so that target identification can be deduced from the map of transformed coordinate positions. Alternatively, targets in the group can be given different colours so that the point identification number can be deduced from

the locations of these differently coloured points. The majority of coding schemes are challenged by large volumes and long oblique sight lines.

6.2.2 Target systems

6.2.2.1 Centring targets

Targets that have to be observed from very different viewing directions or have to be suitable for measuring systems with different target requirements need a reproducible mechanical centring. Examples are nests for laser trackers that can accommodate both reflectors and photogrammetric targets (spherical mounted reflector, SMR, Fig. 6.8a, b), adapters with exchangeable targets (Fig. 6.8c) or tilt & turn targets for terrestrial laser scanning (Fig. 6.105).

a) SMR target (API)

b) Photogrammetric target (Hubbs)

c) Rotatable and replaceable target (Meteriss, Rotbucher)
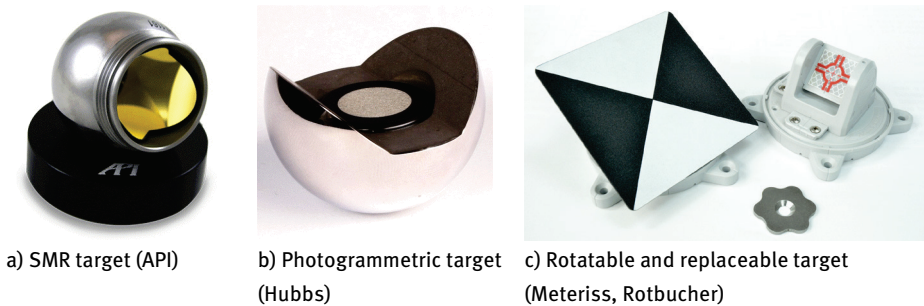
Fig. 6.8: Examples of centring targets.

6.2.2.2 Probes

Targeted probing devices have been developed for measuring "hidden" areas on an object which are not directly visible to the camera system, although it is also convenient to use them for general-purpose measurement. Fig. 6.9 shows probes with an offset measuring point P which is determined indirectly by the 3D coordinates of a number of auxiliary points Q_i attached to the probe in a known local configuration. These points can either be arranged on a line (2–3 points) or spatially distributed (≥3 points) with known 3D reference coordinates in the probe's local coordinate system. The points Q can be either passive retro-reflective targets or active, self-luminous LEDs. Alternatively, point P can be sharp point, regarded as a single touch point, or a ruby sphere such as those found on CMM touch probes. In such a case, a correction must be made in software for the offset caused by the significant radius of the sphere. Hand-held probes are used in many industrial applications where they are tracked by at least two cameras simultaneously, with a new image pair being taken for each point measured by the probe. Such a system is often termed an online system.

The determination of the spatial position and orientation (pose, 6DOF) of a group of locally measured object points (locator) is carried out in single-camera systems via the inverse space resection (section 4.2.5). In multi-camera systems, the locator points are determined by spatial intersection. This is followed by a 3D similarity transformation of the locator points into the object coordinate system.
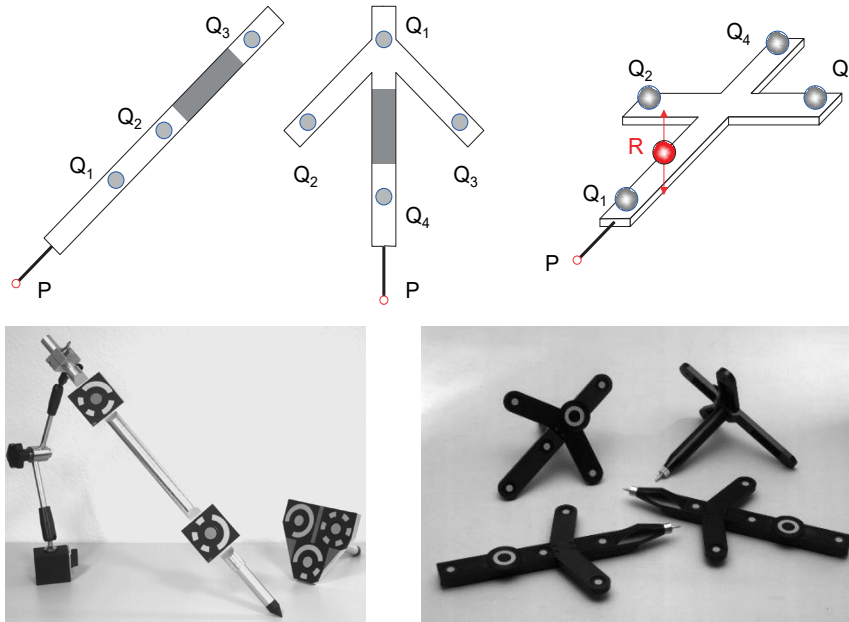
Fig. 6.9: Principle and examples of targeted measuring probes.

If the probe is provided with an additional measuring point R (Fig. 6.9), which can be shifted relative to the reference points Q, then this can act as a switch, for example to trigger a measurement. If R is on a lever which is depressed by the operator, it is shifted slightly away from its nominal position. The photogrammetric tracking system can register this movement and trigger a recording of the measuring point position.

6.2.2.3 Hidden-point devices

In order to measure the centre of drilled holes or edges of components, specially designed target adapters provide a fixed relationship between measurable targets and the actual point or edge to be measured. Manufacturing tolerances in the adapter directly contribute to the overall accuracy of the indirectly measured object point.

For high accuracy applications, individual adapters can be characterized in precise jigs to determine individual variability. This concept allows for a wide range of bespoke adaptors to be manufactured cost effectively.
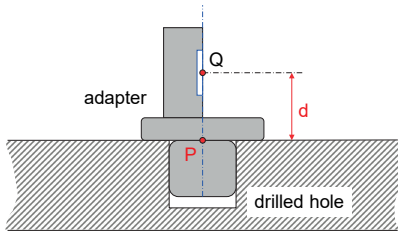
Fig. 6.10: Principle of adapter for drilled hole measurement.

Fig. 6.10 shows the principle of an adapter for indirect measurement of the centre of a drilled hole. The visible target is mounted such that the measured point Q is located on the drilling axis at a known offset d to the hole centre P.
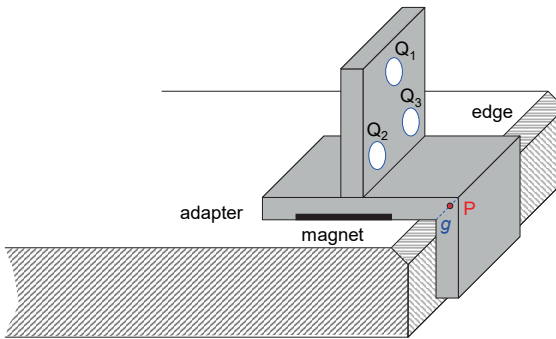
Fig. 6.11: Principle of an edge adapter.

The edge adapter displayed in Fig. 6.11 has three target points Q_i, with a known local relationship to the straight line g which, on attachment of the adapter, represents an edge point P. Such adapters can be equipped with magnets for rapid mounting on metallic surfaces. If the points Q define a coded target, the photogrammetric system can identify the adapter type directly and calculate P automatically. However, multiple-point adapters with coded targets require significantly more surface area, so adapter size increases.
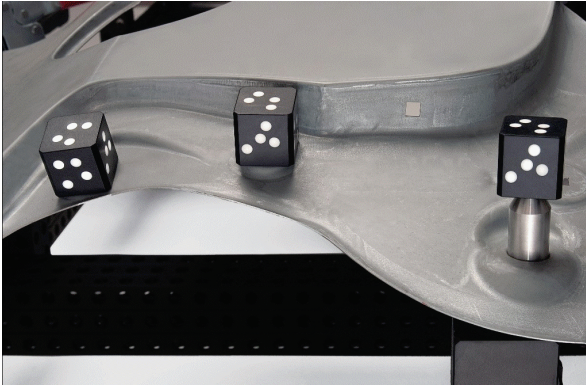
Fig. 6.12: Hexagon adapter cubes.

The adapter cubes developed by Hexagon and Witte Barskamp (Fig. 6.12) have a coded target on each of 5 faces, as well as a measurement point or edge on the sixth side. The points in the coded targets, and the measurement point, are manufactured to have a relative reference coordinate accuracy of around 10 μm. By measuring the points automatically and analysing the spatial distances between them, a particular cube can be identified and its measurement point located in 3D.

6.2.2.4 6DOF targets

While the target types shown above are used exclusively to define the position of a point in space, 6DOF targets can also be used to measure spatial positions and angles.

Fig. 6.13a shows the ArUco target that is widely used in computer vision. 6DOF information is calculated by inverse space resection using one camera and the four corner points of the target. The dimensions of the corner points have to be calibrated in advance. The interior pattern serves for the automatic identification of the target number.

Fig. 6.13b shows a special target from the company Photron, which consists of a point coding, reference marks for the target position, two orientation patterns for the total rotation and two moiré patterns. With the help of the four reference points and the point coding, the 6DOF parameters of the target in the camera coordinate system can be calculated from one camera position by inverse space resection. The rotation κ around the local Z-axis and the XY-coordinates are of higher accuracy than the Z-coordinate and the angles ω and φ around the local X- and Y-axes. The moiré fields are made of lenticular glass and change the position of the central bar depending on the viewing direction of the mark, so that the rotations around X and Y can be inferred with higher accuracy.

The design of the target requires a significantly larger area than conventional targets. The accuracy of the 6DOF parameters depends, among others, on the size of

the reference point field and the imaged area of the points, respectively. Highly accurate measurements cannot be expected when using only one camera, according to the configuration for space resection (section 4.2.3). Point fields (locators) with measured reference points observed from several cameras also provide 6DOF parameters (see section 6.2.2).

a) ArUco marker   b) Photron marker
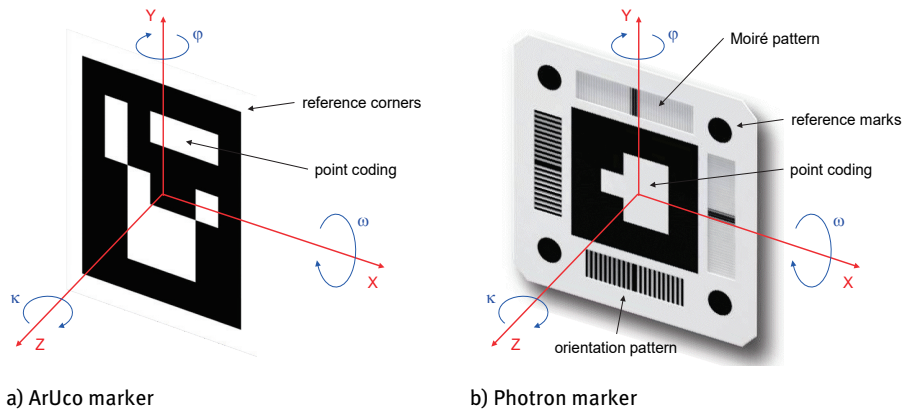
Fig. 6.13: 6DOF targets.

6.3 Realization of reference systems

6.3.1 References

6.3.1.1 Definition of scale

Since photogrammetry is a triangulation method where image points represent spatial directions, at least one reference length for absolute scaling is necessary. Scale can be defined by given distances between distinct object points, scale bars, calibrated target systems or reference points with known coordinates. Alternatively, calibrated multi-camera systems with given relative and absolute orientation between the cameras provide absolute base lines which define scale in object space.

Fig. 7.16 shows examples of photogrammetric scale bars consisting of coded or uncoded targets. Scale bars must be calibrated with an accuracy that fulfils the overall accuracy demand of the application. Since scale bars usually represent the standard Metre unit, traceability to the national standard should be guaranteed. See section 7.2.3.3 for detailed information about reference artefacts and traceability.

Physical scale definitions are sensitive to environmental conditions. Besides mechanical stress like deformations or bending of the scale bar material, thermal expansion is the most critical issue that effects scale. The resulting length difference is given by

$$\Delta L = \alpha \cdot L_0 \cdot \Delta T \tag{6.2}$$

with

α   thermal expansion coefficient

L_0 reference length (usually at 20° C)

ΔT temperature difference

Table 6.1 shows different materials used for photogrammetric scale bars and their impact on scale at a distance of 1m for 1° C and 5° C temperature changes.

Table 6.1: Thermal expansion of different materials for $L_0 = 1$ m.

Material	Thermal expansion coefficient [10^{-6} K^{-1}]	Effect Δ*L* for 1° [mm]	Effect Δ*L* for 5° [mm]
Aluminium	23	0.0230	0.1150
Steel	11.5	0.0105	0.0525
Polyethylene	200	0.2000	1.0000
Acryl	90	0.0900	0.4500
Glass	8.5	0.0085	0.0425
Carbon fibre	-0.1	-0.0001	-0.0005
Invar	1.5	0.0015	0.0075

If temperature difference and expansion coefficient are known, eqn. (6.2) can be used to correct the length of a reference scale bar. If the material of the scale bar and actual object are the same, a length compensation might be negligible. In addition, note that especially for carbon fibre composite material, a change in relative humidity can cause significant length changes.

The influence of a change in reference length on a photogrammetric measurement result is initially linear, i.e. all coordinates are scaled in the same way. This assumes that the measurement volume and in particular the measurement object are isotropic, i.e. the object does not consist of different materials with different coefficients of expansion. Since this is rarely the case in practice, the influence of a change in length can only be predicted to a limited extent. The standard deviation of a photogrammetrically measured point expands due to an uncertain reference length:

$$s_{XYZ} = \sqrt{s_P^2 + \left(\frac{D}{L_0}\right)^2 s_L^2} \tag{6.3}$$

where

s_P measurement uncertainty of a 3D point (without scaling error)
D principal diagonal of measurement volume
L_0 reference length
s_L uncertainty of reference length

This means that the ratio D/L_0 controls the influence of an uncertain (changed) reference length, i.e. with large measuring volumes and short scale bars the influence of an incorrect scale can be serious. If reference length and diagonal are equal, the uncertainty of a measuring point increases only by the value of the change in length. From eqn. (6.2) it follows that the reference length should be as large as possible, but in practice this quickly comes up against technical or financial limits.

Usually (certified) scale bars are calibrated by coordinate measurement machines or interferometric measurements, optimally by a certified calibration service. Scale bars can also be calibrated photogrammetrically together with a calibrated reference scale. However, it must be noted that the accuracy of derived lengths is less than the reference accuracy. The number of required scale bars and their position in object space are discussed in section 7.1.5.2.

6.3.1.2 Definition of reference coordinate systems

In many applications, a workpiece or object coordinate system is required in which the final 3D information from photogrammetry is registered. A common way to achieve this is to incorporate reference points (control points) measured by a system with equal or better accuracy than the photogrammetric survey requires. They then serve, for example, as ground control points for indirect georeferencing in the bundle adjustment (section 6.9.1), for the orientation of camera-based scanning systems (example in section 6.9.2.2) or for the registration of measured which have point clouds (section 6.9.6). If local or global 3D reference points are available, scale information is inherently given. Where reference points are used, they must be visible, and individually identifiable, in the images.

Fig. 4.13 and Fig. 6.23 show examples of a reference tool that incorporates pre-calibrated coded targets. During the photogrammetric process these targets are measured automatically, hence providing a local coordinate system. However, if object coordinates are required in a well-defined object coordinate system, a subsequent 3D transformation into the target system is necessary. The accuracy and distribution of reference points can be critical. As a rule of thumb, 3D reference coordinates should never be less accurate than the final required accuracy of the photogrammetric survey. Since photogrammetry is often the method of highest accuracy available, the generation of adequate control points can be difficult, time consuming or expensive. It must be noted that the spatial distribution of control points should represent the whole measurement volume in order to avoid any extrapolation effects. It should also be noted that control points should be integrated

into the photogrammetric calculations (bundle adjustment) with their realistic a priori accuracy or statistical weight. Finally, the system for measuring control points should identify the same physical object point as it is done by photogrammetry. That means, that the physical and the optical centre of a reference target must be identical. One option is offered by high-quality adapters (nests) which, for example, can support either a photogrammetric target or a retro-reflector for laser trackers.

6.3.2 Measurement systems

Diverse geodetic and industrial systems are typically used to measure reference points. These include total stations, multistations and laser trackers, and their use will depend on the required accuracy and their availability. Note that GNSS methods are particularly used in UAV photogrammetry (section 6.11.2), but they are not discussed further here.

6.3.2.1 Video total stations and multistations

Video total stations (video tacheometers) are based on conventional electronic total stations with a digital video camera integrated into the optical path. They can be used either with prism reflectors or reflectorless. The instruments incorporate an absolute distance meter with accuracies up to 1mm±1ppm and are nowadays available with motorized rotations. While electronic angular read-out is used for acquiring horizontal and vertical directions, the video image can be digitized and processed to enable automatic target detection and point measurement.

Fig. 6.14: Total station with integrated camera (Topcon GPT-7000i).

Fig. 6.15: Multistation (Leica Nova MS50).

If object targets are used, and detected by automatic, high-precision image processing (see section 6.2.1), angular accuracies of 0.15 mgon (5 arc sec) can be achieved. As an illustration, Fig. 6.14 shows the Topcon GPT-7000i total station. This uses a colour camera with 640 x 480 pixel and optics that can be operated in zoom or wide-angle mode.

Multistations extend the functionality of a video total station by adding a laser scanning unit and an optional GNSS/GPS receiver. Tasks like control-point measurement and laser scanning can be combined on site, and incorporated within a global coordinate system by GNSS. Fig. 6.15 illustrates this with the Leica Nova MS50 multistation. This incorporates a wide-angle camera and a co-axial camera both with a 5 Mpixel sensor whose interior and exterior orientations are calibrated to enable photogrammetric processing.

6.3.2.2 Laser trackers

Laser trackers use a dynamically directed laser beam to track and measure the 3D coordinates of the centre of a retro-reflecting corner cube (see Fig. 3.149). This is typically mounted in a spherical housing (spherically mounted retro-reflector - SMR) and used as a manual touch probe to measure object points and surfaces (Fig. 6.16). Angle encoders on the two deflecting axes provide horizontal and vertical angle values, and range to the reflector is determined either by interferometry or, now more commonly, by an absolute optical range-measuring technique. Servo drives on the rotation axes, and return-beam sensing, keep the beam on track as the reflector is moved, enabling a continuous, dynamic path measurement. Laser trackers are used for the measurement of individual object points, scanning large free-form surfaces, as large scale deflection gauges, and for the dynamic tracking of moving objects.
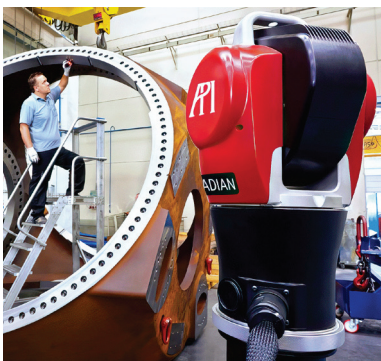
Fig. 6.16: Object measurement by laser tracking (API).

Fig. 6.17: Absolute Tracker Leica AT 960.

Fig. 6.18: Laser tracker Faro Vantage S6.

High-precision angle and range measurement results in measured 3D coordinate accuracies of better than 0.1 mm, at object distances of up to 40 m. Target holders, also called nests, can be used as fixed mounting locations for the SMRs. These allow the reflectors to be rotated towards tracker stations with minimal positional error. In addition, exchangeable photogrammetric targets can also be used in the same nests, re-locating to the reflector position with an uncertainty of about 5–10 μm.

There are currently three manufacturers of 3D laser trackers: Leica Geosystems (part of Hexagon MI), Faro Technologies Inc., and Automated Precision Inc. (API). Fig. 6.17 and Fig. 6.18 show examples of these systems.

Fig. 6.19: Left: Laser tracker combined with a digital camera for measurement of position and orientation of a hand-held probe (Leica T-Probe); right: hand-held probe with internal orientation sensors (API vProbe).

All three manufacturers also offer an enhanced system which tracks a target probe in all six degrees of freedom (6DOF). The Leica system is illustrated in (Fig. 6.19). Here the probing device (T-Probe) has an embedded retro-reflector which is tracked in 3D by the laser beam as normal. LED targets, known with the reflector in a local probe coordinate system, surround the reflector and are imaged by a digital zoom camera (T-Cam) mounted above the tracker head or integrated within it (depending on model). By means of a standard photogrammetric space resection (section 4.2.3) the angular orientation of the probe can be calculated. With known 3D location and angular orientation, an offset touch point, used for measuring object points, can be located in 3D. In addition to a hand-held probe there is also a hand-held surface scanner (T-Scan, see also Fig. 6.47) and a special probe for robot and machine control (T-Mac).

API's vProbe uses internal sensors to determine the probe's pose. These are assumed to be tilt sensing for roll angle and a pinhole camera view of the laser beam behind the retro-reflecting prism for sensing pitch and yaw. The Faro probe is a relatively new development and its operational concept is not yet known, but appears to have some operational similarities with the Leica probe.

6.4 Interactive multi-image processing systems

6.4.1 Programs with CAD functionality

Interactive, multi-image processing systems permit the measurement of image points in more than two images. Here the operator checks point identification and correspondence. As well as pure coordinate measurement, these systems are particularly used for the reconstruction of graphical features, such as lines, or the production of virtual 3D models. The data acquired are usually transferred to a CAD program for further processing.

Typical characteristics of interactive multi-image processing systems are:

– Project management:
All transformation parameters (interior and exterior orientations, datum definitions) and measurement data are stored, and are therefore still available when the project is next loaded.

– Image processing:
The contrast, sharpness and brightness of the images can be enhanced[1] for ease of use. Image rectification or orthophoto production can be integrated into the processing system.

– Automation:
Functions for the detection and measurement of point features not only allow automated interior orientation but they also expand interactive systems to photogrammetric online or offline systems.

– Superimposition of graphical information:
Digital images and graphical information (vector data) can easily be superimposed, so permitting improved interactive control of the measurement.

– Integration into CAD environment:
Display and measurement of images and 3D CAD processing of the resulting object data can take place within one closed system.

The PHIDIAS (Phocad) interactive photogrammetric system (Fig. 6.20) demonstrates the integration of digital photogrammetry with the 3D CAD environment. The system is embedded in the Microstation CAD program (Bentley) which therefore makes available all the standard CAD functions for the photogrammetric reconstruction of the object. The system is well suited to the complex graphical reconstruction of objects such as occur in architectural applications or in the as-built documentation of industrial plants. Integrated bundle adjustment with self-calibration, as well as functions such as mono and stereo plotting, are included within the scope of the system.

---

1 Manipulating images can alter the imaging geometry.
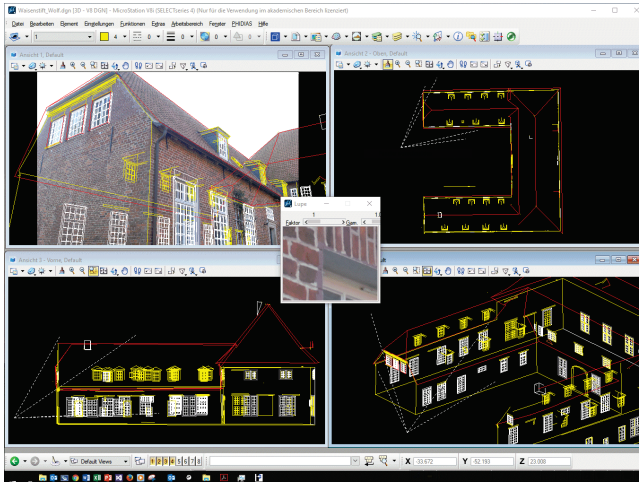
Fig. 6.20: PHIDIAS-MS interactive multi-image system (Phocad).
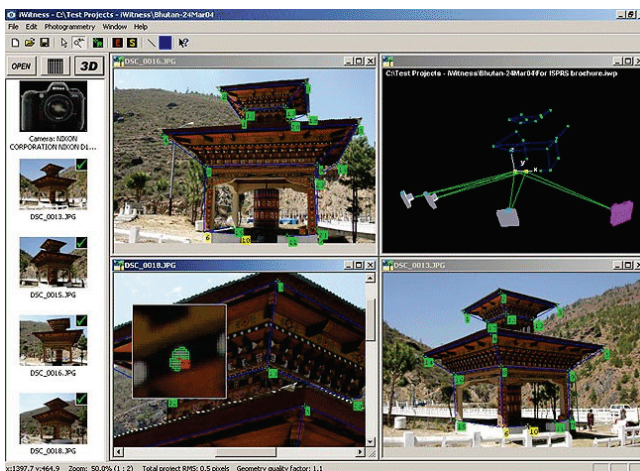
Fig. 6.21: iWitness interactive multi-image system (Photometrix).

Low-cost systems, with a basic range of construction tools for the simple generation of 3D models, have a much more restricted scope and performance. Examples of such systems, which have become widespread in recent years, include Photomodeler (EOS Systems) and iWitness (Photometrix). Here the level of user guidance and automation in the orientation process ensure that even non-specialist users can deliver usable results after a short learning period.

Although the iWitness system (Fig. 6.21) is mainly designed for accident and crime-scene recording (see section 8.8.1), it can essentially be used to record any

scene of choice. It supports a database of commercially available, off-the-shelf digital cameras which is constantly updated. Notable features are the integrated camera calibration which makes use of colour-coded targets (see section 6.2.1.4), as well as robust orientation algorithms which can handle the often unfavourable imaging configurations typical in forensic measurement (see Fig. 8.88).

6.4.2 Structure-from-motion programs

Various software packages are based on the principle of structure-from-motion (section 5.5.2.2) and all provide similar functions:
– processing of arbitrary image configurations, acquired using one or more (uncalibrated) cameras;
– simultaneous calibration of the imaging cameras;
– tie point measurement by interest operators and feature descriptors (section 5.4.3), often based on the SIFT operator;
– automatic matching of image features (tie points) as input for a dense point cloud;
– inclusion of control points and scale information by manual or semi-manual point measurements;
– robust pre-orientation and subsequent bundle adjustment;
– post-processing to generate dense point clouds;
– filtering and segmentation of point clouds;
– triangular meshing of the point cloud to provide a surface model;
– texture mapping of the surface model;
– calculation of orthophotos and image mosaics.

Measurement and matching of image features (tie points) requires objects with a reasonable degree of texture. Interactive control and manipulation by the user are often highly restricted since fully automated processing by non-skilled users is the approach normally favoured. The particular camera calibration model used for interior orientation is not always documented. These programs may well generate results even if camera construction is not stable or imaging configurations are weak. As a result, this can lead to deformations in the object space which are difficult to detect at first glance. Assistance via output logs and analysis functions for quality assessment are often reduced to a minimum.

Typical software packages with the functionality listed above include: Agisoft Metashape (Fig. 6.22), RealityCapture, PIX4Dmapper, ContextCapture, MicMac, COLMAP and Meshroom.
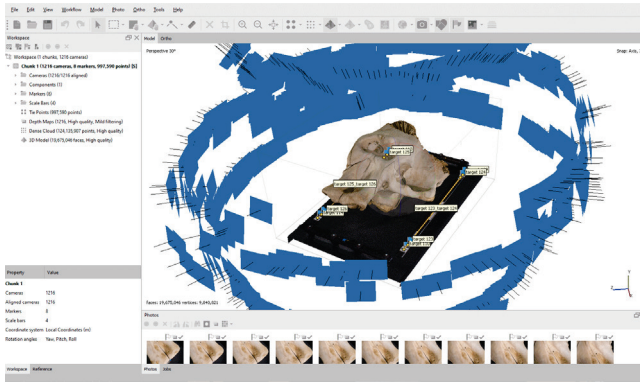
Fig. 6.22: Structure-from-motion software as represented by Metashape (Agisoft).

6.4.3 Offline processing systems for industrial applications

Software packages are available for industrial applications in optical metrology which accept multi-image configurations using both simple and coded targets. These normally provide the following functionality:

– fully-automatic measurement of target points to sub-pixel accuracy (0.02 to 0.05 pixel);
– fully-automatic identification of measured points identified by coded targets;
– bundle adjustment with self-calibration;
– robust detection and elimination of outliers;
– tools for deformation analysis;
– graphical output of measurement results;
– data interfaces to CAD and industrial metrology software.
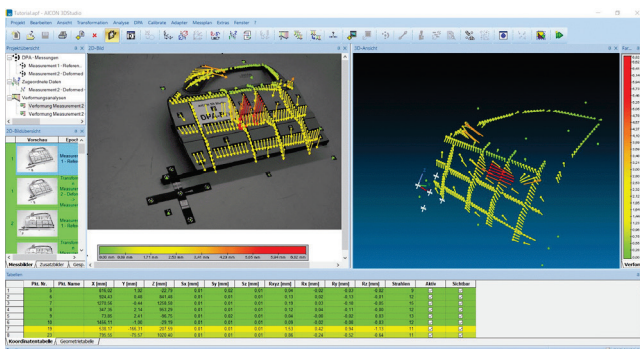
Fig. 6.23: User-interface of AICON 3D Studio (Hexagon).

In most cases these packages are used for the calculation of the 3D coordinates of targets and for camera calibration. They are often associated with offline photogrammetry (see section 6.6.1.1). They are also used to calibrate and orientate online photogrammetry systems.

Commercial industrial processing systems, in different configurations, are available, for example, from Hexagon (3D Studio, originally from AICON Fig. 6.23), ZEISS IQS (TriTop), GSI (VSTARS), Photometrix (Australis) and VMS.

6.4.4 Educational software

Software solutions are available for educational purposes with emphasis on procedure testing, parameter variation, simulations and detailed analysis of results rather than automation of image measurements or 3D reconstructions. Some examples are given below.

As part of an ISPRS initiative, the GRAPHOS software package (Fig. 6.24) mainly adopts structure-from-motion approaches (section 5.5.2.2) and is a joint development by several international universities. It implements a range of algorithms for interest operators, feature descriptors and feature matching and includes a bundle adjustment program for calculation of image orientations and 3D points. In addition, dense point clouds can be generated. Individual process steps can be parametrized, and results can be analysed graphically or numerically.
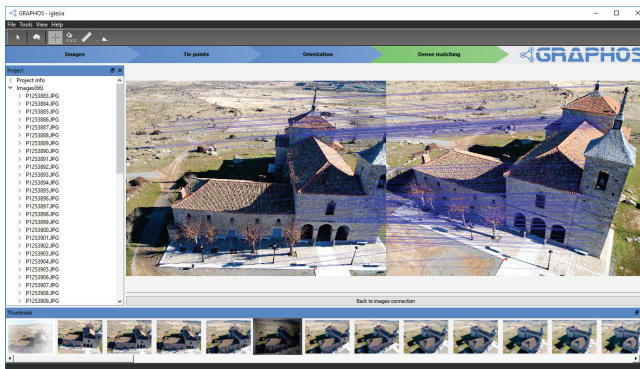
Fig. 6.24: GRAPHOS user interface (ISPRS).

The PhoX program (IAPG Oldenburg, Fig. 6.25) is mainly directed at the analysis of images and their calibration and orientation parameters. Different methods for image measurement are provided. Functions such as space resection and DLT, intersection, monoplotting or 2D and 3D coordinate transformations allow for a number of 3D

calculations. Diverse image processing methods, as well as methods for rectification and orthophotos, are included. PhoX allows for the generation of synthetic images as well as different Monte-Carlo simulations. Different exercises allow for detailed testing of fundamental photogrammetric tasks in order to generate a deeper knowledge of complex methods.
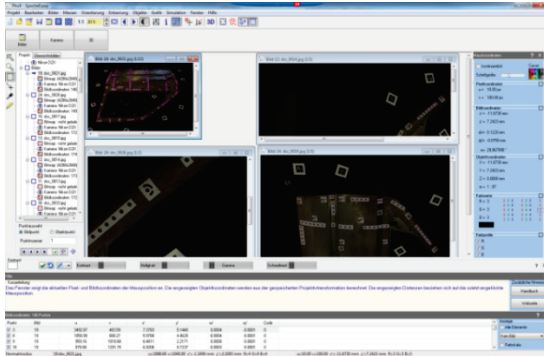
Fig. 6.25: PhoX user interface.

The VRscan3D programme (Fig. 6.26) is a simulator for terrestrial laser scanning within a programmable game environment (Unreal Engine). Users can interactively move in 3D scenes, select scanner and target positions and scan the virtual object. By intersecting the moving laser beam with the object scene, point clouds are created that correspond to the specifications of real scanners. In addition, the programme generates a panoramic image for each point of view. Other tools for generating synthetic point clouds from virtual scanners include Blensor, Blainder or Helios.
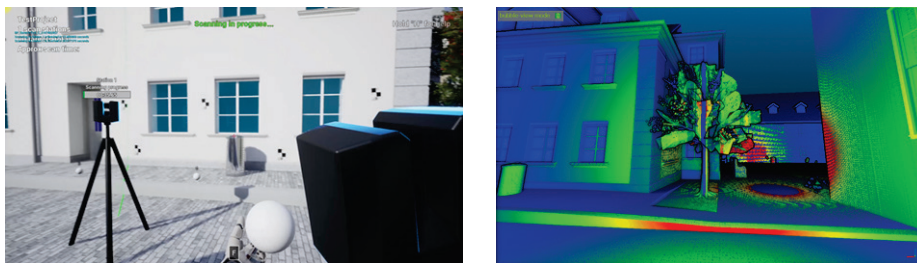
Fig. 6.26: Left: virtual scanning in VRscan3D; right: resulting point cloud (IAPG).

6.5 Tactile probing systems

6.5.1 Measurement principle

This section presents optical 3D systems where the object is measured by mechanical probes or adapters (examples in Fig. 6.27 and see also section 6.2.2.2). The spatial orientation (pose) of the probe within the measurement space is determined by photogrammetric means every time a measurement is triggered. The probe (locator) is provided with at least three reference targets whose 3D coordinates, together with the 3D coordinates of the touch point, are known in a local probe coordinate system by prior calibration. If there are more than three reference targets visible on the probe, over-determined measurements are possible which contribute to an increase of accuracy and reliability in the measurement. The image-based measurement of the probe can be achieved using a single camera (section 6.5.2) or a multi-camera system (section 6.5.3). The spatial position of the probing tip, which is not directly targeted, can be transformed subsequently from the local probe system into the object coordinate system (Fig. 6.33). The actual contacting element on the probe is normally a ball, often a ruby sphere. A correction is normally made for the radius of the ball in order to determine the corresponding point on the object itself.

a) Probe for single camera system (Metronor)   b) Probe for multi-camera system (GSI)
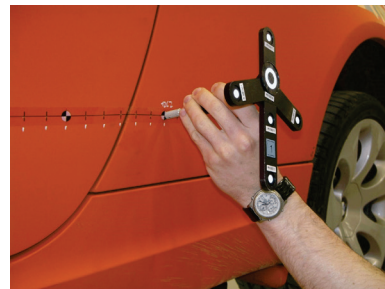
Fig. 6.27: Photogrammetric probes.

Depending on probe type, measurement recording is made:
–   by depressing a button, or activating a switch on the probe, which remotely triggers image recording and measurement;
–   by the use of a touch-trigger probe, commonly seen on coordinate measuring machines, which remotely triggers an image recording and measurement when the probe tip is touched lightly against the object surface;
–   by mechanical displacement of one of the probe's reference points, which is then detected by the cameras and used to record a measurement (see section 6.2.2.2);
–   by operation of an external switch, mouse click, voice or keyboard command.

A computer usually controls the camera, illumination (flash) and image transfer and also implements image processing and coordinate calculations. The computer can be located outside (laptop, host) or inside the camera. In the latter case, image processing (point measurement) is done within the camera and only a few image coordinates are transferred to the host (example in Fig. 3.121b).

6.5.2 Single-camera systems

6.5.2.1 Camera with hand-held probe

Based on the principle of (inverse) space resection (single image, 6DOF calculation, see section 4.2.5), a single camera can determine the current spatial position of a measurement probe in the camera's coordinate system or another reference system (Fig. 6.28). The reference targets can be distributed spatially, in a plane or simply arranged on a straight line. Depending on the chosen arrangement, the number of degrees of freedom which can be calculated drops from 6 (spatial) to 5 (straight line). The 3D coordinates of the touch point are generated according to eqn. (4.43).
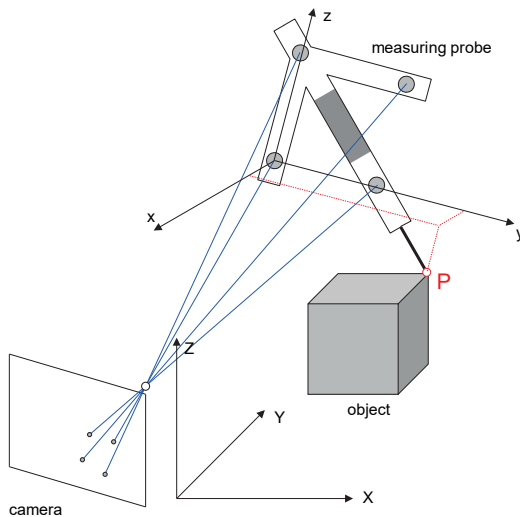
Fig. 6.28: Principle of a single-camera system with a touch probe for object measurement.

Fig. 6.29: Example of a single-camera system with 3D touch probe (Metronor).

It is not possible with a single-camera system to identify the probe targets on the basis of their measured 3D coordinates, as is possible with stereo or multi-camera systems (section 6.6.1.2). Point identification must therefore be achieved by using coded

targets (section 6.2.1.4) or unique target separations, e.g. enabling the use of cross ratios in the image (see section 2.2.1.6).

Fig. 6.29 shows the SOLO system (Metronor). It consists of a factory-calibrated camera and a hand-held touch probe with 5 LED targets arranged on a plane. The operator records a measurement by depressing a switch on the probe. The accuracy of a point-pair separation (distance measurement) in a measuring volume of 1.5 x 1.5 x 1.5 m^3 is 0.12 mm (2-sigma).

6.5.2.2 Probing system with integrated camera

Self-locating cameras automatically determine their own orientation with respect to a fixed or moving set of reference targets.

The Humanetics ProCam Crash system is a touch probe which incorporates a self-locating camera. The probe consists of a digital camera, a recording button and a bayonet connector for attaching a range of exchangeable probe tips which are calibrated with respect to each other in a local xyz system (Fig. 6.31, Fig. 6.30). A system of reference points with known positions in a superior XYZ coordinate system is established on the borders of the measuring volume. If the measuring probe is positioned in such a way that the camera images a minimum of three reference points, the exterior orientation of the probe can be calculated by space resection. Coded targets provide automatic identification of the reference points.

Fig. 6.30: Probe with integrated camera (Humanetics ProCam Crash).
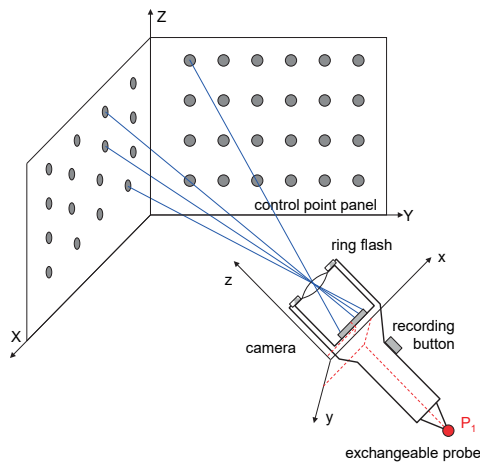
Fig. 6.31: Principle of a probing device with integrated camera.

The measuring probe is connected to the control computer and a synchronized image is acquired when the operator records a measurement. The images of the reference points are detected and measured automatically. A measuring accuracy of about

0.1 mm can be achieved in a variable measuring volume that is limited only by the dimensions and accuracy of the reference point field. The reference field can be mounted in a fixed position, or it can be mobile. By adapting the size of the reference points, the distance to the reference field can be freely chosen.

Fig. 6.32: Mobile probing in measurement cabin (Humanetics).

Compared to the usual online systems, the principle of the self-locating measuring probe has the advantage that the measuring accuracy is independent of the intersection angles of image rays. In addition, any number of object points can be measured without re-arranging the camera configuration, provided a sufficient number of reference points are visible. For large objects to be measured from all sides, special measurement cabins can be constructed which contain a large number of reference points (Fig. 6.32). Prior to online use, these reference points are coordinated by an offline photogrammetric system.

6.5.3 Stereo and multi-camera systems

Multi-camera systems with tactile probing use the method of intersection to calculate the 3D coordinates of the probe's reference points. The orientation and calibration of the cameras must either remain constant or be continually adjusted by observing a reference point field fixed relative to the object.

Fig. 6.33 shows the principle of a dual-camera system. Firstly, the targeted reference points on the probe are calculated by intersection in the camera's coordinate system $X_M Y_M Z_M$. The coordinates of the probing tip P, given in locator system xyz, are then transformed into the coordinate system $X_M Y_M Z_M$ by 3D similarity transformation. If the final coordinates are required with respect to a workpiece coordinate system, an additional 3D transformation into the coordinate system XYZ

is required. The workpiece or object coordinate system can be established, for example, by the 3-2-1 method (see section 4.4.3.2).
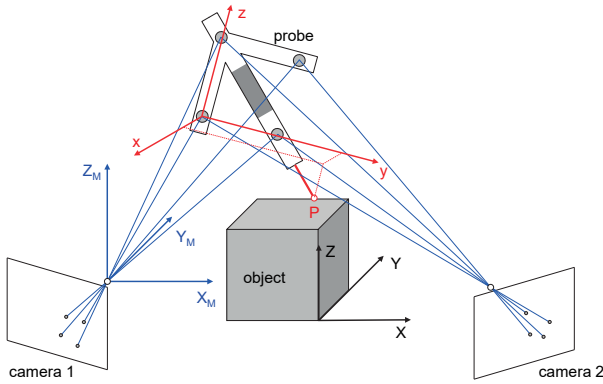
Fig. 6.33: Measuring principle of a multi-camera system with tactile probing.

Fig. 6.34 is an example illustration of the set-up of a dual-camera online system with tactile probing. If the cameras can be positioned individually, there is flexibility in adapting the imaging configuration to the measurement application. However, the cameras must then be oriented on site. If local object reference points are available (e.g. targets around door in Fig. 6.34) cameras can be re-oriented continuously to allow for the fact that relative movements between cameras and object cannot be avoided. For systems with cameras fixed in a housing (examples in Fig. 3.127 or Fig. 6.45), the baseline and the measuring volume are pre-defined and re-orientation is unnecessary. Measurement can therefore start immediately.

Fig. 6.34: Online metrology system with manual tactile probing (GSI).

The accuracy can be estimated according to eqn. (3.49) for multi-camera systems and eqn. (4.88ff) for stereo systems. The calibration uncertainty of the reference points on the probe can be propagated through to an uncertainty at the probing tip. In addition,

manual probing itself creates a certain measurement error, e.g. due to the manual handling. The typical measurement accuracy of tactile probing systems is about 0.2mm at a distance of up to 3m.

The available commercial systems differ mainly in terms of the operating procedure and the strategies for orientation and error detection. Mobile online systems for point-by-point measurement are available, for example, from ZEISS IQS (PONTOS), Metronor (Duo), Hexagon (MoveInspect XR8), GSI (VSTARS-M), AXIOS 3D (CamBar).

6.6 Industrial measuring systems for single point features

6.6.1 Mobile industrial point measuring-systems

6.6.1.1 Offline photogrammetric systems
An offline photogrammetric system has these characteristics (see section 3.2.1.1):
– photography of the object with at least two images from one or more cameras;
– subsequent orientation of the set of images, with simultaneous camera calibration and 3D point determination by bundle triangulation.

In principle, the above two steps are separated in time and, possibly, location (offline, see Fig. 3.37) even if some systems allow for the direct transfer of images for first calculations in parallel. As in much of photogrammetry, there are no restrictions in terms of imaging sensors, object targeting and image configuration (number and position of images).

Offline photogrammetry is very common in industrial applications where the use of digital cameras and simple or coded retro-reflective targets is an integral part of the process. As a result, automatic orientation and image measurement are possible.

Object measurement with industrial offline systems has the following characteristics:
– sequential image acquisition with high-resolution digital cameras with internal image storage and, possibly also, processing within the camera; optional image transfer via WiFi and ring flash;
– object marking with circular retro-reflective targets;
– coded targets for automated generation of approximate values and image orientation;
– calibrated reference tools which establish a local 3D object coordinate system;
– bundle triangulation with self-calibration;
– digital point measurement with sub-pixel accuracy (0.02–0.05 pixel);
– typical number of images between 10 and several hundred (no limit in principle, number also depends on the requirements for camera calibration);

– typical number of images per object point between 6 and 20 (no limit in principle);
– typical object dimensions between 1 m and 15 m (no limit in principle);
– typical duration for object recording and processing between about 10 min and 60 min;
– achievable relative accuracy about 1:100 000 to 1:250 000 (RMS 1-sigma) or 1:50 000 to 1:100 000 (length measurement error, see section 7.2.3).

Fig. 6.35: Object recording using a photogrammetric offline system (GSI)
(lower left: detail enlargement showing coded targets).

Fig. 6.35 shows the recording of an object which is fitted with a number of simple targets (tie points) and coded targets. The use of coded targets ensures that all target types can be reliably identified and correlated across all images and this, in turn, enables fully automatic image orientation (compare with section 4.4.4.2). Suitable software packages for the related analysis are presented in section 6.4.3.

Fig. 6.36 shows the MaxShot 3D camera system from Creaform which uses projected laser lines to show directly on the object the area being imaged (field-of-view). The camera and processing computer are connected by cable so that the current image is immediately oriented and image points can be continuously measured. If the image has an insufficient overlap with the previous image, the projected frame appears in red, otherwise in green. A 3D point cloud of targets is thereby calculated sequentially online in a process which ensures there is a sufficient number of images and suitable intersection geometry. After object recording is complete, a final bundle adjustment is calculated using all images.

The achievable accuracy in object space is strongly dependent on the imaging configuration and on the stability of the camera during image acquisition. To achieve the highest accuracies, the configuration should be chosen so that the camera can be

simultaneously calibrated during measurement (see section 7.3.1). Object points of interest should have good and multiple ray intersection angles and be uniformly distributed over the image formats. The points themselves should mostly be identified by simple artificial targets (signalized), or coded targets, and features such as holes and edges identified by targeted mechanical adapters (see section 6.2.2.3).

Fig. 6.36: Object measurement using an offline photogrammetry system with active projection of the field-of-view (Creaform).

Offline photogrammetry is also changing in concept. With the widespread availability of digital cameras and digital image processing, the conventional separation of image acquisition at object location, and data processing in the lab, is vanishing. Instead, the entire object measurement can be completed on-site. Cameras with integrated image processing and direct computer links enable a seamless transition from classical offline photogrammetry to online photogrammetry, as presented in the following sections.

6.6.1.2 Online photogrammetric systems

Online photogrammetric systems enable the direct measurement of 3D object coordinates much in the manner of a coordinate measuring machine. In the majority of systems, at least two synchronized digital cameras are used for image acquisition, each with known calibration values and pre-determined orientation with respect to an established coordinate system. Online systems designed for tactile object probing have already been discussed in section 6.5.3. This section presents systems which offer contactless coordinate measurement of multiple (targeted) points in real-time,

and which directly transfer the 3D coordinates to a further process. Area-based probing systems are discussed in section 6.7.

Camera operation, illumination (flash) and image transfer are controlled by a networked computer which also provides image processing and coordinate calculations. The computer can be located externally (laptop, computer trolley) or within the camera. In the latter case, image processing (point measurement) is performed directly within the camera, so that only a small number of image coordinates are transferred to the post-processing computer (example in Fig. 3.121b).

6.6.1.2.1 Concepts

There are various configurations and applications for online systems, such as:

- dual-camera system with manually tracked touch probing (as described earlier, see Fig. 6.34);
- dual-camera system for the guidance of medical instruments or other sensors (Fig. 8.85);
- three-camera system for tracking dynamic processes, for example the spatial movement of a robot (Fig. 6.88);
- multi-camera system for analysing human motion (MoCap - motion capture), typically for applications in medical science and entertainment industries (6.10.3, Fig. 6.98);
- multi-camera systems for measuring machine oscillations, e.g. fitness equipment;
- multi-camera system for deformation monitoring by repeated measurements of a set of reference points (example application in Fig. 6.37).
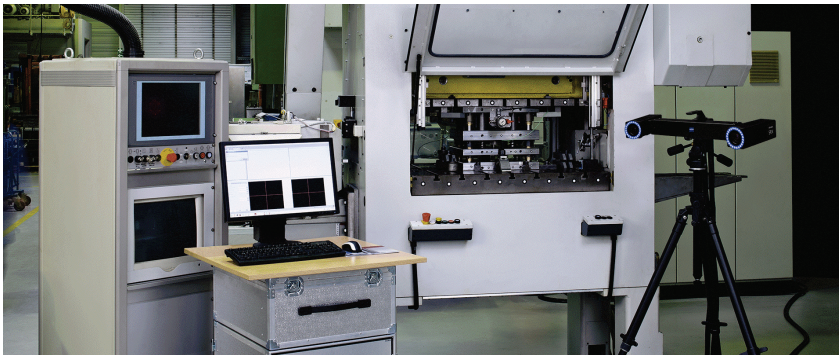
Fig. 6.37: 3D online system with fixed set-up for machine control (ZEISS ARAMIS).

6.6.1.2.2 Calibration and orientation

Calibration and orientation of online systems can be performed in various ways according to the following schemes (Fig. 6.38):
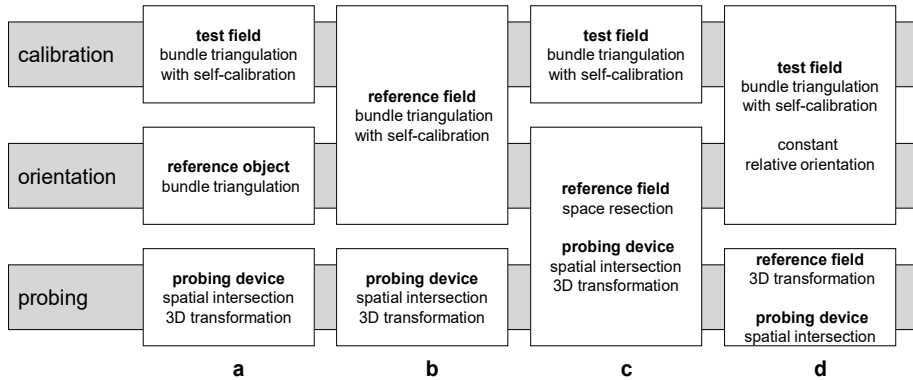
Fig. 6.38: Strategies for calibration, orientation and point measurement for online systems.

a) Separate pre-calibration, orientation and point measurement:
   The parameters of interior orientation of the cameras are determined separately from the process of object measurement, for example by test field calibration. All cameras are oriented by simultaneous imaging of a reference field that defines a local 3D coordinate system. Subsequently the object is measured point by point, by intersection and 3D transformation of the target points and the probing tip.

   This approach requires stable interior orientation of the cameras for the complete period between calibration and completion of the measurement. In addition, the exterior orientation of the cameras must be kept constant, for example by the use of fixed mechanical camera mountings or stable tripods.

b) Calibration and orientation using a reference object added to the measuring object or measurement space:
   A calibrated reference field is positioned on or around the object to be measured and provides local 3D control points or reference lengths. After setting up the cameras, their positions and orientations can be calculated from the reference data by space resection (separately for each camera) or by bundle triangulation. Under some circumstances, the interior orientation can be determined simultaneously if the reference field provides a suitable and sufficient number and distribution of control points, or if it is moved and photographed in a number of different spatial positions. Point measurement follows as in (a).

c) Integrated orientation and point measurement:
   The orientation of the cameras can be integrated with the point measurement itself. For each new measurement of the probe position, the cameras also record

a reference field which is used to check or recalculate the exterior orientations. Such a method is essential if stability of the cameras cannot be ensured.

d) Stable relative orientation and 3D transformation:
    Camera calibration and relative orientation are done by bundle adjustment. If the relative orientation is sufficiently stable, the probe is then located in any convenient coordinate system at the same time as a stable set of reference coordinates is measured. The probe tip coordinates are then transformed into the target system by a 3D similarity transformation.

6.6.1.2.3 Measurement accuracy

Assuming that the system components are the same in each case, the measuring accuracy of online systems will be less than that of offline systems, mainly for the following reasons:

– small number of images for each point measurement (default: 2);
– possibly different measuring accuracies in X, Y and Z (see section 4.3.6.2);
– the strict stability requirements for interior and, especially, exterior orientation are difficult to meet;
– synchronization errors of cameras lead to coordinate errors in dynamic measurement (section 6.10.1);
– if an additional measuring sensor is incorporated into the online system, e.g. a touch probe sensor (section 6.5.3) or a surface scanning sensor (section 6.7.2), the orientation uncertainty of the added sensor will increase the overall online system uncertainty;
– contact of the probe with the surface, and the triggering of the cameras, must occur simultaneously; unless a self-triggering type of probe is being used, this is dependent on the operator's skill.

With these points in mind, commercial online systems offer an accuracy in the object space of about 0.1–0.2 mm for a measuring distance up to 2 m (assumptions: image scale $m = 100$, base length $b = 2$ m). This corresponds to a relative accuracy of about 1:10 000 to 1:20 000, or an image measuring accuracy of 1 μm. Stereo systems based on off-the-shelf industrial cameras provide significantly lower accuracies of the order of 0.1 to 1 mm. They are mainly used in medical applications (section 8.7.2).

6.6.1.3 Stereo vision systems

Stereo vision systems are dual-camera systems in a fixed mounting which provide an online 3D measurement space for applications such as the following:

– location of manual touch probes and surface sensors for object measurement;
– navigation of autonomous vehicles (driver-assistance systems, detection of obstacles);
– 3D navigation for computer-controlled surgical operations;

– control of manufacturing robots (location of workpieces);
– mobile mapping (3D reality capture of indoor and outdoor scenes from mobile platforms);
– tracking people and movement;
– examination using stereo endoscopy (medicine, material testing).

Fig. 6.39 shows a robot with a stereo camera that is controlled by a micro-controller (Raspberry Pi). Fig. 6.40 shows a stereo camera which has a base of 60 mm and is used to assist in neuro-navigation and related techniques in dental and aural surgery (see also section 3.5.4).

Fig. 6.39: Robot with stereo cameras (Boredom Projects).

Fig. 6.40: Stereo camera system for small measurement volumes (AXIOS 3D).

Diverse low-cost stereo cameras are available for applications at lower accuracies (see section 6.7.6). In contrast to the examples of very stable camera constructions shown in Fig. 6.40 or Fig. 6.45, low-cost systems have simpler mechanical designs for the camera housings and fixtures. They must therefore either be calibrated over shorter time intervals or significantly lower accuracies must be accepted. A stereo camera's measurement volume is determined by its base, focal length and direction of camera axes. It is clear from Fig. 6.41 that the height-to-base ratio h/b can be improved by a suitable choice of convergence angle between the camera axes. It is therefore possible to reduce the object distance and improve the accuracy in both XY and Z (towards object), see also section 4.3.6.2.

Based on the properties of digital video cameras, the measurement frequency of stereo cameras lies between 10 and 50 Hz. Higher frequencies can be achieved by the use of high-speed cameras which can be configured to measure in the same way as video cameras. Using a stereo mirror attachment (section 3.4.3.9), a stereo system can be created with a single camera.
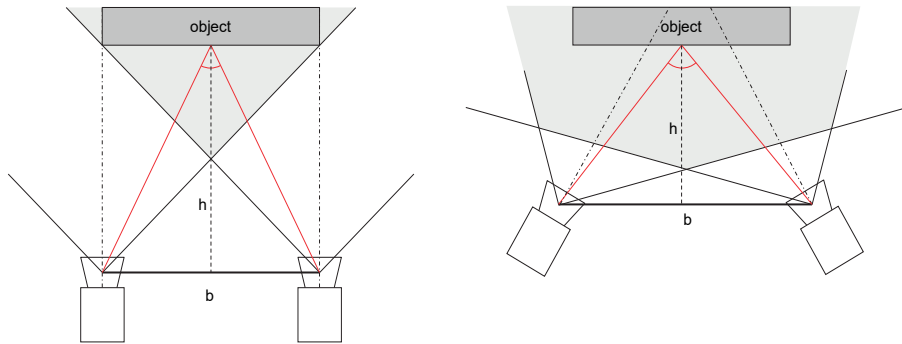
Fig. 6.41: Measurement volume and height-to-base ratio for different stereo configurations.

Example 6.2:

A stereo camera system ($c = 12$ mm) is used in two configurations according to Fig. 6.41. In the convergent setup, the cameras can be positioned closer to the object with a larger base. For an image or parallax measuring accuracy of 0.5 μm, according to eqn. 4.88 and 4.91, the following accuracies are achieved in object space:

Imaging parameters:	$b = 600$ mm	$h = 600$ mm	$b = 700$ mm	$h = 420$ mm$m$
$= 50$	$h/b = 1$	$m = 35$	$h/b = 0.6$	
Measurement accuracy:	$s_x = 0.025$ mm		$s_x = 0.017$ mm	
	$s_z = 0.025$ mm		$s_z = 0.011$ mm	

With convergent geometry a significantly higher measurement accuracy can be achieved, by as much as a factor of 2 in the viewing direction.

6.6.2 Static industrial online measuring systems

Some specialized and repetitive measuring tasks are well suited to the development of tailored solutions with a high degree of automation. In contrast to mobile systems, these are static systems which occupy a semi-permanent position within a factory. They can be designed so that the photogrammetric measuring components (such as cameras, projectors, devices for rotating the object) can remain calibrated for a long period and the cameras can be oriented automatically using a fixed and known reference point field. Such a system is in a permanent state of readiness for repeating the same or similar measurements within a fixed, limited volume and with a considerable degree of automation.

Multi-camera photogrammetric systems can also be integrated directly into production lines for purposes of quality control. The measuring problem usually concerns a limited variety of parts, so that a fixed arrangement of cameras, light sources and control points is possible. The function of such a system is real-time

measurement of the objects followed by quality analysis and the transfer of results to the manufacturing control system.

Stationary online systems using pattern projection for surface measurement are presented in section 6.6.2.3. Flexible systems incorporated into production processes, e.g. for robotic inline manufacturing, are discussed in section 6.9.5.

6.6.2.1 Tube inspection system

Hexagon's TubeInspect offers a solution for the automatic 3D measurement of formed tubes, rods or wires, a typical example being hydraulic brake lines for the automotive industries. This non-contact approach has replaced the previous use of gauges.
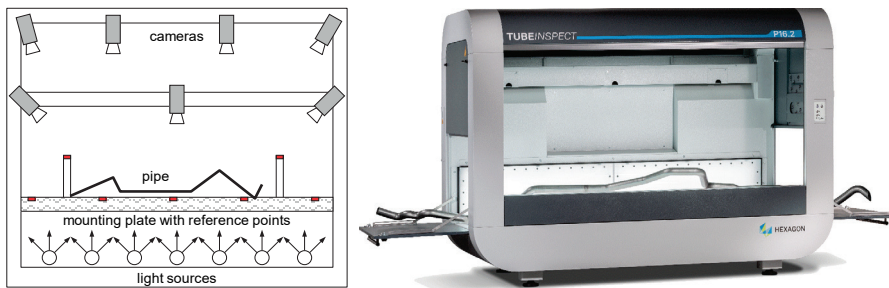
Fig. 6.42: Online system for check measurement of tubes (Hexagon).

The system is arranged on several horizontal planes distributed vertically (Fig. 6.42). The upper two planes contain up to 16 CMOS industrial cameras. The lower two planes consist of an illumination array of light sources and a transparent mounting plate with spatially distributed reference points which are used for camera orientation. The tube to be measured is placed on the mounting plate in an arbitrary position.

In each image, the tube appears as an easily identified black contour against a bright background (Fig. 6.43). Bending points, bending angles, straight line sections and arcs can be determined fully automatically by digital multi-image processing. The system also calculates correction data with respect to given nominal values that are directly transferred to the bending machine for adjustment of the manufacturing process (Fig. 6.44). The measuring accuracy has been reported as 0.3–0.5 mm for pipe diameters between 4 mm and 200 mm and within a measuring volume of about 2.5 x 1.0 x 0.7 m^3.
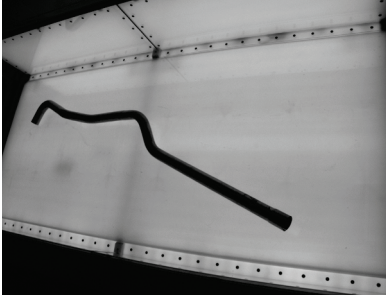
Fig. 6.43: Backlit image of a tube (Hexagon).
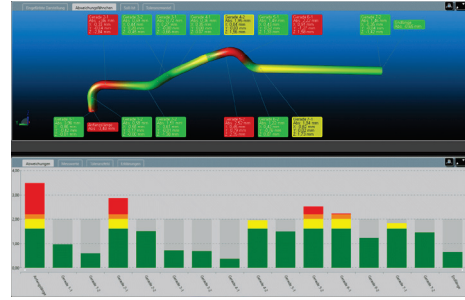
Fig. 6.44: Comparison with nominal values in tube measurement (Hexagon).

6.6.2.2 Steel-plate positioning system

A stereo camera system can be applied to the task of three-dimensional online positioning of components in manufacturing and assembly plants. This is illustrated by the following example of an online stereo system (AXIOS 3D CamBar B2) used to position steel plates for ship construction, with dimensions up to 20 m x 30 m, on laser cutting machines.
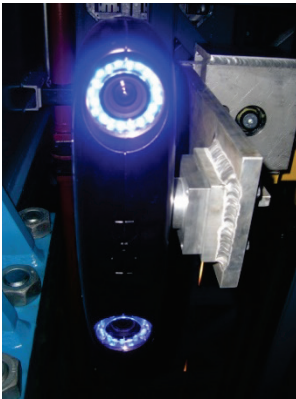
a) Stereo camera
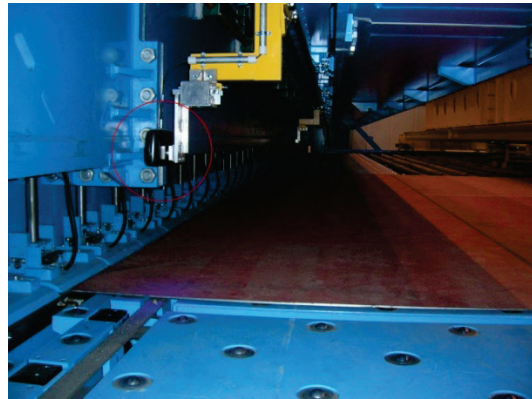
b) Camera attached to positioning system (red circle)

Fig. 6.45: Online measuring system for positioning steel plates (AXIOS 3D).

The system itself only has a measurement volume of 280 mm x 360 mm and an offset distance of 600 mm. It is mounted on a coarse positioning device which places it over an area to be measured on the plate. Cross-shaped marks are etched on the surface of the plate and have known (nominal) positions in the component's local coordinate system. The stereo system can identify the marks automatically, even under highly variable illumination conditions, and can determine their 3D locations relative to

known reference drill holes. Depending on the quality of the target marks, measurement accuracy can be as high as a few hundredths of a millimetre. The entire plate can be positioned to an accuracy of 0.1 mm in the machine coordinate system. Fig. 6.45 shows the stereo system in place in the factory.

6.6.2.3 Multi-camera system with projected point arrays

A non-targeted, textureless free-form surface can be measured point by point, by projecting an array of targets onto the surface (section 3.6.3) and imaging these by a number of cameras synchronized with the projection device. The projector's orientation needs not be known as the point coordinates are calculated by intersected rays from oriented cameras, as is the case with conventional targets. However, if the projector's relative orientation is known then the target points can be projected onto pre-programmed locations.
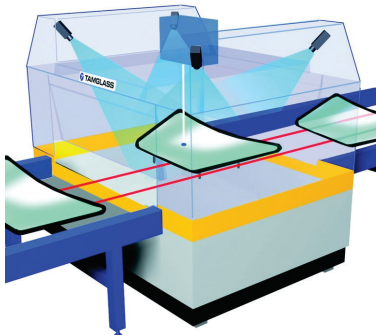
a) Schematic construction   b) Windscreen positioned for measurement
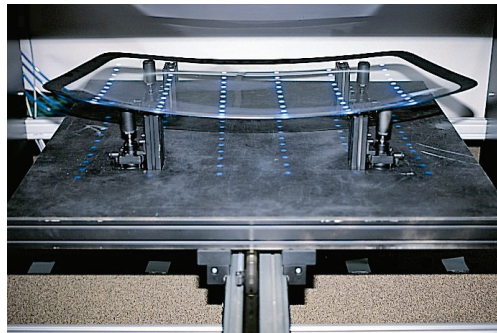
Fig. 6.46: Multi-image measuring system with laser point projection for the 3D measurement of windscreen surfaces (Mapvision).

Fig. 6.46 shows a multi-camera measuring system with target point projection, used to measure car windscreens (Mapvision). The scanner's projector emits ultraviolet light rays which cause fluorescence when they penetrate the glass. The resulting light spots on the surface of the glass are imaged by cameras sensitive to the relevant wavelength. The measuring accuracy of the system has been reported as 0.2 mm.

6.7 Systems for surface measurement

6.7.1 Overview

6.7.1.1 Active and passive systems

For the optical 3D measurement of free-form surfaces, different methods are in use depending on object properties, required measuring time, accuracy und point density. The optical systems described here all offer contactless measurement using either a pattern projected onto the surface or the surface texture itself (natural or artificial). Active projection techniques include fringe projection (phase measurement), raster projection and laser scanning. Purely photogrammetric (passive) methods utilize stereo or multi-image configurations with image matching methods to analyse surface textures or projected patterns (see section 5.5). Hybrid systems also exist in which these technologies are combined.

A different approach is represented by range cameras. These are based on a imaging sensor which can measure distances (section 6.7.6) but they are not used for precise measurement of surface form and are not included in this section.

In analogy with sampling theory (section 3.1.5), the resolution of surface detail requires a density of measured object points sufficient to capture its geometric complexity. This in turn applies to the textures and projected patterns which also require a sufficient resolution so that the corresponding measuring algorithms can deliver the necessary surface points which are independent of neighbouring points (uncorrelated). Surface textures must be unique within larger neighbourhoods, therefore not homogeneous or a repeated pattern. High contrast gradients are required for efficient matching results.

The term "active systems" covers all solutions which use structured light for surface measurement. Here the illumination pattern is an integral part of the measuring principle, i.e. measurement is not possible without the active projection of light. Active structured light systems have a number of properties which are relevant to their practical application:

- Surfaces without texture can be measured if they are diffuse (Lambertian) reflectors.
- A very high point density can be achieved, e.g. one 3D value per pixel.
- Scanning or sequentially measuring methods can handle dynamic processes only if the changes in object space are significantly slower than the measurement rate.
- Fringe projection systems can achieve very high accuracies. However, they are sensitive to extraneous light and their maximum working distance is limited to a maximum of around 1–2 m.
- Laser scanning methods are largely independent of environmental light conditions and can be used at very long ranges which, depending on design, can exceed 1000 m.

Passive photogrammetric methods reconstruct the surface using at least two images and the matching of corresponding image features. Appropriate applications are those which benefit from the following operational features of passive photogrammetric systems:

- Contact with the object is not necessary if it has a natural texture suitable for measurement.
- The measuring volume can be scaled almost arbitrarily and does not depend on the illumination power of a projector.
- In principle, the number of cameras is unlimited.
- Area-based matching methods (section 5.5.3) can achieve subpixel resolution which, in turn, can lead to high measurement accuracy in the object space.
- Real-time generation of large point clouds is possible with suitable image matching procedures and appropriate computing power.
- There are no safety restrictions such as the eye-safety requirements with laser projection.
- Dynamically changing scenes and surfaces can be recorded with multiple synchronized cameras.
- It is possible to track multiple physical points on surfaces, for example where the mechanical characteristics of specific points must be analysed.
- Camera-based systems can be assembled at very low cost.

However, it is emphasized that passive systems require the measured object to have a surface texture which provides the appropriate resolution, contrast and uniqueness for the system to generate the surface form to the specified accuracy.

6.7.1.2 Surface textures for area-based measurement

Area-based imaging systems are often used for measuring the 3D shape of arbitrary free-form surfaces. These image-based techniques require an object's surface to have a texture suitable for accurate and reliable matching of homologous features. The resolution of the texture must match the geometric complexity of the surface and the required resolution or density of the final point cloud. Appropriate contrast and gradients in the texture are also necessary in order to achieve accurate matching results. Textures in use are covered by the following categories:

- Natural surface texture:
  If the object surface has a suitable texture of its own, it can be measured directly without further preparation. The texture should consist of unique patterns which permit reliable matching of homologous areas.
- Physical surface marking:
  If the natural surface texture does not support image matching, artificial textures can be provided by etching, paint and powder sprays or adhesive patterned sheets, as explained next. This physical marking of a surface may not always be permitted but has the advantage that physically defined points can be tracked

through image sequences, e.g. for deformation analysis or strain measurement. There is the further advantage that the applied texture can be adapted to a particular measurement task.

– Etching:

For applications such as strain measurement of metal sheets, a physical texture can be applied by chemical or laser etching of random features into the surface. In this case the texture is of uniform thickness, predefined contrast and resolution and can be optimized for the subsequent analysis of the measurement, e.g. strain analysis. Since the texture is burned into the object material, it cannot be removed later.

– Paint and powder sprays:

It is very common to spray white powder or paint onto the surface and apply brush strokes to provide a random pattern. The object surface then has a structure useful for image correlation. However, manual surface preparation may lead to non-uniform textures, inhomogeneous structures or different thickness of applied material. The texture layer can be removed from the surface after measurement. Recently available sprays provide a matt surface for area scanners with active projection (see below) and then evaporate after a short time.

– Adhesive patterned sheets:

Thin adhesive sheets with a printed surface pattern can be attached to the target surface. Fig. 4.38 shows a test object to which such a textured target film has been added for purposes of evaluating surface measurement accuracy. The adhesive target film has a thickness of around 0.1 mm and can be printed with any convenient random pattern. Depending on application, surface analysis may require the film thickness to be taken into account.

– Pattern projection:

It may be convenient, or there may be a requirement, for fully non-contact measurement. In this case, a regular or irregular feature array can be projected onto the object surface, e.g. by LCD devices or diffraction grids (section 3.6.3). If the illuminated object is observed by at least two cameras, no calibration or orientation of the projector or pattern is required. The pattern simply provides a visible structure defining the surface. Active pattern projection provides high contrast object textures for surfaces with few natural textures. The method is not suitable when using sequential images to measure deformation at specific object points (see section 6.7.5.3). Discrete surface points cannot be tracked through a sequence of images (see also section 5.5.7).

The approaches above can be combined. For example, a fast active projection method can be used for coarse surface measurement which is then refined using a number of cameras and passive illumination of texture on the surface. They can be further combined with standard, point-by-point photogrammetric measuring methods.

6.7.2 Laser triangulation

Fig. 6.47 shows the principle of the laser triangulation (light-section) method. A laser beam is projected through a cylindrical lens to generate a light plane (see Fig. 3.128). An array imaging sensor is arranged with an offset to the laser diode in order to form a triangle with known (calibrated) geometric parameters. The projected laser plane on the object is deformed in the image as a function of object shape and distance.
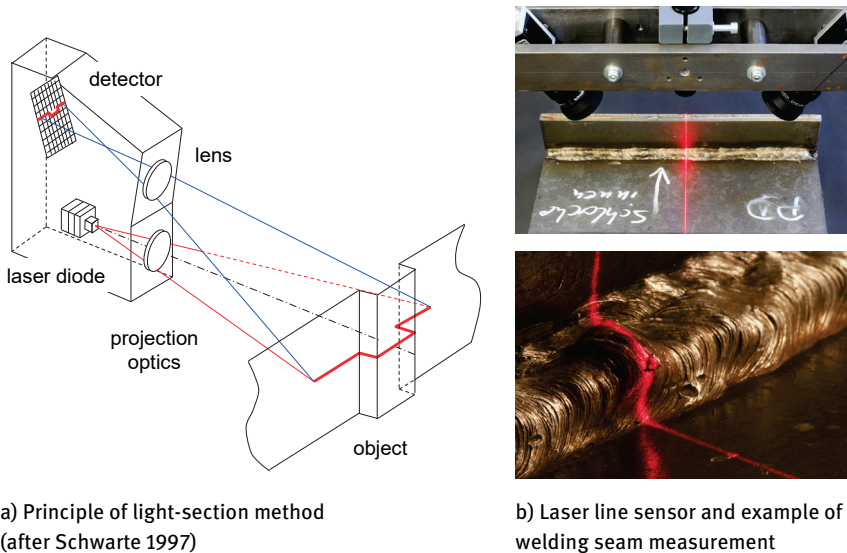
a) Principle of light-section method
(after Schwarte 1997)

b) Laser line sensor and example of
welding seam measurement

Fig. 6.47: Laser triangulation method.

A full 3D measurement of an object surface can only be achieved in combination with a scanning mechanism, e.g. if the triangulation sensor is mounted on a CMM arm (Fig. 6.86b) or positioned by a laser tracker (Fig. 6.17, Fig. 6.18). Laser triangulation sensors have typical accuracies for distance measurement in the order of 0.1 mm at distances of up to 500 mm. They allow for typical measurement rates of up to 1000 Hz.

6.7.3 Fringe projection systems

6.7.3.1 Stationary fringe projection
Methods of stationary fringe projection are based on a fixed grid of fringes generated by a projector and observed by one camera. The grid has a cyclical structure, normally with a square or sine-wave intensity distribution with constant wave length λ.

Fig. 6.48 shows the principle of phase measurement for parallel fringe projection and parallel (telecentric) image acquisition at an angle α. The wave length λ corresponds to a height difference ΔZ_0:

$$\Delta Z_0 = \frac{\lambda}{\sin \alpha} \tag{6.4}$$

The height difference ΔZ with respect to a reference plane corresponds to the phase difference $\Delta \varphi$:

$$\Delta Z = \frac{\Delta \varphi}{\tan \alpha} \tag{6.5}$$

Phase difference is only unique in the range $-\pi \ldots +\pi$ and so only provides a differential height value within this range. Absolute height change also requires counting the fringe number to determine the additional whole number of height units. The measuring method is therefore only suitable for continuous surfaces which enable a unique matching of fringes. A resolution in height measurement of about λ/20 can be obtained.
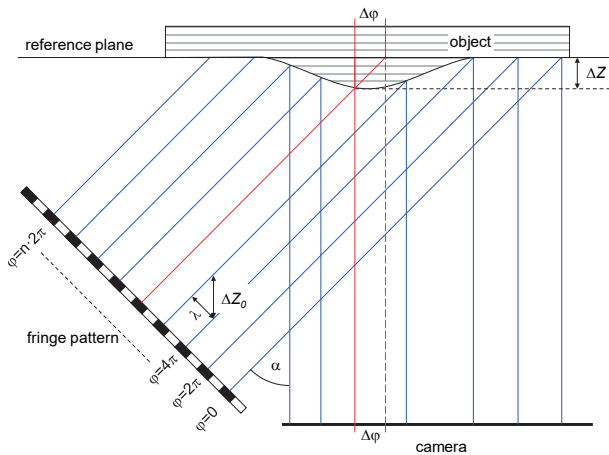
Fig. 6.48: Fringe projection with phase measurement by telecentric imaging.

The telecentric configuration illustrated in Fig. 6.48 limits the size of measured objects to the diameter of the telecentric lens. If larger objects are to be measured the method must be extended to perspective imaging. Stationary fringe projection can also be applied to dynamic tasks, for example the measurement of moving objects.

6.7.3.2 Dynamic fringe projection (phase-shift method)

Phase measurement can be performed directly using intensity values in the image. For this purpose the projected fringes are regarded as an interferogram. For the intensities of an interferogram at a fringe position n, the following equation applies:

$$I_n(x,y) = I_0\left(1 + \gamma(x,y)\cos(\delta(x,y) + \varphi_m)\right) \tag{6.6}$$

where

I_0:   constant or background intensity
$\gamma(x, y)$:   fringe modulation
$\delta(x, y)$:   phase
φ_m: phase difference

The equation above contains the three unknowns I_0, $\gamma(x,y)$ and $\delta(x,y)$. Hence at least three equations of this type are required for a solution. They can be obtained by m sequential shifts of the fringes by the difference φ_n (Fig. 6.50a):

$$\varphi_m = (n-1)\varphi_0 \tag{6.7}$$

where

m: number of shifts
$n = 1 \ldots m,$ where $m \geq 3$
$\varphi_0 = 2\pi / n$

The measuring principle is known as the *phase-shift method*. When $m = 4$, samples are taken at equally spaced intervals of $\pi/2$ (Fig. 6.50a). The phase of interest $\delta(x, y)$ then reduces to:

$$\delta = \arctan\frac{I_2 - I_4}{I_3 - I_1} \qquad\qquad \delta = \delta(x,y); \qquad\qquad I_n = I_n(x,y) \tag{6.8}$$

Finally, the height profile is given by:

$$Z(x,y) = \frac{\lambda}{2 \cdot 2\pi}\delta(x,y) \tag{6.9}$$
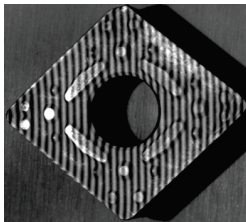
a) Object with projected fringes   b) Phase image with phase discontinuities   c) Demodulated height model
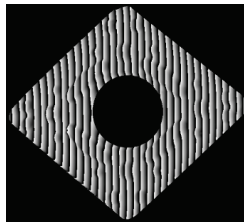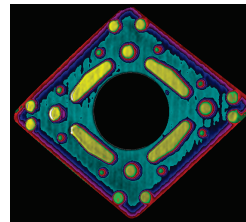
Fig. 6.49: Phase-shift method (Breuckmann).

As with stationary fringe projection, the result is unique only in the interval $-\pi\ldots+\pi$, so that integer multiples of 2π must be added for a complete determination of profile. This process is known as *demodulation* or *unwrapping* (Fig. 6.49).

Discontinuities in the object surface lead to problems in the unique identification of the fringe number. A unique solution is possible by using fringes of varying wavelength (Fig. 6.50b) or Gray-coded fringes (section 6.7.3.3).

a) Phase shifts with $m = 4$ b) Sinusoidal fringes of varying wavelengths
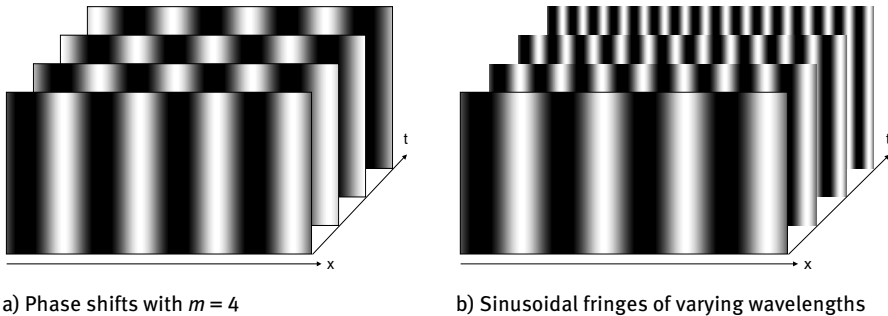
Fig. 6.50: Sinusoidal patterns used in the phase-shift method.

The accuracy of height measurement is about $\lambda/100$. The interior and exterior orientation of projector and camera must be found by calibration. Each pixel (x, y) is processed according to eqn. 6.8. The computations can be solved using fast look-up table operations, with the result that height measurements can be processed for all pixels in an image (for example 780 x 570 pixel) in less than one second. Examples of measuring systems are presented in sections 6.7.3.4 and 6.7.3.6.

6.7.3.3 Coded light (Gray code)

Solving ambiguities is a major problem for fringe projection methods, especially for discontinuous surfaces. In contrast, the *coded-light* or *Gray code* technique provides an absolute method of surface measurement by fringe projection.

The projector generates m coded fringes sequentially, so that perpendicular to the fringe direction x_p a total of 2^m different projection directions can be identified by an m-digit code word (Fig. 6.51 shows an example bit order 0011001 for $m = 7$). A synchronized camera observes the fringe pattern reflected and deformed by the object surface. The m images acquired are binarized and stored as bit values 0 or 1 in a bit plane memory which is m bits deep. Hence, each grey value at position (x',y') denotes a specific projection direction x_p from O_p.

This procedure requires known orientation parameters of camera and projector but otherwise requires no initial values related to the object. It is relatively insensitive with respect to changing illumination and the reflection properties of the object. Continuity of the surface is not a requirement. The accuracy of the method is limited

to about 1:500. It is therefore mostly used for fast surface measurements of lower accuracy, for example as a preliminary to subsequent measurement by phase-shift.
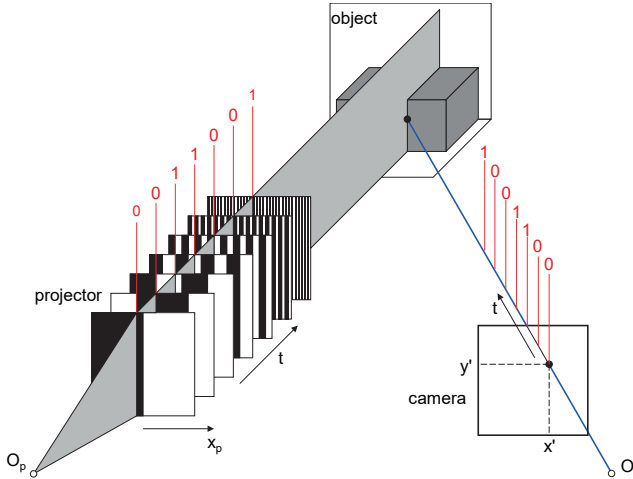
Fig. 6.51: Coded-light approach (after Stahs & Wahl 1990).

6.7.3.4 Aperiodic fringe projection

As an alternative to the individual or combined phase-shift and coded-light approaches, aperiodic sinusoidal fringes offer another way of applying sequentially projected patterns. Here N fringe patterns (typically $6 \leq N \leq 12$) are projected with an intensity distribution defined by:

$$I_{proj}^{(i)}(x,y) = a^{(i)}(x) + b^{(i)}(x) \cdot \sin[c^{(i)}(x) \cdot x + d^{(i)}(x)] \qquad i = 1...N \qquad (6.10)$$

where

a: offset
b: amplitude
$2\pi/c$: period length
d: phase shift

The parameters must be selected such that the resulting functions $I_{proj}^{(i)}(x,y)$ are continuous. The projected pattern sequence is observed by two cameras.

Fig. 6.52a shows an example of a set of aperiodic fringes. A significant advantage of aperiodic sine patterns compared with phase-shifted sine patterns is that there is no 2π ambiguity. Neither the projection of an additional pattern, e.g. Gray code (section 6.7.3.3), nor a phase unwrapping step is therefore required. Hence in general, a fewer number of projected fringe patterns can generate 3D point clouds to a comparable accuracy and greater object coverage. The maximum correlation

coefficient successfully determines the stereo correspondence between both cameras (Fig. 6.52b). A fringe projection system based on this approach is presented in section 6.7.3.6.

a) Sequence of projected aperiodic sine patterns   b) Matching of image points by correlation
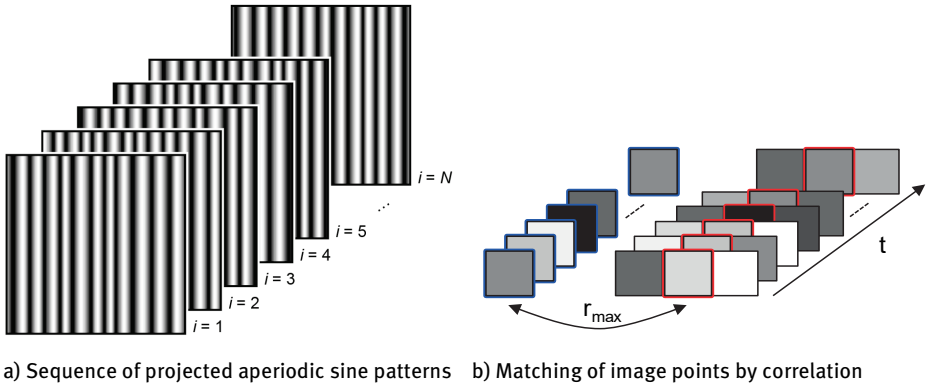
Fig. 6.52: Sequentially projected aperiodic sine pattern (Fraunhofer IOF).

6.7.3.5 Single-camera fringe-projection systems

Measuring systems which operate according to the phase-shift method consist of a fringe projector and a camera with a base separation appropriate for the measurement volume (larger base for bigger volume, but with accuracy decreasing with measurement volume). Camera and projector must be calibrated and spatially oriented to one another. Because of the ambiguities of phase measurement inherent in eqn. 6.8, the systems are frequently combined with a coded light technique (section 6.7.3.3) which provides approximate values for the object surface. Alternatively, fringes of varying wavelength can be projected.

The projector is normally designed using a liquid crystal display (LCD), liquid crystal on silicon (LCOS) or micromirror array (DMD), see section 3.6.3.2. As a rule, the camera is a digital video camera with up to 2000 x 2000 pixels. The performance of this method is determined mainly by the reflective properties of the surface. Usually, homogeneous, diffusely reflecting surfaces are required. Specular reflections and hot spots must be avoided by preparation of the surface (for example by dusting with white powder) and provision of suitable ambient lighting conditions.

Dynamic fringe projection is applicable only to static objects. Depending on camera and projector used, as well as the required measured point density, measurement frequency lies between 0.2 Hz and 50 Hz. The basic configuration of just one camera typically results in occluded areas of the object which are not therefore measured. The most important advantage of the phase-shift method is the fast measurement of a few hundreds of thousands of surface points. At the same time, however, the problem of thinning and structuring of the measured point cloud arises, as is necessary for further processing of the 3D data, for example for the control of production tools.
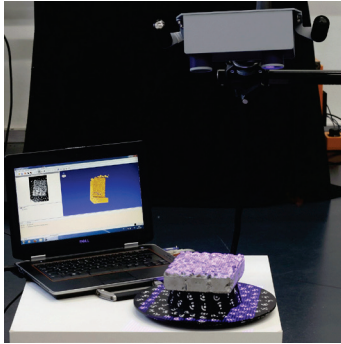
a) Vialux fringe projection system

b) Projection with large fringe separation

c) Projection with small fringe separation
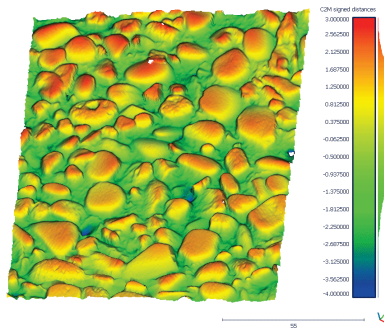
d) 3D measurement result

Fig. 6.53: Scanning an object with a single-camera, fringe-projection system.

A measuring volume between 0.1 x 0.1 x 0.1 m^3 and 1 x 1 x 1 m^3 per scan is typical for fringe projection systems. In order to measure larger objects, mobile fringe projection systems are used in combination with photogrammetric methods or mechanical positioning systems for spatial orientation, or by techniques for registration of individual point clouds (see section 6.8). Relative accuracies of 1:8000 can be achieved (see section 6.7.3.6). Single camera fringe projection systems are offered, amongst others, by ZEISS IQS, Creaform, Minolta and Vialux. Fig. 6.53 shows fringe projection at different wavelengths and the resulting 3D point cloud.

A variant of fringe projection is the method of *deflectometry*. It does not work with diffuse reflected fringes but is designed for use with polished and highly reflective surfaces such as glass. From the distorted reflection of the fringe pattern in the surface, the surface shape can be deduced.

6.7.3.6 Multi-camera fringe-projection systems

In their basic configuration, the fringe projection methods mentioned above use one projector and one camera. Because the projector serves as a component of the measurement system, uncertainties in its geometry adversely affect the results. In addition, for objects with large surface variations there are frequently areas of occlusion and shadow which cannot be observed by a single camera. Furthermore, smooth surfaces often give rise to over illumination or highlights in the image.

Consequently, a number of advantages are offered by multi-camera systems with active fringe projection:

– reduced measuring uncertainty as a result of greater redundancy (number of cameras);
– fewer areas of occlusion or highlights;
– no requirement to calibrate the projector;
– possibility of measuring moving objects using synchronized multi-imaging;
– greater flexibility in measurement by variation of the projected pattern and the relative arrangement of projector and cameras.

Calibration and orientation of multi-camera systems with active illumination follow the principles of test-field calibration (section 7.3.1.1). If required, the orientation parameters of the projector can also be determined in this process since the projector can be regarded as an inverse camera. 3D coordinates are usually calculated by spatial intersection (section 4.4.7.1). Robust methods for matching homologous image patterns utilize epipolar constraints and aperiodic fringe projection or variable wavelength fringe projection. Fringe patterns are therefore aligned approximately perpendicular to the epipolar lines.
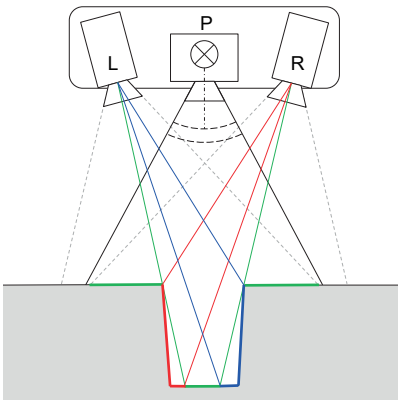
Fig. 6.54: Measurement of a surface with occlusions using two cameras and one projector.

With a calibrated projector, the number of usable observations can be increased from four to six for a dual-camera system. This improves the accuracy of surface measurement. In addition, occlusions can be handled by using the projector in combination with only one camera. Fig. 6.54 illustrates a situation where a part of the surface (green) can be measured from both cameras. The partially occluded areas (red and blue) are observed by only one camera but can still be reconstructed using the calibrated projector.

Fig. 6.55 shows a mobile dual-camera system with a pattern projector (ZEISS ATOS). Two convergent cameras are arranged on a fixed base and a pattern projector is mounted between them. The typical measuring volume of a single recording lies between $12 \times 7 \times 8 \, mm^3$ and $750 \times 500 \times 500 \, mm^3$ depending on the system configuration. However, the ATOS LRX model, equipped with a more powerful light source, is also suitable for detecting large-volume components under natural ambient light. A relative accuracy between 1:2000 and 1:10 000 can be achieved.

a) Two-camera system with projector
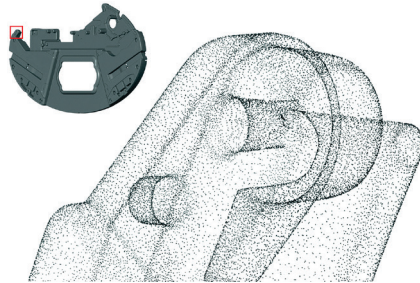
b) 3D point cloud
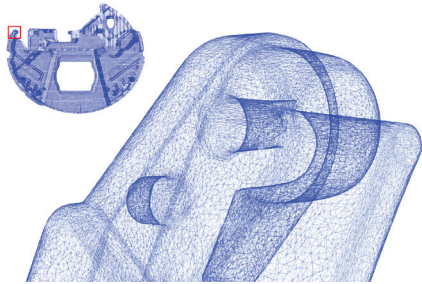
c) Triangular mesh
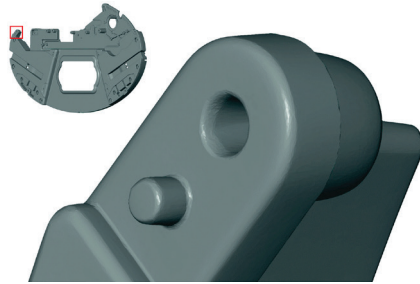
d) Shadowing

Fig. 6.55: Measurement of a design model by a photogrammetric fringe projection system (ZEISS).

Fig. 6.56 shows a structured light sensor using a GOBO projector (section 3.6.3.2) to project an aperiodic sine pattern (section 6.7.3.4). This delivers a fast surface measurement at up to 36 Hz. The projection illumination in the near infrared at $\lambda = 850$ nm, and an additional RGB camera, are used to record true colour information.

Even higher 3D measuring rates of up to 10 kHz can be achieved using special high-speed cameras (section 3.5.3).

Dual-camera fringe projection systems are manufactured by companies such as ZEISS IQS, Hexagon, Faro and Fraunhofer. Low-cost, dual-camera projection systems are also available, for example, from Hewlett Packard (DAVID SLS series). They are significantly less powerful than the systems described above for industrial metrology.

a) Sensor   b) Measurement example
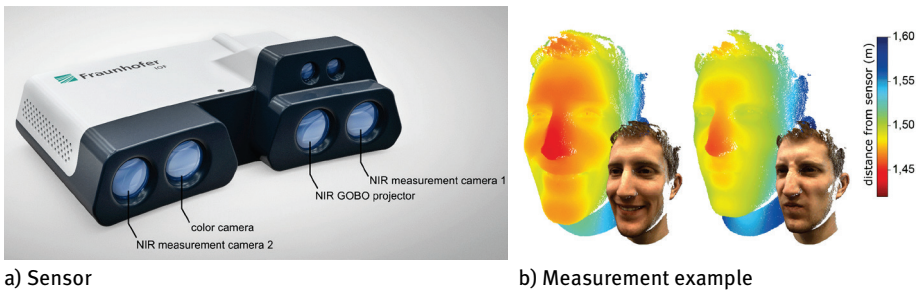
Fig. 6.56: Dual-camera structured light sensor for real-time 3D measurements (Fraunhofer IOF).

6.7.4 Point and grid projection

6.7.4.1 Multi-camera systems with target grid projection
Grid projection methods are used mainly in materials testing for mapping displacement and strain. The surface shape can be reconstructed by analysing the deformation of the raster with respect to a reference position. The reference grid can be generated in different ways:
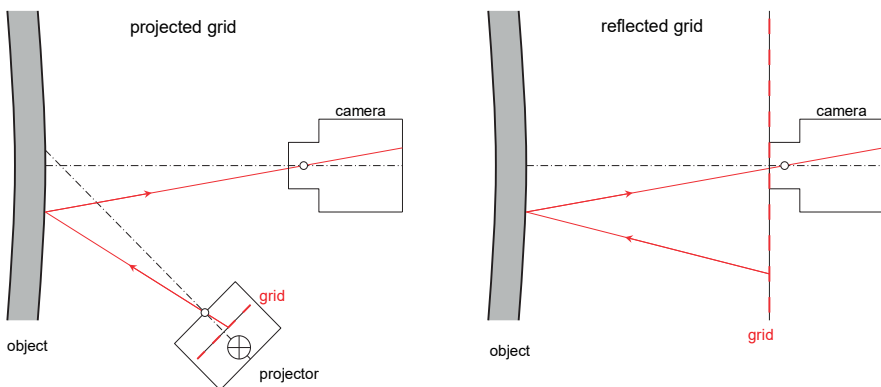
Fig. 6.57: Target grid measuring methods (after Ritter 1995).

–   by a grid that is physically created on the object surface;
–   by projection of a grid from a position to the side of the viewing camera (Fig. 6.57 left);
–   by reflection of a grid (Fig. 6.57 right).

The configuration for grid projection corresponds to that of stationary fringe projection (section 6.7.3.1). Using only one camera requires a calibrated and oriented projector. For multiple cameras, the projector is only used to generate structured light. The reflection method assumes virtually specular surfaces and only allows measurement of surface inclination.

6.7.4.2 Multi-camera system with grid projection

The AICON ProSurf system, now out of production (see Fig. 6.58), used an analogue projector to place a regular grid structure on the object surface. This structure was imaged with at least three cameras, all oriented to one another and arranged around the object.

Measurement of the object surface was done by measuring the grid intersection points in the images and locating them in object space by multiple ray intersection. Epipolar geometry was used for the correct assignment of corresponding rays. With a minimum of three images this leads to a unique solution (see section 5.5.5.2). To measure the entire grid a recursive search strategy was implemented which ensured that all measurable points were located, even when points were obscured or otherwise missing. Object point location accuracy was around 0.1 mm in a measurement volume of approximately 1 m x 1 m x 0.5 m.

In a similar way, the GSI Pro-Spot system presented in section 3.6.3.2 projects a regular array of target points onto an object surface. If projector and object have no relative movement, then the targeted object surface can also be measured by a standard offline photogrammetric system (section 6.6.1.1).
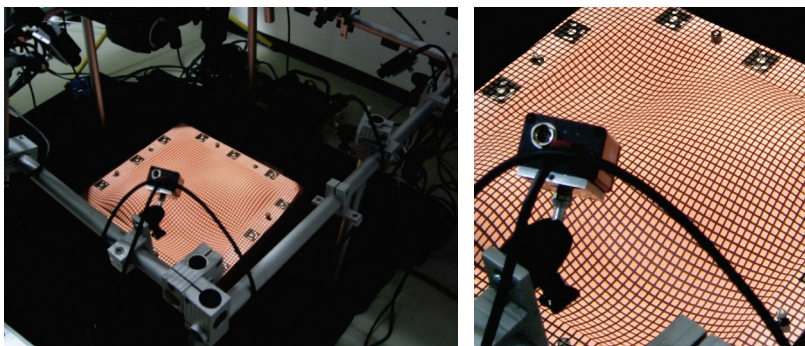
Fig. 6.58: Multi-camera system with grid projection (AICON).

6.7.4.3 Multi-camera system with regular etched patterns

In applications where it is permitted to give a surface a permanent marking, very fine structures, typically grid patterns, can be applied by electro-chemical or laser etching. Since the grid points are physically marked, if the object is observed by at least two cameras then individual grid points can be measured and tracked through a deformation process such as strain resulting from stress loading, deep drawing or thermoforming.
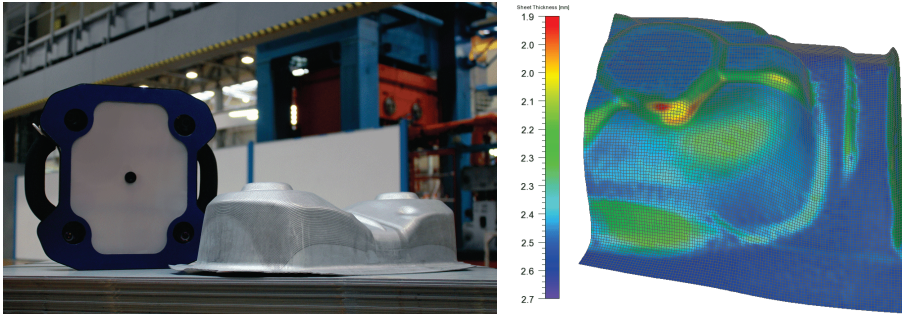
a) Camera head and patterned object   b) Deformation plot

Fig. 6.59: Four-camera system with recorded metric image of etched grids (Vialux).

Fig. 6.59a shows the AutoGrid comsmart, a portable measuring system from Vialux which has four synchronized 5-Mpixel cameras. It can be hand-held and directed at a surface prepared with a grid pattern and a minimum of three local reference points for the definition of a workpiece coordinate system. Fig. 6.59b shows the resulting deformation measurement of a thermoformed object. Accuracy is reported to be around 0.04 mm for an object size up to 1 m.

6.7.5 Systems utilizing random patterns

Digital image correlation (DIC) and matching can be used to measure free-form surfaces if there is enough surface texture information. Typical image measuring accuracy is of the order of 0.1 pixels. Lateral and depth accuracy further depend on the image configuration, e.g. the base-to-height ratio of cameras. The approach has already been used in the early stages of digital photogrammetry, e.g. by the Zeiss Indusurf system (see next section).

The texture must have a structure that allows for reliable and accurate image matching in two or more images. Texture resolution, structure and contrast must be designed for the required 3D result and take into account object curvature, camera resolution and image scale, illumination conditions and object reflectance properties.

6.7.5.1 Dual-camera system with projection of random patterns

The Zeiss Indusurf system, was developed in the early 1980s for the off-line measurement of car body surfaces. At that time it consisted of an analogue stereometric camera system, the Zeiss SMK 120, and a slide projector for the projection of a random dot pattern (Fig. 6.60). The analysis was carried out using an analytical plotter equipped with CCD cameras for digital image acquisition. Parallax measurement was performed using least-squares matching and a measuring accuracy of 0.1mm was achievable. Around 1990 the system was adapted to use digital cameras but is now no longer in production.
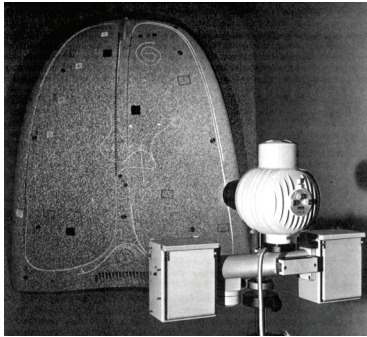
a) Stereometric camera and pattern projector   b) Random pattern
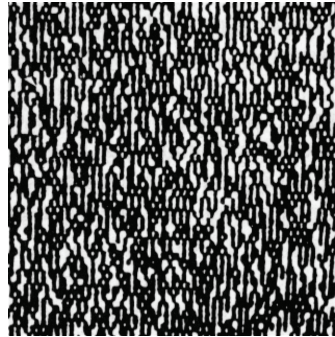
Fig. 6.60: Zeiss Indusurf for the measurement of car body surfaces (from Schewe 1988).
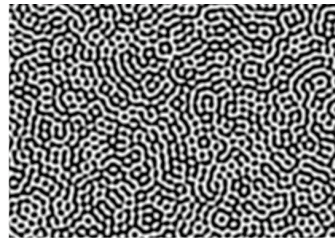
a) Cognex A5030   b) Random pattern for image correlation

Fig. 6.61: Stereo metric camera with structured light projector.

A current example which makes use of this basic principle is the Cognex 3D-A5000 system. This has two fixed cameras with a speckle pattern projector positioned in between them (Fig. 6.61). The system uses a band-limited projected pattern similar to that in Fig. 6.61b. Depending on the baseline of the cameras, between 50mm and 1200mm, and their focal length, a measurement volume from 6.5mm x 4.6mm x 1.2mm up to 1.3m x 1.8m x 0.1m (L x W x H) is provided. The recorded pattern is

analysed by a correlation algorithm. Measuring frequencies of up to 10 Hz, and repeat accuracies of approx. 0.01 mm, can be achieved.

6.7.5.2 Surface measurement with textured adhesive film

Textures of any kind can also be printed onto thin adhesive films which are then attached to an object's surface. Fig. 6.62 shows an example of a test artefact covered by a textured film which is about 0.1 mm thick.

Textures should be designed for hierarchical image matching, e.g. in different image pyramids or scales (see section 5.1.3), and multiple pattern resolutions can be overlaid (example in Fig. 6.62b). High resolution textures are necessary if high resolution is required for surface reconstruction (example in Fig. 6.62c).
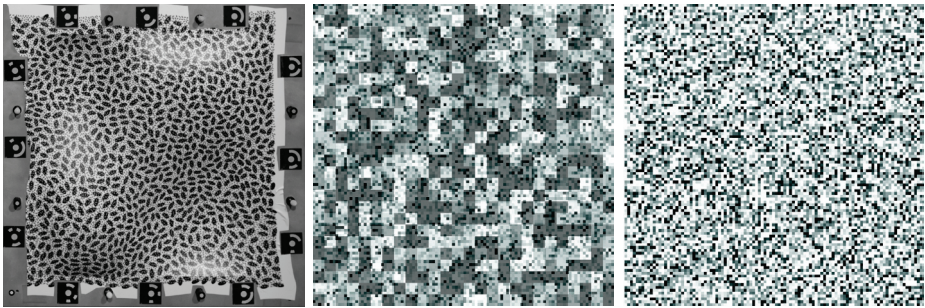
a) Texture with blobs   b) Texture suitable for image pyramids   c) High resolution texture

Fig. 6.62: Different texture patterns suitable for digital image correlation (DIC).

Image analysis is by stereo or multi-image photogrammetry. Typical image measuring accuracy is of the order of 0.1 pixels if least-squares matching is applied (section 5.4.2.4). Lateral and depth accuracy further depend on the image configuration, e.g. the base-to-height ratio of the cameras. As an example, 3D points on the test artefact in Fig. 6.62a (side length 500 mm) can be measured to an absolute accuracy of some 0.1 mm using industrial cameras with sensor resolutions of around 1300 x 1000 pixels.

6.7.5.3 Measurement of dynamic surface change

It is only possible to model dynamic surface changes in 3D if the measurement frequency is higher than the object's rate of deformation. Usually, therefore, methods based on surface scanning or sequential pattern illumination (fringe projection) are unsuitable and optical surface capture with simultaneous multi-image recording is required (an exception is discussed in section 6.7.3.4).

Recording object surface points with respect to the object or component coordinate system can be done in one of the following ways:

– Regular XY grid:

Changes in surface position can be measured using a regular XY grid of points, i.e. at every XY location, and for every deformation state (epoch), a new Z value is determined. A suitable measurement technique for this purpose is the principle of the *vertical line locus* (VLL, see section 4.3.6.3).

– Irregular distribution of points:

By applying feature detection algorithms (see section 5.4.3), the location of irregularly distributed surface points is calculated for every epoch by spatial intersection in which new XYZ values are determined every time. The result is an irregular 3D point cloud.

– Target tracking:

Surface points which are determined at the start, e.g. distinctive features found by interest operators (section 5.4.3), are tracked through the deformation epochs. The result is a spatial curve (trajectory) for every measured surface point. Tracking requires prediction of the location of the surface points in the next epoch, which can be done by Kalman filtering (section 5.5.7.5).

Fig. 6.63 illustrates the recording and evaluation of a surface deformation. The surface was marked with a random pattern and its deformation recorded by a stereo camera. The images are processed in a multi-level image pyramid. This starts with a coarse measurement using normalized cross correlation (section 5.4.2.3) and finishes with a fine measurement using a least-squares adjustment (section 5.4.2.4). The derived deformation vectors are shown in Fig. 6.64, in one case on a fixed XY grid and in another as space trajectories of discrete surface points.

Epoch 0   Epoch 15   Epoch 30
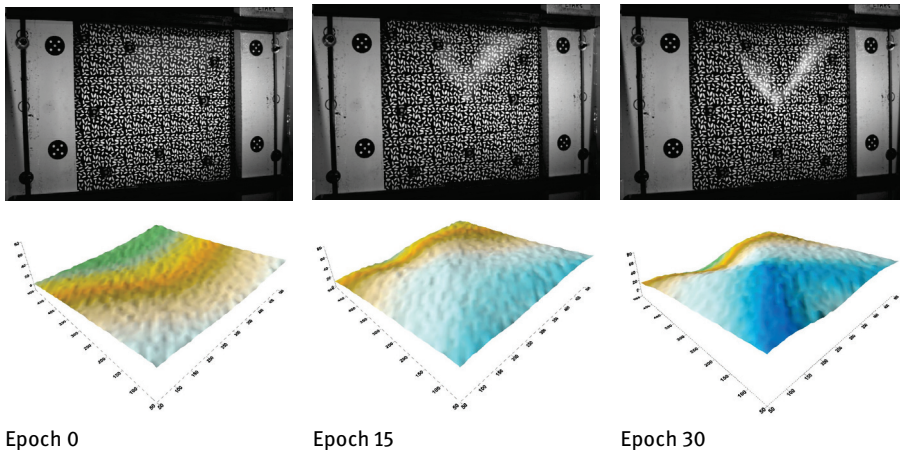
Fig. 6.63: Image sequence (left hand image only) and the corresponding model of surface deformation (IAPG).
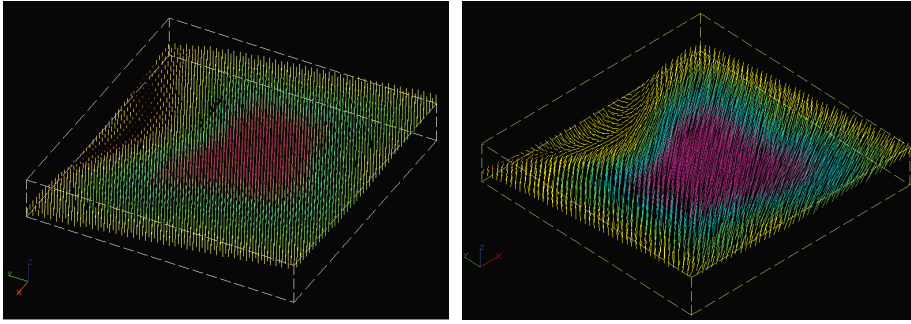
a) Deformation vectors on a regular XY grid

b) Trajectories of discrete surface points

Fig. 6.64: Deformation vectors between epochs 0 and 25 (IAPG).
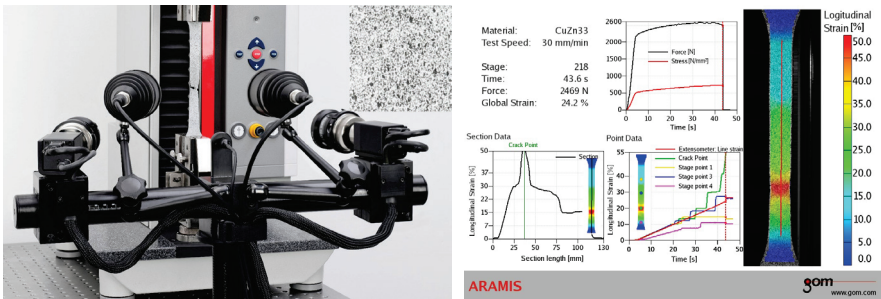
a) Strain measurement using sprayed texture

b) Result of measurement

Fig. 6.65: Digital image correlation system (DIC) for strain analysis (ZEISS ARAMIS).

Fig. 6.65a shows an example application of a digital image correlation (DIC) system for strain analysis (ZEISS ARAMIS). In this case a tensile test specimen, whose length in the field of view is about 125 mm, is mounted on a tensile testing machine which applies controlled strain to the object. The metal surface has been textured by spraying with a random pattern (see upper right corner in Fig. 6.65a). The three-dimensional deformation of the test material is observed by two synchronized cameras and physical surface points are tracked through the image sequence by digital correlation techniques (see section 5.5.7, Fig. 6.65b). The achieved accuracy is reported to 2 μm in XY and 4 μm in Z (= height).

Another application based on digital image correlation is shown in Fig. 6.66. A wind-turbine rotor blade is measured over a length of 10 m. Here the natural texture of the unfinished surface provides sufficient structure for stereo correlation. Fig. 6.66 shows the result of a bending test where the blade deformation in the Z direction is measured. The achieved accuracy is of the order of 0.1 mm in Z (=out of plane).
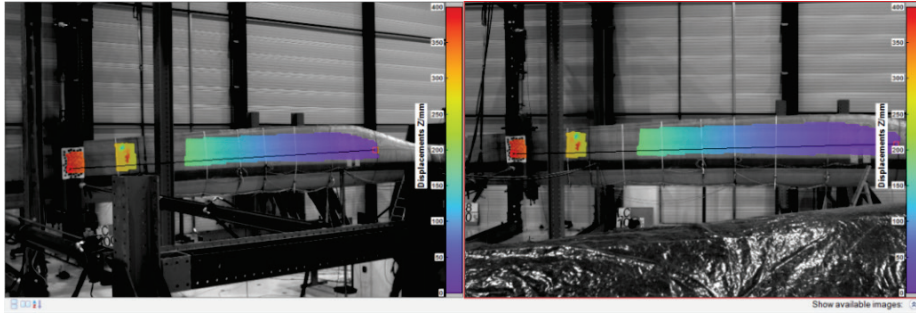
Fig. 6.66: Rotor blade deformation measurement with 3D DIC system (LIMESS).

6.7.6 Range cameras

Here the term range camera is used for a sensor that delivers range images (see section 2.3.3.3) directly from the sensor. All required computations are performed on-board the camera. Such range cameras are frequently used in robotics, at security gates, in industrial manufacturing processes and for consumer products, e.g. gaming consoles, vacuum cleaners. As there is no need for additional processing on a powerful host computer, range cameras are ideal for embedded systems and for mobile systems due to lowered power consumption. Range cameras do not introduce fundamental new principles for determining 3D coordinates but are based on the principles already described previously, namely triangulation and time-of-flight. However, they are implemented using semi-conductor components which are as simple as possible. Mass-manufacturing allows a price per sensor that is several orders of magnitude below high-grade surveying or metrology instruments based on similar principles.

6.7.6.1 Kinect

Originally, demand from the gaming and entertainment industry boosted the development of low-cost consumer-grade range cameras. The prime example is the Microsoft Kinect, first released in 2010 and with subsequent sales of over 30 million units.

The first version of the Kinect (Fig. 6.67a) was based on 3D technology developed by PrimeSense, with further technology modifications by Microsoft to make it suitable for gaming. The 3D technology is essentially triangulation. Here a structured light approach is adopted which uses a fixed pattern of speckle dots generated by a diffractive optical element and a near-infrared laser diode. The projected pattern (Fig. 6.67b) provides high light intensity and is observed by a laterally offset camera. Camera and pattern image are configured according to the normal case of stereo

photogrammetry (section 4.3.1.4). Proprietary algorithms are assumed to be behind the identification and location of the elements of the pattern, and hence the 3D model. The camera resolution is low, e.g. VGA (640 x 480 pixels). However, at the measuring frequency of 30 Hz, almost 10 million points per second can be acquired. The maximum working range of this class of sensor is short, in the range 1-6 m. Coupled with a limited accuracy of 5 to 50 mm, it therefore does not meet the usual requirements for many engineering measurement tasks.

a) Kinect 1 based on PrimeSense ChipSet

b) Projected point pattern

c) Kinect 2 based on TOF principle

d) Range image
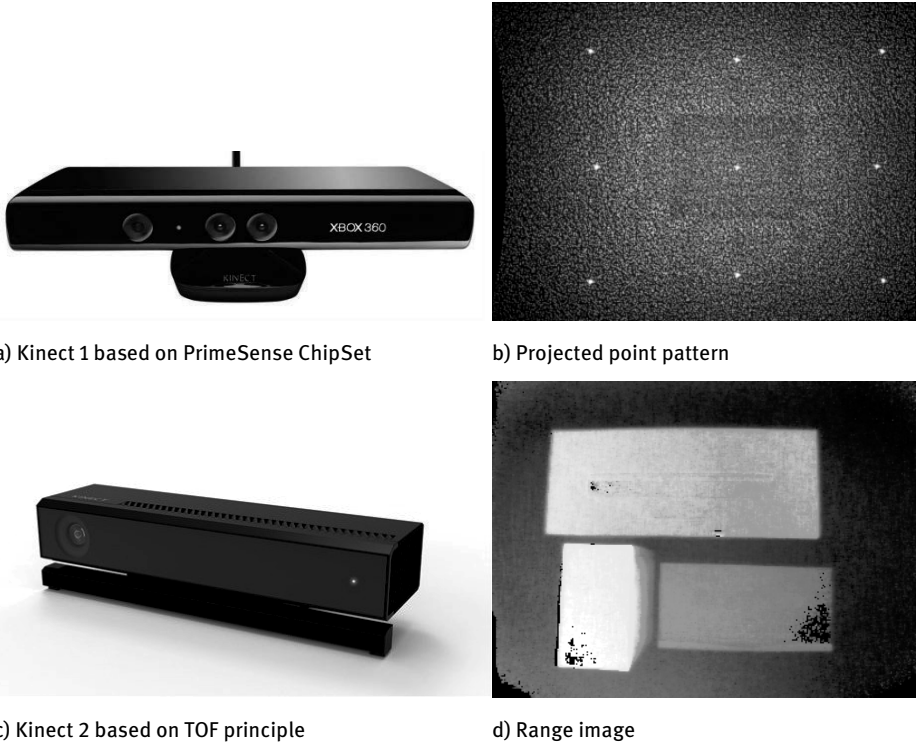
Fig. 6.67: 3D camera Microsoft Kinect.

The Kinect v2, launched in 2014, was the next generation of the Kinect and is based on a time-of-flight principle using amplitude modulation (compare 6.8.1). An array of near-infrared LEDs illuminates the complete scene. The light is amplitude modulated and the light reflected from the scene is received by a camera sensor. The outgoing signal and the received signal are correlated and the time difference is determined from the signal shift. Measuring range and accuracy are comparable to the Kinect 1. Fig. 6.67c shows the Kinect 2 and Fig. 6.67d a range image recorded by the system. Both versions of the product are now discontinued. The technology of the Kinect v2 is

still used in the Microsoft HoloLens and the Azure Kinect which is aimed at the professional market.

6.7.6.2 Current generation range cameras

Since 2020 a time-of-flight range camera has been included in Apple's mobile phones and tablets (Fig. 6.68a). The sensor uses a laser light source and a diffractive lens (see section 3.6.3.2) to emit a sparse dot matrix of a maximum of 576 points on which time-of-flight is measured. Proprietary algorithms are used to up-sample this sparse matrix to a full-frame range image of 192×256 pixels. While the sensor was mainly developed for Augmented Reality, measurement apps are already available from companies such as DotProduct, Pix4D and Polycam. Independent tests on small objects have shown a standard deviation of around 1 cm when the point cloud is compared to a terrestrial laser scanner.

a) Apple iPad Pro

b) Intel RealSense D435

c) Intel RealSense L515

d) Intel RealSense T265

Fig. 6.68: Range cameras.

With many range cameras now only available as embedded systems, Intel's series of 3D sensors under the RealSense brand have become a popular choice at a significantly low price. The RealSense D435 (Fig. 6.68b) is a stereo camera with a baseline of 50 mm and a field of view of 87 x 58 degrees. It delivers a depth matrix with 1280 x 720 pixel (see Fig. 6.69). With the small baseline, the sensor is recommended for use up to 3 metres. Independent tests estimate the ranging error at up to 5 mm at 1-meter distance. The sensor has an integrated RGB camera for colour.

The depth camera has an optional laser pattern projector for featureless surfaces. Different variants of the senor exist with integrated IMU and wider baselines.

The RealSense L515 (Fig. 6.68c) is a solid-state LIDAR using a MEMS mirror to guide a laser beam across the scene. It has a field-of-view of 70 x 55 degrees and outputs a depth matrix of 1024 x 768 pixels at 30 frames per second which corresponds to 23 million points per second. It is intended for use up to 9 metres and is quoted with a nominal accuracy of 5-14 mm. The compact sensor has an integrated RGB camera to colour the point cloud.

The RealSense T265 (Fig. 6.68d) is an unusual range camera as it is not intended to output a depth matrix, but only its 3D position and orientation (6DOF). It is intended as a visual tracking device. However, it is based on a stereo camera principle and features two fish-eye cameras with a baseline of 64 mm. All visual odometry computations (section 5.5.7.6) are incorporated in the sensor.
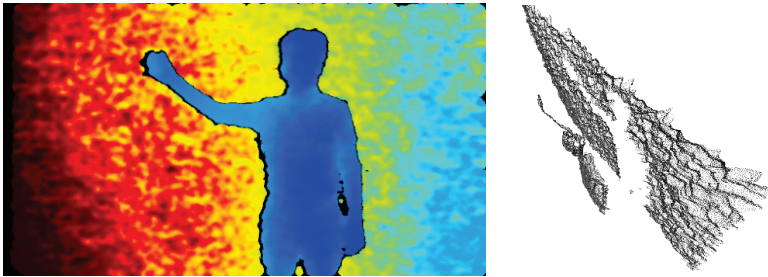
Fig. 6.69: Colour-coded range image from D435 (left) and corresponding point cloud (right).

6.7.6.3 Light-field cameras

Light-field or plenoptic cameras use an array of microlenses positioned at a certain distance between imaging sensor and lens (Fig. 6.70). Depending on the object distance, the image of an object point is collected by number of microlenses and then projected onto the sensor. The resulting image has a multi-facet pattern whose repetitions are used for distance-dependent parallax calculations from which distances can be derived. In principle, the microlens array can be positioned either in the focal plane (see examples P' and Q'), or in front of it or behind it.

Fig. 6.70 shows the image of an object point recorded by a light-field camera at two different distances. It can be seen that the point at the greater distance is imaged with fewer microlenses (compare px_P for the far point with px_Q for the near point).

Light-field cameras are suitable for short imaging distances of a few metres. The achievable accuracy can reach 1/10 mm for very short distances (<10 cm). For larger imaging distances (>1m) the accuracy rapidly drops to about 5–10 mm. Note that the size of the stereo base is defined by the number of microlenses which enable

correlation of an object pattern. Hence, distance accuracy decreases quadratically with distance (see section 4.3.6.2) and is traded off against spatial resolution.
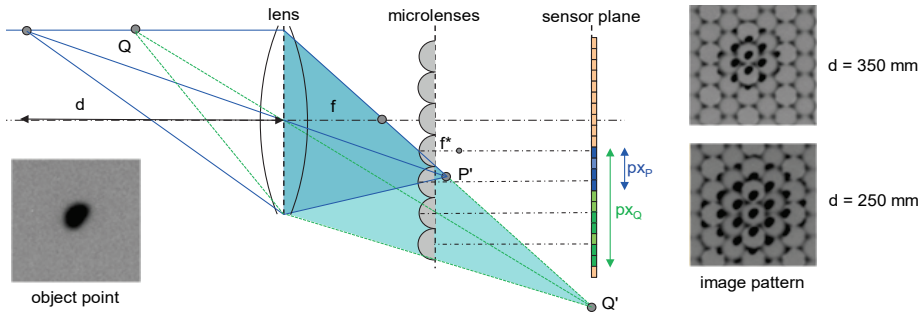
Fig. 6.70: Principle of a light-field camera.

Fig. 6.71 shows two examples of light-field cameras with sensor resolutions up to 4 Mpixel. Using the recorded distance map, images can be refocused, or the depth of field adjusted, by post-processing.

a) Raytrix

b) Lytro Illum

Fig. 6.71: Examples of light-field cameras.

6.8 Laser-scanning systems

6.8.1 3D laser scanners

3D laser scanning describes the three-dimensional measurement of the surface of an object or scene by distance, typically based on time-of-flight (TOF), and angular measurement using a laser beam which is scanned horizontally and vertically over the object or scene. This generates a 3D point cloud of the object. The scanning

mechanism delivers the individual cloud points in polar coordinates (see section 2.2.5.1). These are defined by the scanning instrument's two angle measurements and one range (distance) measurement which have their origin at the centre of the scanning system. If the reflected intensity is also registered, an object image is recorded where each measured 3D point is associated with an intensity or colour value dependent on the reflectance of the object surface to the scanning laser wavelength.

The scanning concept has been implemented in systems operating at both short and long ranges, from a few metres to 100s of metres. In airborne applications the technology is also known as light detection and ranging (LiDAR). Today light-weighted LiDAR sensors also enable drone scanning applications (see section 6.11.2).

Laser scanners are also increasingly used in close-range and industrial applications, now ranging from small volume applications such as weld checking up to larger volumes such as the recording of building interiors and the measurement of complex pipework structures. In addition, they are also used to meet rapidly growing demands for measurement of the built environment from moving platforms (mobile mapping) and for control of autonomous vehicles.

Commercially available systems differ mainly in terms of physical principle, measuring frequency, measuring accuracy, range of operation, beam diameter and costs. Some scanner designs have a restricted horizontal and vertical view, less than a hemisphere, and are also known as camera-view scanners. In contrast, panoramic scanners provide all-around 360° measurement from a single location.

Fig. 6.72 shows the schematic design of a panoramic scanner. A laser diode emits a laser beam which passes through a semi-transparent mirror onto a second mirror which is rotated about vertical and horizontal axes. The vertical displacement is implemented by a horizontal axis rotating continuously at high speed. The horizontal displacement is achieved by rotating the whole scanner head about the vertical instrument axis.

There are various techniques for the laser measurement of distance to the object, one of which is phase measurement. Here a modulated laser beam is continuously transmitted and the distance calculated from the phase differences between transmitted and reflected signals. The distance D can be measured by the phase difference as follows:

$$D = 1/2(n \cdot \lambda + \Delta\lambda) \tag{6.11}$$

where

λ: wavelength
n: number of total wavelengths (period)
$\Delta\lambda$: phase difference

The principle works at high measurement rates of up to two million points per second, but distance measurements are unique only within a certain range. Maximum ranges are around 360 m with a corresponding distance accuracy of about 1 mm.
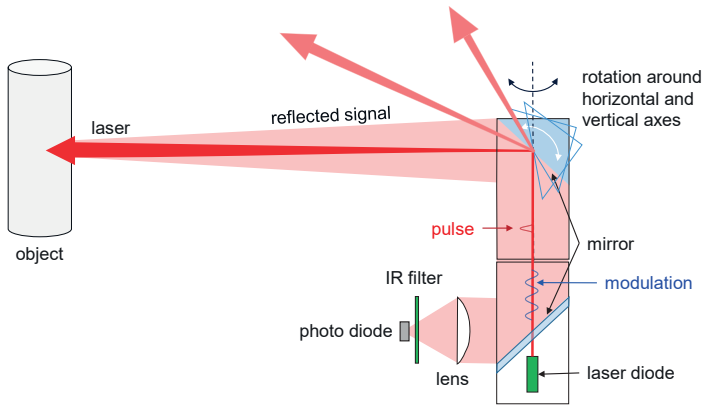
Fig. 6.72: Principle of laser scanning.

Other commercially available laser scanners pulse the laser beam and calculate range by measuring the time-of-flight of the pulses:

$$D = 1/2 \cdot c \cdot t \tag{6.12}$$

where

c: propagation speed of light in medium

t: time of flight

Pulse measurement is less sensitive with respect to object surface variations and is therefore generally preferred for reflectorless surface measurement at longer distances (>400 m). Measurement rates are about one order of magnitude slower than for phase-difference measurement. Pulse measurement methods allow for full waveform analysis in which a sequence of return signals is stored and analysed in a post-processing mode. Typical distance measurement noise of 3-8 mm can be achieved.

Time-of-flight measurement can be combined with phase-difference measurement in a *Wave Form Digitizer* (WFD). The WFD system constantly evaluates, digitizes and accumulates the waveforms of all reflected signals for precise recognition and extraction of the measurement pulses. This provides time-of-flight measurements with the phase-difference characteristics of high measurement frequency and low noise.

In addition to the distance measurement, the system accuracy depends on the angle measurement. Typical angular measurement accuracies are 0.004-0.01 degrees corresponding to 0.8-2 mm at 10 m distance. Further influences on the accuracy are given by the calibration of all system components, the positional stability of the instrument and the atmosphere (refraction, light propagation speed).

Manufacturers of laser scanners with geodetic quality include Leica, Zoller & Fröhlich, Riegl, Faro, Trimble, Topcon and Teledyne Optech. Fig. 6.73 shows three

typical instruments. Well-known manufacturers of simpler laser scanners are, for example, SICK, Livox and Ouster Velodyne, which are often used in security areas or in the control of autonomous vehicles.

Newer solid-state LiDAR sensors function without any mechanically moving components (see also 6.7.6.2). Here, an expanded diode laser beam is projected onto the object via a micro-mirror array and a lens. The achievable accuracies are significantly lower than those of geodetic scanners.

a) Zoller & Fröhlich Imager 5016   b) Leica RTC 360   c) Riegl VZ 600i

Fig. 6.73: Examples of terrestrial laser scanners.

Regardless of technique, if recorded laser-scanned data ($X,Y,Z,intensity$) are formatted as a regular array, two data sets can be derived:

1) A range image (see section 2.3.3.3) where the value (displayed as grey or false colour) assigned to each pixel is proportional to the range of the object surface from the scanner (Fig. 6.74a,b).

2) An intensity image where the value assigned to each pixel is a function of the strength of the return signal from the surface of the object (Fig. 6.74c).

The intensity image appears similar to a photographic image. However, it only records reflection values for the wavelength of the scanner's laser.

The principal feature of laser scanning is the fast three-dimensional measurement of a (large) object surface or scene with high point density. A typical scanning system has a 3D coordinate accuracy of the order of 3-5 mm at a distance of 10 m. In addition to angle and range measurement, the accuracy is also dependent on the stability of the instrument station, for example a tripod. Terrestrial laser scanners can be combined with a digital camera in different ways:

a) Grey-value range image

b) Colour-coded range image

c) Intensity image

d) RGB image overlay

Fig. 6.74: 3D laser scanning of a sculpture.

– A digital camera within the scanner viewing through the mirror system, or mounted onto the rear of the mirror, collects imagery either during scanning or in separate imaging passes of the scanner. This method has the advantage of minimizing parallax between the image and point cloud, but the disadvantage of reduced image quality due to utilization of the scanning optics. Limitations on image quality can be alleviated partially by mounting a small camera on the reverse of the mirror.

– External cameras may be incorporated in or attached to the scan head at known calibrated offsets. Conventionally the imaging is done in a second path. Limitations are in accurate determination of the camera-to-scanner geometry for

pixel-to-point-cloud mapping and, more fundamentally, to differences in parallax caused by the separate projection centres.

– An external camera used separately from the scanner. This allows very high-quality images to be obtained, either from a system placed on a tripod in nominally the same location as the scanner or by mapping a network of photogrammetric images into the scanning geometry. Here the most suitable images can be selected to colour the point cloud, the choice depending on local image magnification at each scanned surface, angle of view to the surface and image content. In this case, arbitrarily high-resolution images can be created, e.g. by eliminating occlusions or where orthogonal viewing directions for building facades are required. In addition, high dynamic range (HDR) images, multispectral or thermal images can be generated and assigned to the scanned point cloud. However, this approach requires a combined network adjustment (see section 4.4.2.3).

If the camera has been calibrated, and oriented with respect to the scanner's internal coordinate system, point clouds and images can be combined and/or separately processed as follows:

– Coloured point clouds:
Since all scanned points have 3D coordinates, they can be back-projected into the images using the collinearity equations (4.10). From their position in the image, colour values can be interpolated and assigned to the points (examples in Fig. 6.74d and Fig. 6.75a).

– Registration:
The 3D transformation of single point clouds into a common coordinate system (registration) can be supported by image information, e.g. through the detection of targets or corresponding object features. Fig. 6.73b shows a laser scanner that uses five cameras looking in all directions to determine the scanner position from one location to the next using visual odometry techniques (section 5.5.7.6).

– Monoplotting:
A feature of interest is measured in the image. The corresponding image ray is intersected with a surface model derived from the point cloud (triangular mesh) to give the 3D coordinates of the corresponding object point (see section 4.2.7 and example in Fig. 4.23).

– Orthophoto:
A true orthophoto can be derived directly when the surface model is orthogonally projected onto a reference plane and filled with colours from the images (Fig. 6.75b, and compare with section 4.2.8.2).

– Central projection:
The original images can be used to create a combined image with central projective properties (tangential image, section 4.5.5.2).

– Panoramic image:
   From multiple overlapping images, a distortion-free panoramic image can be calculated if perspective centre and rotational centre are identical (Fig. 6.75c, and compare with section 3.5.6.2).

a) Complete registered point cloud generated using individual scans from a laser scanner

b) Orthophoto generated from a 3D point cloud and additional RGB images

c) Panoramic image generated from HDR images of an integrated camera

Fig. 6.75: Processing of point clouds and digital images.

Since measurements are made by the reflection of a laser beam from a surface, 3D laser scanners can record large numbers of points without any need for targeting. Like photogrammetry, laser scanning requires lines of sight. Complex objects must

therefore be recorded from several instrument stations whose individual point clouds must be transformed into a common coordinate system by means of common points or matching procedures such as ICP or Visual SLAM (see section 6.9.6). As with other area-based scanning methods, unstructured data acquisition is a major disadvantage. Usually point clouds must be manually or automatically thinned and structured in order to prepare the data for further processing.

6.8.2 2D and 1D laser scanning

In some cases, a 3D laser scanner can be operated with a vertical beam deflection only, i.e. there is no horizontal rotation of the instrument about its vertical axis. In this 2D mode only vertical profiles are scanned. Scanners are also specially designed for this purpose. They are usually mounted on moving platforms and a full 3D scan is created by the motion of the platform. Profile scanners are available which offer a complete 360° scan (panoramic view) as well as a restricted field of view (camera view). Profile scanners usually provide medium accuracies, typically 5–20 mm range accuracy up to about 100 m range, which are suitable for tasks such as the control of autonomous vehicles, security systems and body scanners. Fig. 6.76 shows an example of a 3D profile scanner used for security systems, e.g. robot collision avoidance.

Fig. 6.76: 2D profile scanner (SICK).

Fig. 6.77: Sensor with four distance meters (IAPG).

In 1D laser scanning there is no horizontal or vertical beam deflection. In this case, the device is essentially a distance meter, like a laser tape measure, typically used for monitoring distance to a small surface area which might be moving in the direction of measurement. Operation of these devices may be restricted. In the absence of beam rotation, the beam is essentially static and, depending on its laser class, could cause damage to the eyes. Fig. 6.77 shows an experimental set-up of four distance meters

from Z+F 5006 laser scanners. These are arranged to form a planar fan shape to measure four points synchronously, and in situ, on a wind turbine's individual rotor blades.

6.8.3 Panoramic imaging systems with laser distance measurement

By combining a pan-and-tilt camera with a simple laser distance meter it is possible to create a sequentially recording camera system which, under certain circumstances, enables photogrammetric 3D measurements from only one standpoint.

Fig. 6.78: Panorama camera with distance meter (Rodeon, Dr. Clauß).

The camera is mounted onto a pan-and-tilt unit and adjusted so that the optical perspective centre coincides with the nodal point of the panoramic image (see section 3.5.6). A laser distance meter, such as from Leica or Bosch, is mounted with an offset to the camera (example in Fig. 6.78). In this way, images can be generated with the additional distance information necessary to provide the scale information required for photogrammetric processing. The typical distance accuracy is around 1–2 cm at up to 30 m range. Current commercial systems are primarily designed for the measurement of rooms but can be used for many panoramic interiors.

There are two different measuring strategies for data acquisition:

– Panoramic image recording followed by distance measurement:

Firstly, a panoramic image is stitched together from a sequence of single images in a local 3D coordinate system. After creating the panorama, a manual measuring mode is selected in which the user selects object points of interest in the image. The measuring system drives to the selected points whose range is then measured by the distance meter. In combination with the two angles corresponding to the image location, full polar coordinates are available for these target points. In this mode the object is measured sequentially in a similar way to measurement by total station.

– Panoramic recording with integrated distance measurement:

In this mode, a panoramic image is recorded in the same way as above, but simultaneously one distance measurement is registered per image. The actual 3D model is created offline by combining the assumed parallelism and orthogonality of object planes with the polar coordinates of image and distance measurements. In this way, onsite object recording is accelerated and modelling is done later at another location. Fig. 6.79 shows an example of a model created by this process.

Panoramic image
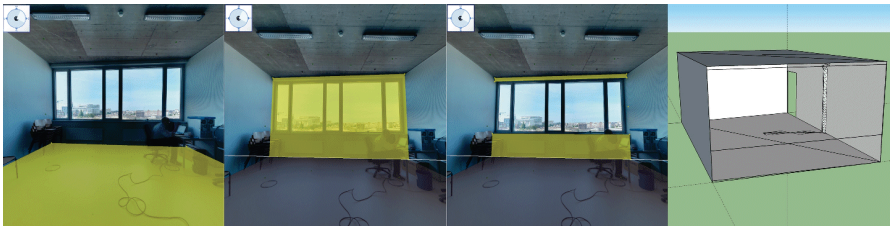
Modelled interior

Fig. 6.79: Measurement of interior with panoramic camera and distance meter (HCU Hamburg).

In principle, any type of panoramic camera can be used to acquire images from different positions which are then processed in a common coordinate system by photogrammetric bundle adjustment. Additional object points can then be measured by intersection. Techniques based on structure-from-motion (section 5.5.2.2) can also be applied if the panoramic imaging model is integrated into the process and the object offers suitable surface textures.

6.9 Registration and orientation of images and scans

When the size or form of the surface to be measured is such that it cannot be covered in a single measurement stage, then the surface must be measured in parts which are subsequently transformed into a common coordinate system to create a single 3D representation of the object. When 3D point clouds are transformed in this way, the process is also known as *registration*.

Registration can be achieved through one of the following methods:
– photogrammetric multi-image acquisition and bundle adjustment;
– SLAM-based evaluation of image sequences;
– measurement of sensor orientation (pose) using photogrammetric control and tie points;
– measurement of sensor orientation by a separate navigation system;

- deriving sensor orientation from the positioning values of a mechanical manipulator;
- connecting individual point clouds using common distinctive features.

6.9.1 Multi-image photogrammetry

The object to be measured is captured in the conventional manner by a multi-image setup and oriented by bundle adjustment. The necessary tie points are acquired automatically. Dense point clouds are usually obtained by SfM approaches and dense image matching. The concept follows the principle of offline photogrammetry, i.e. the surface is only reconstructed in 3D after the overall data acquisition has been completed. Object coordinates or scale definition are usually done via measured reference points or reference scales.

a) Object on rotary table   b) 3D comparison with CAD model
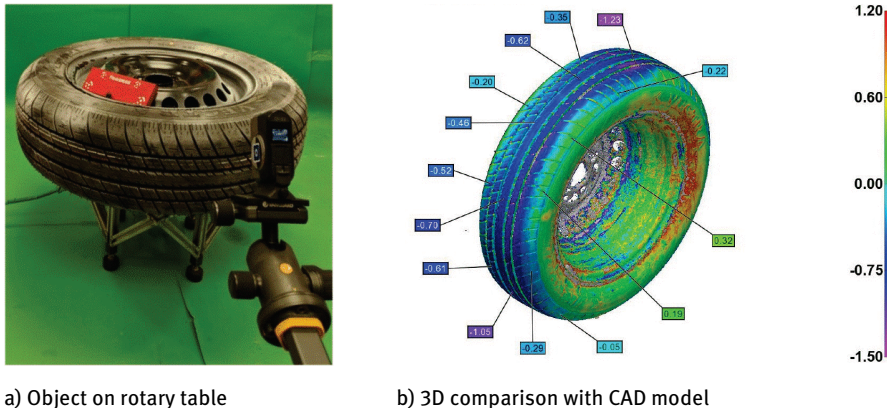
Fig. 6.80: Measurement of complex surfaces by SfM (PhotoGauge).

Fig. 6.80 shows a low-cost system from PhotoGauge, which captures an object lying on a turntable using a smartphone. The background is uniformly coloured so that the images can be automatically masked (see section 5.5.2.3). The resulting all-round image set is processed cloud-based by SfM (Agisoft Metashape) and made available as a dense surface model (mesh) for further processing, e.g. comparison with a CAD model. If there is no natural surface texture, the object must be artificially textured. The smartphone camera is simultaneously calibrated. The measuring accuracy according to VDI/VDE 2617 part 2 is approx. 0.05mm in a measuring volume of approx. 0.5m x 0.5m x 0.3m.

Fig. 6.81 shows the CultArm3D system (Fraunhofer IGD, Verus Digital). A high-resolution camera (PhaseOne, 100 Mpixel) is combined with a polarizing LED ring light and attached to a collaborative robot arm. The object to be measured is placed

on a rotary table and scanned using a combination of table rotation and arm movement. The robot control provides approximate external orientation data for the camera and captured images are evaluated sequentially via SfM. This enables a preliminary 3D model to be displayed during object capture. After image acquisition is finished, a complete model is calculated. Image matching is performed using natural surface features. Depending on the configuration, the resolution (GSD) on the object is up to 15-20 µm in a measuring volume of up to approx. 1m x 1m x 2.0m. The scale is defined via a calibrated marker board next to the object.
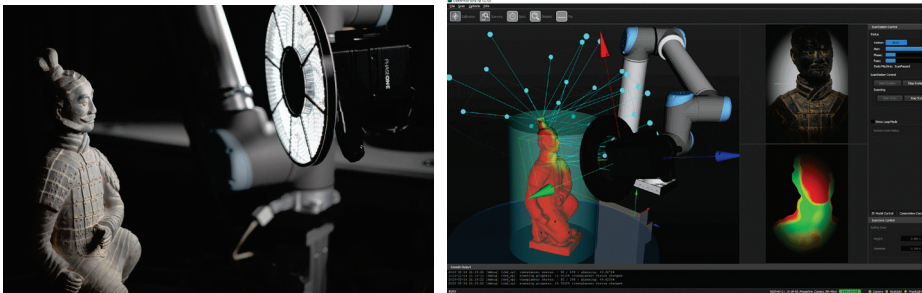
a) Robot-guided camera with light source   b) Robot control and 3D data processing

Fig. 6.81: Robot-based measurement of complex objects by SfM (Fraunhofer IGD).

SfM-based concepts allow high flexibility and are in principle suitable for objects of any size and complexity. Further examples are the recording of anatomical objects (section 8.2.1.4) and insects (section 8.8.2.3).

Several measuring systems use their integrated camera(s) for real-time Visual SLAM (see section 5.5.7.6), optionally combined with other sensor data such as gyroscopes or accelerometers. As examples, the technique is used in mobile devices (section 6.11) or in laser scanners (section 6.8.1). Since the overall accuracy is much lower than for SfM or other precise approaches (see below), the extracted 6DOF values of sensor orientation are usually used as approximations for a final least-squares optimization (e.g. bundle adjustment).

6.9.2 Orientation with object points

6.9.2.1 Orientation with unknown reference targets

Targets can be attached to the object which are suitable for photogrammetric measurement but initially have unknown coordinates. Cameras in surface scanners use these tie points to orient themselves and their scan data relative to the object. If the object structure provides unique surface features, e.g. holes or corners, these

features can be used instead of targets. Local 3D point clouds, scanned from different locations around the object, can then be transformed into a complete model via these points. The output will be in an arbitrary coordinate system. Scale may be defined via a calibrated stereo baseline. Since subsequent data is built on preceding data, this process can lead to a propagation of measurement and orientation errors. However, in a post-processing step all measured tie points can be recalculated by bundle adjustment in order to optimize the orientation of individual point clouds.

6.9.2.2 Orientation with known reference targets

Accuracies can potentially be improved if the targets on the object are first located in a unified coordinate system using a single camera imaging network plus bundle adjustment (section 4.4 and Fig. 6.82b). During subsequent surface scanning, at every scan position each scan patch can be transformed into this network by 3D similarity transformation provided the patch includes at least three reference points, not all lying close to a straight line. Alternatively, the camera orientation can be calculated by space resection. Each scan is thereby locked directly into the common network and not onto the previous scan, hence the potential for improvement.

The accuracy of surface point measurement is quoted at around a few tenths of millimetres in a measurement volume of a few metres, but this is highly dependent on the quality and distribution of the reference targets.

a) HandySCAN 3D Black   b) 3D surface capture using object reference targets
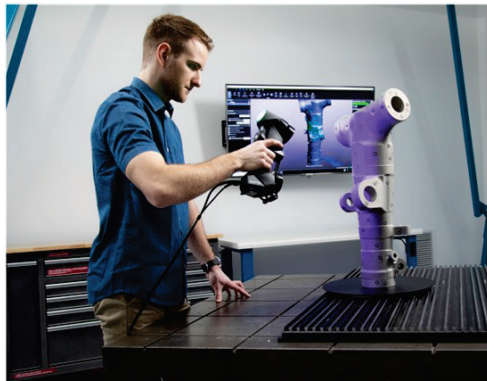
Fig. 6.82: Orientation of a surface scanner using passive reference points (Creaform).

The HandySCAN 3D family of scanners from Creaform incorporates a multi-line projector and dual cameras and operates in a hybrid way (Fig. 6.82). The cameras in the scanner locate the initially unknown reference targets in the camera's coordinate system by triangulation (spatial intersection). Simultaneously, it locates eleven projected blue laser crosses (22 lines) to give a profile of 3D points along the object

surface at a frequency of 60 Hz and an output of up to 1.3 million points per second. When the system is moved to another position, in order to create another laser profile on the surface and so build up a full surface model, the current reference targets in the field of view are again located by the same procedure as before. Although the current camera position is different from the preceding one, the current reference target positions and laser profile points can be added into the previous measurement by a spatial transformation, provided at least three common points have been measured in both camera positions. As new targets come into view, they are transformed into the common reference system and hence extend the spatial reference field.

This is a very flexible procedure in use. For example, if the object is small and light, both object and scanner can be freely moved during scanning, e.g. the object held in one hand and the scanner in the other. The accuracy of surface point measurement is quoted at up to 12μm+20μm/m, but is dependent on the quality and distribution of the targets.

A similar approach is adopted by the Zeiss T-Scan hawk 2 system. Its scanning principle is similar to that of the HandySCAN. The 6DOF orientation of each scan is either performed by reference points in the background of the object, or by calibrated reference points attached to the object. Prior to the actual measurement the scanner acquires a set of images of the object-mounted reference targets which are then processed by bundle adjustment to generate optimal reference coordinates.

These systems are designed to measure geometry without colour information. Fig. 8.18 shows a hand-held scanner which is capable of acquiring colour images using a separate RGB camera to assign a colour value to each measured point.

6.9.2.3 Orientation with known feature points

Instead of photogrammetric targets, unique object features whose 3D position in the reference coordinate system, e.g. a CAD coordinate system, can also be used for sensor orientation. Typical features here, for example, are object corners or drill holes whose reference point (corner, hole centre) can be clearly determined in the measurement images. The achievable accuracy is usually significantly lower than with targets, since detectability and measurement accuracy in the image also depend on effects such as lighting and sharpness of physical edges. However, it is advantageous when there is no need for complex targeting.

6.9.3 Scanner location by optical tracking

In this configuration. the surface scanning system, typically a laser line scanner, has a number of targets (minimum three) attached to its body and these are tracked by an external camera system. Using 6DOF measurement principles (section 4.2.5) the

position and angular orientation of the surface scanner can then be determined. The measurements made by both scanner and tracker must be synchronized and the ability to identify and measure the locating targets must be guaranteed throughout the measurement process. Errors in 6DOF measurement directly affect the quality of the surface measurement.

Fig. 6.83: Optically tracked surface scanner (Zeiss).
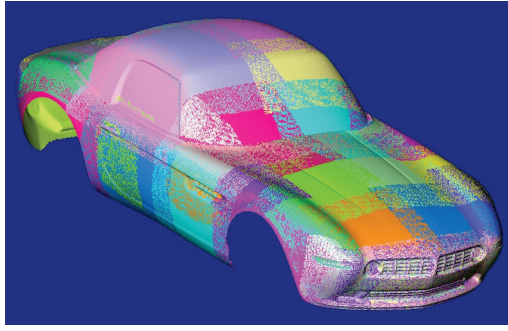
Fig. 6.84: Complete 3D model generated from a number of individual surface scans (Zeiss).

Fig. 6.83 illustrates the concept by showing a three-line camera (section 3.5.4) tracking a laser line scanner. The scanner has a number of LED targets, suitable for this type of camera tracker, which are calibrated in a local scanner coordinate system. These are tracked by the camera which transforms the local LED and laser profile coordinates into its own coordinate system in real time. Fig. 6.84 shows an example of a full 3D model composed of several individual scans. Additional examples are presented in section 6.9.5. See also a similar concept in which a surface scanner is tracked by a laser tracker (section 6.3.2.2).
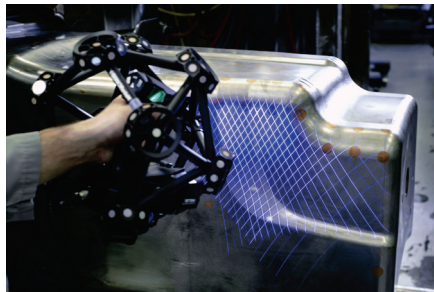
Fig. 6.85: Surface-scanning sensor located by an external stereo camera (Creaform).

The Creaform MetraSCAN 3D (Fig. 6.85) combines optical stereo tracking with a moving (hand-held) surface-scanning sensor based on the HandySCAN principle (see section 6.9.2.2). The scanning sensor is enclosed by a target frame. Multiple reference targets on the frame, which are calibrated in a local sensor coordinate system, are tracked in 6DOF by a stereo camera. The sensor can be moved within the field of view of the cameras and oriented in real-time. If the object is also marked with targets, the stereo camera can be oriented permanently, hence a relative movement between object and camera will be compensated. The scanning sensor can either be hand-held or attached to a mechanical positioning device such as a robot arm. Up to 1.8 million points can be measured per second, to an accuracy of about 0.025 mm.

6.9.4 Mechanical location of scanners

If the surface scanner is connected directly to a mechanical positioning system, then its spatial location and orientation can be determined by the positioning system provided it has the appropriate angle and displacement sensors. Relevant examples of such systems are robots and CMM arms (Fig. 6.86).

To connect the scanner's measurement data with those of the positioning platform, a corresponding calibration is required so that both systems operate in the same coordinate system.

a) Area scanner attached to robot arm (Zeiss) b) Line scanner attached to CMM arm (API)

Fig. 6.86: Orientation of surface scanners by mechanical means.

The accuracy of mechanical orientation with articulated arms (robot, CMM arm) is usually insufficient to transfer the achievable accuracy of the surface sensor to the complete measuring volume of the combined system. Where high accuracies are required in larger volumes, photogrammetric orientation procedures are therefore preferred.

6.9.5 Orientation with external systems and reference points

Robots can be used to drive a surface measuring sensor around complex objects to pre-programmed positions which are relevant for quality control purposes. Here, measurement uncertainties of 1/10 mm or less are a typical requirement. For object volumes greater than ca. 3m x 3m x 3m, the positioning accuracy of a robot is not sufficiently good to deliver directly the exterior orientation of the sensor. In this case, photogrammetric orientation procedures are used to measure the six degrees of freedom (pose) of the sensor with respect to a global coordinate system.

One method requires reference points either on the measured object itself or in the surrounding environment, such as on the floor or a mounting fixture. These are assumed to be in a fixed position relative to the object. By imaging these points, the camera(s) on the scanning device can then apply the method of space resection (section 4.2.3) to determine the 6DOF parameters of each scanning position in the coordinate system of the reference points.
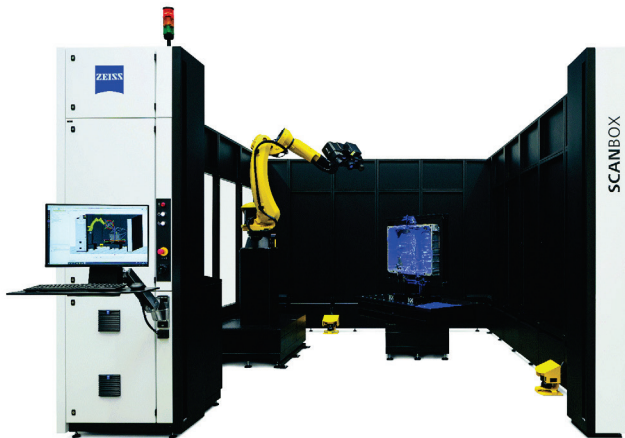
Fig. 6.87: Surface-measuring sensor mounted on a robot arm (ZEISS ScanBox).

Alternatively, separate oriented cameras can be installed around the measuring space. Using spatial intersection (section 4.4.7.1), these can locate targets on the scanning head which have been calibrated with respect to the scanning sensor, hence locating the sensor in full 6DOF with respect to the coordinate system of the cameras.

A final transformation of measured point clouds into the object (workpiece) coordinate system is usually done by identifying specific object features with known positions in the CAD model, for example by using special target adapters positioned at related CAD positions (section 6.2.2.3), by manual probing or by feature matching procedures.

In the ScanBox system from ZEISS (Fig. 6.87), the company's ATOS fringe projection system is mounted on a robot arm. At every measuring position, enough reference points must be visible in the sensor's field of view so that it can be oriented with respect to the global coordinate system. In principle, there is no limit to the size of the measuring volume provided the reference points have a distribution and accuracy appropriate for the application. Since the robot-mounted sensor combination operates independently, several measurement robots can operate in parallel.

C: Cameras   S: Sensors with locator L   R: Robot   P: Reference points
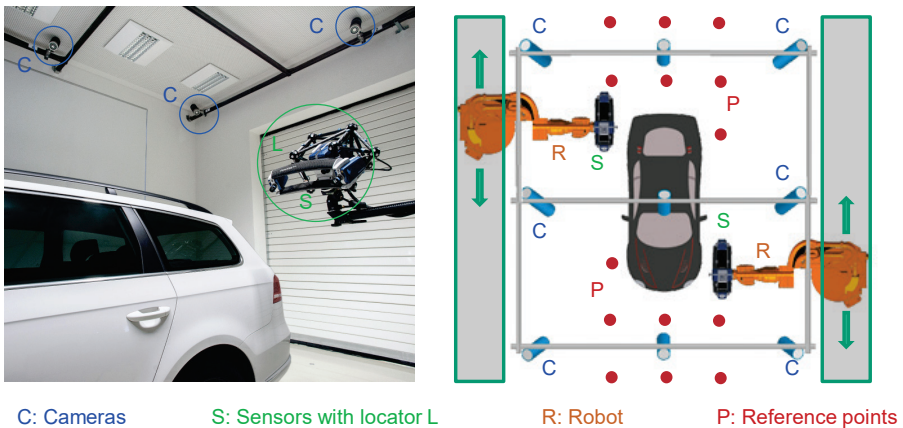
Fig. 6.88: Surface sensor mounted on a robot arm (Hexagon 3D Arena).

Fig. 6.88 shows the 3D Arena system from Hexagon. This has several of high-resolution cameras (see Fig. 3.121b) which detect and measure targets fully automatically. The targets are mounted on stable locations in the space surrounding the measured object and are located in a common coordinate system by prior photogrammetric measurement using the cameras. The surface is measured by one or more stereo fringe projection systems mounted on robot arms. The fringe projection systems are contained within an additional local reference field (locator) which is also observed by the cameras. By determining the 6DOF parameters of the surface measurement sensor relative to the reference points at the moment of surface measurement, there is no requirement for potentially uncertain mechanical positioning information from the robots and the required measuring accuracy can be delivered across the whole measurement volume. A surface measuring accuracy of 0.1–0.3 mm is reported when using 16 cameras in a measurement volume of 5 m x 3 m x 2.5 m.

6.9.6 Connecting point clouds (registration)

If no absolute orientation or transformation parameters are available for individual point clouds, unique object features within the point clouds, and with known positions in 3D space, can be used. These features could, for example, be suitable targets or artefacts for which a distinct reference point can be derived from the neighbouring points in the cloud. In this way, unique tie points are generated for each single point cloud. Alternatively, any other distinct features, such as corners or drill holes, can be used if they can be detected and identified. Where distinct features are not available, orientation is possible using local surface curvatures which, in analogy to interest operators for image matching (see section 5.4.3), enable a unique matching in neighbouring point clouds. Yet another technique is global least-squares optimization which minimizes the distances between two point clouds (ICP, see section 6.9.6.2). In all registration methods, the surface features must contain the information necessary to determine all six transformation parameters. Some point clouds cannot be connected using surface geometries. These include, for example, point clouds which represent single planes or multiple parallel planes, or which represent single surfaces with a rotationally symmetric axis such as spheres, cylinders and cones.

6.9.6.1 Registration with 3D reference targets

For point cloud registration, targets can be used whose reference point can be extracted from the point cloud. The well-known checkerboard-type targets (see section 6.2.1.3) consist of black and white areas which can be detected by the intensity values of the point cloud. Within a best-fit plane of neighbouring points the intersection point of the pattern edges can be defined. Alternatively, spherical targets (usually with known radius) can be used where the centre of a best-fit sphere can be calculated (see section 2.3.2.3) and can then be used as a local reference point for orientation (example in Fig. 6.89). Both techniques are commonly part of the registration process in terrestrial laser scanning. In comparison with photogrammetric reference targets, the overall accuracy of feature point registration is significantly less.

a) Spherical target

b) Measured point cloud and best-fit sphere
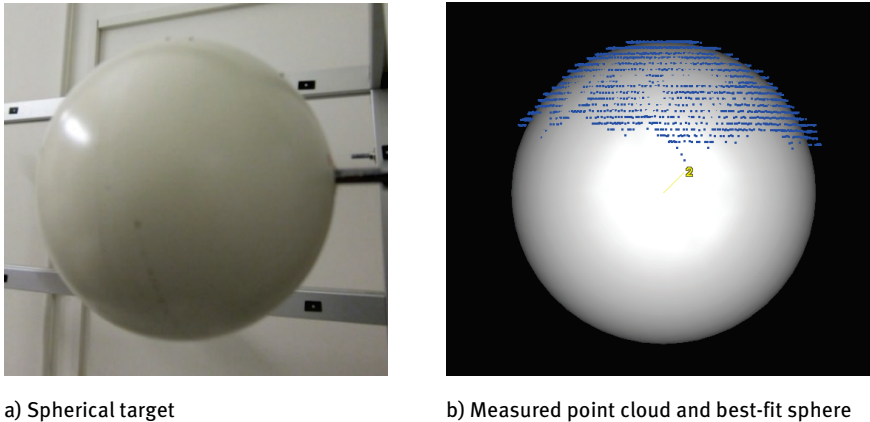
Fig. 6.89: Measurement and analysis of a spherical target.

6.9.6.2 Iterative closest point (ICP)

The most well-known algorithm for registering two point clouds (cloud-to-cloud matching) based on overlapping surfaces is the *iterative closest-point* algorithm (ICP). As the name suggests, it is an iterative algorithm which requires a starting value for the transformation, i.e. the two point clouds must be roughly aligned. The following discussion considers the case of aligning point cloud F to point cloud G, where F consists of n points f_i and G of m points g_j (Fig. 6.90). The ICP starts by searching for the closest point for every point f_i among all points in G. The closest point is the point g_j which has the shortest Euclidean distance to f_i. All closest point pairs (f_i, g_j) are assumed to be corresponding points. From this set of corresponding points a rigid body transformation is computed as described above. This transformation is applied to point cloud F. The Euclidean distance $d_i(f_i, g_j)$ for all corresponding point pairs is computed after transformation. If the overall error, e.g. the RMS of all d_i, is below a pre-set threshold, the iteration is aborted. Otherwise the iteration is continued by searching again for the closest points and establishing new corresponding points. See Fig. 6.90 for a flow diagram of the algorithm.

The computationally most intensive step of the algorithm is the search for the closest point. This problem is also known as the *nearest neighbour search*. A naïve implementation would compute distances $d_i(f_i, g_j)$ for every point g_j of G and select the point with smallest d_i. This would require m computations. Well-known spatial search structures such as the kd-tree can reduce the search time to $\log m$. If a surface meshing is available for point cloud G, the method can be adapted to consider distances between points and surfaces. This typically improves the accuracy of the alignment.
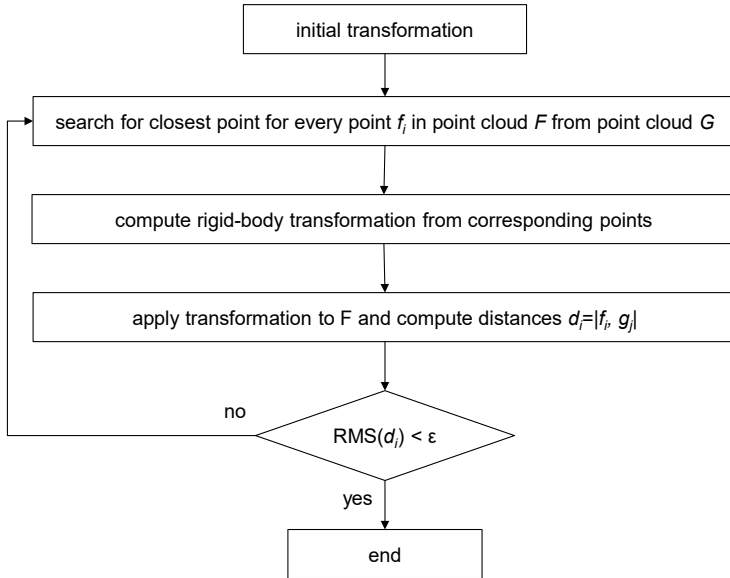
Fig. 6.90: Flow diagram for iterative closest-point algorithm.

In addition to point-to-point distances, the ICP can be modified to make use of point-to-plane or plane-to-plane distances. As an alternative to a simple distance function, it is also possible to use distances to a surface centroid or parallelism of normal vectors of two planes or combinations of these.

In analogy to the least-squares matching of two image patches (section 5.5.3.5) 3D point clouds can be matched (*least squares 3D surface matching*). Under the assumption that two overlapping point clouds have corresponding regions $f(x,y,z)$ and $g(x,y,z)$ which are identical apart from random noise $e(x,y,z)$, then the following relationship can be established:

$$f(x,y,z) - e(x,y,z) = g(x,y,z) \tag{6.13}$$

Reference region $f(x,y,z)$ and search region $g(x,y,z)$ each consist of single surface elements, e.g. discrete 3D points. The geometric transformation between both regions is achieved by a 3D similarity transformation with seven parameters (section 2.2.4). In an iterative Gauss-Markov adjustment (section 2.4.2), the Euclidian distances between corresponding surface elements are minimized. The method requires suitable starting values for the transformation parameters and correspondences between the regions and can be extended to the simultaneous matching of multiple regions.

6.9.7 ICP-based scanner devices

Different suppliers offer hand-held 3D scanners. In contrast to high-end industrial scanners (example in Fig. 6.82) these scanners are usually designed for smaller objects and lower accuracy demands. As examples, Fig. 6.91 shows two systems that permit a hand-held digitization of object surfaces. Both project a point pattern which is observed by a camera with a lateral offset to the projector. 3D coordinates of projected points are calculated by triangulation, i.e. camera and projector must be calibrated and oriented with respect to each other. When scanning the surface, the device must be guided slowly and evenly over the surface. The registration of individually measured point clouds is based on matching of local object surfaces by ICP (section 6.9.6).

a) Artec Leo
b) Mantis Vision F5 Short Range

Fig. 6.91: 3D hand-held scanners.

The Artec Leo system (Fig. 6.91a) projecting a point pattern using blue LED technology. It is designed for a measurement space of about 90 mm x 70 mm at a range between 0.17 m and 0.35 m. The achievable accuracy is reported to be from 0.5 mm to 2 mm depending on the surface properties. In this system, point clouds and surface models can be coloured using colour information collected by a second RGB camera. The Mantis Vision F5 system (Fig. 6.91b) uses infrared LEDs for pattern projection. Depending on system model, the measurement range lies between 0.3 m und 4.5 m. At a distance of about 0.5 m the achievable accuracy lies in the millimetre range.

6.10 Dynamic photogrammetry

Image sequences can provide chronological records of an object's spatial movements (deformations, trajectories, velocity and acceleration curves). In addition to recording

time, suitable measures must be taken to identify and image discrete object points. Examples of dynamic applications are:

– recording crash tests in the automotive industry;
– investigating turbulent flow in fluids or gases;
– surveying structural deformations (buildings, bridges, ...);
– material testing under mechanical or thermal stress;
– vibration analysis;
– calibrating and evaluating robot movement;
– human motion analysis.

Image sequences usually imply a sequential series of multiple images which record an object movement at an appropriate frequency. In object space, a stationary set of reference points is required to which the object movements can be related. It can also be used for image orientation, e.g. with moving camera stations.

Slow movements, such as the deformation of a cooling tower due to changes in direction of the sun's illumination, can be recorded as a sequential set of 3D reconstructions, each made using the offline photogrammetric process i.e. using a single conventional camera followed by a standard photogrammetric object reconstruction.

The recording of high-speed image sequences can be achieved using video cameras (camcorders) and digital high-speed cameras (section 3.5.3). Dynamic spatial modelling can be achieved by using multi-camera systems or a single camera with stereo image splitting (section 3.4.3.9). Cameras with an integrated memory for storing images or videos have a limited maximum recording time. Cameras with integrated point measurement are capable of almost unlimited recording of targets without the need to store images.

Methods for tracking and matching features in image sequences are discussed in section 5.5.7.

6.10.1 Relative movement between object and imaging system

6.10.1.1 Static object
Relative movement between a static object and moving camera occurs in a number of applications:

– hand-held photography;
– photography from an airborne vehicle;
– image acquisition from a moving car;
– image acquisition on unstable ground (oscillations, vibrations).

Stationary objects can be recorded in an offline process by sequential imaging with only one camera. Movements of the camera, during exposure, cause an image blur $\Delta s'$, dependent on velocity, exposure time and image scale:

$$\Delta s' = \frac{\Delta t \cdot v}{m} \tag{6.14}$$

where
Δt: exposure time
v: speed of moving camera
m: image scale factor

Blurring due to image motion results in a decreased modulation transfer in the direction of movement. The maximum permitted image motion can therefore be expressed as a function of resolving power. Investigations of aerial cameras have shown a maximum tolerable image motion of:

$$\Delta s'_{max} = 1.5 \cdot RP^{-1} \tag{6.15}$$

where
RP: resolving power of the sensor [L/mm]

For applications in close-range photogrammetry, a maximum image blur of 1 pixel might, for example, be tolerated.

The maximum permitted exposure time in a given situation can be derived from (6.14) as:

$$\Delta t_{max} = \frac{\Delta s'_{max} \cdot m}{v} \tag{6.16}$$

Example 6.3:

A row of houses is imaged from a moving car ($v = 30$ km/h = 8.33 m/s) at image scale factor $m = 2000$. The maximum image blur should not exceed 1 pixel (e.g. 6 µm). It is necessary to find the maximum exposure time Δt and the resulting blur in object space:

Solution:

1. Permitted image motion: $\Delta s'_{max} = 6$ µm

2. Maximum exposure time: $\Delta t_{max} = \dfrac{0.006 \cdot 10^{-3} \cdot 2000}{8.33} = 0.0014\,s \approx 1/700\,s$

3. Blurring in object space: $\Delta S = 2000 \cdot 0.006 = 12$ mm

---

In practice, exposure times must also be selected on the basis of illumination conditions, available lens aperture, required depth of field and film or sensor sensitivity (ISO number).

6.10.1.2 Moving object

The relationships discussed above for blurring due to camera motion are also valid for a moving object. However, 3D measurement of a moving object requires at least two synchronized cameras.

Depending on object velocity, synchronization errors lead to positional errors $\Delta s'$ which are proportional to the corresponding distance moved ΔS.

$$\Delta s' = \frac{\Delta t \cdot v}{m} = \frac{\Delta S}{m} \qquad (6.17)$$

where

Δt: synchronization error
v: object velocity
ΔS: distance moved
m: image scale factor

If the object is moving parallel to the baseline between two cameras, the positional error is effectively an x-parallax error $\Delta px'$. According to the standard case of stereo photogrammetry shown in Fig. 6.92, it is clear that the movement ΔS between times t_0 and t_1 results in a measurement of virtual point P^*. The corresponding error ΔZ in the viewing direction is given by the following expression (compare with eqn. 4.88):

Fig. 6.92: Lateral and range errors caused by synchronization error.

$$\Delta Z = \frac{h}{b} \Delta S = \frac{h}{b} \cdot m \cdot \Delta px' \qquad (6.18)$$

where

h: object distance
b: stereo base
$\Delta px'$:   x-parallax error
m: image scale factor

In the direction of movement, the lateral error ΔX is as follows:

$$\Delta X = \frac{x'}{c} \Delta Z \qquad (6.19)$$

where

x':   image coordinate in left image
c: principal distance

Conventional cameras can be synchronized by an electrical pulse linked to the shutter release. For the synchronous image input from multiple video cameras there are frame grabbers with multiple parallel A/D converters, e.g. RGB inputs, whose

synchronization signal is used to control the cameras. The synchronization pulse can also be generated by one camera in master/slave mode.

Example 6.4:

For the application in example 6.2 (v = 30 km/h = 8.33 m/s; m = 2000) stereo images with two digital video cameras are also to be recorded (b = 1.5 m, c = 8 mm, h = 16 m, x' = 4 mm). Technical limitations in the installation require a synchronization error of Δt = 1/500 s to be taken into account:

1. Distance moved: ΔS = 1/500 · 8.33 = 0.017 m

2. Lateral image error: $\Delta px'$ = 0.017/2000 = 8.5 · 10^{-3} m = 8.5 μm

3. Error in viewing direction: ΔZ = 16/1.5 · 2000 · 8.5 · 10^{-3} = 181 mm

4. Error in direction of movement: ΔX = 4/8 · 181 = 90 mm

The example demonstrates the serious effect of a synchronization error on the quality of object coordinates. The lateral image error of 8.3 μm is of the order of 1–2 pixels.

6.10.2 Recording dynamic sequences

6.10.2.1 Camera system for robot calibration

In general, robot calibration is used to check the robot arm position with respect to its nominal position, and to derive correction data for angle and distance adjustments. This requires the determination of the 6DOF parameters (pose) of the robot's end effector, a task which can be solved by online or offline photogrammetry.

a) Robot calibration with three-line camera (Nikon Metrology)
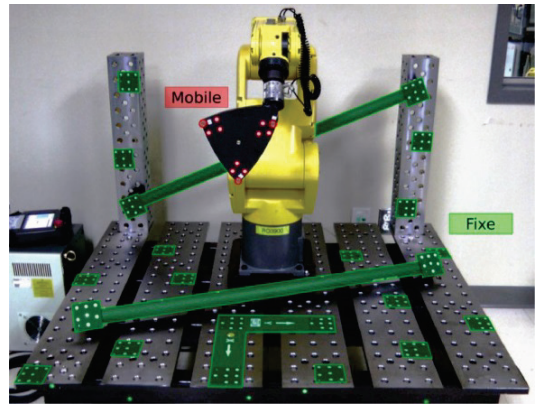
b) Robot calibration by offline photogrammetry (Filion et al. 2018)

Fig. 6.93: Photogrammetric robot calibration and adjustment.

A possible solution is to have one or more digital cameras fixed in such a way that they observe the workspace of the robot. The robot arm is fitted with a calibrated reference object which is moved to various positions in space. By space resection (section 4.2.3), the camera(s) determine the position and orientation of the reference object and hence of the robot tool point.

The 6DOF parameters between the robot's target frame, and an additional reference field if necessary, can be calculated if both point fields are visible to the camera at the same time (see section 4.2.5). The robot can be calibrated in motion, i.e. its kinematic properties are taken into account in the calibration. Fig. 6.93a shows a target frame equipped with LEDs which is tracked and measured by a high-frequency three-line camera (compare with Fig. 3.128). The typical workspace (measuring volume) is about $3 \times 3 \times 2m^3$ within which an object accuracy of 0.1 mm can be achieved.

Fig. 6.93b shows a set-up for photogrammetric robot calibration using offline photogrammetry. Again, typically a single or stereo camera takes multiple images of a moving target frame attached to the robot's end effector. The robot moves to different spatial positions and at each position a separate set of images is recorded. The achievable accuracy is higher than for online measurements, but more time is required for the measurement process and it does not take into account all the robot's kinematic characteristics.

Robot-based systems like those presented in section 6.9.5 can use external measurements based on control points as part of a feedback loop which enables them to drive to more accurately known locations. This type of measurement feedback control is also implemented using laser trackers (section 6.3.2.2).

6.10.2.2 High-speed 6 DOF system

Fig. 3.121b shows a high-speed camera system used to record the dynamic changes in position and angular orientation of a moving object. The camera incorporates an internal FPGA processor with software which can automatically locate bright circular targets in the image. An example application is the AICON WheelWatch (see Fig. 6.94), used to monitor wheel movements on a moving car, for example to analyse the steering mechanism or the effect of road impacts.

The principle of operation is again based on a space resection, as described in section 4.2.3. By monitoring two separate sets of targets, one on the car body around the wheel arch and one on the wheel itself, the movements of the wheel relative to the body can be determined in real-time and at high speed (up to 50 m/s). This means that a dynamic measurement sequence can be of unlimited duration since there is no image recording but instead a continuous output of 6DOF values.
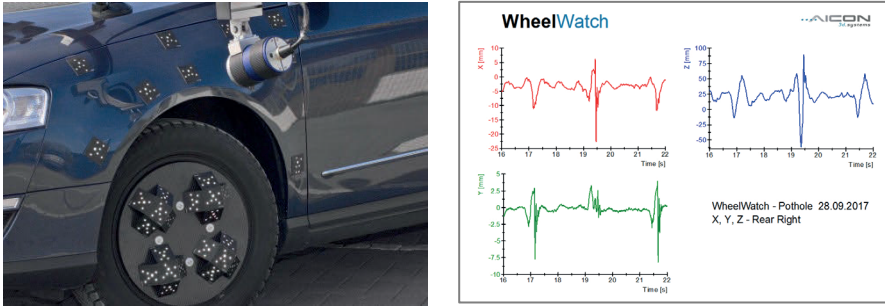
a) High-speed camera and reference targets   b) Processing example

Fig. 6.94: 6DOF real-time measurement of wheel movement (AICON WheelWatch).

t = 0 ms   t = 5 ms   t = 25 ms

t = 50 ms   t = 75 ms   t = 100 ms

t = 125 ms   t = 150 ms   t = 200 ms
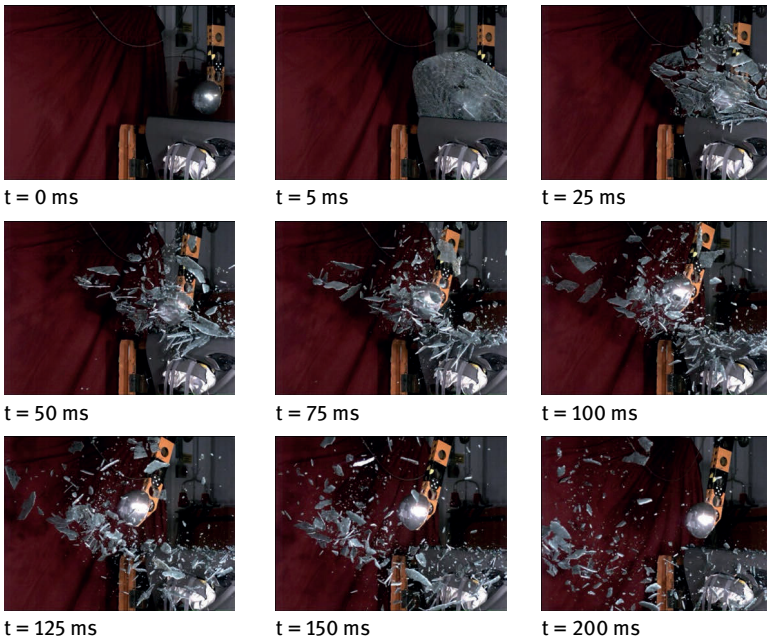
Fig. 6.95: High-speed video sequence (Porsche).

6.10.2.3 Recording with high-speed cameras

The recording of high-speed image sequences is also known as *cinematography* for which digital high-speed cameras are used (section 3.5.3). Image sequences for 3D analysis can be captured by multiple synchronized cameras or a single camera fitted with a split-beam mirror for stereo viewing (section 3.4.3.9). Fig. 6.95 shows part of a high-speed video sequence to determine the path of glass fragments in the simulation of a head striking a car windscreen. The sequence was recorded with two synchro-

nized cameras (Weinberger Speedcam Visario, 1024 x 768 pixels, 2000 frames per second) in order to determine the three-dimensional paths of the glass particles.

6.10.2.4 Particle Image Velocimetry

Particle Image Velocimetry (PIV) encompasses various methods for measuring the spatial trajectories of multiple particles in media such as air, gas and water. Typical applications are flow analysis inside wind tunnels, gas turbines and high-pressure nozzles or the analysis of stream velocities and directions in fluids. The term *Particle Flow Tracking* is also often applied in this case.

Fig. 6.96 illustrates an experimental set-up for recording spatial trajectories by a multi-camera image sequence. Three video cameras are used to observe particle flow in a water tank. A light-section projector creates a "box" of light in which particles are visible. The image sequence is stored at video rate on a video recorder. The multi-media path of the imaging rays through the transparent wall of the tank is taken into account by the photogrammetric model (section 4.6).
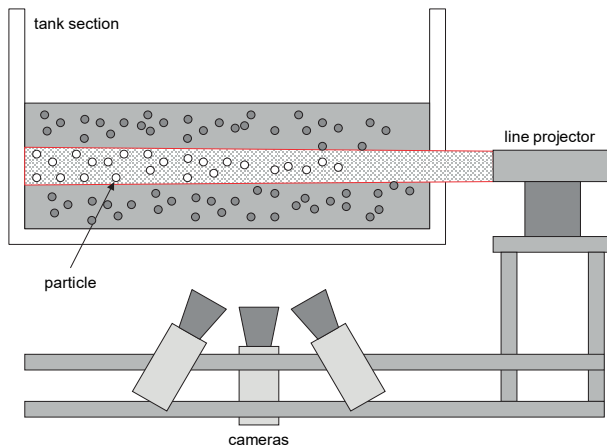
Fig. 6.96: Experimental set-up for multi-image recording of particle flow (after Maas 1993).

For this application the key problem lies in solving the correspondence, at a given point in time, between the three images of a large number of particles which are recorded as bright points (Fig. 6.97). The correspondence problem is solved by an image matching process based on epipolar geometry as outlined in section 5.5.4.

As a result of the photogrammetric analysis, the 3D coordinates of all particles throughout the complete image sequence can be determined. Particle trajectories can subsequently be derived (Fig. 6.97), although this requires solution of the correspondence problem between consecutive images in the sequence.

image sequences   extraction of particles   3D velocity field
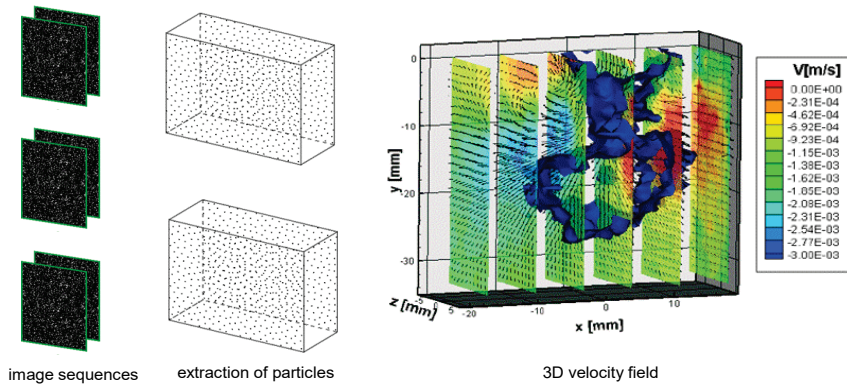
Fig. 6.97: Particle image velocimetry (Dantec Dynamics).

6.10.3 Motion capture (MoCap)

A number of commercial photogrammetric systems (motion capture systems) exist for the 3D recording and analysis of human movement e.g. for applications in sport, medicine, ergonomics and entertainment. They are based either on off-the-shelf digital video cameras or specialized measurement cameras with integrated processors for the online acquisition of marker (target) locations.

a) Freely configurable camera setup   b) Example of 3D motion analysis
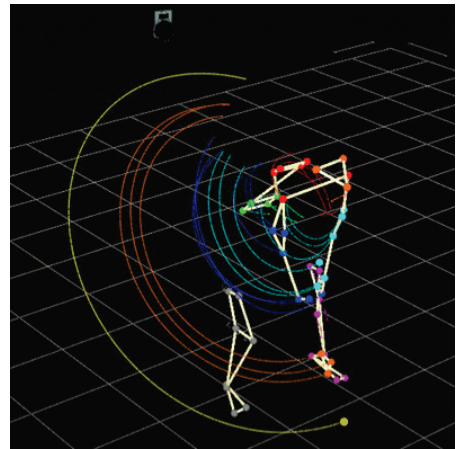
Fig. 6.98: Use of a MoCap tracking system to analyse golf strokes (Qualisys).

The cameras can be freely positioned, e.g. on tripods, attached to a mobile base or mounted in fixed locations around a measurement space (Fig. 6.98). System calibration and orientation is generally made using local reference point arrays or targeted scale bars which are moved around the measurement area.

Retro-reflective targets placed on the body are normally used to record movement. They are illuminated from the camera locations and provide reliable target identification and tracking. The results of a measurement are the spatial trajectories of the target points which typically support a motion analysis or computer animation. Many companies offer commercial multi-camera MoCap systems, for example Qualisys, Vicon and ART.

6.11 Mobile measurement platforms

6.11.1 Mobile mapping systems

6.11.1.1 Outdoor mapping

Mobile mapping systems (Fig. 6.99) are moving platforms equipped with a range of sensors (cameras, laser scanners, radar, sonar, odometers, positioning systems). They are used to measure transport routes (roads, railways, waterways) in both open and built environments and, commonly, the built environment itself. The objective is a continuous sensing of the scene with simultaneous registration of exterior sensor orientation using GNSS and INS.
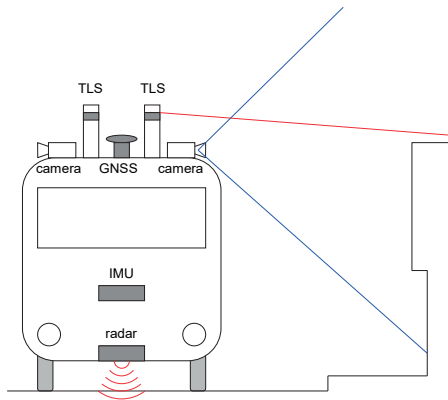
Fig. 6.99: Schematic mobile mapping system.

Fig. 6.100: Mobile mapping system for road modelling (Trimble).

Mobile mapping systems for railway modelling are used to determine rail geometry and areas of damage, as well as bridge and tunnel measurement. Waterways, river beds and river banks can be measured by terrestrial laser scanners on board floating

platforms with additional hydrographic sensors, e.g. echo sounders, on board. To record and check road surfaces and edges, laser and radar systems are typically used (see example in Fig. 6.100). In populated areas it is mostly building façades and 3D city models which are generated. These are used for both tourism and technical purposes, in geoinformation systems and internet portals. For the mapping of areas which are difficult to access and cannot be reached by normal vehicles, systems are available which can be carried as backpacks (examples in Fig. 6.102).

A particular challenge for mobile mapping systems is the calibration and orientation of the individual sensors in a common coordinate reference system. For this purpose, large area reference fields must normally be created. These should provide a large number of reference points which can be recorded by the various sensors such as image and laser systems.

Fig. 6.101 shows an example of a mobile-mapping system which uses cameras only. Such systems use stereo matching in recorded image sequences to reconstruct the 3D model of the environment, typically surface models of roads and embankments.

a) Vehicle with cameras   b) Surface reconstruction by image matching

Fig. 6.101: Mobile mapping with stereo photogrammetry (iNovitas).

6.11.1.2 Indoor mapping

Inside buildings or other closed environments, GNSS signals are not available. Here mobile mapping systems must obtain their orientation parameters by means of image-based and/or laser-scanning techniques. One solution is based on a SLAM algorithm (simultaneous localization and mapping, see section 5.5.7.6) which uses image sequences for the continuous orientation of cameras within an unknown environment. The quality of the recorded data can be significantly improved by means of measured control points (targets), additional sensors on the mapping platform such as odometers, and appropriate measurement strategies for the point clouds, e.g. loop closure. In 3D space the typical resolution (point density) is of the order of 5 to 10 mm, with absolute accuracy of the order of 20 to 50 mm.

a) Leica Pegasus

b) Gexcel Heron

Fig. 6.102: Mobile mapping in backpack.

Fig. 6.102 shows examples of backpack systems which are equipped with one or more 2D laser scanners, cameras, GNSS and IMU modules.

a) NavVis Trolley
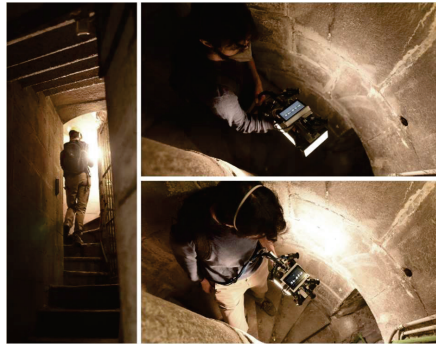
b) Fisheye system for narrow spaces (Politecnico di Milano)

c) Hand-held laser scanner (Leica)

d) Autonomous robot platform with laser scanner (Boston Dynamics, Leica)

Fig. 6.103: Mobile mapping for indoor measurement.

Fig. 6.103a shows an example of a commercial indoor mapping system, in this case the NavVis M6 device comprising multiple laser scanners and cameras. Fig. 6.103b is an example of a hand-held system based on five fisheye cameras that is specially designed for narrow spaces such as stairways, tunnels or caves. Fig. 6.103c and d show a mobile laser scanner that can be used as a hand-held device as well as on robots or other platforms such as drones. The displayed examples can also be applied in outdoor environments.

6.11.2 Close-range aerial imagery

Images from low flying aerial platforms are used, as example, in the following applications:

– archaeological surveys;
– volumetric measurement of spoil heaps and waste disposal sites;
– roofscape mapping;
– large-scale mapping of inaccessible areas;
– test site mapping for geographical investigations;
– biotope monitoring;
– accident recording;
– inspection, monitoring and photographic documentation of construction sites;
– reconnaissance.

The acquisition of low-altitude aerial images is, in the first instance, a problem of flight technology (choice of sensor platform, navigation, flight authorization). Photogrammetric processing is based on standard methods.

6.11.2.1 Aerial systems

Manned aircraft are suitable for large-scale aerial images if image blur at their minimum speed is tolerable. For example, with an exposure time of $\Delta t = 1/1000$ s, flying speed $v = 180$ km/h and image scale factor $m = 1000$, image blur is approximately 50 μm according to eqn. (6.14). Compensation for image blur is widely implemented in aerial cameras by *forward motion compensation* (FMC) or *time delayed integration* for digital cameras. Low-speed ultra-light aircrafts (gyrocopters, microlights, $v_{min} = 40$ km/h) provide advantageous flight properties but lower payloads (<100 kg). Manned helicopters are seldom used due to their high operational costs.

For large scale mapping of smaller areas, a variety of unmanned platforms are available to deploy cameras and scanners. These include balloons, blimps and kites but the most commonly used is the continually and rapidly developing technology relating to pilotless aircraft and helicopters. These are commonly known as *drones* but have various other names such as *remotely piloted vehicles* RPV, *unmanned aerial*

vehicles UAV, *unmanned aerial systems* UAS and *remotely piloted aircraft systems* RPAS.

Table 6.2: Unmanned platforms for low-altitude aerial imagery.

Type	Typical (max.) altitude [1] [m]	Payload [kg]	Range [2] [m]	Number of operators	Max. flying time [3] [min]
Fixed-wing aircraft	50–120 (500)	0.2–15	5000	1 – 2	30 – 120
Helicopter	10–120 (150)	1–20	5000	1 – 2	10 – 120
Multicopter	10–120 (150)	0.3–7	3000	1	up to 60
Hot-air balloon, blimp	10–100 (1000)	10–50	5000	3 – 5	60
Gas balloon, blimp	10–100 (1000)	10	stationary	2 – 3	unlimited [4]
Kite	50–100 (300)	10–50	stationary	2	unlimited [4]

[1] limited to 120 m in Europe; [2] usually required to stay in view; depending on sender/receiver; [3] payload dependent; [4] weather dependent.

A particularly strong development area is represented by remotely controlled helicopters. With automatic flight stabilization and navigation, they can carry a payload of up to 18 kg which permits, for example, the simultaneous use of digital cameras and airborne laser scanning systems. Some also use compression-ignition motors which allow for longer flying times. High-end multicopters with four to eight rotors (example in Fig. 6.105) have been designed for higher payloads, but these are usually powered by batteries and have a more limited maximum flying duration. As an alternative, low-cost multicopters are available with reasonable flight stability but more limited payloads and flight durations (see Fig. 6.106).

Fig. 6.104: Helicopter (Aeroscout).

Fig. 6.105: Quadcopter with laser scanner (Riegl).

Fixed-wing UAVs (example in Fig. 6.107) are suitable for aerial coverage of larger regions or for following traffic routes. They are more stable than multicopters in strong winds and are easy to manoeuvre. As drawbacks, they generate more vibration and roll, although roll can partly be compensated by designing the flight path to ensure the imaging has a higher side overlap (usually 80%).

The flight ranges which are technically possible depend not only on the maximum flying time but also on the data transmission systems used and can be up to several kilometres. Whether a visual line of sight (VLOS) or beyond visual line of sight (BVLOS) is permissible is regulated by law.

Fig. 6.106: Quadcopter (DJI Phantom 4 RTK).

Fig. 6.107: Fixed-wing UAV (SenseFly).

In practice, multicopters are preferred since their requirements for takeoff and landing are very flexible and they can hover in fixed locations. The advantages and disadvantages of various unmanned aerial systems are summarized in Table 6.3.

Aerial platforms have also benefitted from advancements in sensor technology which, amongst others, help to improve stabilization and navigation. Useful advances include miniaturized GNSS receivers, gyroscopes, altimeter, powerful rechargeable batteries and existing electronics for wireless image and data transfer. These complementary technologies have helped to create a mass market for UAVs which, in turn, has opened up numerous new fields of application for photogrammetry. This widening spectrum of applications has created a parallel need for regulations, with national laws now in place in many countries to control flight permissions and pilot certification.

6.11.2.2 Sensor technology

For widely used low-cost drones, the on-board sensor is often a light and inexpensive fixed-focus digital camera, supplied in different designs (RGB, multi-spectral, thermal). Image data is usually stored within the camera, because remote data transfer at typically required resolutions (usually > 20 Mpixel) is costly and demanding. Most of the relevant off-the-shelf cameras have only limited suitability

for metric purposes, as their lower construction quality, coupled with normal UAV flight patterns, ensures that calibration can only achieve moderate accuracy levels. In addition, geometric quality is also affected by vibration, motion blur and rolling-shutter effects (section 3.4.2.2).

Table 6.3: Pros and cons of different UAV platforms.

	Fixed wing	**Multicopter**	**Helicopter**	**Balloon / Blimp**
Technology				
Handling / calibration	simple	simple	complex	simple
Controlability	++	++	++	–
Sensitivity to wind	o	+	–	++
Robustness	++	+	+	++
Sound level	low (e-motor)	low (e-motor)	high (kerosene)	low
Transportation	trunk	trunk	trunk	small trailer
Flying duration	++	o	+	++
Costs				
Purchase	+	– to +	++	++
Maintenance	o	+	++	+
Repair after crash	o	++	+	O
Application				
Velocity / area efficiency	++	o	+	–
Minimum flying height	+	–	–	–
Pointing ability	–	++	+	++
Gimbal	rarely	default	default	default
Payload vs. weight	o	+	++	+
Take-off and landing	landing site	everywhere	everywhere	landing site

In addition to the photogrammetric camera, the aerial platform must carry additional components such as remote control, navigation devices, video camera or camera housing. The camera suspension (gimbal) ensures a stable camera alignment by motorised compensation of any platform movements. It also permits adjustment of the viewing direction, e.g. to acquire oblique images.

Even with skilled operators, navigation remains a practical problem. UAVs usually have a video transmission with which the aircraft can be driven into position,

both automatically and manually. Barometric altimeters can be used in order to ensure a constant altitude. UAVs with GNSS receivers can, in principle, be navigated autonomously, i.e. within the GNSS accuracy they can fly a predefined course. Note that GNSS methods with real-time kinematics (RTK) require a reference station within receiving range and thus allow the reduction of control points. As for aerial image acquisition with real aircrafts, all on-board components must be calibrated with respect to each other (boresight calibration). For professional surveying flight, a skilled (certified) pilot is usually required to control takeoff, landing and emergency cases whilst the actual flight path is run automatically according to a pre-defined program.

For some time now, airborne laser scanning (ALS) systems have also been available with a weight and measuring distance corresponding to typical UAV requirements (example in Fig. 6.106). The ALS systems can additionally be combined with an RGB camera. For the accurate determination of the external orientation of each laser signal, the systems are equipped with high-precision IMU components.

6.11.2.3 Flight planning

Aerial flight planning is usually implemented with software tools provided by the UAV or third-party suppliers. The most relevant planning criteria are:

– Ground sampling distance (GSD):

The ground sampling distance GSD, or pixel size on ground, is primarily defined by the application. It is calculated from the physical pixel size of the imaging sensor and the image scale (see below). Physical rules such as sampling theorem, or modulation transfer function, should be taken into account. With current systems, the typical GSD ranges from 5 to 20 mm. Achievable accuracies in object space lie in the region of 0.5–1 pixels in XY and 2–3 pixels in Z (height).

– Choice of camera and lens:

Apart from scale-dependent parameters, the choice of camera is very dependent on the maximum payload of the platform. Additional selection criteria are memory capacity, control of aperture and exposure time, spectral range and light sensitivity. The focal length selected will be a function of GSD, image scale, flying height and overlaps.

– Flying height:

Determination of a suitable flying height is dependent not only on scale requirements but also on other technical and legal side issues. For example, the maximum flying height of UAVs can differ considerably and the operational requirements of the aviation industry must be followed.

– Image scale:

The image scale resulting from eqn. (1.1) is a function of flying height and focal length, or GSD and sensor pixel size.

– Oblique images:

For special applications, oblique images may be necessary. It should be noted that image scale and GSD become less favourable with increasing distance from the nadir direction, depending on the angle (see Fig. 6.108).
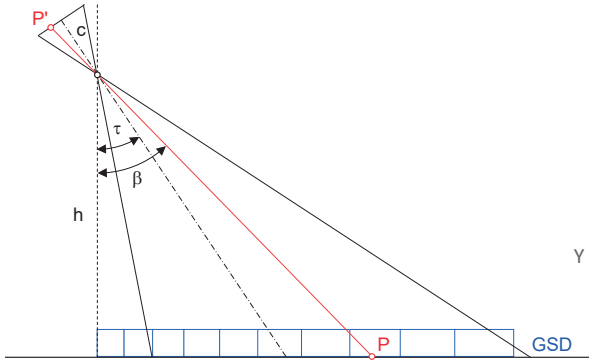
Fig. 6.108: Imaging geometry for oblique images.

The image scale is calculated by

$$m = \frac{h \cdot \cos(\beta - \tau)}{c \cdot \cos \beta} \tag{6.20}$$

where

β: imaging angle to object point P
τ: tilt angle (τ=0 for nadir image)
h: flying height above ground
c: principal distance

– Forward and side overlap:

In contrast to flight patterns used in conventional aerial photogrammetry, forward and side overlaps have a different meaning for UAV flights. Here it is not an objective to acquire a minimum of images for subsequent stereo processing but to acquire a dense set of images with minimum occlusions. UAV parameters are around p = 80 % (forward overlap) and q = 60 % (side overlap). For fixed-wing UAVs, which normally fly at higher speeds, an 80 % side overlap is usually flown (see above). Large overlaps are also appropriate when processing with structure-from-motion techniques. For complex terrains with greater height differences, cross strips are recommended where the area is recorded twice at 90° directions, optionally also at another flying height. Cross strips at different heights are also helpful for simultaneous camera calibration. The total number of images is less important.

– Height-to-base ratio:
The height-to-base ratio h/b (see section 4.3.6.2) is the primary factor determining the achievable height-measuring accuracy. The base b between two image is given by:

$$b = (1-p) \cdot S = (1-p) \cdot m \cdot s' \tag{6.21}$$

where
p: forward overlap
S: imaged side of terrain in flying direction
s':   image format of the sensor in flying direction
m: image scale number

The distance between adjacent flight strips follows in an analogous way using the given side overlap.

– Image time interval:
For fixed-wing UAVs, the time interval between two sequentially recorded images is given by the base b and velocity v:

$$\Delta t = b / v \tag{6.22}$$

If necessary, the image time interval must be adjusted to accommodate the maximum frame rate of the camera.

– Exposure time and image blur:
Maximum permitted exposure time and resulting image blur are defined according to the relationships described in section 6.10.1.

– Targeting and measurement of reference points:
Reference points (ground control points) are usually measured by GNSS or other geodetic methods (see section 6.3.2). Targeting is realized through typical chessboard patterns or circular targets. Fig. 6.109 shows a turn-and-tilt target which can be used in parallel for terrestrial measurements. Alternatively, targets with an integrated GNSS receiver are available which do not require any control point measurement (Fig. 6.110).

– Position and accuracy of reference points:
The number and distribution of (geodetic) reference points are a function of the specified accuracy on the ground and in height, the quality of the camera in use, overlap parameters and issues of economy. UAV applications with a small number of reference points, e.g. only at the corners of the mapped region, tend to show deformations in the 3D reconstruction, e.g. systematic height errors. With on-board GNSS RTK, the number of ground control points (GCP) can be reduced significantly. For accuracy requirements >1 dm, RTK may not require any control points at all, but this method of direct georeferencing does not allow independent accuracy control. For applications with higher accuracy, control points are usually required at least in the corners of the flight area.

Fig. 6.109: Tilt & turn target (Zoller & Fröhlich).   **Fig. 6.110:** GNSS target (Propellor Aerobotics).

6.11.2.4 Photogrammetric processing

In most UAV applications, photogrammetric processing is done using structure-from-motion software (SfM, section 6.4.2) which also allow for the orientation of unstructured image configurations with uncalibrated cameras, Usually, natural surface texture is sufficient to match feature points reliably (application example in Fig. 6.111).
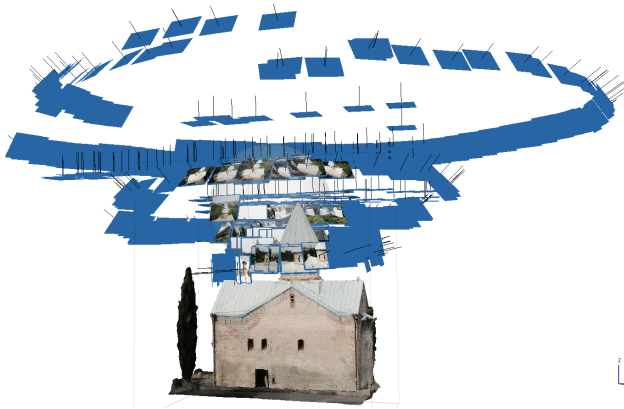
Fig. 6.111: UAV flight and object modelling with SfM (see also Fig. 6.112).

If unstable cameras are used for flights above flat terrain, camera calibration can be problematic (see section 7.3.5). Weak camera calibrations often remain undetected in SfM solutions because the overall result seems to be correct. To avoid this, UAV flights should be configured with some depth in object space, for example flying at two different altitudes. Ideally, a stable, high-quality camera should also be employed and camera functions such as autofocus or internal distortion correction should be switched off for a good photogrammetric result.

Fig. 6.112: Use of a quadcopter to record a church (see also Fig. 6.111).

Fig. 6.112 shows the use of a quadcopter to acquire imagery of a church tower (Lurdji monastery, Tbilisi). Here, several flight paths following circular and meandering patterns have been combined with images taken from the ground and terrestrial laser scanning. Fig. 6.113 shows an orthophoto generated from UAV images used to create a cadastral plan.
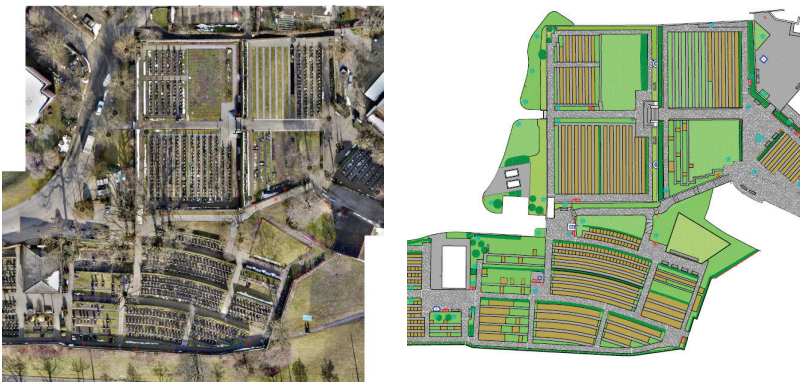
Fig. 6.113: Orthophoto (left) generated from UAV images used to create a cadastral plan (right) (Land Surveying Office, City of Winterthur, Switzerland).

6.12 Visualisation systems

6.12.1 Digital stereo viewing systems

In digital stereo workstations, images are displayed and moved digitally on a suitable stereo monitor. Each image may be moved relative to a superimposed measuring mark. The measuring marks, the images, or both simultaneously, can therefore be moved.

The (interactive) measurement of 3D coordinates uses the principle of the 3D floating mark (section 4.3.6.3). For this, both images must be oriented in the target coordinate system. From given 3D positions in space, the corresponding image coordinates are calculated and the images are displayed for visual stereoscopic control at the corresponding position. The measurement procedure corresponds to the functional principle of earlier analytical plotters (example in Fig. 6.60).

Stereoscopic viewing is achieved through either optical or electronic image separation. Essentially, the following techniques are used:

– Split screen:

  Both images are displayed side by side on a standard monitor and are viewed stereoscopically using a simple optical system similar to that of a mirror stereoscope. The images move which requires a fixed floating mark. The operator is required to sit in a fixed position in front of the monitor and the system can only be used by one observer.

– Anaglyphs:

  Usually, left and right images are presented in two different colours, e.g. red/green, red/cyan, blue/yellow, green/magenta, and viewed through glasses with corresponding colour filters. Depending on the selected colour combination, some significant colour loss may occur in the model. Stereoscopic viewing is possible over large viewing distances and angles, and for a number of people simultaneously.

– Alternating image display:

  Left and right images are alternately displayed at high frequency (>100 Hz). The operator wears wireless glasses which have synchronized shutters triggered by an infrared signal so that only the currently displayed image is visible to the corresponding eye. Within the limitations of the infrared receiver, stereoscopic viewing is possible for a number of people simultaneously in a large workspace in front of the monitor.

– Polarized image display:

  An LCD filter is mounted in front of the monitor. The filter switches polarization direction as the left and right images are alternately displayed (see section 3.1.1.5). The operator wears glasses with correspondingly oriented polarizing filter in order to present the correct image to the correct eye. Electronic control of

these glasses is not required. Parallel stereo viewing by multiple users is possible in a space that is larger than for shutter glasses.

a) Digital stereo monitor (Schneider Digital)   b) Stereo glasses (Vuzix)

Fig. 6.114: Examples of digital stereo viewing systems.

–   Beam splitting display:
    A beam splitter (semi-silvered mirror) is located at the bisecting angle between the two displays for left and right images (Fig. 6.114a). Polarized light emitted from the bottom monitor is transmitted through the mirror while polarized light emitted from the top monitor is laterally inverted upon reflection off the beam splitter. Polarized glasses are used for stereoscopic viewing.
–   LCD monitor glasses:
    The operator wears glasses with two separate LCD mini-monitors. The monitors can either display digital stereo images or can be connected to two mini cameras which observe the scene instead of the human eye. This kind of stereo viewer is used mostly for augmented reality applications (example in Fig. 6.114b).
–   Stereo projection:
    Two digital projection devices (see section 3.6.3.2) can be used to display a stereo image on a large screen. Stereo viewing is usually enabled by shutter glasses that are synchronized with the projectors. This solution is mostly adopted in applications of virtual and augmented reality, where observers interact with a virtual scene. In this case, the position of the observer must be measured by 3D tracking systems, by the VR glasses itself and/or handheld controller devices (see section 6.10.3 and Fig. 6.116).
–   Auto-stereoscopic display (micro-prisms):
    In this case the stereoscopic image is displayed in separated image columns (even columns for the left image and odd columns for the right image). A vertical micro-prism system is mounted over the monitor surface and deflects the column images in the desired spatial viewing direction. The stereoscopic effect may be observed without any additional aids. The resolution of the stereoscopic image

corresponds to half of the monitor resolution. An automatic image and prism adjustment (eye finder) compensates for movements of the operator, so creating a large workspace in front of the monitor.

6.12.2 AR/VR systems

AR/VR (augmented reality, virtual reality) describes various user-controlled systems and applications in which digital (virtual) objects are visualised, enriched with further information or combined with real image information. In detail, a distinction is made between:

– Virtual Reality (VR):

VR refers to the representation and simultaneous perception of reality and its physical properties in a real-time computer-generated, interactive virtual environment. In contrast to the pure visualisation of 3D data, the user interacts directly with the environment and thus receives an immersive experience, i.e. the virtual world is perceived as real. The reference coordinate system is usually given by the digital 3D data.

– Augmented Reality (AR):

AR is an application in which additional information from the real or virtual world is combined and visualised in real time, e.g. by overlaying the real environment with 3D data (see Fig. 6.118). The reference system is the real object coordinate system, i.e. visualisation systems such as projectors, stereo glasses or head-mounted displays (HMD) must be calibrated and continuously oriented in the reference system.

– Mixed Reality (MR):

Mixed reality refers to the mixed perception of virtual and real objects, i.e. it is an amalgamation of AR and VR. The distinction between the terms is fluid, and their use is often not very precise.

The connection to photogrammetry is that a three-dimensional reference between visualisation system, projection system, real world and virtual data must be established. For this purpose, real-time multi-camera systems, as represented by motion capture systems, are often used (section 6.10.3). The systems are able to track and position all relevant system components in space.

6.12.2.1 VR systems

VR systems are usually interactive programming environments with high graphics performance, such as those provided by game engines like UnReal or Unity. These sophisticated environments allow the integration of CAD or BIM models as well as point clouds or meshes, which can be enhanced by further effects such as material

and reflection properties, lighting effects or moving objects. One application example is the development of a virtual laser scanner (section 6.4.4).

a) Virtual reconstruction of the "Peking"

b) Historical 3D model of Stade city anno 1620

Fig. 6.115: Example VR applications (HCU Hamburg).

Fig. 6.115a shows the virtual reconstruction of the "Peking", a four-masted barque from Hamburg, for an immersive VR application in Unreal. The 3D information used in the construction is derived from historical construction plans. Fig. 6.115b is an example of a reconstructed historical 3D city model (Stade, 1620), whose 3D model data is derived from a photogrammetric recording of a physical city model 1:500 (see also section 8.3.1).

a) Hololens (Microsoft)

b) Cave (Jade Hochschule)

Fig. 6.116: AR/MR systems.

6.12.2.2 AR/MR systems

The Microsoft HoloLens (Fig. 6.116a) is one example of an AR system. It is a wireless HMD with stereo glasses, cameras, positioning sensors and computer unit. The stereo glasses allow simultaneous viewing of the environment with superimposed data (graphics, text). With the help of the built-in cameras and other sensors, the real

environment and the user's own 6DOF position are recorded in real time using SLAM approaches.

Fig. 6.116b shows an MR cave. The cave is a walk-in room with several projection walls over which a virtual 3D environment is displayed. One or more people can move around the room at the same time and interact with the virtual environments. Movements and actions of the cave visitors are continuously recorded via a multi-camera system and integrated into the virtual environment.

6.12.3 Industrial 3D projection systems

In many industrial applications, surfaces are measured with the objective of detecting or analysing local defects (dents, surface damage), checking for completeness of manufactured parts or measuring the shape and position of functional elements (drill holes, cutouts, brackets etc.). In such cases it is often helpful to have positionally accurate projection, directly onto the surface, of relevant object data, e.g. written instructions or measurement results compared against a CAD model (see Fig. 8.66).
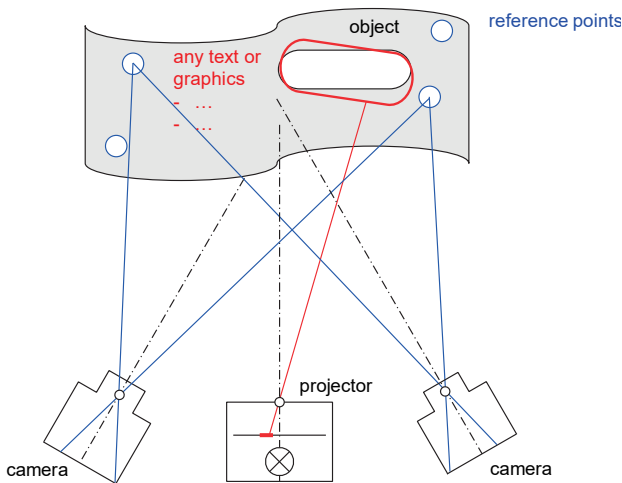
Fig. 6.117: 3D projection.

If the measuring system includes a calibrated and oriented projector, arbitrary 3D data can be transformed into the image plane of the projector by the collinearity equations (4.10) and then accurately re-projected onto the surface. This is also possible when either object or projector are moving, provided that the system's cameras can continuously measure reference features on the object, such as corners, edges and artificial reference marks. This enables the 3D coordinate transformations

and exterior orientations of cameras and projector to be updated continuously, hence stabilizing the projections in their correct locations (Fig. 6.117).
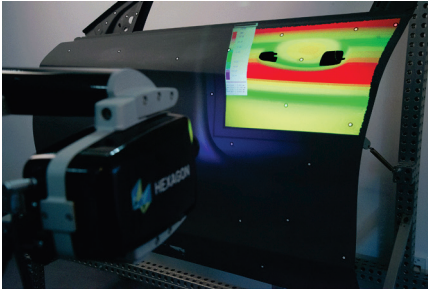
a) Projection of surface deformation results (Hexagon)

b) Projection of object features (Fraunhofer IGP)

Fig. 6.118: 3D projection onto an object surface.

Fig. 6.118a shows an example of a fringe projection system (section 6.7.3.6) used to superimpose previously measured surface deformation data at the correct surface position. This can directly help to make any necessary repairs and adjustments. Fig. 6.118b illustrates an application where a laser projects shape and design features onto the object, e.g. for visual inspection or visualisation of a hidden cable.

7 Measurement design and quality

This chapter considers photogrammetric projects from the perspective of planning, accuracy and optimization. In most practical cases the required measurement accuracy in object space is critical to the design and configuration of a measurement process. Generally speaking, accuracy is therefore always closely connected with costs and commercial viability of a system solution. However, there are a number of additional planning criteria to be taken into consideration. These are summarized in section 7.1.1.

The planning of a photogrammetric configuration and strategy is a complex process requiring due regard for a wide range of issues. Here the question of camera calibration plays a particular role and requires careful consideration, depending on the selected imaging technology and required accuracy. The measurement network must be so configured that constraints imposed by local conditions, the specified accuracy, camera calibration and technical effort required are all taken fully into account.

7.1 Project planning

7.1.1 Planning criteria

The planning of a photogrammetric project includes the description of the actual measuring task, the concept for a solution und the presentation of results. It should be carried out in close co-operation with the client as planning errors are often detected at a late stage when they are then difficult to correct. Planning the imaging configuration is one aspect of the complete project plan which should include, in addition to metrology issues, economic aspects such as staff and time management, use of instruments, cost management etc.

The initial project plan should specify the following features of the measuring task:

– number and type of object areas to be measured, including a description of the object, its situation and the measuring task requirements;
– dimensions of the object;
– specified accuracy in object space (tolerances, measurement uncertainty);
– smallest object feature (resolution of fine detail, GSD);
– environmental conditions (variations in temperature, humidity, pressure and the presence of any vibration);
– options for object targeting;
– definition and implementation of the object coordinate system;
– determination of scale and reference points (geodetic measurements);

- alternative or supplementary measuring methods;
- online or offline measurement;
- static or dynamic measurement;
- processing software;
- acceptance test procedure or verification of accuracy;
- available times for on-site work;
- maximum permitted time for analysis and evaluation;
- output of results (numerical and graphical, data formats, interfaces, relationship to processes).

In the subsequent detailed planning, an appropriate concept for solution can be developed. In particular, the image acquisition system and imaging configuration, as well as the type of image and data processing, must be defined. Apart from purely technical considerations such as accuracy, the choice of components used also depends on the availability of instruments and personnel.

The following criteria should be defined in detail:
- estimation of average image scale;
- estimation of average spatial resolution (GSD);
- selected processing system (analogue/digital, monoscopic, stereoscopic, multi-image);
- camera stations (number of images, network design, ray intersection geometry);
- required image measuring accuracy;
- selected imaging system (image format, focal lengths);
- method of camera calibration (in advance, simultaneously);
- optical parameters (depth of focus, resolution, exposure);
- amount of memory for image data (type and cost of archiving).

7.1.2 Accuracy issues

Besides economic aspects, meeting accuracy requirements is generally of the highest priority in practical project planning. Approximate accuracy estimates can be based on the relationships given in section 3.3.1 which depend on three primary parameters:
- image measurement accuracy;
- image scale;
- design factor of the imaging configuration.

Image measurement accuracy depends on the performance of the camera (stability and calibration), the accuracy of the image processing system (target image quality, measurement algorithm, instrumental precision) and the positioning capability (identification, targeting). These criteria must be balanced with respect to each other and given appropriate weight in the planning process. The camera selected for any given project is particularly important as it not only defines the quality of image

acquisition and processing but, through choice of lens, also defines image scale and configuration.

Digital imaging systems can reach image accuracies of 0.2–1 μm (1/50–1/10 pixel) depending on the mechanical stability and signal transfer type. Digital processing systems providing sub-pixel operators for signalized targets can yield image accuracies between 1/100–5/100 pixels. Methods for detection and matching of natural features provide typical image accuracies of about 0.5–2 pixels.

As explained in section 3.3.1, large image or sensor formats are advantageous for photogrammetric accuracy. It is essentially also the case that a larger number of pixels on a sensor will, in the first instance, improve resolution and therefore indirectly also the image measuring accuracy. However, achievable accuracy at the object will really only be improved by higher pixel numbers if the quality of the optical imaging components (resolving power of the lens), mechanical stability of the camera and image noise are consistent with the properties of the imaging sensor. It is particularly the case with sensors which have very small pixels (<2.5 μm), that an image measurement accuracy of better than 1/10 pixel (< 0.25 μm) cannot be expected due to limits in construction (camera mechanics, sensor mounting) and optics (diffraction).

In applications where there is essentially no limitation on the number of images, the achievable accuracy can be increased by taking additional images from locations which improve imaging geometry. The accuracy enhancement due to an increase of images at every station k, as expressed by eqn. (3.49), essentially only results in an improvement in precision. In this case, the increased numbers of observations going into the bundle adjustment improve the standard deviation of unit weight σ_0 and thereby the statistical precision of the calculated object coordinates. If object targeting permits fully automatic image measurement, then an increase in number of images has no significant disadvantages.

Specification of measurement accuracy should be done in close consultation with the end user or client. It must be made clear which accuracy parameters are to be used, the justification for the specified limits, and the method of verifying that the accuracy has been achieved. Quality measures obtained from various adjustment methods are explained in section 2.4.3 and 7.2.1. Metrologically defined parameters of accuracy (measurement uncertainty) are presented in section 7.2.2.

7.1.3 Restrictions on imaging configuration

A generally applicable geometric configuration for photogrammetric measurement cannot be defined because it always depends on circumstances specific to the object. A compromise must normally be found between different and partly incompatible restrictions:

– Image scale:

The image scale is influenced by object distance, focal length (principal distance) and usable image format (eqn. 1.1, see Fig. 3.40). A larger image format enables shorter object distances for the same imaged object area. It not only leads to a larger scale (and higher accuracy) but also to a smaller number of images (economic benefits in reduced data processing and storage). It must be remembered that for complex object structures and highly convergent images, image scale can vary significantly within an image or from image to image (see Fig. 3.41).

– Image quality and resolution:

The ability to detect and measure object details is again a function of image scale. The size of imaged object structures should lie between certain typical limits as follows:

  – visual digital processing: 3–10 pixel
  – automatic digital point measurement: 6–10 pixel
  – automatic digital surface measurement: 11–25-pixel window size

Eqn. (1.2) gives the relationship between pixel size in object space (ground sampling distance, GSD) and pixel size in image space.

– Object environment:

The selection of suitable camera stations is often restricted by inaccessible object areas. It is therefore often necessary to use either additional lenses, or to increase the number of images or to dispense with optimal ray intersections. Additional camera/lens combinations may also increase the effort required for camera calibration. This is particularly applicable to zoom lenses with adjustable focal lengths (see section 3.4.3.5).

– Depth of field:

The available depth of field (section 3.1.2.3) is mainly a function of image scale and f-stop. It restricts the choice of camera stations, especially for large image scales and under difficult lighting conditions. If applied targets are measured automatically, a slightly defocused image can be accepted if image contrast and point diameter are sufficiently large.

– Imaging angle:

The imaging angle β at which an object is photographed should not be less than 20° for critical object features, and not less than 45° for retro-reflective targets, in order to achieve suitable image sizes and contrasts (Fig. 7.1). Furthermore, extremely oblique views have an eccentricity effect on the centre of circular targets as explained in section 6.2.1.1. Spherical targets permit all-round imaging but note that they are also elliptical in shape when imaged off axis (see Fig. 6.3).
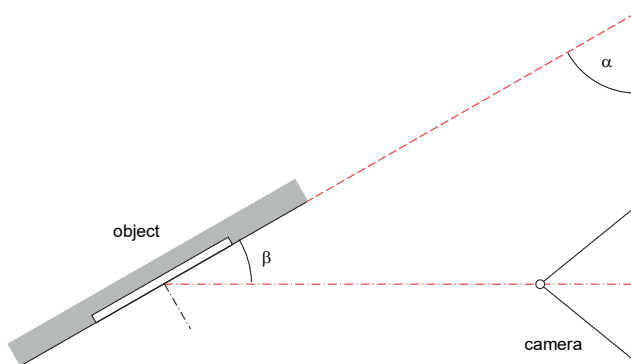
Fig. 7.1: Imaging angle.

– Number and distribution of image points:
  The total number of object points has little effect on the total redundancy of the bundle adjustment. The quality of bundle triangulation depends more on the number and configuration of camera stations which, in addition to creating a reasonable intersection geometry (see below), should ideally utilize the full image format (see Fig. 7.39).
– Intersection angle:
  Good intersection angles are critical to the accuracy of point measurement (see Fig. 4.43). For the (graphical) reconstruction of approximately flat object surfaces (building facades) a reduced accuracy in the viewing direction can often be tolerated and in these cases it is possible to work with poor intersection angles (or insufficient height-to-base ratios).

  Engineering or industrial projects often require an object accuracy which is equal in all directions. Optimal ray intersection angles are around 90°–100°. In practice, intersection angles between 45° and 120° are sufficient if at least 4 to 6 images contribute to the measurement.
– Field of view and visibility:
  The field of view of a camera is defined by the format angle (section 3.4.3.2). If all object points can be imaged in all photos, then the imaging configuration is simplified and the number of images required is reduced. At the same time, the total redundancy of the bundle triangulation increases. Since this assumption is only valid for simple objects (e.g. test fields for calibration purposes, see Fig. 7.28), objects with occluded areas must be recorded with additional photos which only include a small portion of all object points. To ensure network stability, these additional images should contain a reasonable number of well distributed tie points. For simultaneous calibration of distortion, it is advantageous to have an irregular distribution of object points across the image format, with many

occupying the image corners. In different images, it is an additional advantage to have the same points appear in different image areas.

– Image analysis and use:

Depending on subsequent image analysis and presentation of results, additional images in diverse arrangements may be required. For image rectification and texture projection onto 3D models, useful images have few occlusions and view the object surfaces perpendicularly. For visual stereoscopic viewing, images with a good height-to-base ratio and parallel camera axes are required.

7.1.4 Accuracy estimation by simulation

Numerical simulation can be used to estimate the expected accuracy of a photogrammetric project provided that the imaging configuration, e.g. stereo or multi-image, parameters of interior and exterior orientation and the distribution of object points are defined in advance.

In this process, the simulation applies error propagation to derive the effect of noisy data, typically image coordinates with associated measurement uncertainties, on the required output parameters (orientation parameters, 3D coordinates). Simulation results are valid only if the assumptions about uncertainties in input values are correct and if the actual project conforms closely to the simulated data and its configuration.

Almost all photogrammetric calculations are based on a least-squares adjustment (overdetermined Gauss-Markov model) in which the observations are modelled as a function of the required unknowns. These unknown parameters are therefore determined indirectly from (simulated) measurements (see eqn. 2.180 ff):

$$\hat{\mathbf{L}} = \mathbf{L} + \mathbf{v} = \boldsymbol{\varphi}\,(\hat{\mathbf{X}}) \tag{7.1}$$

The simulation's required input values can be generated according to the scheme presented in section 4.4.6.1. The covariance matrix of the unknowns can be derived from equations (2.194) and (2.213) and used as a measure of their expected precisions. This provides a realistic estimation of quality if the functional adjustment model is sufficiently well represented in linearized form, there are no significant systematic errors and observations are free of outliers.

7.1.4.1 Variance-covariance propagation

The conventional law of variance-covariance propagation, also known as error propagation, is based on the linearization of a Taylor series expansion which is truncated after the first term (first order derivative). In applications where there is an explicit relationship $\mathbf{y} = \mathbf{f}(\mathbf{x})$ between input values $\mathbf{x}$ and output values $\mathbf{y}$ then, in accordance with eqn. (2.185), the following is given:

$$\mathbf{y} = \boldsymbol{\mu}_y + \boldsymbol{\varepsilon}_y = \mathbf{f}(\mathbf{x}) \approx \mathbf{f}(\boldsymbol{\mu}_x) + \mathbf{J}(\mathbf{x} - \boldsymbol{\mu}_x) \tag{7.2}$$

where

f: vector of functions

J: design matrix (Jacobian matrix) with first partial derivatives

$\boldsymbol{\mu}_y$: expected value of **y** (compare with eqn. 2.232)

$\boldsymbol{\varepsilon}_y$: noise value for y (random error)

$\boldsymbol{\mu}_x$: expected value of **x**

Non-linear functional relationships are not sufficiently well modelled by first derivatives. Extending the modelling to the second derivative gives:

$$\mathbf{y} = \boldsymbol{\mu}_y + \boldsymbol{\varepsilon}_y = \mathbf{f}(\mathbf{x}) \approx \mathbf{f}(\boldsymbol{\mu}_x) + \mathbf{J}(\mathbf{x} - \boldsymbol{\mu}_x) + \left[\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_x)^T \mathbf{H}_i (\mathbf{x} - \boldsymbol{\mu}_x) \right]_i \tag{7.3}$$

where

$\mathbf{H}_i$: Hessian matrix with the second derivatives of the functional equations i

$[\ldots]_i$: Symbolic notation for a vector $[v_i]_i$, where v_i is its element i; by analogy $[m_{ij}]_{ij}$ denotes a matrix where m_{ij} is the element (i,j).

For a classical photogrammetric calculation, such as a space intersection based on collinearity equations (section 4.4.7.1), vector **x** would contain all contributing parameters on the right side of the equation (3D coordinates, interior and exterior orientations) whilst the vector **y** would contain the measured image coordinates. In this example, the explicit relationship would express, for example, the effect of orientation parameters on the image coordinates.

Applying linearized variance-covariance propagation, the expected value of observations **y** and their uncertainty (covariance matrix $\boldsymbol{\Sigma}_{yy}$) is given as follows (compare with eqn. 2.225):

$$\boldsymbol{\mu}_y = \mathbf{f}(\boldsymbol{\mu}_x)$$
$$\boldsymbol{\Sigma}_{yy} = \mathbf{J} \cdot \boldsymbol{\Sigma}_{xx} \cdot \mathbf{J}^T \tag{7.4}$$

Extending to the second derivatives gives:

$$\boldsymbol{\mu}_y = \mathbf{f}(\boldsymbol{\mu}_x) + \frac{1}{2} \left[\text{trace}(\mathbf{H}_i \cdot \boldsymbol{\Sigma}_{xx}) \right]_i$$
$$\boldsymbol{\Sigma}_{yy} = \mathbf{J} \cdot \boldsymbol{\Sigma}_{xx} \cdot \mathbf{J}^T + \frac{1}{2} \left[\text{trace}(\mathbf{H}_i \cdot \boldsymbol{\Sigma}_{xx} \cdot \mathbf{H}_j \cdot \boldsymbol{\Sigma}_{xx}) \right]_{ij} \tag{7.5}$$

The comparison of eqn. (7.4) with eqn. (7.5) makes clear that both the uncertainties and the expected value itself change when the second derivative is taken into consideration. If non-linearity can justify the additional effort of calculation using second derivatives, eqn. (7.5) delivers results which are close to those derived from the numerical approaches below. The procedure described here can, in principle, be transferred to overdetermined systems of equations with both explicit and implicit

functional relationships. However, due to the complexity of the presentation, it will not be discussed in detail here.

Although typical photogrammetric calculations are non-linear, higher derivatives are neglected in practice as they are often complex to calculate. However, numerical simulation methods such as the Monte Carlo method (next section) can generate estimates free of linearization errors if the number of repeated calculations with randomized errors is high enough.

7.1.4.2 Monte Carlo simulation

A Monte Carlo simulation is a computational method for statistically analysing results from complex systems of calculation. The input data is randomized or altered in a structured way and its effect on the calculated output data observed. Every input parameter can be introduced with its own statistical distribution and variance, as well as a function to describe systematic deviations. For every calculation the input data is modified with the aid of a random number generator and the calculation is repeated many times until a statistically meaningful output sample has been obtained (Fig. 7.2).
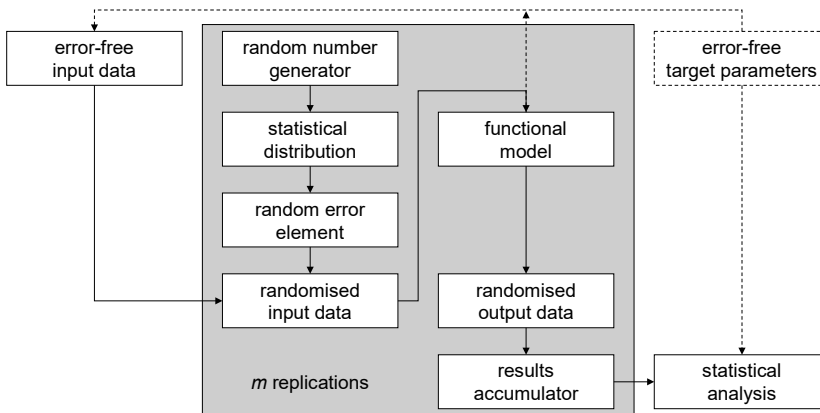
Fig. 7.2: Monte Carlo simulation.

The total number of repeated calculations (replications), m, should be in the thousands, possibly up to ten thousand, in order to ensure a homogeneous distribution of random variations across the relevant data spectrum. It is usual to assume that the random variations follow a normal distribution. These can, for example, be generated by applying the Box-Müller method to the uniformly distributed random numbers $\varepsilon_x^{(i)}$, $i = 1, m$, which are often available from a built-in

random number generator. numbers. A different set of random variations must, of course, be taken from this distribution for each replication.

Depending on application, it may be sensible to limit the noise on the input data, for example to 2 or 3-sigma, in order to avoid outliers in the simulated data sets. Alternatively, by targeted modelling of possible error sources, the effect of outliers on the estimated parameters can be analysed and worst-case scenarios investigated.

For simulated input values $\mathbf{x}$, the expected values $\boldsymbol{\mu}_x$ and their uncertainty $\boldsymbol{\Sigma}_{xx}$ must result from the m replications:

$$
\begin{aligned}
\boldsymbol{\mu}_x &\approx \boldsymbol{\mu}_x + \frac{1}{m}\sum_{i=1}^{m}\boldsymbol{\varepsilon}_x^{(i)} = \frac{1}{m}\sum_{i=1}^{m}\left(\boldsymbol{\mu}_x + \boldsymbol{\varepsilon}_x^{(i)}\right) \\
\boldsymbol{\Sigma}_{xx} &\approx \frac{1}{m-1}\sum_{i=1}^{m}\boldsymbol{\varepsilon}_x^{(i)}\boldsymbol{\varepsilon}_x^{(i)T}
\end{aligned}
\tag{7.6}
$$

Results from all replications are accumulated and statistically analysed at the end. This enables the generation of a histogram of the output parameters from which their standard deviations can be derived. It is also possible to make a direct comparison with defined, error-free target parameter values, from which error-free input data can be simulated using the functional model (dotted line connection in Fig. 7.2).

The expected values of $\mathbf{y}$ and their uncertainty $\boldsymbol{\Sigma}_{yy}$ is given by:

$$
\begin{aligned}
\boldsymbol{\varepsilon}_y^{(i)} &= \mathbf{f}\left(\boldsymbol{\mu}_x + \boldsymbol{\varepsilon}_x^{(i)}\right) - \boldsymbol{\mu}_y \\
\boldsymbol{\mu}_y &\approx \frac{1}{m}\sum_{i=1}^{m}\left(\boldsymbol{\mu}_y + \boldsymbol{\varepsilon}_y^{(i)}\right) \\
\boldsymbol{\Sigma}_{yy} &\approx \frac{1}{m-1}\sum_{i=1}^{m}\boldsymbol{\varepsilon}_y^{(i)}\boldsymbol{\varepsilon}_y^{(i)T}
\end{aligned}
\tag{7.7}
$$
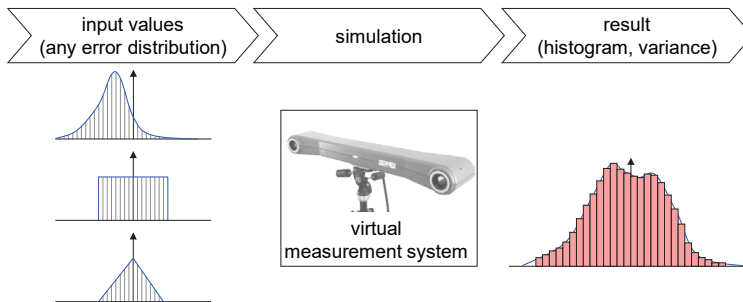
Fig. 7.3: Principle of a virtual measuring system by Monte Carlo simulation (after Schmidt et al. 2008).

The Monte Carlo method permits the computational modelling of a measuring system which, in this form, can also be regarded as a virtual measuring system. A necessary

condition is accurate knowledge of the sources of error in a system and their mathematical description. It is then possible to create a realistic simulation of complex measuring tasks, and various influences on the measurement method, without the need to have a system physically available (Fig. 7.3).

Through application of the complete software chain of a system, for example bundle adjustment followed by spatial intersection followed by 3D transformation, the complete propagation of uncertainties is modelled. In contrast to the linearized variance-covariance propagation, linearization errors do not occur, hence a strict error propagation is performed.
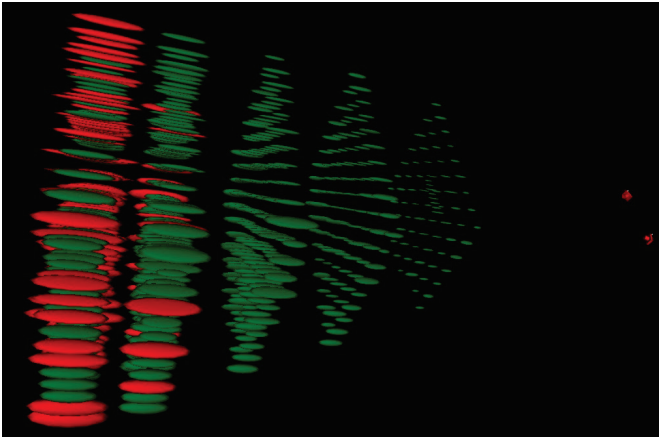
Fig. 7.4: Result from a Monte Carlo simulation of a stereo camera showing error ellipsoids at object points in various distances.

Fig. 7.4 shows the result from a Monte Carlo simulation of the expected measurement accuracy of a stereo camera. Using defined, error-free object coordinates, as well as parameters for the interior and exterior orientation of each camera, error-free image coordinates are calculated according to eqn. (4.10). The defined object points are representative of the complete measurement volume of the system. In the subsequent Monte Carlo simulation, positive and negative random deviations, according to their respective expected measurement uncertainties, are added to the error-free image coordinates, camera parameters and orientation parameters. Using this randomized data, object point intersections are repeatedly calculated and their spread analysed.

In addition to numerical values for the calculated standard deviation and maximum value of the deviations, Fig. 7.4 also shows, at every object point, an error ellipsoid to illustrate the spatial distribution of the expected measurement uncertainty.

7.1.4.3 Unscented transformation

As an alternative to the classical Monte Carlo simulation, the *unscented transformation* is a non-iterative numerical approach to approximate parameters and related uncertainties. Whereas the Monte Carlo simulation needs randomized observation values, the unscented transformation requires a discrete set of points as observations. These are often called sigma points (Σ points). The number of Σ points of the Standard Unscented Transformation described here is $2n+1$, where n is the dimension of the $\boldsymbol{\mu}_x$ vector:

$$
\begin{aligned}
\mathbf{X}_0 &= \boldsymbol{\mu}_x & , w_0 &= \frac{\kappa}{n+\kappa} \\
\mathbf{X}_i &= \boldsymbol{\mu}_x + \left(\sqrt{(n+\kappa)\boldsymbol{\Sigma}_{xx}} \right)_i & , w_i &= \frac{1}{2(n+\kappa)} \\
\mathbf{X}_{i+n} &= \boldsymbol{\mu}_x - \left(\sqrt{(n+\kappa)\boldsymbol{\Sigma}_{xx}} \right)_i & , w_{i+n} &= \frac{1}{2(n+\kappa)}
\end{aligned}
\tag{7.8}
$$

where
$$\kappa = \alpha^2(\lambda+n) - n$$
α: noise of Σ points
λ: control parameter

For normally distributed input values, $\alpha = 10^{-3}$ and $\lambda = 0$ can be assumed. The square root of $\boldsymbol{\Sigma}_{xx}$ can be calculated by spectral decomposition (eqn. 2.253). Using the Σ points $\mathbf{X}_j$, where $j = 0...2n$, the resulting output values are:

$$
\begin{aligned}
\mathbf{Y}_j &= \mathbf{f}(\mathbf{X}_j) \\
\boldsymbol{\mu}_y &= \sum_{j=0}^{2n} w_j \mathbf{Y}_j \\
\boldsymbol{\Sigma}_{yy} &= \sum_{j=0}^{2n} w_j (\mathbf{Y}_j - \boldsymbol{\mu}_y)(\mathbf{Y}_j - \boldsymbol{\mu}_y)^T + (1+\beta-\alpha^2)(\mathbf{Y}_0 - \boldsymbol{\mu}_y)(\mathbf{Y}_0 - \boldsymbol{\mu}_y)^T
\end{aligned}
\tag{7.9}
$$

with $\beta = 2$ for normally distributed input values

The unscented transformation estimates the expected value and its variance-covariance matrix from a fixed number of Σ points without the use of a Taylor expansion. In comparison with the Monte Carlo simulation, the usage of a discrete set of observations provides repeatable results and reduces the numerical effort. This makes the method attractive for solving non-linear real-time applications e.g. in the framework of Kalman filters (section 5.5.7.5). In the case of a space intersection using two images, with uncertainty in the parameters of interior orientation (2×10 parameters) and exterior orientation (2×6 parameters) then $n = 32$. Here only $2n+1 = 65$ replications are required whereas the Monte Carlo method needs at least 5000 to 10000 repeated calculations. The unscented transformation offers a reasonable compromise between setting up the Hessian and Jacobian matrices and the required computational effort of the Monte Carlo method.

7.1.4.4 Bootstrap simulation

In contrast to the above approaches, *bootstrap simulation* does not use simulated input values but n real measurement samples, e.g. n independent measurements of a space intersection. The method is therefore less suitable for accuracy estimation prior to measurement and more suitable for post-measurement analysis.

In comparison to the Monte Carlo method, the bootstrap simulation does not rely on a particular statistical distribution or pre-defined uncertainty values. Instead, it is assumed that all statistical information is contained within the empirical data of the input sample. This is also known as a plug-in principle or plug-in estimation. Whilst Monte Carlo is based on a synthetic simulation of random numbers according to a given probability function, the bootstrap method uses sub-samples generated by an m-times resampling of the input sample (Fig. 7.5).
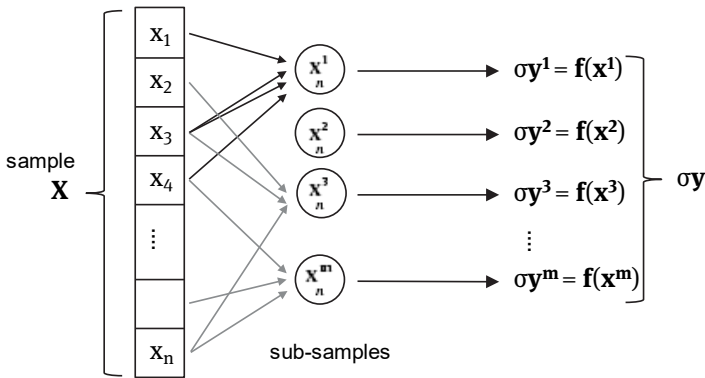
Fig. 7.5: Principle of bootstrap simulation.

Each sub-sample has the same population count n as the input sample and is created by multiple random sampling with replacement. As a simple example of this process, if **X** contains $n=5$ observations, i.e. **X** = [1 2 3 4 5], then possible random sub-samples with replacement could be $\mathbf{X}_1$ = [2 2 1 4 5], $\mathbf{X}_2$ = [4 3 4 5 5], $\mathbf{X}_3$ = [1 5 2 4 1], $\mathbf{X}_4$ = [1 3 5 4 3].

Clearly, in this process no further information is introduced into the estimation. The sub-samples are then analysed in an analogous way to the Monte Carlo method which leads to an estimation of the expected value and its uncertainty using eqn. (7.7). The number m should lie between 50 and 200.

7.1.5 Design of the imaging configuration

7.1.5.1 Network design

The restrictions mentioned in section 7.1.3 limit the choice of camera stations and viewing directions. In practice, a well configured network which surrounds the object, as indicated in Fig. 3.39, is impossible for many applications because of object restrictions or economic circumstances, e.g. the maximum number of images. Image scales and intersection angles can therefore vary widely within a project, so that eqn. (3.49) is only valid for average accuracy estimations. Weak areas in an imaging network cannot be detected by this approach.

The image configuration can be simulated by bundle adjustment (section 4.4.6.1) if there are sufficient a priori object coordinates to represent object geometry. In this process the principal distance and exterior orientation of all cameras are iteratively varied until the following criteria are optimized:

– Maximum accuracy (minimum standard deviations of object coordinates):

In general, object point accuracy and associated derived quantities (lengths, distances etc.) are of greatest importance in practice. In order to assess the expected accuracy it is helpful to split up the vector of unknowns in the bundle adjustment and express eqn. (4.104) in the form:

$$\begin{bmatrix} \hat{\mathbf{x}}_1 \\ \hat{\mathbf{x}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1^T\mathbf{PA}_1 & \mathbf{A}_1^T\mathbf{PA}_2 \\ \mathbf{A}_2^T\mathbf{PA}_1 & \mathbf{A}_2^T\mathbf{PA}_2 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \mathbf{A}_1^T\mathbf{Pl} \\ \mathbf{A}_2^T\mathbf{Pl} \end{bmatrix} \quad \text{and} \quad \mathbf{Q} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_2 \end{bmatrix} \qquad (7.10)$$

where

$\hat{\mathbf{x}}_1$: exterior orientation and additional parameters
$\hat{\mathbf{x}}_2$: object point coordinates

In order to optimize the object point accuracy, only $\mathbf{Q}_2$ need be analysed according to the following considerations:

– Maximum reliability (ability to control and detect outliers):

Statistical measures for reliability and robustness of an adjustment can be derived from the cofactor matrix of corrections $\mathbf{Q}_{vv}$ (section 2.4.3.4, eqn. 2.239).

– Maximum economic efficiency:

The objective is to achieve the specified accuracy and reliability for minimum instrumental and personal effort.

The optimization of photogrammetric imaging configurations (network optimization, network design) can be considered in four stages:

1. Zero order design: definition of datum (object coordinate system):

Standard deviations of object coordinates are directly influenced by the definition of the object coordinate system (section 4.4.3). Potential network deformation can be avoided by a datum definition without constraints. Optimal standard deviations are obtained by a free net adjustment which minimizes the

trace of the covariance matrix. In most cases the covariances of the object coordinates of a point i can be estimated by

$$\mathbf{C}_{2i} \approx s_0^2 (\mathbf{A}_2^T \mathbf{P} \mathbf{A}_2)_i^{-1} \tag{7.11}$$

2. First order design: optimization of the observation configuration:
   The purpose of optimising the configuration is to define an observation network whose design matrix $\mathbf{A}$, given a predefined weight matrix $\mathbf{P}$, results in a covariance matrix $\mathbf{C}_2$ corresponding to the specified accuracy. This is primarily a question of achieving good intersection angles at the object points. Assuming an appropriate minimum configuration which does this, $\mathbf{C}_2$ can be estimated by:

$$\mathbf{C}_2 = \frac{s_{x'y'}^2}{k} (\mathbf{A}_{2B}^T \mathbf{P} \mathbf{A}_{2B})^{-1} \tag{7.12}$$

where

$\mathbf{A}_{2B}$: basic configuration design matrix
$s_{x'y'}^2$: standard deviations of image coordinates
k: number of additional, symmetrically arranged camera stations

The design factor q and the image scale number of eqn. (3.49) are reflected here in the matrix product.
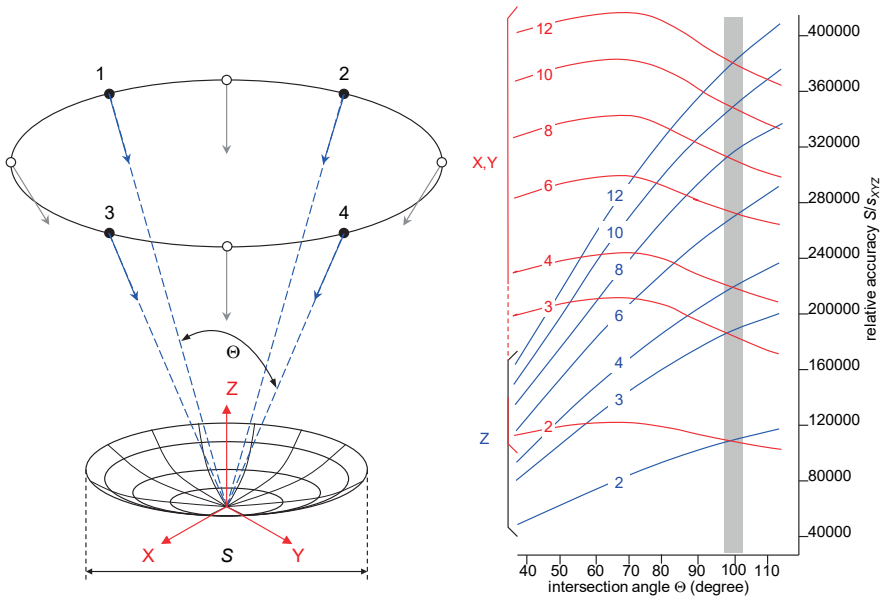
Fig. 7.6: Imaging configuration for antenna measurement and resulting object accuracies after variation of intersecting angles and number of camera stations (after Fraser 1996).

Fig. 7.6 illustrates the imaging configuration for an antenna measurement. The basic configuration consists of four camera stations. Additional images are added symmetrically at equal object distances. The intersection angle Θ is defined between the viewing directions of diametrically opposite camera stations. The diagram shows the resulting relative object accuracies S/s_{XYZ} as a function of the intersection angle and the number of camera stations. The relative accuracy increases with an increasing number of images. Equal values in all coordinate axes are obtained at $\Theta \approx 100°$.

3. Second order design: definition of observation weights (image measuring accuracy):
   The normal system of equations can be readily manipulated by appropriate choice of observation weights. Generally, the a priori standard deviations of image measurements are defined according to the precision of the image measuring device. Normally all image observations are given equal weights. Unequal weights are only justified if different measuring devices or algorithms are employed. In such cases an analysis of variances enables groups of observations of similar weights to be defined (see section 2.4.4.3).

4. Third order design: optimization of point density (object points):
   Where self-calibration of the camera is not required, around 20–50 object points are sufficient to achieve a stable network geometry, optimized according to the above criteria. The quality of point measurement does not significantly improve if the number of object points is further increased. However, if the camera must be calibrated (self-calibration), targets must have a good distribution and density in the images in order to determine distortion parameters reliably across the entire image format (Fig. 7.39).

A relatively high effort is required for realistic simulations because, in many cases, object coordinates are either not available or must be generated manually. In addition, the selection of camera stations and viewing directions is usually performed interactively. Photogrammetric multi-image processing systems connected to 3D CAD systems offer an efficient basis for the simulation of imaging configurations if the object can be represented as a CAD model and if the camera stations can be edited graphically.

7.1.5.2 Scale

Every photogrammetric configuration is based on measured image coordinates, i.e. only directions in space are determined. The intersection of homologous image rays leads to object coordinates (point clouds) that are true in shape but without absolute scale. The required scale information can be provided in one of the following ways (see also section 6.3.1.1):

- exterior orientations are known with respect to a metric coordinate system;
- the unscaled set of object points can be transformed subsequently into a metric target coordinate system (absolute orientation with a 7-parameter transformation, see section 2.2.4);
- the configuration is a stereo or multi-camera system with known baseline;
- at least one known distance is observed in object space (scale bar, distance between two reference points);
- at least three reference points are provided in object space for the absolute definition of datum.

Distances or reference coordinates are usually introduced into a bundle adjustment as additional constraints or observation equations (see section 4.4.2.3). Their specified weight must be appropriate to their calibration or measuring accuracy. False weighting leads to deformation in object space. Net strains can result from overdetermined datum definitions (see section 4.4.3.2).

The number and spatial distribution of reference distances, e.g. scale bars, have a significant impact on the accuracy of measured object points and on the calculated statistical quality values. In general, reference lengths should be as long as possible and ideally extend across the whole measuring volume. For high-accuracy industrial applications, there should be at least one reference length for each major coordinate direction in order to ensure equal scaling in all directions. For applications where only one scale bar is available, it should be positioned within the inner third of the measuring volume. Generally, a longer but less accurate reference distance is preferred to a shorter but highly accurate reference length (see eqn. 6.3).
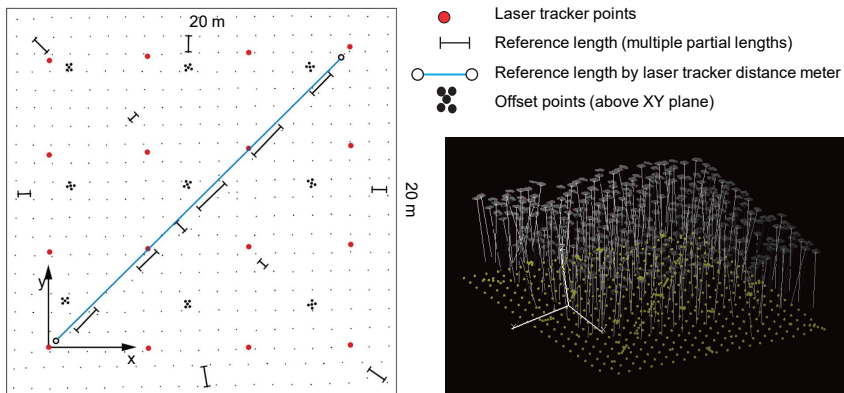
Fig. 7.7: Experimental set-up for investigating different configurations of reference scales.

Table 7.1: Effect of different scaling configurations.

Configuration	Number	Length	A priori st.dev.	RMS 1σ	LME
1)	1	21 m	20 µm	26 µm	155 µm
2)	116	0.15 – 2.2 m	10 µm	29 µm	164 µm
3)	5	1.4 – 2.2 m	10 µm	34 µm	204 µm
4)	1	2 m	10 µm	54 µm	193 µm
5)	16	XYZ	0 µm (fixed)	24 µm	110 µm

Fig. 7.7 shows the experimental set-up of a large-volume project in which number, position and accuracy of reference distances have been investigated. Photogrammetric targets at different heights, reference distances and laser tracker points (nests) were placed in a 20 m x 20 m area. Image acquisition was achieved using a digital medium-format camera (ALPA 12 WA metric) arranged in a pattern similar to an aerial survey. A total of 2200 images was recorded from which 480 were processed.

Table 7.1 shows the effect of different scaling configurations on the resulting RMS values of object points and the maximum length measuring error LME (section 7.2.3.4). Where scale bars are used, it can be seen that a long but less accurate reference length (1) generates the lowest RMS and the lowest LME. A similar result can be achieved with a number of much shorter distances distributed uniformly over the object area (2). Several short scale bars arranged in a line (3) do not provide any significant improvement in comparison to one short distance in the centre of the object space (4). The best result is obtained with a large number of reference points (5) which, however, is not usually achievable in practice. The results of this investigation may not be transferred one-to-one to other imaging configurations but they demonstrate the sensitivity of photogrammetric processing with respect to the chosen scaling method.

7.2 Quality measures and performance testing

The parameters and methods summarized in the following sub-sections are important criteria to use in planning and optimizing a measurement task, as well as for the

subsequent verification of the accuracy achieved in object space. Here there is a distinction between statistical quality parameters, which are generally derived from the stochastic model of an adjustment process, and standardized metrological parameters defined by national and international organizations responsible for industrial metrology.

7.2.1 Statistical parameters

7.2.1.1 Precision and accuracy parameters from a bundle adjustment

The quality parameters which can be derived from the statistics generated by an adjustment process, reflect how well the measured values (observations) fit the functional model defined by the chosen geometrical configuration (design matrix). If there are no systematic errors in the measurements, the parameters of precision describe the random variations in the measured values. In this case they also represent estimates of the measurement accuracies. Precision parameters include the following (see sections 2.4.3 and 4.4.5):

– Standard deviation of unit weight:
  The a posteriori standard deviation $\hat{s}_0$ (sigma 0) is derived from the observation residuals and the redundancy $r = n - u$ (eqn. 2.209). By increasing the redundancy (through an increased number of observations) $\hat{s}_0$ can almost be reduced to any level required. s_0 a priori and $\hat{s}_0$ should be of a similar size (compare with eqn. 2.258).

– Averaged residuals of image coordinates:
  The averages of residuals of all image coordinates are a measure of the quality of image measurement. In addition to the quality of point recognition, the quality of camera calibration (interior orientation) and camera stability are also covered by this measure.

– Standard deviations of object coordinates:
  Following an adjustment process, every unknown is assigned a standard deviation according to eqn. (2.227) (section 2.4.3.1). If the unknowns relate to 3D coordinates, the average standard error of measured points (average point error or mean Helmert error) is given by:

$$\hat{s}_{XYZ} = \sqrt{\hat{s}_X^2 + \hat{s}_Y^2 + \hat{s}_Z^2} \tag{7.13}$$

This, in turn, depends on the standard deviation of unit weight s_0 (see eqn. 2.251). The average standard deviation of n adjusted object coordinates is given by the RMS values:

$$RMS_X = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\hat{s}_{X_i}^2} \qquad RMS_Y = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\hat{s}_{Y_i}^2} \qquad RMS_Z = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\hat{s}_{Z_i}^2} \tag{7.14}$$

The standard deviations of object coordinates depend on the chosen datum and, depending on the positions of points in the object coordinate system, they can provide a false picture of the accuracy achieved. A more homogeneous distribution of standard deviations is achieved by using a free-net adjustment (section 4.4.3.4). Using confidence ellipsoids according to Helmert, adjusted coordinates can be associated with confidence regions where the directions of the semi-major and semi-minor axes are independent of the selected datum (section 2.4.3.5).

- Intersection residuals:
  The residuals resulting from ray intersections (section 4.4.7.1) are a measure of the precision of object point determination. In the stereo case (e.g. using a stereo camera as shown in Fig. 3.127) the intersection residual (y-parallax) can be zero, even when intersection error is present, due to the error being directed along the epipolar line.

7.2.1.2 Accuracy

The accuracy of a photogrammetric measurement can only be assessed by the use of independent references which have not been incorporated into the previous calculation process. References can be artefacts with known dimensions or control points with coordinates measured by a system with superior accuracy to the one under evaluation. The resulting accuracy includes all systematic and random effects due to the measurements.

Typical procedures for determining accuracy include:

- Checks of independent reference lengths:
  From the measured coordinates of end targets on reference lengths, distances between them can be calculated and compared with the calibrated lengths. Reference lengths for close-range testing are relatively easy to manufacture and their use follows normal practice in technical applications where coordinates are the source of derived elements (distances, deformations, surfaces, etc.). Furthermore, reference lengths permit the calculation of standardized length comparisons and the associated traceability back to the SI unit of the metre. The precondition is that at least one known length, e.g. a scale bar, is included in the photogrammetric network. The test procedure is an industrial standard defined by the German guidelines VDI/VDE 2634/1 and can also be used as an evaluation test for camera calibration (see section 7.2.3).

- Comparison with independent reference points:
  If some of the object points used in the bundle adjustment have independent reference coordinates of higher accuracy, then a direct comparison between bundle coordinates and reference coordinates can be made.

  This situation is difficult to achieve in practice. Firstly, comparison points which can be accurately identified, and have higher accuracy reference

coordinates, are difficult to create in industrial situations. Secondly, independence of the data sets will not be achieved if the same measurement method, such as photogrammetry, is separately used to create both (bundle and reference) as this will result in correlations between them.

– Comparison with independent intersections:
Following a bundle adjustment, it is more effective to use the separate measurement (by intersection) of additional image points which correspond to object reference points or which define object reference lengths. These intersections utilize the interior and exterior orientation parameters derived from the bundle adjustment and, as the target points are not included in the actual bundle adjustment process, their measurements are independent of it. They are therefore suitable for estimating the accuracy of a photogrammetric process which includes all uncertainties in the bundle adjustment. It also corresponds to the situation in practice where computed orientation data (from a bundle adjustment) are often used to make further object measurements.

For industrial photogrammetric systems, the achievable accuracy is often given in the form $x\,\mu m \pm y\,\mu m/m$. Here x is the minimum absolute uncertainty of a measurement. Parameter y is a length-dependent uncertainty which, in principle, relates to the image scale.

7.2.1.3 Relative accuracy
The term relative accuracy, introduced in section 3.3.1, is used in photogrammetry to represent the performance of a measurement system independently of the dimension of its measurement volume. It is a dimensionless number, e.g. in the form 1:100 000, 10^{-6} or 10 ppm. It is usual in close-range photogrammetry to quote the achieved object measurement accuracy in relation to the maximum extent of the object.

If precision measures are used, e.g. RMS value of the object coordinates, then the derived relative value still remains a precision value. Since they are often based on a simple standard deviation, they then also only relate to a 68 % coverage of the measurements.

7.2.2 Metrological parameters

7.2.2.1 Measurement uncertainty
According to the International Dictionary of Metrology (VIM – Vocabulaire international de métrologie), and related publications such as DIN 1319-1, *measurement uncertainty* describes the range within which the *true value* of a

measurand[1] lies. The true value is itself never known. value obtained from measurements, i.e. external conditions.

Measurement uncertainty is a characteristic value obtained from measurements. Additionally, external conditions, influences due to the system and application and empirical values also contribute to this. Measurement and measurement uncertainty are estimated values and a complete measurement result only exists when both are specified. The measurement and its uncertainty must be traceable back to a reference standard, e.g. the SI unit of the metre.

The statement of measurement value and measurement uncertainty is typically given in the following form:

$L = 1533.162$ mm ± 0.015 mm

The measurement uncertainty encompasses all the unknown systematic and random error contributions to the measured value (Fig. 7.8). The extent defined by the uncertainty is itself uncertain. To deal with this in practice, a *coverage factor k* is introduced which is used as a multiplier of the coverage in the form:

$L = (1533.162 \pm 0.030)$ mm, $k = 2$

For a factor $k = 1$, the uncertainty value is a *standard uncertainty*, designated u and expressed in the form $\pm u$. In the example above, it is ± 0.015 mm. When a coverage factor greater than 1 is applied, the uncertainty is then known as the *expanded uncertainty* and designated U where $U = k \cdot u$ and is expressed in the form $\pm U$. In the example above with $k = 2$, $U = \pm 0.030$ mm.

If the probability density function which characterizes the measurand is a normal distribution, then k defines the confidence interval, e.g. a 95 % confidence interval for $k = 2$, and the standard uncertainty limits correspond to the standard deviation.

7.2.2.2 Reference value

Reference values are frequently used in practice in order to provide a comparison for measurements and hence an estimate of their quality. A typical example is a reference length defined by the end targets on a scale bar which is measured by a system under test.

Reference values are themselves derived from measurements. Values measured by a system of higher order accuracy, or supplied with an officially recognized calibration certificate, are acceptable as reference values if their own measurement uncertainties, within the context of the application, are sufficiently small that they can effectively be regarded as true or error-free values. A reference uncertainty 5–10

---

[1] The quantity to be determined, for example, the length of a scale bar at 20° C. Effects such as correction for measurement at a different temperature would influence the uncertainty.

times smaller than achievable by a system under test would generally be regarded as good.

7.2.2.3 Measurement error

The term *measurement error* defines the departure of an assigned value, derived from measurements of a quantity such as length or coordinate, from a reference value.

Fig. 7.8 illustrates that measurement errors have systematic and random components. If the systematic errors are known, they can be determined by means of calibration and largely removed by applying corrections to the measurements. Unknown systematic errors in the measurements add to the random components to give the resulting measurement uncertainty.
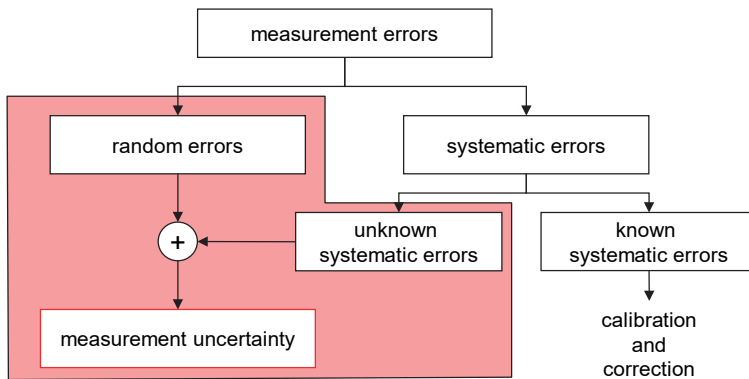
Fig. 7.8: Measurement error and measurement uncertainty.

7.2.2.4 Accuracy

Accuracy describes the closeness of agreement between a measurement result and a measurement standard or accepted reference. Higher accuracy implies closer agreement but the term is strictly only qualitative, not quantitative. A statement about accuracy can only be made after a comparison is made with an independent, higher order reference value.

7.2.2.5 Precision

Precision describes the statistical spread of a measured quantity as derived from repeated measurements or an adjustment process. It is normally expressed as a standard deviation or RMS value. Precision is a measure of relative accuracy. If a quantity is measured multiple times under repeatable conditions, it indicates the internal spread of a measurement result. In an adjustment calculation, precision is

calculated as a standard deviation. Estimates of precision should always be provided with the coverage factor, e.g. 1-sigma.

7.2.2.6 Tolerance

Tolerance is a parameter used in manufacturing to define the permissible limits to a feature's dimensions. Relative to a *nominal* value it can have different positive or negative values.

In a typical industrial measurement process, it is necessary to measure critical dimensions on a part and decide if it is within tolerance, and therefore accepted, or out of tolerance and therefore rejected. Rejection is costly as parts must then either be discarded, recycled or reworked to bring them within tolerance.

Fig. 7.9 shows the relationship between the accuracy of a measuring device or procedure, and the extent to which it can decide if critical part dimensions are in tolerance (green) or out of tolerance (red). Depending on the ratio between tolerance and measuring uncertainty, e.g. 10:1, there is an area (yellow) where no clear decision can be made.
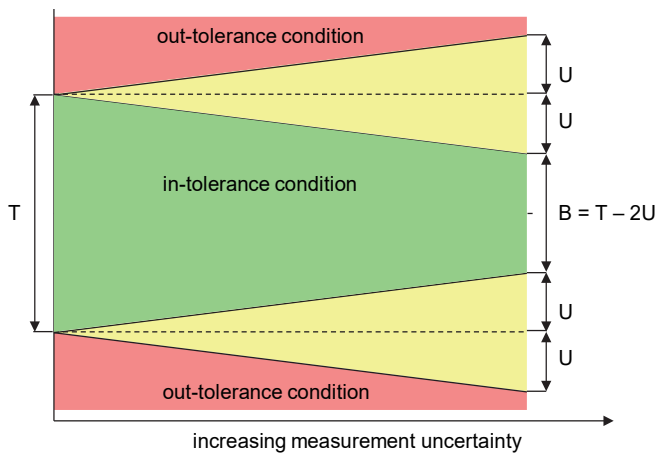
Fig. 7.9: Suitability of testing device or procedure for in-tolerance evaluation.

The device's accuracy is indicated by an agreed expanded measurement uncertainty $\pm U$ which corresponds, for example, to a 95 % confidence level ($k = 2$). The diagram shows a measurement made within the tolerance limits (dotted lines). On the extreme left, a perfect measuring device with zero uncertainty can reliably determine if a feature lies anywhere within the entire tolerance band T. As the device uncertainty increases towards the right, a feature measured at the tolerance limits could, due to device uncertainty indicated by the yellow areas, be either in tolerance or out of

tolerance. To be sure (within the confidence level) that the part is good, only measurements within the green area are acceptable. In effect, an increasing measurement uncertainty reduces the effective size of the tolerance band. In the extreme case (right hand side, not shown) where $U = T/2$, the device cannot determine if a part is in or out of tolerance.

The problem in practice is to select a compromise between measuring device or system with high accuracy (increased cost) giving rise to reduced rejections (lower cost).

7.2.2.7 Resolution

The *resolution* of a measurement system is the smallest increment which it can display or store, for example as represented by the most significant decimal place in a digital display. Resolution therefore defines the smallest change in the quantity to be measured which produces a significant change in the measurement signal, i.e. one which is above the noise level of the measuring system. In contrast, the *resolving power* (section 3.1.5) defined in optics and photogrammetry denotes the capability of an optical system to transmit a threshold frequency with sufficient contrast or modulation.

Fig. 7.10 illustrates the relationship between reference value, resolution, precision and measurement error, defined above. The measured or displayed values are spread according to their precision or repeatability, but may depart significantly (measurement error) from the reference value.
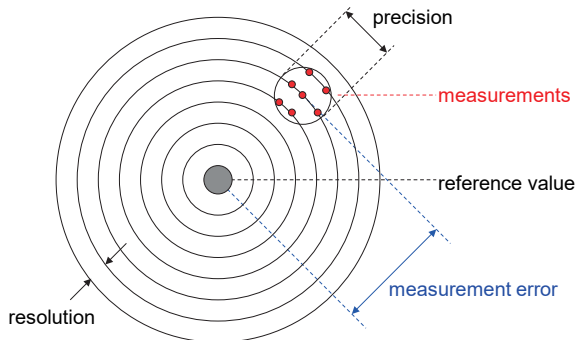
Fig. 7.10: Resolution, precision and measurement error (after Hennes, 2007).

7.2.3 Acceptance and re-verification of measuring systems

The checking of achievable accuracy is of fundamental importance in industrial metrology. In the field of mechanical coordinate measuring machines (CMM), long-

established and standardized methods (VDI/VDE 2617, ISO 10360-2, JCGM/GUM) define parameters and procedures for acceptance, re-verification and monitoring of measuring accuracy which are generally accepted and implemented in practice. They are applicable to both CMMs with touch probes and those with optical sensing heads. Photogrammetric systems employing touch probing can also be evaluated according to these guidelines. Since the year 2000, evaluation of optical non-contact 3D measuring systems has been covered by the German guideline VDI/VDE 2634. This recommends procedures for the acceptance and re-verification of systems based on point-by-point probing and area scanning. In the coming years, a new binding framework for optical 3D measurement systems will come into force with the international standard ISO 10360-13, replacing older guidelines.

7.2.3.1 Definition of terms

– Acceptance test:
  An *acceptance test* is the procedure for acceptance and approval of a measuring system after installation at the customer's site. The selected acceptance procedure is usually incorporated in the delivery contract. The goal of the test is the proof of the specified measuring accuracy under defined conditions. The acceptance test is usually performed jointly by system supplier and customer.
– Re-verification test:
  Re-verification or monitoring is the periodic checking of the measuring system after commissioning to ensure that it conforms with specifications. Compared with an acceptance test, re-verification can be simpler. As a rule, it is carried out by the user, who also defines the time interval between such checks.
– Traceability:
  Traceability is the establishment of a link between the measured quantities, with their uncertainties, and a measurement standard, for example the standard metre. For this purpose, a continuous chain of comparative measurements up to the national representation of the standard must be provided (Fig. 7.11). The standard used, for example a reference scale bar, must be calibrated and certified by a recognized calibration service.
– Characteristic parameter:
  A *characteristic parameter* is a measured or calculated value (threshold, maximum permitted value) which characterizes the performance of a system or its individual components (for example probing error, see below).
– Probing error:
  Probing error is the value that describes the precision of probing of a single measured point. This parameter is used mainly with respect to CMMs in which the active probe, as distinct from the length measuring system, contributes to the total system accuracy. For example, the probing deviation can be determined by repeated measurement of geometrically known reference objects (for example a

sphere). Residuals of single measurements with respect to the surface of the reference object indicate the probing error. Determination of probing error for area scanning systems is discussed in section 7.2.3.5.
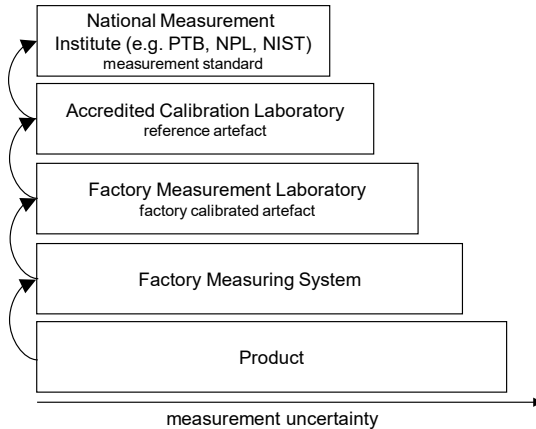
Fig. 7.11: Traceability to national standards.

– Length error:
  The three-dimensional length error E (also called *length measurement error LME*) is defined as the difference between a measured (displayed) length L_m and the calibrated reference length L_r:

  $$E = L_m - L_r \qquad (7.15)$$

  The *length measurement error* is usually derived from the measurement of two single probings (for example on a gauge block). Alternatively, it can be determined from the distance measured between two spheres if the probing error can be eliminated.

  The length measurement error is used to analyse the accuracy of length measurement. Calibrated reference lengths can easily be established (for exceptions see below), and they can be traced back to a standard. Uncertainty of length measurement implicitly includes the probing uncertainty. The maximum permitted positive and negative limit of length error E is the maximum permissible error (*MPE*). It is defined as a length-dependent value that may not be exceeded in an acceptance or re-verification test. The *MPE* of length error E is shown graphically in Fig. 7.15 and expressed analytically as:

  $$MPE(E) = A + K \cdot L \leq B \qquad (7.16)$$

where
E: length error
A, K: machine-specific constants
L: measured length
B: maximum permitted deviation of length measurement

Since the error of photogrammetric length measurement does not necessarily depend on the length itself, the constant *K* may be zero.

– Sphere-spacing error:
Sphere-spacing error indicates the capability of a system to measure the separation between the centres of two spheres which are derived from measurements on the spherical surfaces. The probing error is not explicitly included in the sphere-spacing error as multiple sampling of the surfaces averages it out. This parameter is particularly applicable to area scanning systems which cannot always directly determine a length measurement error. The length measurement error can be estimated from the sphere-spacing error if the probing error is known.

7.2.3.2 Differentiation from coordinate measuring machines (CMMs)

With respect to mechanical CMMs, photogrammetric 3D measuring systems have fundamentally different properties which are apparent in the procedures and parameters which characterize them:

– Image-based measurement of a large number of points:
Optical 3D measuring systems, being image-based, enable the simultaneous registration of large numbers of object points (in the limit, each pixel). In contrast, touch-probe CMMs measure only one point per probe, although optical and line-scanning probing is also possible.

– Triangulation principle:
Photogrammetric and fringe projection systems are based on triangulation, which leads to accuracies of object points which are dependent on scale and configuration. A homogeneous accuracy cannot, therefore, be expected within a specified measuring volume.

– Mobility:
Optical 3D measuring systems are mobile and can be brought to the object. Consequently, from time to time their calibration data may change, or they operate under changing environmental conditions, or their imaging configuration may vary.

– Flexible configurations:
Non-stationary photogrammetric systems allow a free choice of camera stations. Users themselves therefore determine the number and distribution of the images, the selection of cameras and lenses and the type of object targeting or probing.

– Unlimited measuring volume:
  In principle, the measuring volume of photogrammetric systems is unlimited in so far as depth of focus and field of view allow. If scale-dependent resolution of object details is taken into account, arbitrary object dimensions can be measured. Conversely, the measuring volume of mechanical CMMs is always limited.

7.2.3.3 Reference artefacts

A *reference artefact* is a physical object with known (calibrated) geometrical parameters. It should be economical to manufacture and easy to handle. As indicated in section 7.2.2.2, the calibration accuracy for a reference object should be approximately 5–10 times higher than that of the measuring system to be checked. Acceptance and re-verification procedures usually require a calibration certificate for the reference object.
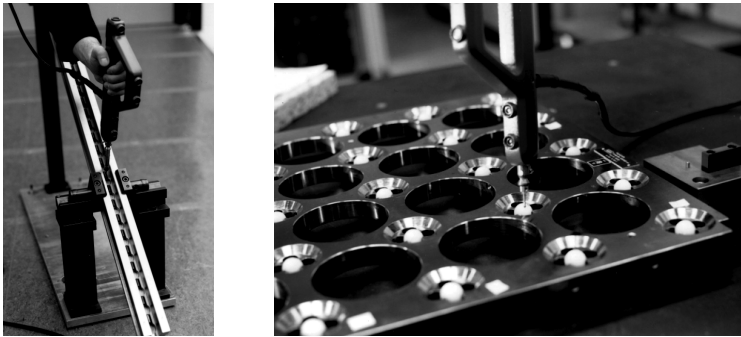
a) Step gauge   b) Ball plate

Fig. 7.12: Standard reference objects for 3D coordinate measuring machines.

If the measurement uncertainty of a system under test is determined using an artefact, the result is not only affected by random and systematic measurement error by also by the artefact's own calibration uncertainty. The measurement uncertainty cannot therefore be better than the calibration uncertainty.

The type of artefact used depends mainly on measurement volume, method of probing, availability and cost. Tactile probing implemented in coordinate measurement technology makes use of gauge blocks, step gauges and ball plates (Fig. 7.12) designed for a measurement volume with a typical diagonal length of 2 m. These are only suitable for photogrammetric systems using touch probing (example in Fig. 6.27 and Fig. 6.29).

Fig. 7.13a shows an arrangement of reference scale bars for testing optical 3D measuring systems according to VDI/VDE 2634 part 1. The scale bars are provided with the type of targets which are otherwise used by the system under test. Fig. 7.13b

shows a dumbbell target which has a calibrated separation of its ball centres (see also Fig. 7.24). This type of artefact can be used to evaluate the sphere-spacing error of area scanning systems according to VDI/VDE 2634 part 2.
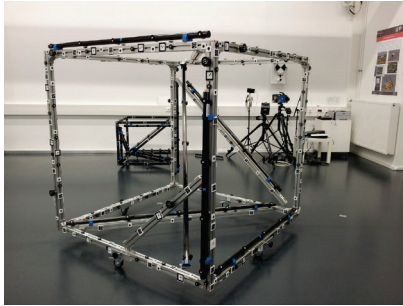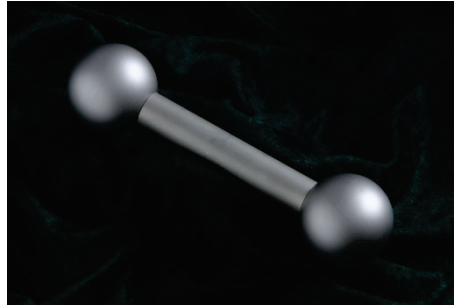
a) Arrangement of reference scale bars   b) Dumbbell artefact

Fig. 7.13: Test artefacts for 3D measurement systems.

Supplying reference objects for acceptance and re-verification tests for photogrammetric systems can be problematic. It is difficult to find scale bars suitable for testing large measuring volumes (dimensions >3 m). For smaller measuring volumes (<1 m³) optical 3D systems achieve measuring accuracies of the order of 10 μm (1:100 000) requiring high-precision reference objects with suitable targets.

For larger measuring volumes, such as required to measure parabolic antennas or ships, measuring accuracy is often checked using laser trackers (section 6.3.2.2). Differences in coordinates or in computed distances enable a comparison of accuracies. However, several points should be noted:

– high accuracy photogrammetric systems almost match the performance of laser trackers;
– laser tracker measurements are costly in time and personnel;
– laser trackers require geometric stability over the period of the observations;
– multi-purpose targets are required which are measurable by both systems without additional significant loss of accuracy.

7.2.3.4 Testing of point-by-point measuring systems

Part 1 of the VDI/VDE 2634 guidelines recommend parameters and methods for acceptance and re-verification testing of optical 3D systems which operate with point-by-point probing. The single parameter to be tested is the length measurement error and the testing is made by measurement of calibrated length artefacts.

Scale bars manufactured at appropriate lengths can serve as references. They should use the same type of targets as are used for the actual object measurement. The length can be calibrated, for example, by optical CMMs (for shorter lengths), high

accuracy photogrammetry, by field survey (using total stations) or by laser interferometry.

In order to guarantee a sound analysis of the system, the arrangement of reference scale bars in object space should match the measuring task. If equal measuring accuracies are required for all coordinate axes, the scale bars must be arranged in such a way that direction-dependent length measurement deviations can be determined. A possible set-up is illustrated in Fig. 7.13a and Fig. 7.14. Of seven scale bars, three are arranged parallel to the coordinate axes and four along diagonals of a cuboid measuring volume. In order to increase the number of reference lengths of different sizes, individual scale bars can be divided into several sections. In this manner, the scale bar displayed in the foreground of Fig. 7.14 provides six partial distances which can be combined to give 21 different lengths.
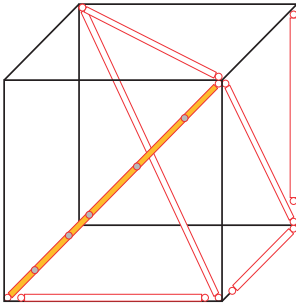
Fig. 7.14: Arrangement of scale bars in the measurement volume.
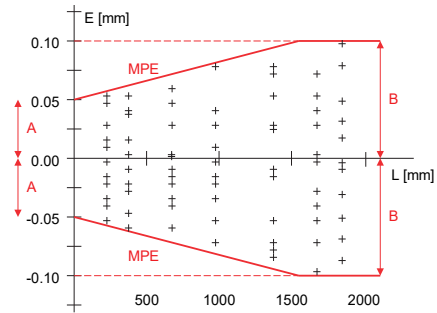
Fig. 7.15: Length measurement errors and limiting bounding box.
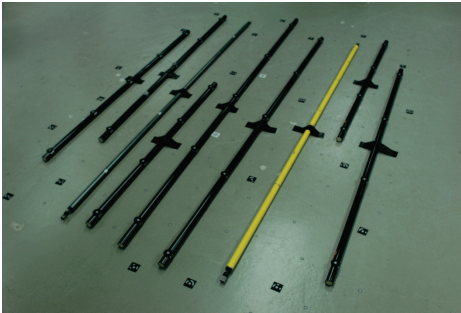
Fig. 7.16: Photogrammetric scale bars.

Fig. 7.17: Imaging configuration for assessment of length measurement error.

As a practical tool for the display and analysis of length measurement errors, a diagram showing the measured differences with respect to the nominal distances can

be used (Fig. 7.15). The maximum permitted limits shown by the red bounding box correspond to eqn. (7.16). A measuring system can be accepted as successful if all length errors lie within the box.

Fig. 7.16 shows examples of photogrammetric scale bars that can be used either as reference lengths to scale the measuring volume, or as check lengths in the procedures described above. Fig. 7.17 illustrates a photogrammetric imaging configuration for measuring a VDI cube.

If the scale-bar arrangement illustrated in Fig. 7.14 is not possible, or the system under test can only view the measurement volume from one direction (e.g. the stereo camera system in Fig. 6.40), then a single scale bar can alternatively be moved to different positions in the volume for measurement. Fig. 7.18 shows the testing of a stereo camera system. This continuously observes a reference scale bar with five retro-reflective ball targets which is recorded at multiple locations and orientations within the measurement volume.
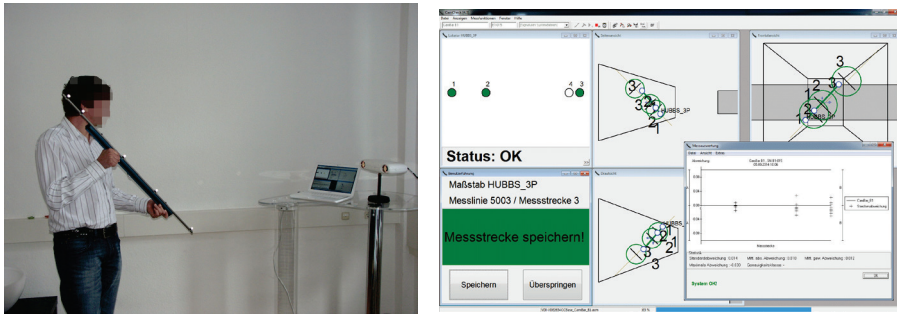
a) Reference scale-bar and stereo camera   b) Evaluation of length measuring error

Fig. 7.18: Evaluation of length-measuring error by moving reference scale-bar (AXIOS 3D).

If a photogrammetric calculation has provided standard deviations of object coordinates, then these can be used to estimate analytically the expected length measurement error. By applying error propagation to the equation for distance measurement:

$$L^2 = (X_2 - X_1)^2 + (Y_2 - Y_1)^2 + (Z_2 - Z_1)^2 = \Delta X^2 + \Delta Y^2 + \Delta Z^2 \qquad (7.17)$$

the variance of the calculated length can be obtained:

$$s_L^2 = \left(\frac{\partial L}{\partial X_1}\right)^2 s_{X_1}^2 + \left(\frac{\partial L}{\partial X_2}\right)^2 s_{X_2}^2 + \left(\frac{\partial L}{\partial Y_1}\right)^2 s_{Y_1}^2 + \left(\frac{\partial L}{\partial Y_2}\right)^2 s_{Y_2}^2$$
$$+ \left(\frac{\partial L}{\partial Z_1}\right)^2 s_{Z_1}^2 + \left(\frac{\partial L}{\partial Z_2}\right)^2 s_{Z_2}^2$$
$$= \left(\frac{\Delta X}{L}\right)^2 s_{X_1}^2 + \left(\frac{\Delta X}{L}\right)^2 s_{X_2}^2 + \left(\frac{\Delta Y}{L}\right)^2 s_{Y_1}^2 + \left(\frac{\Delta Y}{L}\right)^2 s_{Y_2}^2$$
$$+ \left(\frac{\Delta Z}{L}\right)^2 s_{Z_1}^2 + \left(\frac{\Delta Z}{L}\right)^2 s_{Z_2}^2$$

(7.18)

Assuming coordinate standard errors are equal along the individual axes X, Y and Z ($s_{X1} = s_{X2}$ etc.), the following is obtained:

$$s_L^2 = 2\frac{\Delta X^2}{L^2}s_X^2 + 2\frac{\Delta Y^2}{L^2}s_Y^2 + 2\frac{\Delta Z^2}{L^2}s_Z^2 = \frac{2}{L^2}(\Delta X^2 s_X^2 + \Delta Y^2 s_Y^2 + \Delta Z^2 s_Z^2)$$

(7.19)

and if standard errors are equal in all directions ($s_X = s_Y = s_Z$):

$$s_L^2 = \frac{2s_{XYZ}^2}{L^2}(\Delta X^2 + \Delta Y^2 + \Delta Z^2) = \frac{2s_{XYZ}^2}{L^2}L^2 = 2s_{XYZ}^2$$
$$s_L = \sqrt{2} \cdot s_{XYZ}$$

(7.20)

In order for the length measurement error to apply to all lengths measured by one system, a 3-sigma coverage factor (99 % confidence level) should be applied. The theoretical length measurement error is then given by:

$$E = k\sqrt{2} \cdot s_{XYZ} = \sqrt{18} \cdot s_{XYZ} = C \cdot s_{XYZ} \approx C \cdot RMS_{XYZ}$$

(7.21)

In the first instance, this estimation is only valid if it results from a bundle adjustment where all coordinate directions are of equal accuracy. This applies to the homogeneous all-around configuration shown in Fig. 7.17. In addition, camera stability plays an important role since its impact is not always apparent in the resulting standard deviations of object points (precision), although it does affect the final LME (accuracy). Suppliers of high-accuracy, offline photogrammetric systems can achieve a maximum uncertainty of object coordinates of around 4μm±4μm/m (see section 3.5.2). For a VDI cube of 2m side length, a typical RMS$_{XYZ}$ of 4–5μm with an LME of around 21μm is obtained. The constant C in eqn. (7.21) is then around 4–5 (compare with example 7.1). For a DSLR camera (ring flash not mounted around the lens) applied in a similar imaging configuration, the RMS_{XYZ} amounts to around 25μm with an LME of approximately 75μm, i.e. C is around 3. Using a less stable camera, e.g. with plastic housing or ring flash mounted on lens, but which has similar RMS values, may result in an LME of up to 150μm ($C=5$–6). Ideally, the objective is to achieve a small RMS value (high precision) and a small value of C (high accuracy).

Eqn. (7.21) does not hold for online measuring systems, such as stereo cameras, which have been pre-calibrated by bundle adjustment and measure further points by intersection. On the one hand, redundancy is much lower, e.g. only two images. On the other, the physical structure of the camera could change slightly between the time of calibration and the actual time of measurement. There are also differences in measuring accuracies between the XY and Z directions.

$$
\begin{aligned}
s_L^2 &= \frac{2}{L^2}\left[s_{XY}^2(\varDelta X^2 + \varDelta Y^2) + s_Z^2 \varDelta Z^2 \right] \\
s_L &= \frac{1}{L}\sqrt{2s_{XY}^2(\varDelta X^2 + \varDelta Y^2) + 2s_Z^2 \varDelta Z^2}
\end{aligned}
\tag{7.22}
$$

For s_Z and s_{XY} the assumptions in eqns. (4.88) and (4.92) can be used.

Example 7.1:

After bundle adjustment of 144 images of the reference body in Fig. 7.13a, taken by an ALPA camera, the following RMS values of object coordinates were obtained:

RMS (1-sigma)   X: 0.0051 mm   Y: 0.0049 mm   Z: 0.0040 mm

According to eqn. (7.20), the standard error of length is given by $s_L = 0.007$

At a confidence level of 99%, corresponding to a 3-sigma coverage factor, the theoretical length measurement error according to eqn. (7.21) is given by $E = 0.020$ mm. The value agrees with an empirical result from multiple tests which gave a value between 0.019 and 0.025 mm ($C \approx 4$).

Example 7.2:

The artefact in Fig. 7.13a was used to test an AXIOS stereo camera (Fig. 3.127a, bases $b = 200$ mm, measuring volume up to 3.0 m distance). The following RMS values of object coordinates were obtained following system calibration by bundle adjustment (72 images per camera):

RMS (1-sigma)   X: 0.007 mm   Y: 0.007 mm   Z: 0.007 mm

The standard error of length was calculated using eqn. (7.20) and eqn. (7.21) which resulted in $s_L = 0.01$ mm and $E = 0.03$ mm.

In subsequent independent measurements of a 1.63 m long reference scale bar, oriented randomly in space at distances h between 1 m and 3 m (h/b between 5:1 and 15:1), an LME of 0.9 mm was obtained. In this example, the LME in online mode is larger by a factor of 128 compared with the single standard deviation from the system calibration. The example demonstrates the potential large difference between the accuracy estimates from the bundle adjustment and the intersections.

7.2.3.5 Testing of area-scanning systems

Part 2 of the VDI/VDE 2634 guidelines recommends parameters and methods for acceptance and re-verification testing of optical 3D systems using, for example, fringe projection, tracked laser line scanners and image correlation methods.
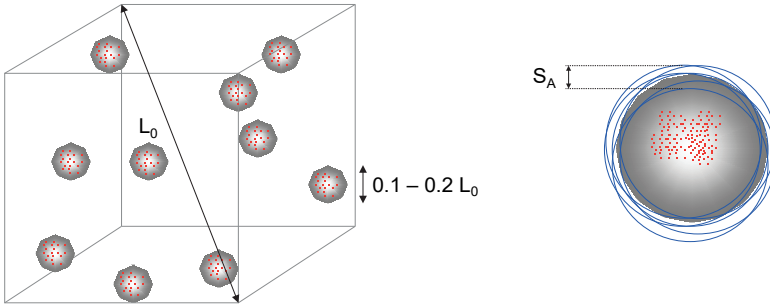
Fig. 7.19: Method of determining probing error.

The *probing error* parameter describes the error effects associated with surface point coordinates in a small measurement volume. It is derived from the measurement of a calibrated spherical surface to which a best-fitting sphere with variable radius is fitted. The range S_A of the measurement deviations from the best-fit sphere defines the probing error. The reference sphere is positioned at multiple locations within the measurement volume as illustrated in Fig. 7.19. The sphere's diameter should amount to around 10–20 % of the diagonal L_0 of the measurement space.
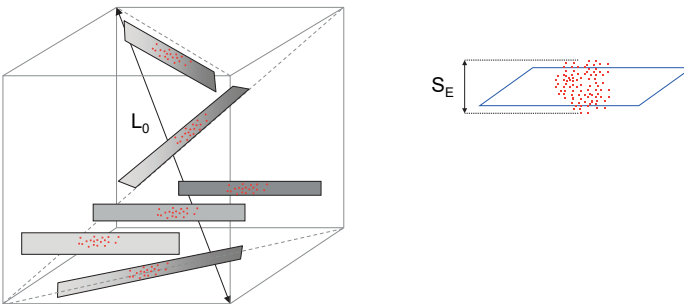
Fig. 7.20: Method of determining flatness error.

The *flatness error* parameter indicates the capability of the test system to measure a plane surface. For this purpose, a reference flat surface is measured at multiple locations within the test volume and the range of deviations of the measurements from a best-fit reference plane is determined at each (Fig. 7.20). The length of the reference flat should be around $0.5 \cdot L_0$.

The *sphere-spacing error* parameter indicates the deviation between the calibrated and measured separation of two spheres whose surfaces are scanned and centres found using a best-fit sphere with a given radius (Fig. 7.21). The sphere-spacing error must be representative of the entire measurement volume. The length

of the dumbbell artefact used for this purpose should be around $0.3 \cdot L_0$ with a sphere diameter of $0.1-0.2 \cdot L_0$. Since the probing error is not a part of the test parameter, due to multiple scanning and subsequent best fit of spheres, an ISO-compliant length measurement error is not generated (ISO 10360-2). Generally, the sphere-spacing error is always smaller than the length measurement error.
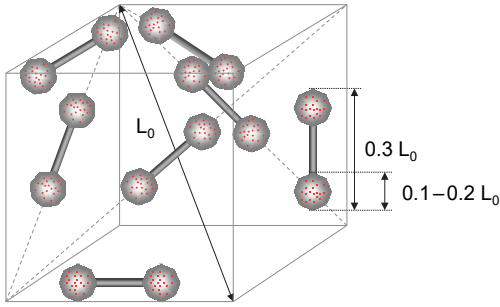
Fig. 7.21: Method of determining sphere-spacing error.

When testing area scanning systems it is generally acceptable to eliminate up to 3 ‰ of the measured points from the raw data. This acknowledges the existence of unavoidable outliers in the data which are caused, for example, by reflective highlights off the object surface. It is also the case that many scanning methods acquire several million measurement points per scan and, as a matter of practicality, a thinning or filtering of the point cloud is necessary.
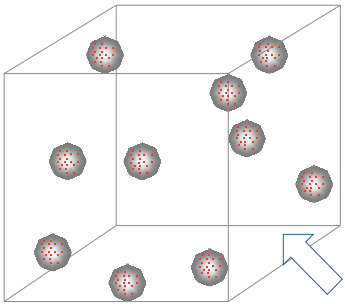
Fig. 7.22: Moving object with fixed sensor position.
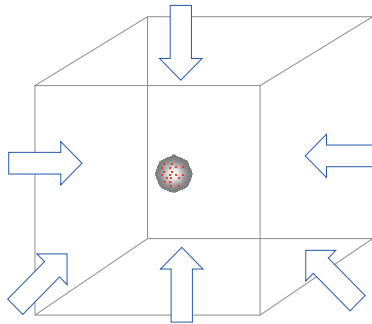
Fig. 7.23: Moving sensor with fixed object position.

Part 3 of the VDI/VDE 2634 guidelines deals with systems that combine area scanning with multiple sensor orientations (multiple point clouds) in order to measure objects

that are larger than the direct measuring volume of a surface measuring sensor. As examples, multiple point clouds can be created by moving the object (Fig. 7.22) and/or the sensor (Fig. 7.23), or by combining multiple sensors into one measurement system. Here it is necessary to deal with the task of merging individual point clouds into a single point cloud (registration), i.e. transforming them into a common coordinate system (compare with section 6.9).

As with VDI/VDE 2634/2, the characteristics of probing error, sphere-spacing error and length measurement error are again evaluated. Practical constraints ensure that the recommended testing procedures of 2634/3 differ from those of 2634/2. In 2634/3, the probing error *Form* describes the span of the radial distances of the measuring points from a calculated best-fit sphere with free radius, while the probing error *Size* is calculated from the difference between the measured diameter and the calibrated diameter of the sphere. It is highly recommended to study the guidelines carefully for specific requirements and parameter definitions.
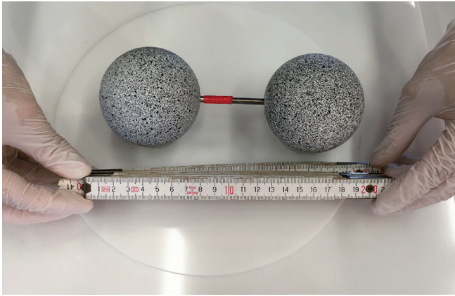
Fig. 7.24: Dumbbell with random pattern.
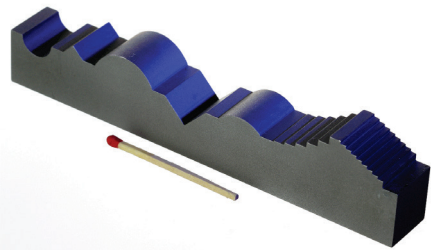
Fig. 7.25: Reference object with calibrated geometric elements.

In addition to the diverse reference artefacts proposed in various guidelines, there are a number of other suggestions for quality checking of area-probing systems. Fig. 7.24 displays a dumbbell with random pattern that is used to check photogrammetric systems based on digital image correlation or SfM. Fig. 7.25 shows an artefact composed of different geometric elements (half cylinder, planes, steps) whose nominal shape is calibrated by a high-accuracy coordinate measuring machine (CMM). Fig. 7.26 illustrates a reference body for the measurement of free-form surfaces. This is also calibrated by a CMM and has its own local coordinate system defined by the reference spheres in the corners. The assessment of measuring uncertainty is done using a nominal-to-actual comparison of the measured surface with the CMM reference surface. This is achieved, for example, by a best-fit which minimizes the distances between both surfaces.
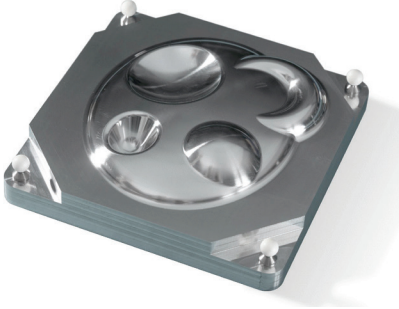
Fig. 7.26: Reference object with free-form surfaces (NPL London).
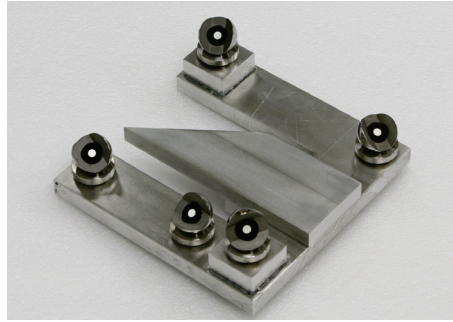
Fig. 7.27: Reference object for determination of spatial resolution.

Fig. 7.27 shows a reference artefact designed for measuring the lateral spatial resolution of an area-based probing system. The wedge located in the centre has sharp edges. As it gets narrower towards the apex, there is a point at which the width is smaller than the resolution of the measuring system. This enables the system's ability to measure edges to be evaluated, as well as determining the minimum object size which the system can sensibly determine.

7.3 Strategies for camera calibration

7.3.1 Calibration methods

The purpose of camera calibration is to determine the geometric camera model described by the parameters of interior orientation (see section 3.3.2):
– spatial location of the perspective centre in the image coordinate system: principal distance and image coordinates of principal point;
– parameters describing image errors: distortion and sensor corrections.

In general, the interior orientation is assumed to be known and constant for metric cameras. The problem of camera calibration therefore mainly concerns those imaging systems, e.g. semi- metric cameras, commercially available digital cameras, whose geometry is subject to variation over time. However, depending on the actual accuracy specifications, even metric cameras may have to be calibrated for the duration of image acquisition.

The calibration of an imaging system is of major importance in many photogrammetric applications where the measurement must be optimized with respect to accuracy and economical aspects. Already at the planning stage it is therefore necessary to evaluate if a pre-calibrated camera can be used, if a pre-calibration can be done immediately before the measurement, if the camera can be

calibrated simultaneously with the measurement or if the camera can be calibrated after the measurement (with the camera settings used on site).

In general, imaging systems can be classified as follows:

- 1 uncalibrated, or approximately calibrated, camera for general network configuration;
- 1 pre-calibrated, stable camera for general network configuration;
- multiple uncalibrated cameras in flexible imaging configuration;
- several calibrated and oriented cameras in a mechanically fixed configuration.

If time and configuration conditions permit, then sufficient images can be taken during object measurement such that the camera can be calibrated simultaneously with the 3D object reconstruction (section 7.3.1.3). This procedure is often selected in offline photogrammetry and normally results in the highest measurement accuracy.

If only a restricted or weak camera network can be configured, then the camera can only be partly calibrated or not at all. In this case, if possible, the camera must be calibrated in a separate process directly before or after object measurement (e.g. using a test field, section 7.3.1.1). Here the validity of the camera parameters directly depends on its mechanical stability.

Today, camera calibration techniques involve a computational solution for camera parameters (camera model) which often cannot be separated from the actual object measurement. Consequently, an understanding of the different approaches to calibration requires a detailed knowledge of photogrammetric orientation and object reconstruction, especially bundle adjustment (see section 4.4).

Two calibration methods can effectively be distinguished. These are characterized by the reference object used and by the time and location of calibration[2]:

- test-field calibration (section 7.3.1.1);
- self-calibration (section 7.3.1.4).

7.3.1.1 Test-field calibration

Test-field calibration is based on a suitable targeted field of object points with optionally known coordinates or distances. This test field is imaged from several camera stations, ensuring good ray intersections and filling the image format. Test fields can be mobile (Fig. 7.28a), or stationary, e.g. a building wall. In general, there is a distinction between plane test fields, such as checkerboards, or spatial test fields where the 3D points are spatially distributed in 3D space (example in Fig. 7.28b, see also section 7.3.2).

---

2 Laboratory calibration with goniometers or collimators is an older concept no longer discussed here.

The parameters of the camera model are then calculated as part of a bundle adjustment in which the parameters of exterior orientation and the unknown 3D object coordinates are also normally calculated. Any known data (coordinates, distances) can be incorporated in different ways and are used to provide scale.

a) Portable plane test field   b) Spatial test field
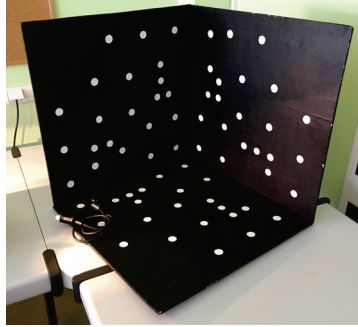
Fig. 7.28: Examples of photogrammetric test fields.

It is not necessary for the target points in a test field to be reference points with nominal coordinates of superior accuracy. In fact, potential errors in reference coordinates, perhaps due to displacements caused by stresses in the supporting structure, can result in an erroneous or inaccurate camera calibration without any indicators in the bundle adjustment. It is therefore better to provide good quality measurement and then determine their 3D coordinates using a free-net bundle adjustment. Scale bars can, of course, also be introduced to provide the bundle with absolute scale. However, cameras can be calibrated without scale information because directions are being corrected and these are independent of scale.
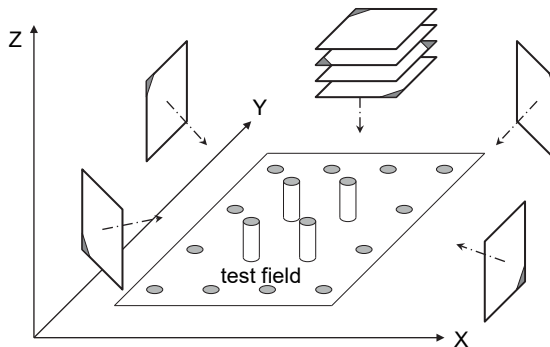
Fig. 7.29: Imaging configuration for test-field calibration.

Fig. 7.29 shows a suitable image configuration for test-field calibration. In order to calibrate the camera, eight images are sufficient in this example. They should image the test field perpendicularly and obliquely and each image should have a relative rotation of 90° around the optical axis (see also Fig. 7.33). The number of object points and the imaging configuration should be chosen such that, considering all the images, the complete sensor format is filled with imaged points. It is additionally advantageous if not all images completely cover the object but only show a part of it. As a consequence, the total number of required images is increased. Fig. 7.30 shows a series of images acquired for test-field calibration. Further imaging configurations are discussed in section 7.3.2.
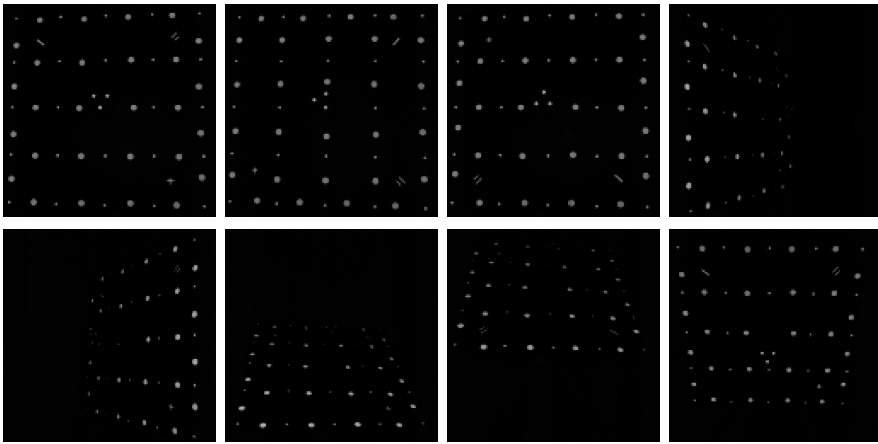
Fig. 7.30: Image series for test-field calibration.

Measured image coordinates and approximately known object data are processed by bundle adjustment to give the parameters of the camera model (interior orientation) as well as the adjusted test-field coordinates and the parameters of exterior orientation.

For test-field calibration, the datum should be defined by an unconstrained technique (section 4.4.3) in order that possible inconsistencies between object point coordinates do not have a negative influence on the calculated parameters. An unconstrained datum can be created by a free net adjustment of the 3-2-1 method.

Numerical calculations can lead to unwanted correlations between the calculated parameters but these can largely be avoided by suitable imaging configurations. It is most important to provide at least one piece of scale information along the viewing direction in order to compute the principal distance. This can, for example, be achieved by a reference distance, by spatially distributed test-field points or by oblique images of a plane test field. Images rotated by 90° around the optical

axis are used primarily to determine the principal point coordinates and affinity parameters. Three-dimensional test fields, e.g. with out-of-plane points, have the advantage over flat test fields that they can offer more points in depth which ensures that parameters are easier to determine and have smaller correlations (see section 7.3.2.2).

Normally, test-field calibrations are done when a simultaneous calibration as part of an object measurement is not possible or an accuracy evaluation of a camera is required. In general, the design of the test field should represent the actual object to be measured. The number and distribution of image points are of major importance for an accurate determination of distortion parameters (see also section 7.3.5). In order to preserve the calibration parameters, there should never be any changes made to the camera (focusing, different lens) between test-field measurement and object reconstruction.

7.3.1.2 Plumb-line calibration

The plumb-line method uses a test field with several straight lines, created for example by vertically hanging wires (plumb lines, Fig. 7.31). Since, in theory, the projection of straight lines is invariant for perspective geometry, all departures from this condition must be caused by distortion effects. The deformed test field lines can only be used to determine distortion parameters and are insufficient to determine also principal distance and principal point. The calculated distortion parameters are not correlated with the further parameters of interior orientation or the exterior orientation parameters.
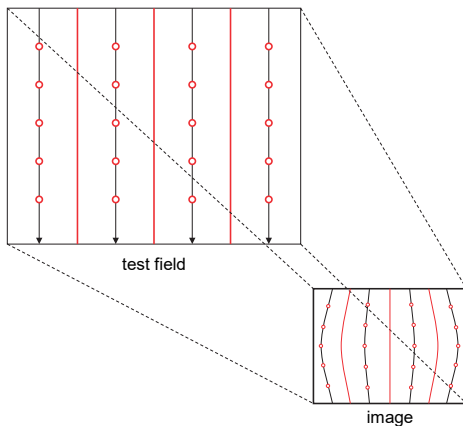
Fig. 7.31: Plumb-line method for test-field calibration.

In a practical implementation of a plumb-line test field, targets can be added to the lines and these individually measured in the image. The lines can also be continuous features, such as thin white plastic cords set against a dark background for enhanced contrast. This arrangement facilitates automatic line following at high point densities. Alternatively, natural straight-line object features such as building edges can be used for calibration.

Plumb-line calibration can be sensibly applied in cases where a pre-calibration of distortion parameters is desired, e.g. if lenses with high distortions such as fish-eye lenses are used and the measured image coordinates are to be corrected for distortion prior to a system calibration (see section 7.3.1.5). Distance-dependent changes in distortion can also be determined by the plumb-line method.

7.3.1.3 On-the-job calibration

The term *on-the-job calibration* is often used where a test-field calibration (recording of a known point field) is combined with the actual object measurement. This approach is reasonable, for example, if the measuring object itself does not provide suitable geometry to enable self-calibration (see section 7.3.1.4).

A simple solution is provided by a portable frame consisting of several spatially distributed scale bars, positioned beside the measuring object and photographed simultaneously with it. The local coordinate system of the test field can be used as a three-dimensional object coordinate system and further reference points are not required.

7.3.1.4 Self-calibration

An extension to on-the-job calibration is *self-calibration* which simultaneously uses the images acquired for the actual object measurement also for the calibration. In effect, the test field is replaced by the object itself which must be imaged under conditions similar to those required for test-field calibration (spatial depth, tilted images and suitable ray intersections). Fig. 4.57 illustrates multiple imaging of a targeted car door which calibrates the camera as well as calculating the configurations of the ray bundles and the target coordinates.

The essential advantage of self-calibration is that the parameters of interior orientation are determined simultaneously with measurement of the object, so providing the highest of accuracies in object reconstruction. Any remaining numerical correlations between the calculated parameters, in particular between those of interior and exterior orientation, have a lower influence in self-calibration analyses. With regard to the actual adjustment process, the observations generate the best possible result for the parameters and the calculated set of object points (point cloud) is in conformance with those parameters.

Self-calibration does not require coordinates of known reference points. The parameters of interior orientation can be calculated solely by the photogrammetric

determination of the object shape, i.e. by incorporating only image information and intersection conditions for unknown object points. If employed, reference points can be used to define a particular global coordinate system for the parameters of exterior orientation. In order to define scale, it is sufficient to measure a single reference length in object space although it is good practice to measure multiple reference lengths.

In many applications that are evaluated using Structure-from-Motion, the camera is also calibrated simultaneously. Essentially, the same requirements for camera calibration apply as described above and in the following sections. Deficiencies in the mechanical stability of the camera or in the imaging configuration are often not directly visible. This is because the number of images and pixels are frequently large and this obscures the influence of the interior orientation (smearing effect, see section 4.4.6.2). If present, calibration deficiencies can be reflected as errors in the object space, e.g. as deformations. It should be further noted that most SfM programs state the parameters of interior orientation in pixel units which, if required, must be converted into equivalent metric units, e.g. mm (see section 3.3.6).

If the object to be measured does not permit a suitable image configuration, or if a multi-camera online system is used, then a test-field or on-the-job calibration must normally be performed.

7.3.1.5 System calibration

The expression *system calibration* is generally used for the determination of all geometric parameters of a complete measurement system, i.e. the interior and exterior orientation parameters of all the system components. System calibration is relevant to digital multi-camera systems that are either mobile and can be freely configured, for example dual camera online systems, or are mounted in a fixed position, such as a 16-camera system for the inspection of pipes (see section 6.6.2.1).

For dual-camera online systems it is possible to calibrate each camera individually in advance. Alternatively, self-calibration can be applied to a set of images which have been acquired with both cameras simultaneously.

During operation, multi-camera systems on fixed mountings require particular care in monitoring and calibration. Such mechanical restrictions can cause problems, for instance, by not permitting convergent or tilted images. In general, exterior orientation parameters can be monitored on a regular basis by the use of reference points and, if necessary, can be recalculated by bundle adjustment or spatial resection. However, the interior orientation parameters can only be determined by object fields with a suitable distribution of object points. Fig. 1.42 shows an example of a multi-camera system which can be oriented and calibrated by a motor-driven rotating test field and where the orientation of the rotary table itself is simultaneously calculated.

7.3.2 Imaging configurations

The following imaging configurations are principally designed for self-calibration by bundle adjustment. The illustrated configurations of point fields and camera stations are a limited selection from many possibilities. Modifications and combinations are possible and often unavoidable. For stationary objects the camera is suitably positioned at locations around the object in order to obtain a good spatial distribution of images. For a stationary camera a mobile object is observed which is placed in different positions and angular orientations within the camera's field of view.

7.3.2.1 Calibration using a plane point field

If only a plane point field (flat test field or measurement object) is available, several convergent images are necessary. The minimum number of images depends on the availability and distribution of reference points with known coordinates. If reference points are not provided, known distances (scale bars) in object space can also be used for calibration.

a) 2 camera stations above test field with reference points

b) 8 camera stations above test field with no reference points

● reference point   ○ target point
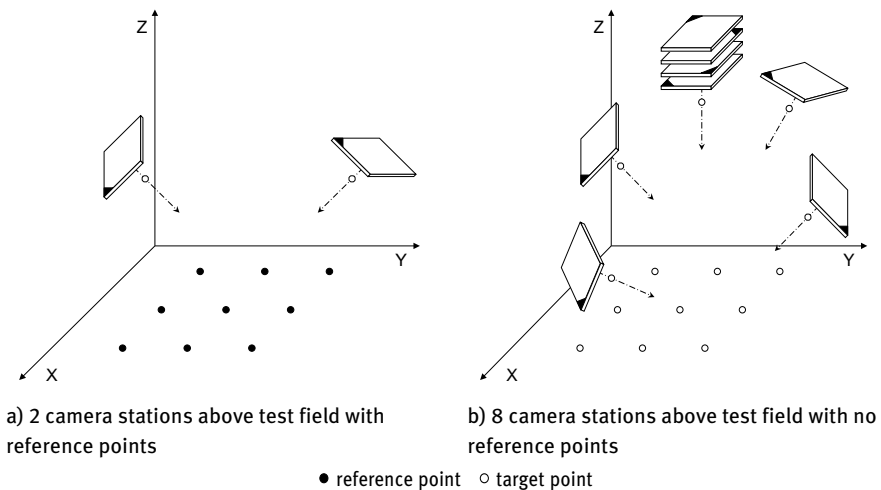
Fig. 7.32: Calibration configurations for plane test fields (after Wester-Ebbinghaus 1983, 1985).

Fig. 7.32a illustrates a minimal image configuration for a plane test field with known reference points and an invariant camera interior orientation. The points are obliquely imaged with convergent camera axes and different roll angles (rotations about the camera axis). For a test field without control points, Fig. 7.32b shows a configuration of 8 images which generates better ray intersections, higher redundancy and improved use of the image format (more reliable determination of distortion).

7.3.2.2 Calibration using a spatial point field

Self-calibration is more reliable if object points are spatially distributed in three dimensions. Spatial point fields area preferable if the measuring task permits.

a) 1 image above spatial test field with reference points

b) 2 different cameras C_1, C_2 above test field with 5 scales and no reference points

c) 4 images with arbitrary roll angles above spatial test field with no reference points

d) 8 images above spatial test field with 3 scales and no reference points

● reference point   ○ target point   ○━━━━○ known distance
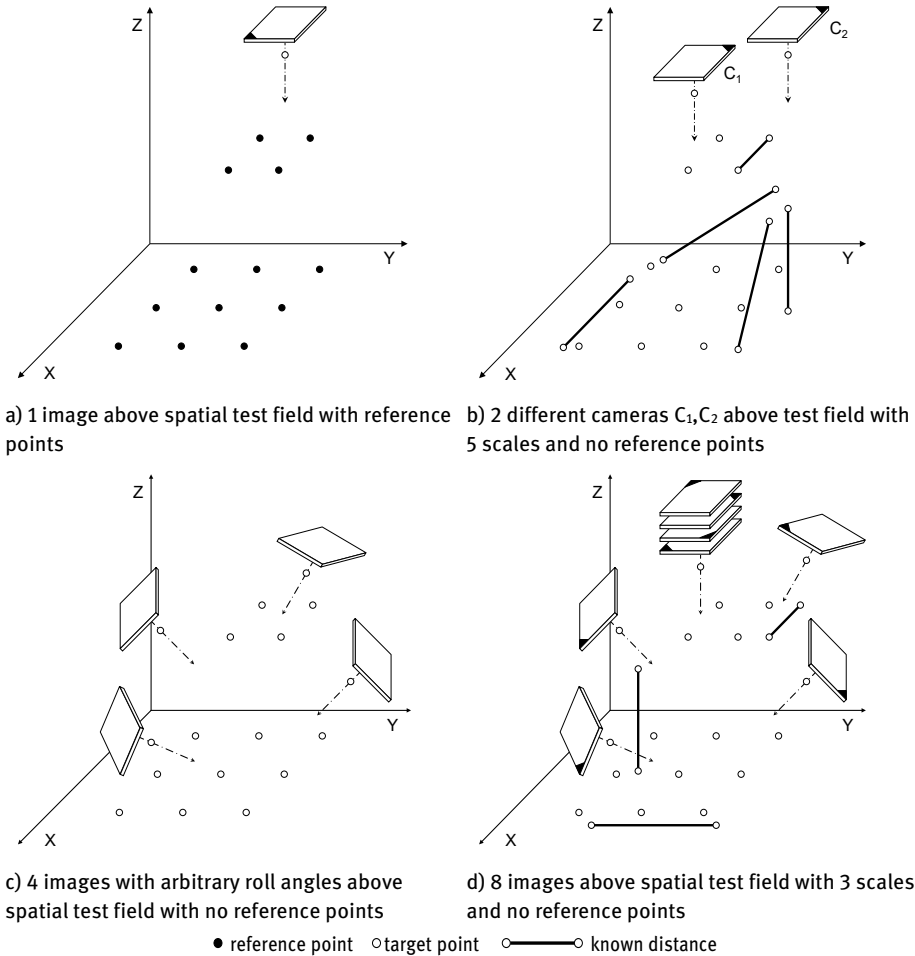
Fig. 7.33: Calibration configurations for spatial test fields (Wester-Ebbinghaus 1983, 1985).

Fig. 7.33a shows an example of single image calibration using a known 3D test field (see section 4.2.3.2). The camera can be calibrated by means of an extended space resection or linear projective methods (see section 4.2.4). Fig. 7.33b illustrates the minimal configuration for a system composed of two different cameras (image-variant interior orientation), for example an online dual camera system. Explicitly rolled images are not necessary if at least four convergent images of a spatial point field are available (Fig. 7.33c). Finally, Fig. 7.33d displays the most demanding but

also most reliable imaging configuration comprising eight images tilted with respect to each other and arranged above a spatial point field containing known distances. Similar configurations with around 16–25 images are recommended, although the total number is, in principle, unlimited.

In principle, all the imaging configurations shown above can be created by placing a portable test field in multiple positions with respect to the camera to be calibrated, such that the same perspective conditions between test field and camera are generated.

7.3.2.3 Calibration with moving scale bar

The principle of on-the-job calibration described in section 7.3.1.3 assumes that one or more cameras can be moved around the object in order to generate a sufficient number of convergent ray bundles. Alternatively, where cameras are in fixed locations, a test field can be moved into various positions in the object space. In this case, exterior orientations are initially based on the individual test field locations and must subsequently be transformed into a common coordinate system.

An alternative technique for calibrating and orienting sensors in fixed locations uses observations of a scale bar which is moved to multiple positions in the object space and is therefore measured from different directions (Fig. 7.34). This creates a set of unknown object points and at every scale bar position there is a length observation between a pair of points. Datum definition is achieved by a free net adjustment (six degrees of freedom) and scale information is derived from the length observations. In addition to the simple implementation, the method has the further advantage that the object space is completely defined by known lengths.
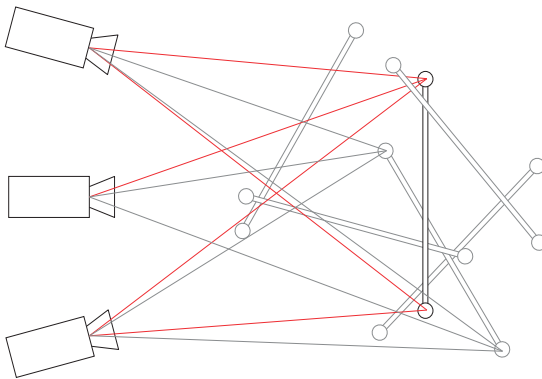
Fig. 7.34: Calibration of an imaging system with a moving scale bar.

The method is particularly applicable for multi-camera systems which record image sequences and can automatically track and measure the targets on the scale bar. This type of system can be found in medical navigation (Fig. 8.85) or applications of motion capture (Fig. 6.98).

7.3.3 Calibration of multi-camera and non-standard imaging systems

7.3.3.1 Calibration of stereo and multi-camera systems

Imaging systems with two or more cameras mounted in a fixed relative position are normally calibrated using the multi-image schemes described above. In addition to the interior orientation of each camera, their relative orientation, or exterior orientation in a machine coordinate system, must also be determined. The following calibration strategies can be identified:

– Separate camera calibration and subsequent relative or exterior orientation:
   Prior to mounting on a system frame, the interior orientation of each camera is calibrated independently of the others by a test-field calibration designed to minimize correlations with exterior parameters. After assembling into a mechanical device, the exterior orientations are separately determined using fixed interior orientation parameters, e.g. by bundle adjustment without self-calibration.

– Simultaneous camera calibration with orientation (full system calibration):
   The fully assembled measuring system is calibrated as a complete unit by test field calibration, i.e. all parameters of interior and relative or external orientation are calculated simultaneously. Since it is assumed that the relative camera positions remain constant in subsequent applications, constraints on the relative orientation of pairs of cameras can be introduced into the bundle adjustment as outlined in section 4.4.2.3. This approach leads to a more accurate system calibration.

7.3.3.2 Calibration of fisheye cameras

Camera-lens combinations with very large field angles (>110°) usually have large distortions which do not correspond to the conventional central-projective imaging model (see section 3.3.7). The following approaches are suitable for modelling their imaging properties:

– Spherical imaging model:
   A strictly spherical imaging model can be implemented in a bundle adjustment, e.g. according to eqn. (3.82) to (3.85), and optionally extended with conventional functions for lens distortion.

– Prior correction with a lens-map function and subsequent conventional modelling:
  If the lens manufacturer provides a lens-map function, e.g. as a table of distortion values along the image diagonal, the original image can either be rectified (see below) or the measured image coordinates can be corrected a priori by the given distortion values. Then the remaining distortions can be modelled by conventional distortion methods.
– Correction grid:
  In principle, every imaging device can be modelled with a finite-element correction grid as outlined in section 3.3.4.5. However, the calculation of a correction grid assumes a sufficient number of measured image points distributed over all regions of the sensor in order to avoid negative extrapolation effects.

Due to the wide field of view, it is often difficult to provide a test field sufficiently large to cover a significant part of the imaging sensor and which also ensures that the highly distorted targets can still be measured. Fig. 7.35a shows the original image of a fisheye camera taken from above a plane object with circular targets. The targets are increasingly distorted towards the image edges, to the extent that they then cannot be measured with normal processing software. In contrast, if the image has been rectified according to the manufacturer's lens-map function, the imaged targets are close to ellipses (Fig. 7.35b) and these can be conventionally processed. However, the reduced quality at the image borders cannot be corrected.
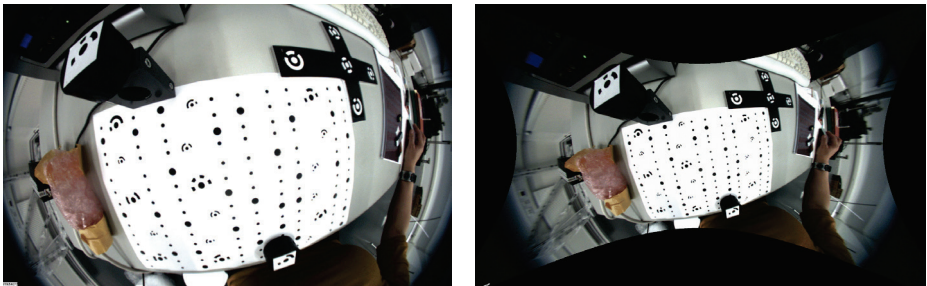
a) Original image with highly distorted targets b) Rectified image with measurable targets

Fig. 7.35: Image taken by a fisheye camera above a plane object.

Lower distortions occur if the fisheye camera images a test field which surrounds its field of view. Fig. 7.36a shows, for example, a test field defined by the inner surface of a hemisphere. The corresponding original image (Fig. 7.36b) shows circular targets in focus across the whole image format. It should be noted that a spherical surface for fisheye calibration has a similar geometric effect to a plane surface used for

calibrating a normal central-projective camera. In these cases, the result is higher uncertainty in determining the principal distance and increased correlations between interior and exterior parameters.
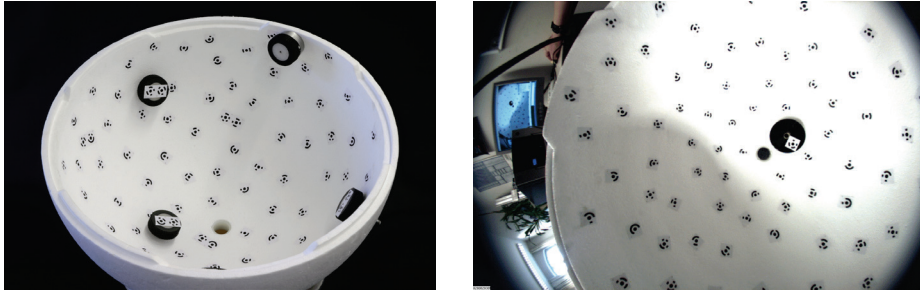
a) Hemispherical test field

b) Original image with non-distorted targets

Fig. 7.36: Hemispherical test field.

7.3.3.3 Calibration of underwater cameras

In principle, cameras designed for use under water, or in media other than air, should be calibrated using the imaging models of multi-media photogrammetry (section 4.6). To a first approximation, cameras with a hemispherical lens housing (Fig. 4.97) can be calibrated using the central-projective model and its associated distortion functions. Similar approaches can be applied if the optical axis of the camera passes perpendicularly through a plane cover glass since then the image rays refracted at the plane interfaces can be modelled by radial distortion functions. In both cases, the calibration should be conducted in situ, i.e. within the media which defines the actual measurement space since its temperature and chemical properties, such as salt concentration, will influence the refraction effects.

For applications which do not permit in-situ calibration, the camera can be calibrated in the lab under similar measurement conditions. However, there may be a loss of accuracy due to potential mechanical instabilities of the camera and optical conditions in the lab which differ from the actual measurement conditions. If need be, interior orientation parameters can be updated in situ using a calibrating space resection (section 4.2.3.2).

7.3.4 Quality criteria for camera calibration

Assessing the quality of camera calibration is complex. In simple terms, camera calibration is good enough if subsequent 3D measurements achieve the required accuracy. Despite the detailed procedures for bundle adjustment (section 4.4) and calibration (section 3.3) it is still often difficult to identify the reasons for a weak 3D

measurement accuracy. Since the camera is generally the most critical component in photogrammetry, objective criteria for the quality of camera calibration should be defined and adopted if possible.

Internal quality measures for precision in camera calibration provide a first hint as to the reliability of the camera model and the strength of the imaging configuration. Usually, the a posteriori standard deviations of principal distance and principal point should be of the same order as the accuracy of image measurement (see also section 4.4.5.5). For the adjusted distortion parameters, the standard deviation should be one order of magnitude smaller than the parameter itself.

More complex quality statements can be derived from the minimum description length *MDL* which describes how efficiently the model reflects the input data:

$$MDL = \frac{k}{2} \cdot \log_2 n + \frac{\mathbf{v}^T \mathbf{P} \mathbf{v}}{2 \ln(2) \cdot \hat{s}_0^2} \; \hat{s}_0^2 \tag{7.23}$$

where

n: number of observations

k: number of parameters modelling the interior orientation

$\hat{s}_0^2$:   standard deviation of unit weight (image measuring precision)

The *MDL* value increases with k (complexity of the model) and the squared sum of corrections $\mathbf{v}^T \mathbf{P} \mathbf{v}$ (noise, precision of the model), and is reciprocally proportional to the image measuring precision $\hat{s}_0^2$. In theory, a better image measuring accuracy would allow for a more complex imaging model.

The quality of self-calibration can also be assessed by the quality of the adjusted 3D coordinates of the object points. Here it should be checked if the standard deviations of object coordinates (see section 4.4.5.4) are of the same magnitude as the specified accuracy. It should be noted that network strains introduced by reference points can affect the calculated camera parameters without this being apparent in the bundle adjustment results. Camera calibration is therefore typically carried out using a free-net adjustment.

The most thorough method of quality assessment is, however, an independent comparison of measured values with reference values of superior accuracy. The length measuring error described in section 7.2.3 is the preferred parameter since it is obtained under real and reproducible measurement conditions.

7.3.5 Problems with self-calibration

Practical problems with camera calibration typically arise in the following cases:

– Inadequate modelling of the interior orientation:
Limitations in the mathematical modelling of the interior orientation usually lead to residual errors in the self-calibration. This can result in a lower precision

(sigma 0), in systematic residual patterns within the image measurements or in systematic errors in object space. Fig. 7.37a shows a systematic distribution of residuals after bundle adjustment, which indicates an incomplete model (here: aspherical lens, see section 3.1.4). In contrast, Fig. 7.37b shows a distribution without recognisably systematic errors, so indicating a well-developed model (see also section 4.4.5.3).

a) Systematic pattern   b) Random pattern
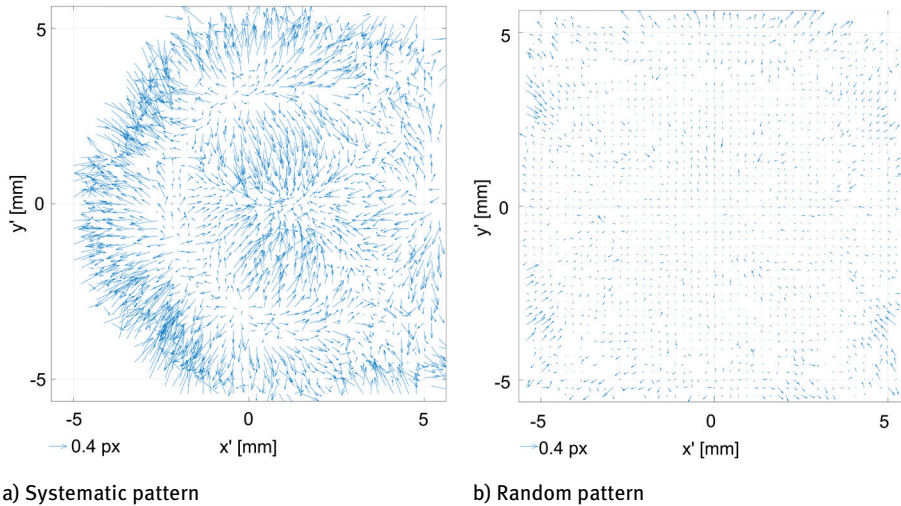
Fig. 7.37: Image measurement residuals after simultaneous self-calibration.

– Correlations between parameters:
  Using a bundle adjustment with self-calibration for the estimation of interior orientation parameters usually results in correlations between adjusted parameters. The presence of any significant correlations can be ascertained from analysis of the covariance matrix. High correlation values indicate linear dependencies between single parameters and should be avoided. Correlations often arise between the following parameters:
  – principal distance, principal point and exterior orientation;
  – A_1, A_2 and A_3 will always be correlated to some extent as they are sequential terms in the radial lens polynomial model;
  – principal point x'_0 and affine parameter C_1 or alternatively y'_0 and C_2.

Fig. 7.38 illustrates the correlation matrix between the parameters of interior orientation. The corresponding measurement network consists of 140 images of the 3D test object shown in Fig. 7.13a. Larger correlation coefficients appear, as expected, between the A parameters, as well as between the principal point coordinates and the B parameters of the tangential distortion.

```
C     1.000
X0    0.017   1.000
Y0   -0.071   0.011   1.000
A1   -0.122  -0.001   0.010   1.000
A2    0.037  -0.002  -0.010  -0.966   1.000
A3   -0.006   0.005   0.010   0.905  -0.981   1.000
B1    0.018   0.930   0.010  -0.003  -0.005   0.006   1.000
B2   -0.034   0.010   0.905   0.004   0.002  -0.002   0.006   1.000
C1    0.147  -0.010  -0.002   0.077  -0.106   0.108  -0.019  -0.006   1.000
C2    0.002  -0.028  -0.028   0.007  -0.009   0.010  -0.014  -0.036   0.000   1.000
        C       X0      Y0      A1      A2      A3      B1      B2      C1      C2
```

Fig. 7.38: Example of correlation between calibrated camera parameters.

In addition, principal distance and principal point are correlated with the parameters of exterior orientation of individual images. All-around configurations of 3D objects help to compensate for these correlations, i.e. they generate no explicit relationship between the exterior orientation of an image with the corresponding position of the perspective centre of the camera. For a set of images taken from one object side only, for example a hemispherical distribution over a plane test object, the correlations between interior and exterior orientations are higher and they appear between explicit parameters, for example c is correlated with Z_0 (Z=viewing direction). In general, a camera calibration which makes use of a spatially distributed point field results in smaller correlations and more reliable camera parameters.

Correlations between parameters can largely be neglected if object reconstruction and camera calibration are calculated in one simultaneous computation, as is the case for bundle adjustment with self-calibration. Nevertheless, parameters with no statistical significance can be detected by suitable test procedures and eliminated from the functional model. If individual interior orientation parameters are correlated and then used in subsequent, separate calculations, they no longer completely represent the chosen mathematical camera model. For example, if cameras in a fixed online measuring system are pre-calibrated using a different imaging configuration, then the subsequent online use of the resulting parameters can lead to errors in the computation of 3D coordinates. As already indicated, calibration against a spatial test field results in lower correlations and more reliable camera parameters. However, calibration as part of a space resection (section 4.2.3.2) or DLT (section 4.2.4.1) can be problematic since interior and exterior orientation are highly correlated, especially for object fields without significant depth.

– Over- and under-parametrization:
Under-parametrization means that the selected mathematical model does not sufficiently represent the real imaging characteristics of the camera. In this case, the resulting errors in camera calibration have a significant impact on subsequent measurements, depending on the specified accuracy level. An example of under-parametrization is the absence of tangential distortion in some calibration approaches used in computer vision.

      Over-parametrization occurs when correlation between selected parameters is close to 1, i.e. in principle the one parameter is completely modelled by the correlated parameter. In the example of Fig. 7.38, the correlation between parameters A_2 and A_3 is -0.981. Here it could be argued that A_3 is not required for this specific camera. However, empirical experience has shown that this kind of over-parametrization does not damage the overall result and has no significantly negative effect on the final measurement result.

– Images without relative roll angles:

Images with relative roll angles (rotations about the optical axis) are necessary for the determination of principal point coordinates, and possible affine transformation parameters, if the test field does not provide a suitable distribution of reference points, or if sufficient convergent images cannot be taken. The coordinates of the principal point are highly correlated with the parameters of exterior orientation if rolled images are not available.

– Missing scale information in the viewing direction:

If scale information in the viewing direction is missing, as in the case of orthogonal images of plane test fields, principal distance and object distance cannot be uniquely determined. Just one known distance in the viewing direction, e.g. through the use of known reference points, or one known coordinate component, e.g. for convergent imagery of a plane test field, is sufficient for the calculation of principal distance. As a simple alternative, a network of convergent images of a planar target field will allow recovery of principal distance, but such a solution requires careful assessment. Another alternative is to set the principal distance to a fixed value, and not determine it in the bundle adjustment. Compensation for any potential scale error arising from this procedure in subsequent measurements can only be made using reference lengths in object space.

– Incomplete use of the image format:

The imaging sequence for camera calibration should be arranged in such a way that, within the full set of images, use of the complete image format is achieved. Only then is it possible to determine distortion parameters which are valid across this whole format (Fig. 7.39). It should also be noted that the optical axis should not, in all images, be directed at the centre of the object. The camera should be pointed in different directions, including if necessary, at only part of the object.
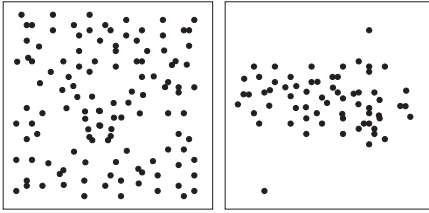
Fig. 7.39: Optimal (left) and poor (right) distribution of image points.
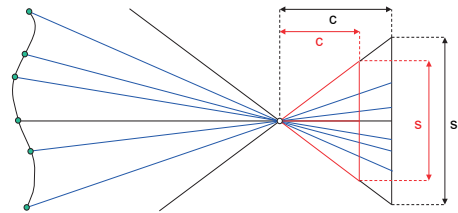
Fig. 7.40: Sensor format and principal distance.

– Use of high distortion lenses:
  Many wide-angle and super wide-angle lenses cause large distortions in the image corners. For these lenses standard distortion models are often insufficient and result in lower accuracies of points imaged in the corners. The problem becomes even more critical if the image format is incompletely used at the calibration stage. In cases where there is any doubt, it is good practice to ignore any image measurements made in the outer 10 % of the image format.
– Lack of camera stability:
  Determination of interior orientation becomes more uncertain if the camera geometry changes from image to image within a sequence, for example due to thermal effects, loose attachment of the lens or unstable mounting of the image sensor in the camera housing. Each variation in this case must be handled by defining a separate camera in the adjustment model. However, many imaging configurations only permit the simultaneously calibration of a small number of cameras. In section 4.4.2.4 an approach to the calibration of image-variant parameters is discussed.
– Unknown pixel size in the image sensor:
  For many simple digital cameras (consumer cameras, mobile phone cameras) there is often insufficient technical data available. If there is missing information to determine the pixel size of the image sensor or the physical sensor format, then an arbitrary pixel size can be set. Although the image coordinate system is also arbitrarily scaled in this case, the calibrated parameters (in particular the principal distance) are determined with respect to the selected sensor scale. The form of the ray bundle defined by the interior orientation will, in fact, remain the same (Fig. 7.40).

8 Example applications

The techniques of close-range photogrammetry provide universal methods and approaches to the geometric measurement of almost any kind of object. As a result, there are a wide range of potential application areas. The following example applications represent only a small selection from the entire spectrum of possibilities. They are restricted to sample images, results and key technical specifications. Examples processed using film cameras can, conveniently and without restriction, be implemented using current digital cameras. References to examples can be found in section 9.8.

8.1 Architecture, archaeology and cultural heritage

8.1.1 Photogrammetric building records

Photogrammetric building records mostly aim to generate 2D and 3D data of buildings and monuments for the following applications:
– preservation and restoration of the building;
– planning conversions, renovations and extensions;
– art historical analysis;
– digital twins;
– geometric documentation.

The essential technical requirements for this field of architectural photogrammetry were already developed in the 19th century (see section 1.4). Today's optical technologies of 3D measurement techniques (photogrammetry, laser scanning) offer a number of advantages for building surveys:
– non-contact measurement avoiding scaffolding on the façade;
– fast on-site image acquisition;
– high accuracy;
– high resolution;
– three-dimensional coordinate measurement;
– measurement of free-form contours, surfaces (ornamental details) and complete 3D object models;
– recording of colours and textures in different spectral bands
– image rectifications (orthophotos, photo maps);
– subsequent object measurement from archived images or historical photos.

It is emphasized that cooperation between photogrammetrists, architects and heritage experts is of major importance. This ensures that imagery and point clouds are also correctly interpreted with regard to architectural and artistic features,

cultural heritage issues and relevant construction methods. In this context, historical, ethical and technical requirements must be brought into harmony, as specified, for example, by national regulations on heritage protection or by internationally recognised recommendations (e.g. London Charter) or organisations (e.g. ICOMOS, CIPA). This also applies to the recording of smaller cultural heritage objects, which is dealt with in section 8.2.

Photogrammetric data acquisition is carried out with suitable digital cameras, as well as with terrestrial laser scanning (TLS) and mobile scanning systems, which usually also employ cameras. Signalized points (targets) are mostly used for control point marking (georeferencing) and image orientation (registration). The actual 3D reconstruction is done using imaged natural features of the object.

Depending on environmental conditions and object shape, photogrammetric image recording can be carried out from locations on the ground, elevated locations (ladder, crane) and drones. As usual, the imaging configuration must be selected such that every object point of interest is recorded in at least two images with a good intersection angle, with recording in three images preferred for structure-from-motion (SfM). This requires many more images when recording interiors or courtyards than when structures are recorded from the outside. For applications using SfM, appropriate image overlaps must be selected (see section 5.5.2.2).

Processing can be achieved using SfM programs (see section 6.4.2), photogrammetric stereo workstations, interactive digital multi-image processing systems with CAD functionality (section 6.4.1) and CAD systems with functions for analysing images or point clouds. Monoplotting functions are also applicable in combination with 3D point clouds from laserscanning or photogrammetry (see section 4.2.7). Furthermore, with digital processing systems it is easy to generate rectified images e.g. for generation of façades (Fig. 8.3) or orthophotos (Fig. 8.36).

The reconstruction of complex building structures is increasingly achieved by processing 3D point clouds which are created using SfM and TLS. Newer technologies for mobile scanning have become more popular but have accuracy limitations. Their advantage is the simple object capture and highly automated registration. The disadvantage is the capture of a large amount of data relating to unstructured object points. This must then be reduced to essential geometric features, typically by manual evaluation. Point clouds can be visualised and processed with numerous software packages, e.g. CloudCompare, 3D Systems Geomagic, Autodesk ReCap or Bentley PointTools. Software packages such as Rhino, AutoCAD or Revit CAD or BIM (Building Information Modelling) applications, are also used for further processing of point clouds.

Data can be stored in standard 3D formats such as E57, OBJ, STL and FBX. Attention is increasingly focused on standardized 3D graphics languages such as X3D, 3DS or CityGML that include textures and which enable system-independent visualisations and interaction with the 3D model. Visualisation examples are shown in Fig. 1.6, Fig. 4.30 and Fig. 8.5, as well as in following sections.

8.1.1.1 Drawings

Drawings include floor plans and elevations, profile sections and deformation-related measurements (as a basis of further as-built and damage mapping) of objects that are captured in two- or three-dimensional form as vector data. They are created by interpreting image data or point clouds, i.e. during analysis the relevant features are visually extracted by the user and, if necessary, are reduced, simplified or generalized. Fig. 8.1 shows an example of a line drawing obtained by interactive stereo evaluation using an analytical plotter. Fig. 8.2 shows contour extraction based on point clouds in an orthogonal view for an HBIM model (Historical Building Information Modelling) (see also Fig. 8.8).
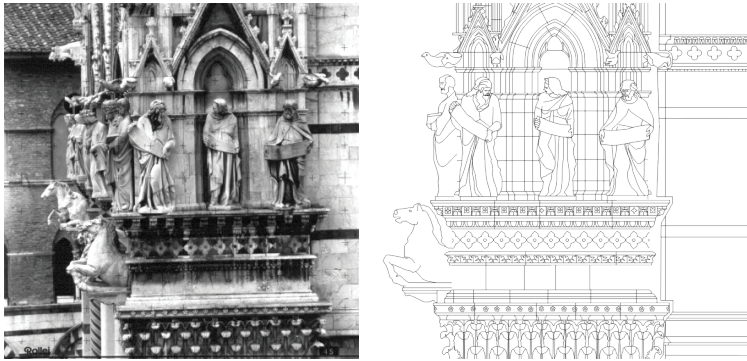
Fig. 8.1: Analogue metric image (UMK) and stereoscopic mapping of façade (Messbildstelle).
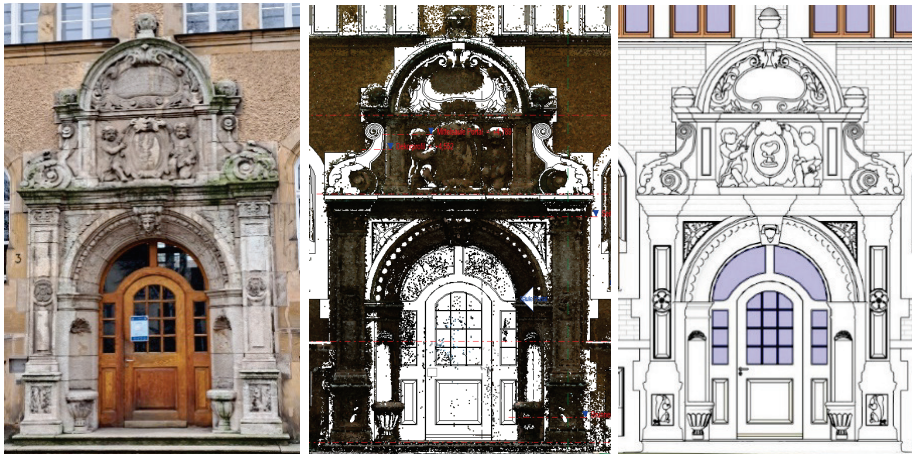
a) Original image   b) 3D point cloud   c) HBIM model

Fig. 8.2: Example of contour extraction based on point clouds (University of Bamberg).

8.1.1.2 3D building models

If a building's 3D CAD data is available in a topologically structured form, e.g. from a photogrammetric process, 3D visualisation methods (section 5.3.3: texture mapping) can generate a photo-realistic representation. In addition to the purely aesthetic effect, these models also have practical application in building planning, facility management or BIM.

3D point clouds are often themselves regarded as 3D models although they simply consist of an unstructured set of surface data. However, when a point cloud is meshed (triangular meshing) a closed surface is generated which is suitable for purposes of visualisation or 3D calculations, such as volumes and deformations. In a narrower sense, 3D models exist if single components are modelled together with sematic or topological information, comparable to a CAD or GIS model. The latter is mainly used for BIM and HBIM (see also section 8.4.1.2 for the Scan-to-BIM approach). Examples of 3D building models are shown in Fig. 8.5, Fig. 8.7 and Fig. 8.11.

8.1.1.3 Image plans and orthophotos

In the place of scaled 2D drawings, there is an increasing demand for rectified metric image plans of building facades. They have a significantly greater information content and they offer an objective representation of the object which has not been interpreted and generalized as in the case of a drawing.

Fig. 8.3 shows a rectified image of a row of buildings. Each building was separately rectified by plane projective transformation (section 4.2.8.1) with the aid of reference points measured on site. Based on the rectifications, façade plans were subsequently derived by manual drawing.

Fig. 8.3: Graphical elevation and rectified image of a building terrace (IAPG).

Fig. 8.4 is an example of a true othophoto of a façade (see also section 4.2.8.2). In this case, a triangular meshing of a laser-scanned (TLS) point cloud was available. Images

were oriented using Agisoft PhotoScan using reference points extracted from the TLS point cloud and converted to orthophotos based on the meshing. Here, there is a specific problem in that the window reveals are not modelled correctly by the meshing process and at those locations, lateral errors appear in the orthophotos. Since laser scanning and image recording were carried out from locations at street level, these errors occur mainly around windows on the upper stories. The resolution of the orthophoto is 5 mm and the absolute positioning error of image points is estimated to be of the order of 7 – 10 mm.

It must be noted that accuracy and completeness of the integrated surface model is of fundamental importance to the quality of the final orthophoto. Depending on the application, TLS point clouds can improve the quality of SfM point clouds, or vice versa. Since modern SfM software enables automatic point cloud generation, meshing and orthophoto production in one closed workflow, this if often the preferred and most cost-effective solution in practice. In general, true orthophotos are the most reliable metric representation of building façades if the required digital surface model and available images are of sufficient quality (see section 4.2.8.2).

Fig. 8.4: True orthophoto of a façade with close-up detail (SPM3D).

8.1.2 Examples of photogrammetric building reconstructions

8.1.2.1 Torhaus Seedorf

Fig. 8.5 shows an example of 3D measurement using SfM. Here the model is of the Torhaus Seedorf which is close to Bad Segeberg (Germany) and was built in 1583. A combination of terrestrial images (Nikon D7000 and Nikon D800) and UAV images (Canon IXUS 125 HS) enabled all exterior walls, and the complex roof structures, to be modelled. A total of 215 images were jointly processed using Agisoft PhotoScan and residuals at 18 reference points were reported to be around 1.2 cm. The absolute accuracy, however, is probably in the range of 2 – 3 cm.
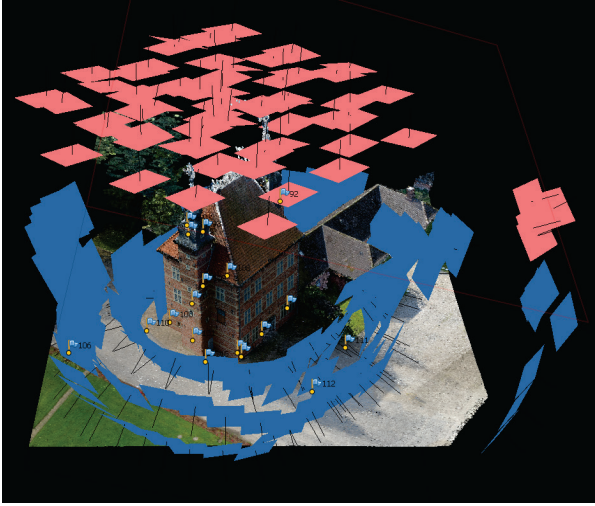
Fig. 8.5: 3D model of Torhaus Seedorf (HCU Hamburg).

8.1.2.2 The Milan Cathedral

The photogrammetric reconstruction of Milan Cathedral is a prime example of the measurement of a complex building in an environment with restricted access. The objective is to produce precise plans, detailed illustrations and 3D data for research purposes and preservation of both the external facade and the interior. Accuracy and resolution have been specified to less than 2–3 cm for map scales of 1:20 – 1:50. The structure is the fifth largest church in the world with dimensions of 158 m x 93 m x 108 m.

Fig. 8.6: Examples of Milan Cathedral's complex areas (Politecnico di Milano).

The project has been running since 2008. Various photogrammetric and geodetic methods are used. The structure of the object requires free multi-image recording with

different camera-lens combinations including fisheye systems (see Fig. 6.103b) for very narrow areas. Since the use of drones was not permitted, the upper areas had to be photographed from a crane (Fig. 8.7). A GSD of 2mm was achieved on the outer façade. In total, more than 12000 images have been taken and evaluated so far. The evaluation is mainly done by SfM, but depending on the local conditions, images also have to be oriented manually. A total of approximately four billion 3D points were captured. Georeferencing was carried out via 122 control points measured by totalstation. Manual line drawings by stereo photogrammetry, automatically generated profile sections (Fig. 8.8), orthophoto plans as well as 3D meshings and BIM models were derived from the raw data obtained.
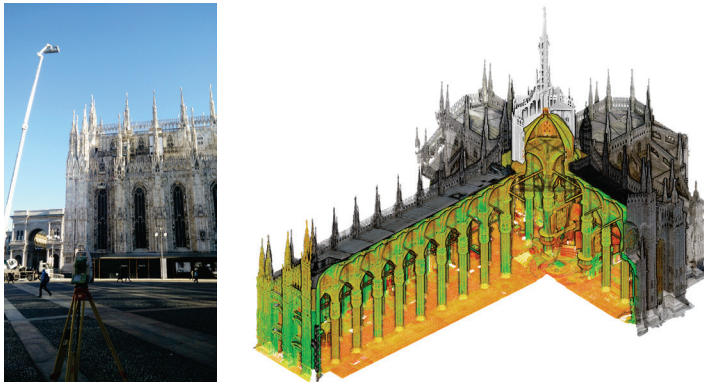
Fig. 8.7: Image recording by crane and final 3D-point cloud (Politecnico di Milano).
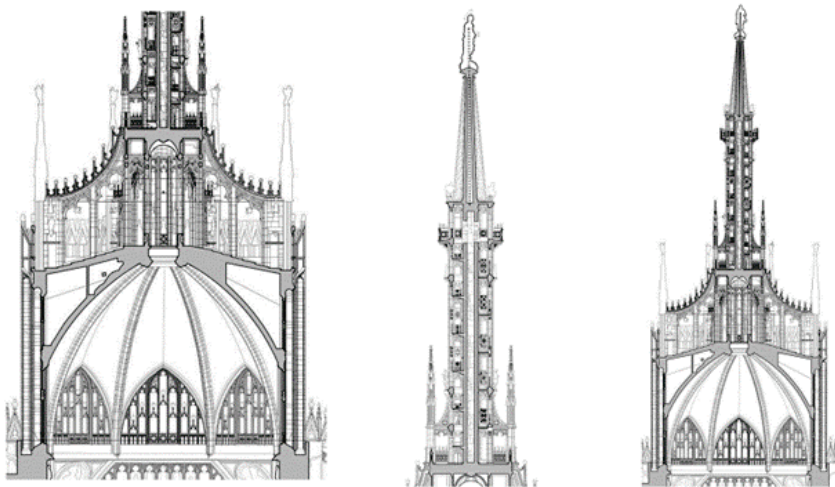
Fig. 8.8: Profile sections extracted from point clouds (Politecnico di Milano).

Fig. 8.9 shows the image recording paths around the cathedral and an example point cloud of a very confined space, acquired with the fisheye camera system shown in section 6.11.1.2. Since the unfavourable image configuration without loop closure entails poor error propagation, the system calibration (interior and relative orientation of the cameras) must be highly accurate so that the subsequent orientation by Visual SLAM (section 5.5.7.6) and bundle adjustment generate the required accuracy. Usually, control points cannot be provided in such environments.
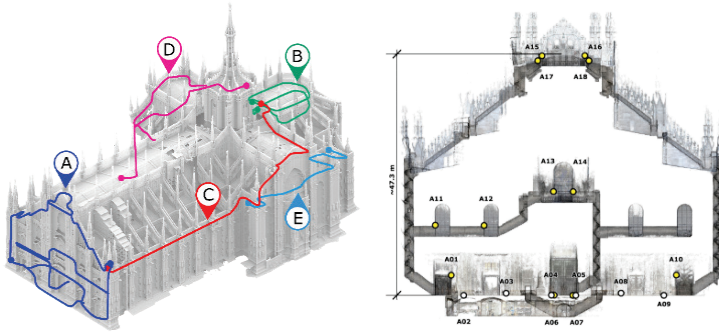
Fig. 8.9: Imaging paths and 3D model of narrow spaces (Politecnico di Milano).

8.1.2.3 Lurji monastery

A combination of terrestrial laser scanning, UAV and terrestrial photogrammetry was used for the 3D recording of historic church buildings in Tbilisi, Georgia, including the Lurji Monastery. The measurements were carried out under conditions made difficult by restricted recording locations and visual occlusions.

Leica BLK360 and Faro Focus 3D X330 scanners were used for terrestrial laser scanning. Registration was carried out using checkerboard markers at signalized points. These were then further used as control points for the photogrammetric evaluation. Resolution and registration accuracy were about 5mm.

A Canon EOS 6D Mark II full-frame SLR camera with 24-50mm zoom lens (26 Mpixel) was used for terrestrial image acquisition. The image overlap was designed to be 80-90% with a GSD of 3-5mm. A total of about 1000 images were taken of the monastery (example photos in Fig. 8.10).

The drone images were taken with a manually controlled DJI Mavic Pro with FC 200 camera and 4.7mm lens (12 Mpixel). The flight path was a meandering stripe pattern with two circular orbits around the church tower. For the drone imagery, a total of 368 photos were taken with GSDs of 5-11mm (examples in Fig. 8.10).

Data processing was done with the RealityCapture software package. This offers a particular advantage in that laser scans and photogrammetric image data can be simultaneously analysed. The registered laser scans are used as coordinate reference.

In RealityCapture, synthetic image cubes are calculated for each laser position (see section 5.5.2.2), and these are linked with the photogrammetric images via feature matching. The result is a high-resolution 3D model with textures that correspond to the quality of the photographic images (Fig. 8.11). The overall accuracy achieved is about 5 mm in areas close to the ground and about 10 mm in areas of the tower that are difficult to access. Detailed views of the 3D model show the high quality of the reconstruction which reproduces some of the smallest object details.

Fig. 8.10: Sample measurement images: top: UAV, bottom: terrestrial (IAPG, Uni Bamberg).

a) 3D model

b) Detailed views of the 3D model

Fig. 8.11: 3D model of the Lurji Monastery (IAPG, Uni Bamberg).

8.2 Cultural heritage artefacts

8.2.1 Art and museum pieces

For art historical tasks, restoration work, provenance research, virtual museums or documentation of objects collections, e.g. sculptures or decorations, digital copies (digital twins) with exact representation of geometry and surface patterns (colour, texture) must be generated. Objects of this kind are often unique, valuable, sensitive and not easy to transport, so must be measured using mobile non-contact systems.

Conventional stereo measurement is suitable for line extraction if the object surface has distinct contours (example in Fig. 8.27). Objects without natural surface texture are usually measured by fringe or pattern projection systems using both high-accuracy industrial systems (section 6.7.3) and low-cost and hand-held systems (section 6.9.6). Smooth and polished surfaces are often sprayed with white powder in preparation for measurement, if permitted. Most projection-based systems have a limited measurement volume, so that larger and more complex objects are recorded from multiple measuring locations. Individual point clouds are registered by means of reference points or other tools such as rotary tables or measuring arms (see section 6.9).

3D modelling with correct (calibrated) colour representation are of increasing importance. Here the measuring process typically incorporates RGB cameras, calibrated light sources, white balance and colour calibration charts in order to generate true colour representation of the measured object.

The raw measured 3D data are passed through the processing steps of triangle meshing, registration, smoothing and thinning after which Bézier or NURBS surfaces can be generated (see section 2.3.3). The derived 3D CAD data can be used, for example, to control 3D printers, NC milling machines or stereo lithography processing in order to generate facsimiles of the objects. They are also applicable in diverse aspects of 3D visualisation, e.g. in virtual museums or animations.

8.2.1.1 Statues and sculptures

Statues and sculptures mostly have complex surface shapes which cannot be interpreted as a 2½D surface and measured from only one side. If the objects of interest can be measured under controlled lighting conditions, e.g. in a laboratory or room with no external light, and they have reasonably bright, diffusely reflecting surfaces, then a fringe projection system is suitable for 3D recording.

For cultural heritage recording, the registration of individual 3D point clouds is generally achieved using the object's own geometry (natural points, distinctive surface features). As an alternative, or in combination, additional photogrammetrically measured control points can be used, either to bridge areas lacking in natural detail, or to enhance or control the overall accuracy. Smaller

objects can make use of a servo positioning, e.g. rotary tables, to generate part scans in a common coordinate system.

a) Terracotta warrior (Xi'an, China) b) Gargoyle (Freiburg Minster) c) Statue of Heracles (Antalya Museum)
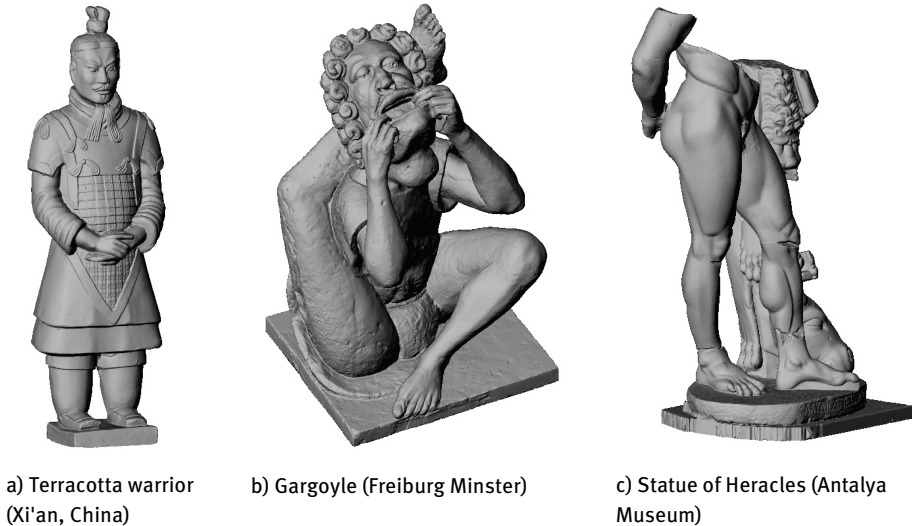
Fig. 8.12: 3D digitizing of statues and sculptures (AICON/Breuckmann).

Fig. 8.12 shows examples of digital, all-round surface measurement of diverse statues. Fig. 8.13 shows the measurement sequence in a fringe projection system.

Fig. 8.13: Projected pattern and image sequence in a fringe projection system (AICON/Breuckmann).

8.2.1.2 Death mask of pharaoh Tutankhamun

Fig. 8.14 shows how the gold death mask of the Egyptian pharaoh Tutankhamun was measured with a GOM ATOS III fringe projection system (resolution 0.25 mm) and photogrammetric reference points (Nikon D800, f = 35 mm). The 3D data will be used for purposes of evaluation and restoration, as well as to manufacture fitted transportation cases for the mask. It was not permitted to prepare the surface with powder, and as a result some gaps and artefacts remained in the 3D model around the

face. By selecting suitable scan parameters and multiple scan stations, the 3D modelling quality was sufficient for CNC milling of the transportation cases and for documenting the restoration work.
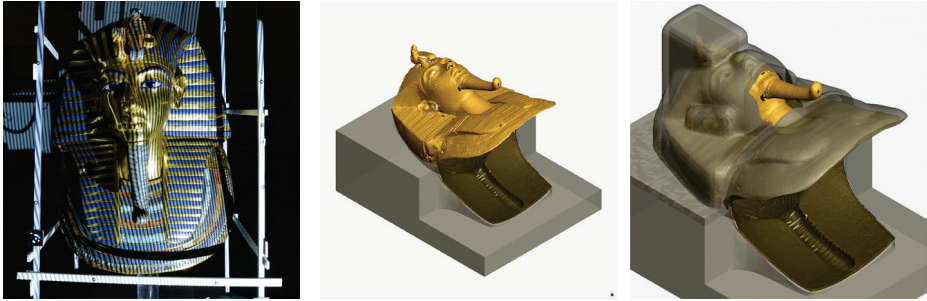
a) 3D scan with projected fringes b) 3D model   c) Transport mask

Fig. 8.14: Reconstruction of Tutankhamun's death mask (i3mainz, RGZM).

The project was a cooperation between the Mainz University of Applied Sciences, the Romano-Germanic Central Museum, the German Archaeological Institute and the Egyptian Museum in Cairo. It was funded by the German Ministry of Foreign Affairs and the Gerda Henkel foundation.

8.2.1.3 Sarcophagus of the Spouses

The Sarcophagus of the Spouses (Fig. 8.15) is a unique heritage masterpiece, an Etruscan anthropoid Sarcophagus dating back to the late 6th century BC. It is made of terracotta and was once brightly painted; it measures about 2m wide, 1.5m high, 1m deep. The sarcophagus was digitized using close range photogrammetry in 2013.

The planning of the photogrammetric image acquisition entailed several factors and constraints such as location, environmental illumination (mixed lighting), limited space around the asset, and the required resolution for the photographic texture of the mesh of 0.3mm. These parameters guided the choice of the photographic equipment, consisting of a Nikon D3X 24-Mpixel camera with a Nikkor 50mm f/1.8 D lens (GSD about 0.12mm). Two photographic lamps placed symmetrically about the camera's optical axis were used to mitigate the effect of shadows. Cross polarization (section 3.1.1.5) of the light was used to give images virtually free from specular reflection components. Colours appear more saturated (pure) and lighting more homogeneous all over the photographic subject (Fig. 8.15 right).

Fig. 8.15: The Sarcophagus of the Spouses; top right: image recorded under regular lighting; bottom right: image recorded with cross polarization (FBK Trento).

Fig. 8.16a shows the imaging configuration with about 520 images consisting of orthogonal, rolled and convergent images. The photogrammetric network was scaled using a 300 mm scale bar placed at 10 different positions around the sarcophagus and verified to have an RMS value of about 80 μm after SfM evaluation. The accuracy of the 3D coordinates of the tie points was about 60 μm (Fig. 8.16b). A high-resolution mesh was created with a spatial resolution of 0.5 mm and a texture of 0.3 mm, resulting in about 50 million triangles (Fig. 8.17).
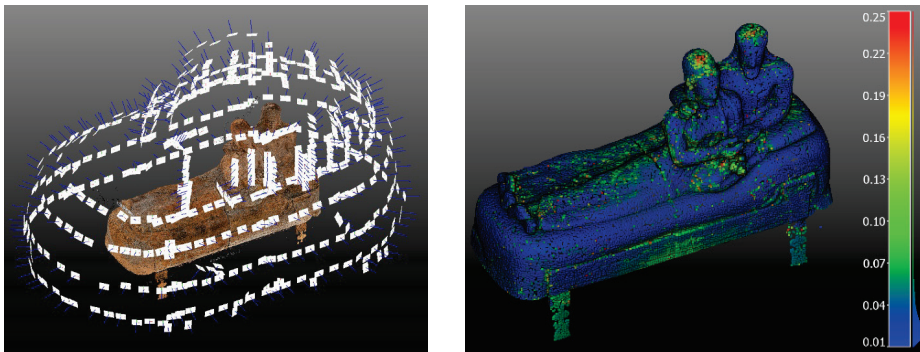
a) Camera network

b) Tie points coloured according to precision

Fig. 8.16: Results of photogrammetric reconstruction (FBK Trento).

Fig. 8.17: Texturized 3D mesh model of the Sarcophagus of the Spouses (FBK Trento).

8.2.1.4 Goethe elephant skull

The historically significant Goethe elephant skull in the Kassel Museum for Natural History was digitised using SfM and hand-held scanners (Fig. 8.18). The skull was examined around 1784 by Johann Wolfgang von Goethe for anatomical studies and consists of three parts (skullcap, main skull and lower jaw). The overall dimensions are approximately 80 cm x 60 cm x 80 cm. The images captured with a DSLR were colour calibrated and then processed in Agisoft Metashape. A Creaform Go!Scan hand-held scanner was used which achieves a measuring accuracy of approx. 0.05 mm at a maximum resolution of 0.2 mm.

Several precision scale bars were introduced to scale the photogrammetric model. Each object part was measured separately and subsequently fused into an overall model using ICP procedures (Fig. 8.19).

Fig. 8.18: Acquisition of the Goethe elephant skull with SfM (left) and hand scanner (right) (IAPG).
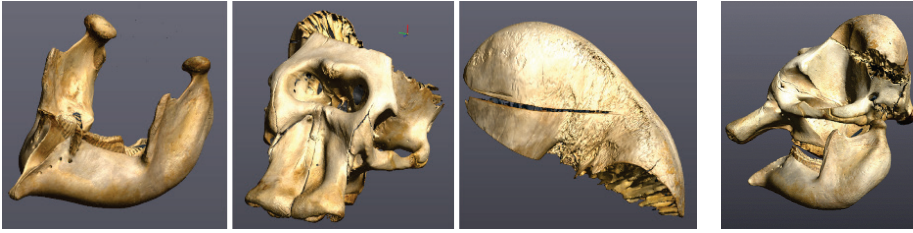
Fig. 8.19: 3D models of the individual object parts and overall model (IAPG).

8.2.2 Archaeological recordings

8.2.2.1 3D record of Pompeii

The complete recording of the excavations at Pompeii illustrates well the use of diverse data sources to create a full 3D documentation of a large, extended object. The approximately 350 finds distributed around the site, as well as the standing walls and their details, must be recorded in a single 3D model. This is required in order to prepare the data at different levels of resolution for purposes of digital conservation, animation (*virtual reality*) and connection to archaeological databases.
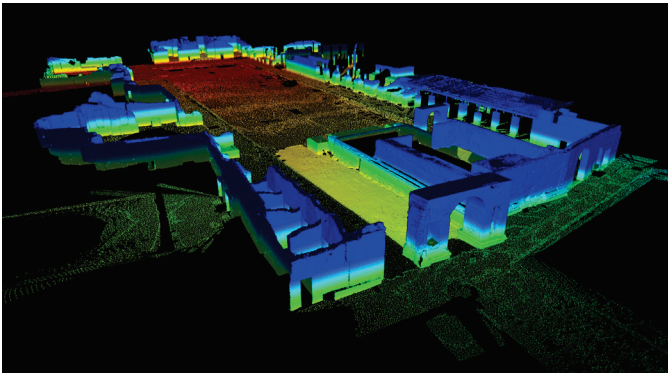
Fig. 8.20: Point cloud derived from terrestrial laser scanning (FBK Trento, Politecnico di Milano).

Aerial and terrestrial imagery, as well as terrestrial laser scanning, are used for data acquisition and, by means of GNSS measurements, transformed into a global coordinate system. Accuracy requirements are in the cm range for aerial images and the mm range for terrestrial measurements.

Fig. 8.20 shows a representative 3D point cloud created by merging multiple individual laser scans. Fig. 8.21 shows the complete model of the Forum in Pompeii and Fig. 8.22 shows detailed sections of the model with texture overlays.
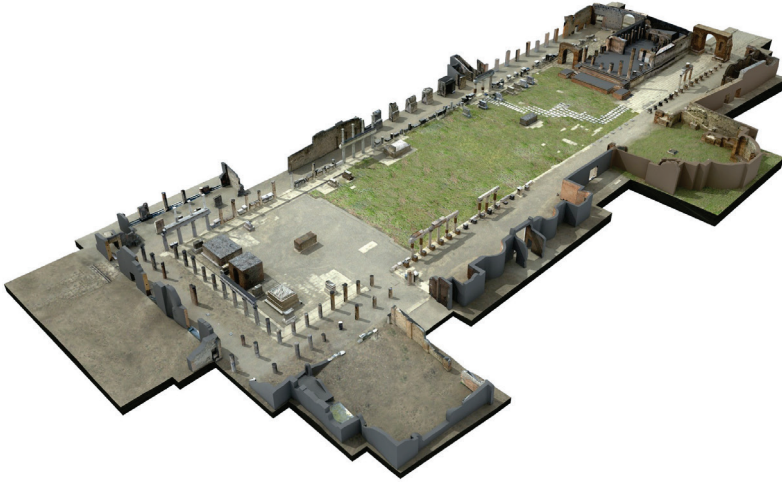
Fig. 8.21: 3D model of the Forum in Pompeii (FBK Trento, Politecnico di Milano).

Fig. 8.22: Detail sections of the 3D model of Pompeii (FBK Trento, Politecnico di Milano).

8.2.2.2 Temple of Hadrian

The 3D reconstruction of the Temple of Hadrian in Ephesus (Turkey) illustrates the measurement of a large, complex object with the aid of fringe projection and photogrammetric orientation of the scanning stations. Some 75 coded targets were attached to the object (Fig. 8.23a) and recorded in a multi-image network using a Nikon D3 digital camera. They were located in 3D using the AICON DPA Pro system. In addition, 35 of the points were intersected by theodolite which was used to reference them to the local geodetic coordinate system. The accuracy of the control points was around 1–2 mm.

a) Temple with photogrammetric
targets
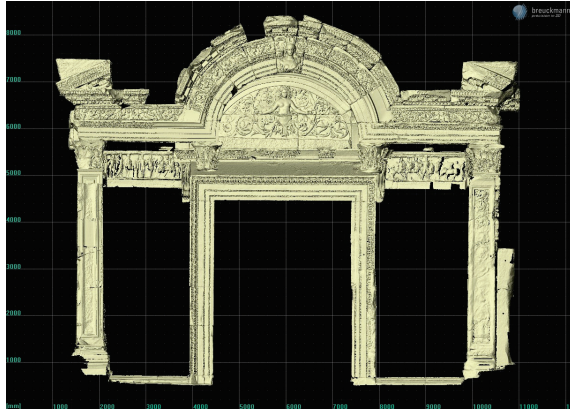
b) 3D model of the main entrance

Fig. 8.23: Recording the Temple of Hadrian using fringe projection and photogrammetry (AICON/Breuckmann, Austrian Archaeological Institute).

a) 3D model with natural texture

b) 3D model without texture

Fig. 8.24: Detail views of the Temple of Hadrian (AICON/Breuckmann).

Because fringe projection systems are sensitive to extraneous light, in particular sunlight, all scanning was performed at night. The temple extends across a volume of around 10 m x 10 m x 8 m and was recorded by around 1800 individual scans which generated a total of more than 1 million surface points. Fig. 8.23b shows the 3D model of the front view of the temple. Fig. 8.24 shows a detail view of the model, with and without texture, from which it can be seen that the model without texture gives a better 3D impression of the object.

8.2.2.3 Otzing burial chamber

The burial chamber in Otzing, Bavaria, dates back to the iron age. Due to the delicate and fragile findings, the site was excavated as a single block (2.3 m x 1.7 m) which was removed layer by layer. A 3D record of two of the exposed layers was made using an GOM ATOS III fringe projection system and a colour camera. For the first measurement in 2012, an area of 1.0 m x 1.0 m x 1.0 m was scanned at a resolution of 0.5 mm. The second measurement in 2015 scanned an area of 0.5 m x 0.5 m x 0.5 m with a pixel resolution of 0.25 mm. Colour information was provided by an additional set of images taken using a Nikon D800 camera with a 35 mm fixed-focus lens. Photogrammetric accuracy was estimated at around 0.03 mm (RMS of corrections at scale bars). This ensured that the surface model generated by fringe projection could be correctly combined geometrically with the colour information.

a) Surface scanning using fringe projection

b) 3D model of two excavation epochs (2012 and 2015) with detail enlargement top right
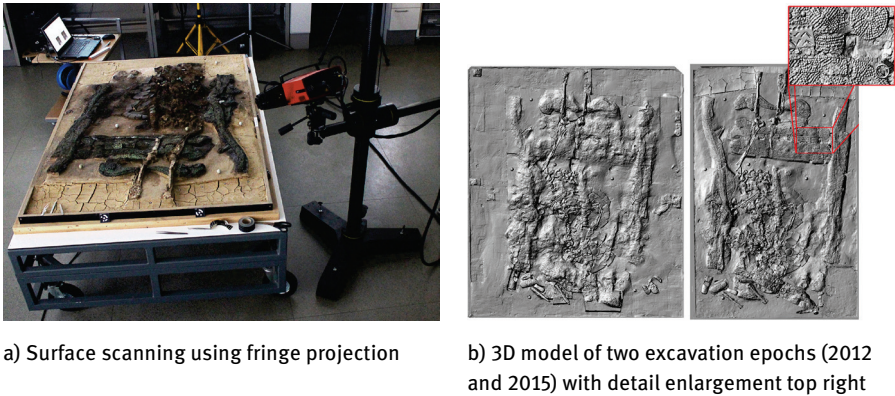
Fig. 8.25: 3D modelling of an excavation block (i3mainz).

Fig. 8.26: Coloured 3D model (i3mainz).

Fig. 8.25a shows the measurement being made with the fringe-projection scanner. The 3D model from two excavation epochs is shown in Fig. 8.25b. Fig. 8.26 shows the3D model combined with colour information. As part of the documentation, the intention is to present the archaeological findings in a multi-layered geographic information system (GIS). This will utilize 2½D/3D data and metric orthophotos. In addition, replicas of specific findings will be produced by 3D printing and an interactive virtual model will be provided.

8.2.2.4 Survey of the Bremen cog

The Hansa cog (a type of ship), found in Bremen harbour and dating from the 14th century, was completely measured by photogrammetry prior to water conservation lasting from 1982–1999. Recording and analysis were done using analogue cameras and analytical stereoplotters. Results were scaled plans with the principal contour lines of the object (Fig. 8.27). Following removal from the conservation tanks and a drying period of several years, a new digital measurement was made in 2003. This provided 3D profiles and models used to document changes in the cog during the conservation process and its subsequent presentation to the German Maritime Museum in Bremerhaven.
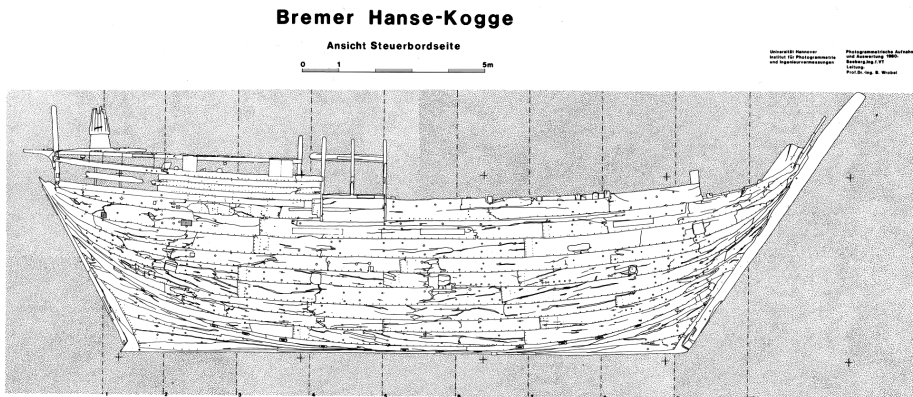
Fig. 8.27: Stereoscopic line extraction (Bremen Cog, IPI Hanover).

The entire set of targeted points and lines were digitally measured in the approximately 100 digital images. The image network was oriented and analysed using PhotoModeler (EOS System). At image scales between 1:100 and 1:900, object point accuracies of around 1–2 mm in XY and 8–10 mm in Z were achieved. Fig. 8.28 shows the digital imaging configuration and an example comparison of profiles.

Starting in 2020, a system for monitoring critical object deformations (>2 mm) has been established. Based on extensive feasibility studies, a suitable concept for the

geometric monitoring of the cog was developed. On the one hand, a set of reference points surrounding the cog was established by laser tracker to an accuracy of 0.04 mm (RMS$_{XYZ}$ 1-sigma). On the other hand, a photogrammetric measurement was conducted.

a) Digital image network configuration

b) Comparison of profiles
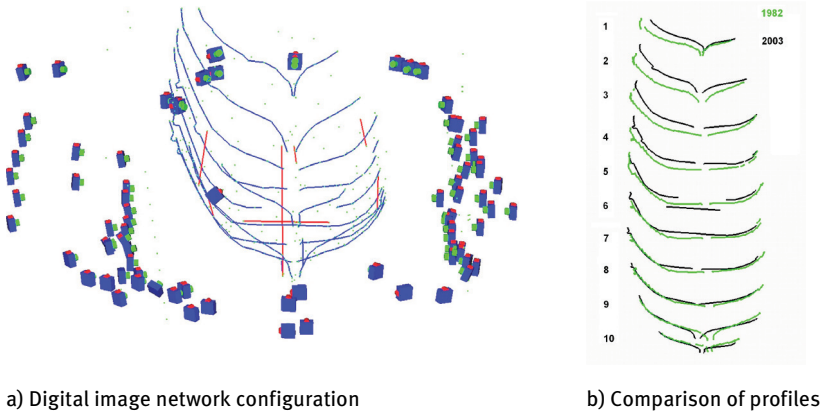
Fig. 8.28: Digital survey of the Bremen Cog (IPI Hanover).

a) Cog in exhibition hall (2017)
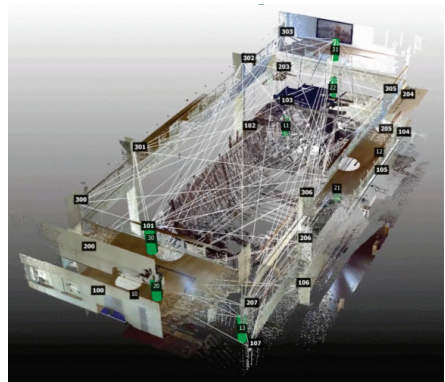
b) 3D laser scan and planning of network

Fig. 8.29: Bremen Cog in 2017 (IAPG).

Fig. 8.29 shows a TLS point cloud of the exhibition hall, used for planning the geodetic and photogrammetric networks. One specific problem for a photogrammetric solution is that the reference points are located outside the ship and significant extra effort was required to measure them photogrammetrically to a sufficiently high accuracy. Fig. 8.30 shows the photogrammetric imaging configuration and the measured object points (superimposed on the laser scan). The

photogrammetric result achieved a precision of 0.07 mm (1-sigma RMS$_{XYZ}$) with maximum uncertainties of 0.15 mm for the exterior reference points.

a) Imaging configuration   b) Measured points
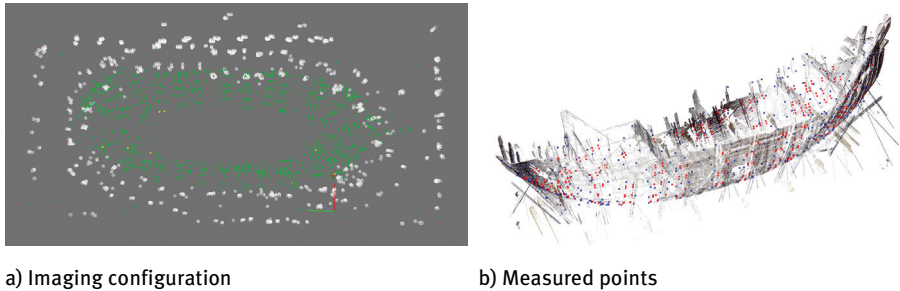
Fig. 8.30: Photogrammetric measurement (IAPG).

8.2.2.5 Megalithic tomb Kleinenkneten

The prehistoric megalithic tomb Kleinenkneten II was excavated in the 1930s and subsequently reconstructed with some rearrangements. To analyse the excavation situation at the time, the current situation was recorded three-dimensionally using terrestrial images, UAV flights and TLS measurements. In addition to the newly created 3D models and orthophotos for documentation and visualisation, uncalibrated historical photographs were digitized (Fig. 8.31). Some of these images could be oriented using the DLT (section 4.2.4.1) or via SfM (Fig. 8.32). This made it possible to locate earlier excavation details and relate them to the present situation. The project was conducted in cooperation with the Oldenburg State Museum of Nature and Mankind (LMNM).
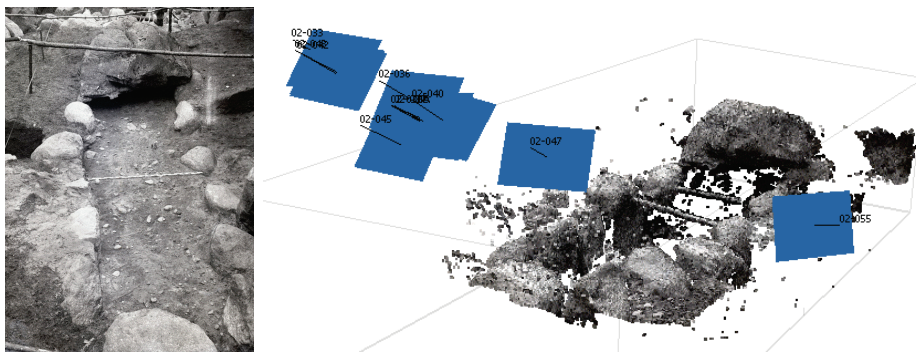
Fig. 8.31: Historical photo (LMNM).   **Fig. 8.32:** 3D point cloud-from historical photos (IAPG).

The UAV flights were carried out with an ultra-light DJI Mavic Mini drone (example in Fig. 8.33). Terrestrial laser scanning was carried out with a Leica RTC 360. Signalized control points measured with GNSS were used for georeferencing. The object resolution was approx. 4-5 mm.

Fig. 8.33: UAV image (IAPG).

On the basis of the 3D data (example section in Fig. 8.34), changes to the layout of the stone grave after excavation could be demonstrated. Furthermore, individual historical photos could be oriented via tie points in the system of the 3D model (Fig. 8.35).

Fig. 8.34: Detailed view of the 3D model (IAPG).
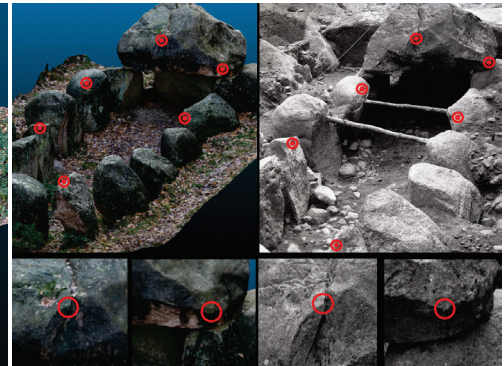
Fig. 8.35: Tie points between 3D model and one of the historical photographs (IAPG).

8.3 3D city and landscape models

8.3.1 City models

The application of 3D city models is becoming widespread, for example in:
– urban planning;
– emissions analysis (sound, exhaust gases);
– planning mobile telephone networks;
– setting up information systems (operational planning for rescue services, transport management);
– tourism (websites);
– three-dimensional city maps and navigation systems.

These applications require the 3D model to be visually appealing, up-to-date and complete, rather than having high accuracy in the geometry and detail. Fast generation of city models can be achieved through:
– aerial photogrammetry (automatic extraction of buildings);
– airborne laser scanning (extraction of buildings from discontinuities in the height model; alignment with ground plan);
– video and laser scanning acquisition from moving vehicles (mobile mapping, section 6.11.1).

a) Addition of snowfall   b) Addition of trees   c) Complex building geometry

Fig. 8.36: Extracts from 3D city models (Oldenburg campus, IAPG).

Fig. 8.36 shows extracts from 3D city models whose components have been derived from aerial photography. Textures obtained from close-range imagery have been applied to the façades and create a realistic impression of street scenes. Different visual impressions can be further achieved by the addition of computer-generated graphics and animations, for example trees or snowfall.

Due to the large area covered by some of the façades, some of the textured images in this example have been created from a number of individual images. For this purpose the separate images are rectified in a common coordinate system and subsequently connected in an image mosaic (Fig. 8.37).
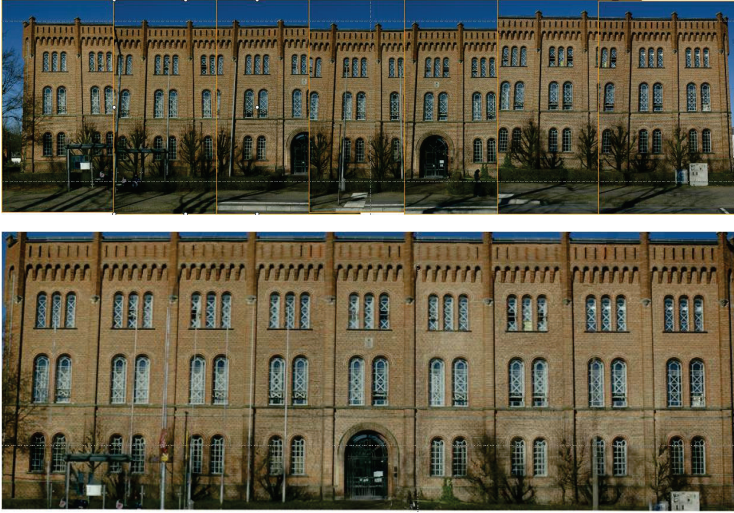
Fig. 8.37: Texture image created from a number of individual images (IAPG).

8.3.2 Generation of terrain models from multi-spectral image data

The output from a photogrammetric terrain measurement is normally a digital surface model (DSM) in which the point cloud could well represent the surface defined by the ground vegetation, but not the ground itself (Fig. 8.38). In contrast, airborne laser scanning can detect the ground by analysing the return signals (first pulse, last pulse), provided that the vegetation cover is sufficiently open for the laser beam, in part, to penetrate through.

Fig. 8.38: Profile section through a DTM including vegetation (Frankfurt UAS).

Photogrammetric terrain measurement of the ground, when vegetation is present, can be achieved using multispectral imaging in the near infrared band (NIR, see Table

3.1). Vegetation indices known from remote sensing, e.g. the Normalized Difference Vegetation Index NDVI, can be used to separate vegetation from other object classes.

Fig. 8.39 shows RGB and NIR images from a UAV survey of a restored waste site. Two Sigma DP2 cameras with Foveon sensors (section 3.4.1.6) were used. One was modified for infrared imaging by replacing the normal IR filter by one which only transmitted wave lengths >720 nm. Ground sample distance (GSD) amounts to 4 cm. A four-channel orthophoto mosaic is derived from a point cloud created by SfM. In order to detect trees and bushes, and ignoring grassed areas (lawns, fields), regions are masked where the NDVI exceeds a given threshold (Fig. 8.40). At these locations, the surface model is filtered, i.e. their related heights are interpolated from neighbouring elements. Finally, at areas of residual vegetation the height model is interpolated on the basis of terrain slope in order to generate the final terrain model of the ground (Fig. 8.41).

Fig. 8.39: RGB and NIR image from a drone (Frankfurt UAS).
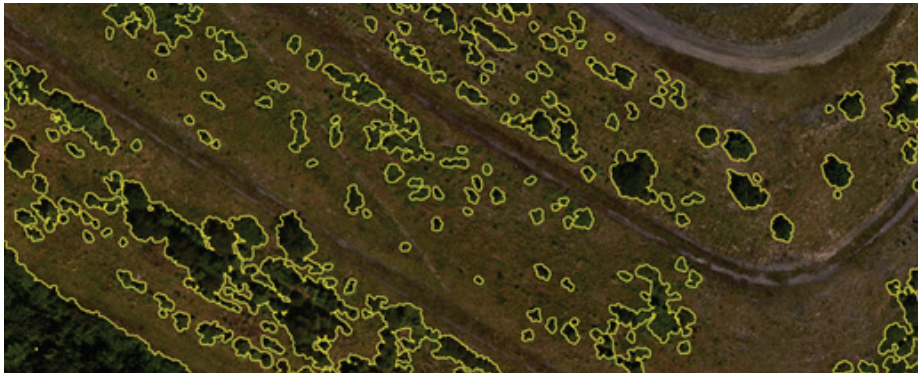
Fig. 8.40: Masked vegetation regions after NDVI analysis (Frankfurt UAS).
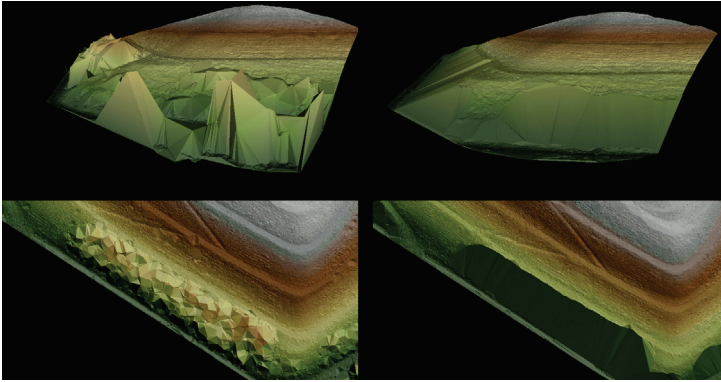
Fig. 8.41: DTM with deviations caused by vegetation (left) and NDVI-filtered DTM (right) (Frankfurt UAS).

8.3.3 Landscape model from laser scanning and photogrammetry

The reconstruction of the Tustan rock landscape close to Lviv (Ukraine) is a good example of the combination of photogrammetry and terrestrial laser scanning. Local tree density and limited measurement locations ensured that the TLS point clouds were incomplete and not registered, but complete photo coverage was available:

– 8 separate point clouds (19 million points) covering approximately 70 % of the object;
– 9 sets of images (6467 images in total) covering 100 % of the object.

Processing had 3 stages:

1. Registration of the TLS point clouds:
   Of the 8 point clouds, only 5 could be oriented using reference points and registered to each other by ICP. As a result, only 50 % of the object surface was covered.
2. Calculation of a photogrammetric point cloud:
   By means of SfM (Agisoft PhotoScan), a dense point cloud was generated which, however, could not be oriented and scaled absolutely due to missing reference points. In addition, a non-metric camera was used.
3. Fusion of point clouds:
   The registration of the TLS point cloud with the unscaled SfM point cloud was performed by a best-fit algorithm (CloudCompare). The mean of the transformation residuals was 3.3 cm.

Fig. 8.42: Terrain model from TLS and SfM (left) and virtual reconstruction of a castle (right) (SPM3D).

The calculated surface model was ultimately used for the virtual reconstruction of an ancient castle (Fig. 8.42).

8.4 Engineering surveying and civil engineering

8.4.1 3D modelling of complex objects

8.4.1.1 As-built documentation

As-built documentation encompasses the measurement and inventory recording of an existing production facility. Most commonly, the complex structures and arrangements of pipes and machinery must be measured three-dimensionally for the following purposes:

– Generation of up-to-date construction plans and CAD models, to centimetre accuracy, for production planning and control, and plant information systems (facility management).
– Provision of precise geometric data (millimetre accuracy or better) to enable the replacement of large components which are manufactured off site to fit existing mechanical interfaces.

For power plants, and especially for nuclear facilities, on-site measurement time must be minimized in order to reduce the danger to personnel and avoid interruption of processes in operation. Components such as pipe sections, heat exchangers and boilers are replaced during regular shut-down periods and the necessary geometric data must be available in advance. Using targets attached to the objects, accuracies are specified to about 0.5 mm for object dimensions of 10–20 m.
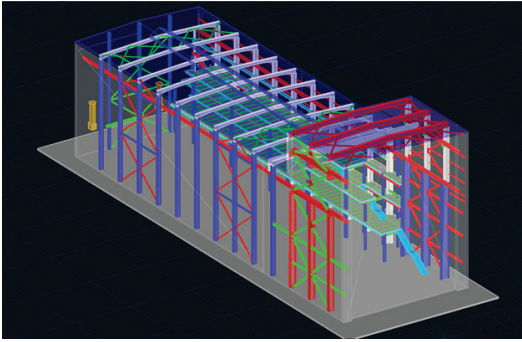
Fig. 8.43: 3D model of a production facility (SPM3D).

The documentation of complex pipework is of considerable importance. Relevant sites such as chemical plants, oil refineries and water works are characterized by difficult environmental conditions, bad visibility and complex geometries. The objective of documentation is to produce a correct inventory and plans which in practice can only be processed by 3D CAD systems due the three-dimensional nature of the facility. The data acquired can be used for 2D and 3D views (Fig. 8.43), inventory analysis such as parts lists or the location of pipes, as well as the simulation of production processes.

A purely photogrammetric measurement of industrial facilities is generally too laborious. A more efficient approach is to combine 3D laser scanning with photogrammetric image processing. 3D recording by laser scanner is fast and requires relatively little effort on site. The further processing of the measurements of these complex structures is then simplified using images which have been recorded at the same time and oriented to the laser scanned data in a common coordinate system. For example, monoplotting (section 4.2.7) can be used for visual interpretation and identification of structural elements while the actual 3D measurement is performed by interpolation within the point cloud (see Fig. 4.23). In general, oriented images are always useful for interpreting the measured scene, even if the actual measurement is based on TLS point clouds only.

8.4.1.2 Building Information Modelling

Building Information Modelling (BIM) takes a database approach to the digital documentation of the entire life of a building, from its planning through its construction and on to its monitoring and maintenance. Measurement of the building's 3D geometry, both interior and exterior, is an essential component in this process and one which can be solved by the application of terrestrial laser scanning (TLS), photogrammetry or total stations, alone or in combination. The process of deriving a BIM model from 3D point clouds is known as Scan-to-BIM (Scan-2-BIM).
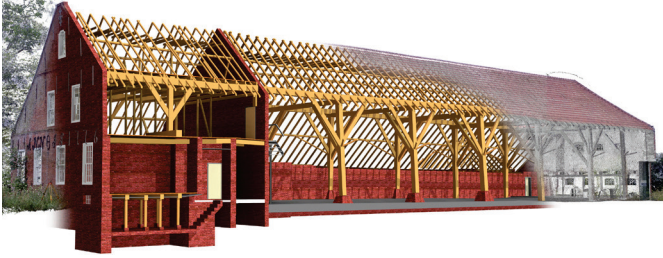
Fig. 8.44: BIM model for purposes of renovation and redesign (DhochN).

BIM does not usually store individual 3D coordinates but instead defines object elements such as walls and doors, together with their dimensions, topological connections and semantic information. Fig. 8.44 shows the TLS point cloud of a building and derived elements for future renovation and redesign purposes. In this case, data processing was done both manually and semi-automatically (detection of walls, beams etc.) in Autodesk Revit, with parallel use of images in which object materials can often be identified more reliably than from (coloured) point clouds.

Fig. 8.45a shows a dense point cloud derived from UAV images which has been combined with a TLS point cloud in order to generate facade elements for a BIM (Fig. 8.45b). The specified accuracy was only 4 cm for measurement of object details larger than 10 cm and the data were processed in Autodesk Revit.
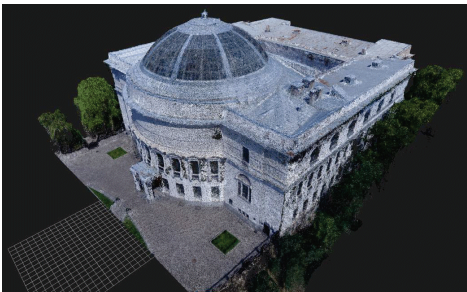
a) Point cloud from UAV imagery   b) 3D building model
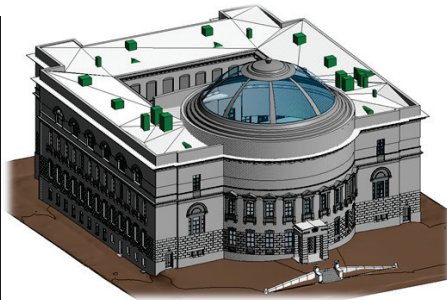
Fig. 8.45: Object modelling for BIM (SPM3D).

Fig. 8.46 shows a photogrammetric system for monitoring construction sites. The hardware system is mounted on tower cranes and it is configured to collect nadir images of the site (GSD ca. 1 cm) every day automatically or on demand. Due to the rotation of the crane, images are arranged in a circular configuration. RTK GNSS and IMU measurements assist the processes of image selection and calculation of exterior

orientations. Photogrammetric image processing is done with Pix4Dcloud Advanced. The end results are an unclassified dense point cloud, an orthomosaic, and a set of oriented images which can be further processed manually. This includes the extraction of elements for BIM or the support of monitoring tasks such as the detection of planned or unplanned changes (change detection). As well as enabling geometric processing, the acquired images represent a valuable documentation of the construction progress which can be analysed at any future time.
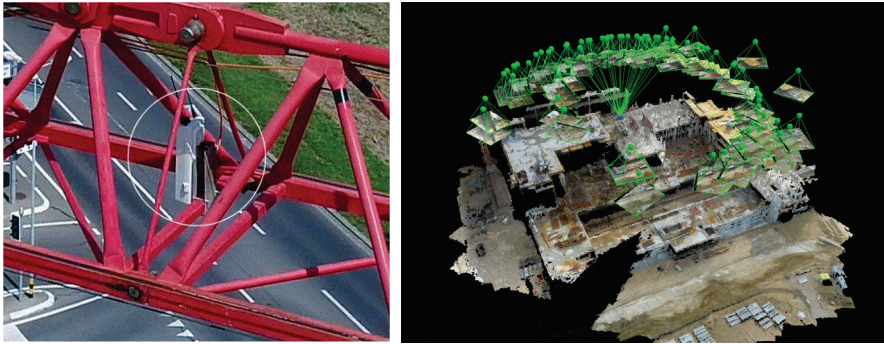
Fig. 8.46: 3D as-built modelling with a crane camera (IGP TU Braunschweig, Pix4D).

8.4.1.3 Stairwell measurement

The three-dimensional recording of stairwells is a complex task due to the usual on-site conditions (accessibility, visibility). The problem can be solved photogrammetrically when the elements to be measured (edges, corners, steps, handrails) are suitably targeted and an appropriate imaging configuration is selected.
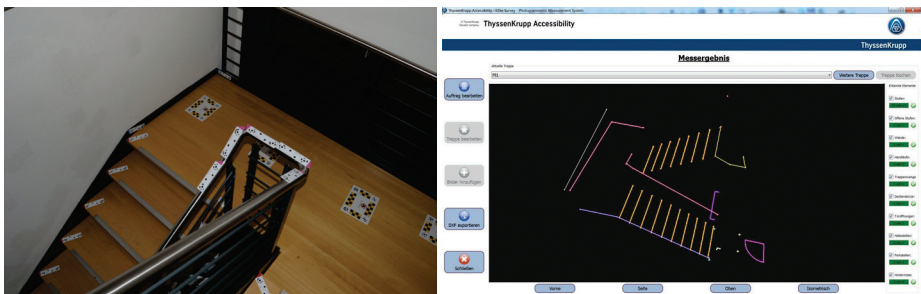
a) Targeting using special target adapters   b) 3D analysis

Fig. 8.47: Photogrammetric stairwell measurement (AICON, ThyssenKrupp).

Fig. 8.47a shows a stairwell targeted with adapters whose measurement points define a unique relationship to the actual point of interest, e.g. the edge of a step (see section 6.2.2.3). The coding integrated into the adapters is designed so that the type of feature (corner, edge) is automatically identified. Measurement of features of interest is typically achieved to an accuracy of around 1 mm.

8.4.2 Deformation analysis

A key task in geodetic engineering surveys is to monitor deformations on buildings exposed to some particular mechanical or thermal loading. For these applications accuracy requirements are typically in the order of millimetres for object dimensions of more than 100m, for example cooling towers, chimneys, wind turbines, dams, bridges, sluices, cranes and historical buildings.

Photogrammetric deformation analysis is usually applied in cases where object or environmental conditions do not allow sufficient time on-site for extensive geodetic point measurements, or a large number of object points are required. Image acquisition additionally provides an objective documentation of the object's state at the time of exposure and, if simultaneous measurements are made, can also record rapid object changes.

High precision photogrammetric object measurement requires high resolution and mechanically stable cameras. Critical object points are targeted and a stable network of reference points is necessary for detecting possible object deformations or movements.

Measurement of the targeted points is done using digital image processing methods (section 5.4.2) with an image measurement accuracy in the region of 1/10 to 1/50 of a pixel. 3D object coordinates are calculated by bundle adjustment with simultaneous camera calibration and the inclusion of any additional geodetic measurements. Absolute scaling is essential and must be carried out either using high-precision measured reference distances or sufficiently precise control points (see section 6.3). Deformation analysis can also be performed within the bundle adjustment, or by separate 3D transformations of the object points in different measurement epochs. In addition to point-based deformation analysis, areal analyses based on point clouds are becoming increasingly important.

8.4.2.1 Deformation of concrete tanks
In this example, the deformation of concrete tanks used for galvanizing and electroplating must be measured under working conditions. The tanks are constructed from a special concrete and have dimensions of approximately 4 m x 10 m x 2 m. In operation they are slowly filled with liquid of total weight 7.5 tons. The tank walls are subject to critical deformations which must be observed

photogrammetrically at 10 minute intervals. Around 325 points must be measured to an accuracy of less than 0.1 mm.
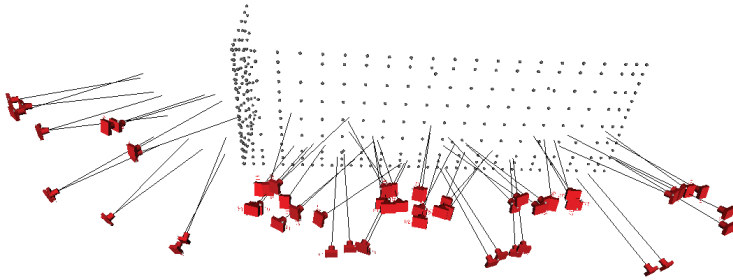
Fig. 8.48: Imaging configuration (IAPG).
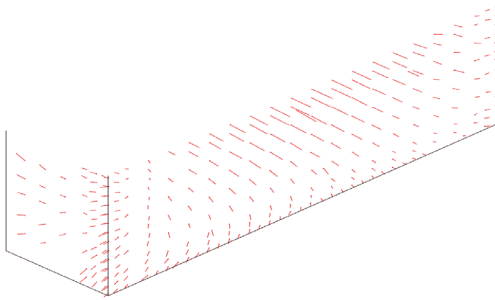
Fig. 8.49: Resulting deformation vectors between two sequential measurements (IAPG).

Due to the very confined object environment, the shortest focal length available for a Fuji FinePix S2 digital camera (f = 14 mm) must be used. Fig. 8.48 shows the measurement configuration which makes use of two reference lengths to define scale. However, points on the smaller object side could not be observed with optimal intersection angles, hence object point accuracy is weaker in those areas. The image scales achieved were between 1:70 and 1:120. The calculated deformations show a systematic behaviour (Fig. 8.49).

8.4.2.2 Measurement of wind-turbine rotor blades

In operation, the rotor blades of wind turbines are subject to various types of deformation such as bending, out-of-plane deflection and torsion. An understanding of these effects is important for optimizing the efficiency of power generation and designing for material properties relating to the blades' strength and durability. However, it is a significant metrological challenge to measure an operational wind

turbine. These generators can have towers up to 200 m in height, blades with lengths >80 m and locations with challenging environmental conditions. In addition, accuracy requirements may be demanding, perhaps a few centimetres. This requirement is further complicated by the dynamic nature of the problem as the rotating blades respond differently to the changing wind conditions around the turbine and their changing position.

Fig. 8.50 illustrates the principle of a photogrammetric solution using retro-reflective targets attached to the components of the wind turbine. Images are recorded at a frequency of 100 Hz by two pairs of synchronized stereo cameras, relatively oriented to one another. Tower, nacelle and rotor blades can be measured point-by-point during operation. This enables the trajectories of the target points to be measured over a longer period of time in all positions of the rotor blades. Measurement accuracy is estimated to be around 5–10 mm at a distance of 110 m.
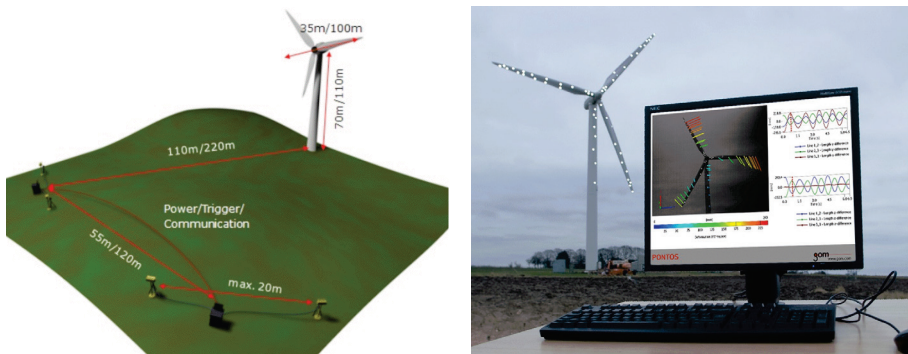
Fig. 8.50: Photogrammetric measurement of rotor blades using individual targets (GOM).

Fig. 8.51: Photogrammetric measurement of rotor blades using textured adhesive film (ForWind, Uni Hannover).

Fig. 8.51 shows an alternative photogrammetric approach where the rotor blade is partially covered by a textured film so that a dense point cloud can be measured by correlation techniques. This approach enables a more accurate determination of torsion than is possible with a small number of individual targets.
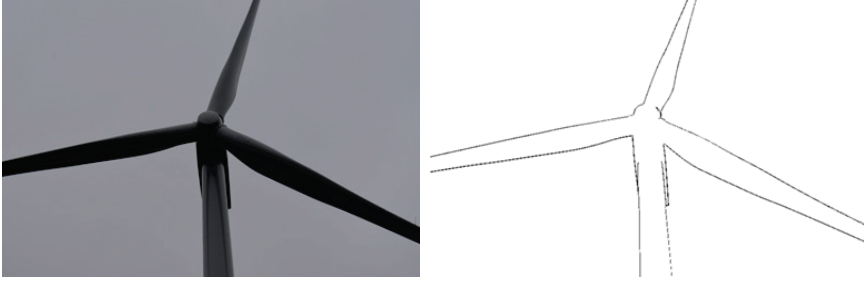
Fig. 8.52: Photogrammetric measurement of rotor blade by contour extraction (IAPG).

A drawback of both the above approaches is that the rotor blade must be signalized. For this, the turbine must be stopped, a process which has logistics requirements and generates financial costs. Targetless measurement techniques have therefore been developed. Fig. 8.52 shows a photogrammetric solution where the outer contours (grey-level edges) of the blade are measured in several images. A 3D model of the blade can be estimated by using the image coordinates along the contours, optionally combined with laser-scanned profiles, together with an integrated finite-element model of the blade's distortion under load.

a) Measuring concept

b) Camera image for measurement of nacelle movement

c) Measured rotor-blade profiles
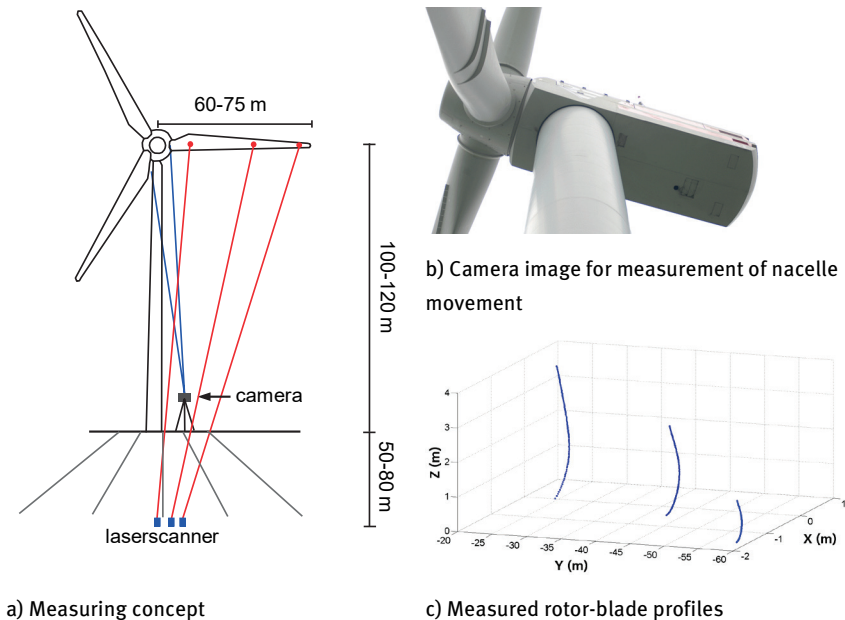
Fig. 8.53: Measurement of rotor blade using distance measurement only (IAPG).

Another targetless method uses either multiple laser scanners in 1D mode or several distance meters (see section 6.8.2). This technique involves only distance (range) measurements (Fig. 8.53). Here surface profiles are generated at locations where the blade rotates through the laser ranging beam. The assignment of profiles to the physical position of the blade is solved by using a camera which continuously monitors the rotation of the nacelle, and by orientation of the scanning system with respect to the turbine's coordinate system.

A general problem is the definition of the turbine's coordinate system, usually at the centre of rotation, and the determination of the exterior orientation of the sensors. Errors in datum definition and/or orientation directly cause positional errors in the measured locations on the rotor blade and this may result in misinterpretation of deformations.

8.4.3 Material testing

8.4.3.1 Surface measurement of mortar joints in brickwork

The following application is an example of deformation analysis in building maintenance and is concerned with the measurement of erosion which is affecting the pointing (mortar joints) in brickwork. Over a period of some 10 years, the progress of erosion will be monitored on site every two years. The test sites are located on a church tower at a height of around 40 m (Fig. 8.54). Each test area is approximately 360 mm x 220 mm and is measured by stereo imagery. Due to the difficult lighting conditions, fringe projection systems cannot be used.
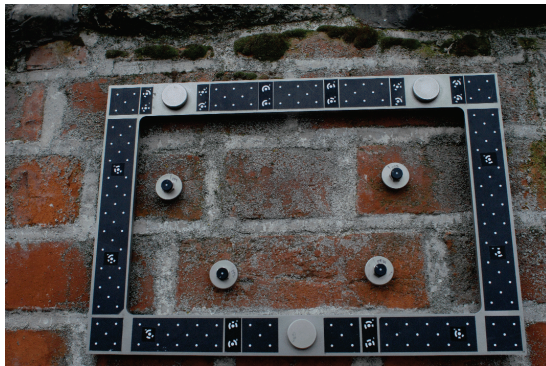
a) On-site location   b) Sample image

Fig. 8.54: Data acquisition for erosion measurement (IAPG).
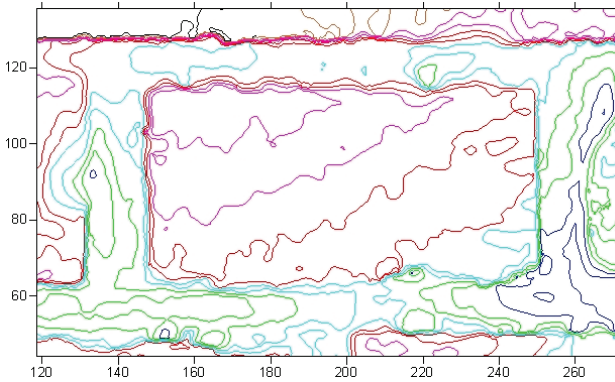
Fig. 8.55: Surface model derived from stereo matching (IAPG).

The accuracy is specified to about 0.1 mm. Four masonry bolts define a fixed object coordinate system. A separate reference frame containing calibrated reference points can be re-attached to a test area in order to deal with repeated measurement. A digital camera is used for image recording. Since the measurements must be made under difficult daylight conditions, a fringe projection system cannot be employed. The natural surface structure provides enough texture for image matching and 3D reconstruction. Results are presented in the form of contour lines and differential height models (Fig. 8.55). Typically, around 100000 surface points are measured.

8.4.3.2 Structural loading tests

The following example describes the photogrammetric recording of deformations and cracks in fibre-reinforced concrete test objects (TU Dresden). The qualitative and quantitative development of cracks and deformations during the loading is of particular interest.

Fig. 8.56 shows the imaging configuration for the measurements. The test load object is recorded in three dimensions using digital stereo cameras. Mirrors positioned to the side provide side views of the test object. Only one of each camera views each side in order to record crack development in those areas. Camera calibration and orientation is done prior to testing using a multi-image network with bundle adjustment and self-calibration.
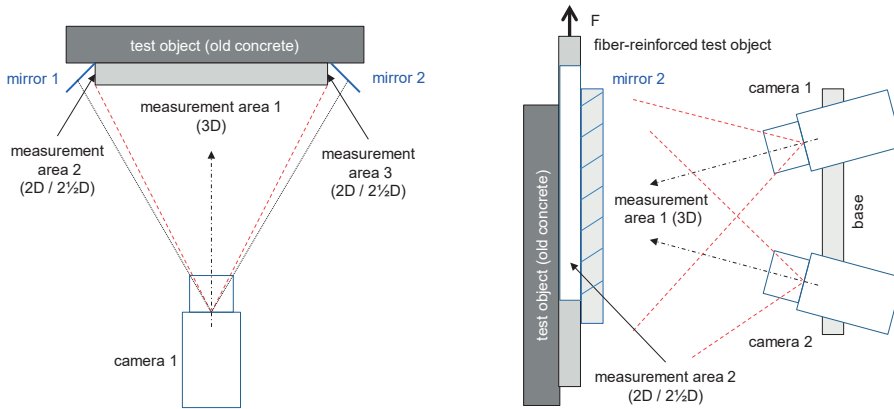
Fig. 8.56: Imaging configuration for displacement and crack analysis (IPF TU Dresden).

Cracks are detected by applying dense area-based matching to consecutive images and analysing the resulting shift of image patches for discrepancies. This way, cracks can be localized with pixel accuracy, and crack widths can be determined to an accuracy of about 1/20 pixel. Fig. 8.57a shows a greyscale-coded visualisation of the matching results, showing the position of cracks for one load stage. Cracks in the material are clearly distinguishable from the areas between them. Fig. 8.57b visualises measured crack patterns for single a load stage, where *dZ* denotes the width of the crack. A subsequent crack analysis (crack location and width) is made, for every load stage, along defined profiles shown in the image. The photogrammetric system permits continuous measurement of object deformation with an accuracy of up to 1 µm, and cracks with an accuracy of around 3–5 µm.
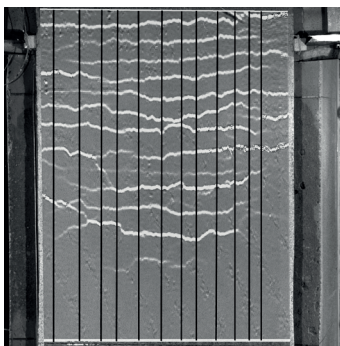
a) Crack pattern with marked profiles   b) Surface evaluation (around 1.8 million points)
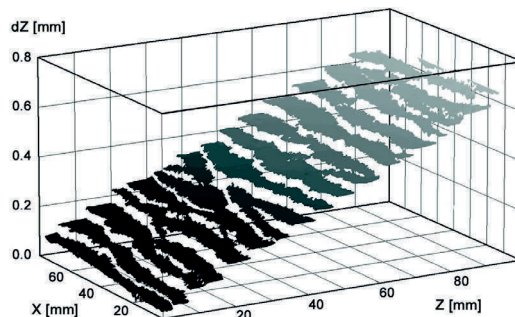
Fig. 8.57: Result of displacement and crack analysis (IPF TU Dresden).

8.4.4 Roof and façade measurement

The measurement of building surfaces is valuable for the following applications:
–   measuring of façade areas for cleaning and painting;
–   measuring of façade areas for calculating quantities of construction and insulation material;
–   roof measurements for tiling;
–   simulation of building modifications and additions, e.g. colour, brickwork and tiles;
–   measurement of windows and doors;
–   roof measurements to plan for solar panels.

Fig. 8.58 shows the result of measuring the façade an apartment block. For this image the plane projective transformation parameters are known (section 4.2.6) so that the side lengths of the façade, and the location and areas of windows and other objects within the façade, can be determined. In this case, transformation coefficients are calculated using the four corner points of a window with known dimensions. This results in extrapolation errors at the edges of the image.

Fig. 8.58: Façade measurement and simulation of new building colour (IAPG).

Fig. 8.59a shows a measurement image with a graphical overlay of a planned solar panel array. Plane transformation parameters are obtained through the use of a calibrated reference cross whose target points can be automatically found in the image (compare with Fig. 4.18). Planning accuracy is around 5 cm. The required solar modules were inserted interactively into the image from a databank and, with texture projection, given a realistic appearance. Using an approximately computed exterior orientation of the uncalibrated image, it is possible to estimate the locations of points which are perpendicularly offset from the plane of the roof, for example in order to visualise the appearance of a raised module (Fig. 8.59b).
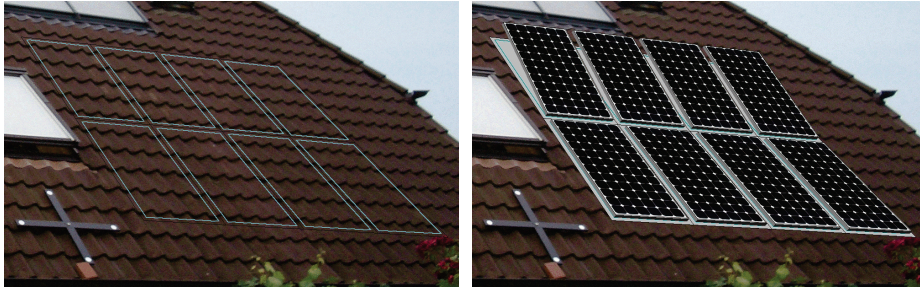
Fig. 8.59: Image analysis for planning the installation of a solar panel array on a roof (IAPG).

8.5 Industrial applications

The range of applications for photogrammetry and other optical 3D measurement methods in the industrial environment is very broad. It includes construction and design, quality assurance (form and position measurements, target/actual comparisons), incoming goods inspection, real-time control of machines and robots (inline measurement), metrological experiments (e.g. wind tunnel, crash test, driving characteristics), driver assistance systems and control of autonomous platforms. The digitalisation of all processes, interaction between man and machine as well as networking and analysis of data, which can be subsumed under the term Industry 4.0, is becoming increasingly important.

8.5.1 Power stations and production plants

8.5.1.1 Wind power stations
There are a number of photogrammetric applications relevant to the construction and operation of wind power plants, for example:
− Measurement of rotor blades:
  Rotor blades are made of fibre composites and core materials such as balsa wood, PET and PVC foams. Both glass and carbon fibres are increasingly used for load-bearing elements. For quality assurance during manufacture, shape and size are partly controlled by physical reference gauges but increasingly also by optical 3D measuring systems. During the validation process, deformations and material strains are measured and bending lines determined from them. Here conventional techniques (wires, strain gauges) are also increasingly replaced by optical methods. Fig. 8.60 shows optical deformation measurement during a fatigue test. The dynamic blade deformation is recorded by a photogrammetry system, using measuring targets on the blade and fixed markers in the hall to

define a reference system. Deformation measurements on real wind energy systems in situ are discussed in section 8.4.2.2.

Fig. 8.60: Rotor blade during a load test (red: photogrammetric targets) (Fraunhofer IWES).

– Measurement of wind turbine tripods and rotor blade hubs:
Checking the flatness and roundness of the base of the turbine tripod (support tower) is done directly on site. This helps to ensure that the turbine is assembled according to plan. Measurement is by offline photogrammetry using a multi-image network which provides three-dimensional coordinates of targets attached to the surface of the tripod flange. Typical accuracy is around 0.1 mm for a base diameter of up to 4.5 m. Fig. 8.61 illustrates the tripod measurement and the results from analysis. Fig. 8.62 shows the measurement of offshore tripods which have heights of up to 60 m.
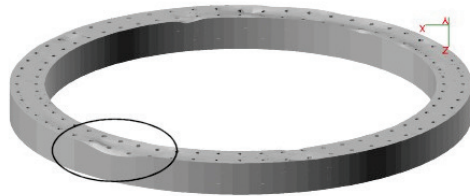
Targeting the flange and form analysis of the tripod base showing areas of defects

Fig. 8.61: Photogrammetric recording of wind turbine tripods (Hexagon).

a) Tripods for offshore wind turbines

b) Targeting with flange adapters

c) Measurement during production

Fig. 8.62: Photogrammetric measurement of offshore structures (Hexagon).
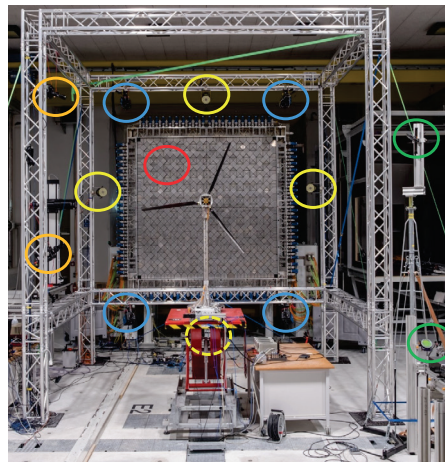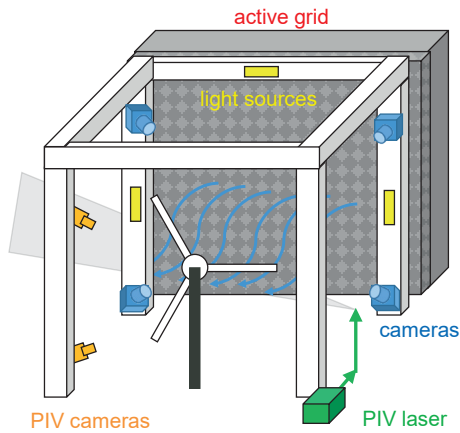
Fig. 8.63: Experimental set-up in the turbulence wind tunnel at Oldenburg University (IAPG, ForWind).

– Measurements in wind tunnels:
  Experiments are carried out in wind tunnels to investigate the behaviour of the turbine under different wind conditions such as turbulent wind flow, or in aerodynamic cases such as fluid-structure interaction. Small-scale wind turbine

models are measured in the wind stream using photogrammetric methods and/or PIV (see section 6.10.2.4). For realistic investigations, considerably higher rotation speeds must be achieved, e.g. a factor 50 compared with a real wind turbine, due to the reduced size of the model, e.g. 1:70. Optical 3D acquisition in the wind tunnel requires the use of high-speed cameras under very high accuracy requirements (approx. 0.5 mm) to avoid negative extrapolation effects. Fig. 8.62 shows a typical test setup in the wind tunnel. In this example, photogrammetric measurement methods and PIV are combined. Turbulent wind streams are created by an active grid to which the cameras are synchronized. From the synchronized data available in a common coordinate system, information on blade deformations and aerodynamics of the rotating system can be derived.

8.5.1.2 Particle accelerators

The photogrammetric measurement of detectors in a particle accelerator at CERN is an example of large object measurement requiring very high accuracy. Fig. 8.64a shows a sub-detector in the ATLAS experiment, part of the LHC (Large Hadron Collider), which has a diameter of 27 m. The detector has 12 sectors, each with 22 chambers and each marked with four targets. The objective of the measurement is the installation and determination of position of the chambers in the sub-detector system, as well as the alignment of the sub detector in the accelerator's global coordinate system.
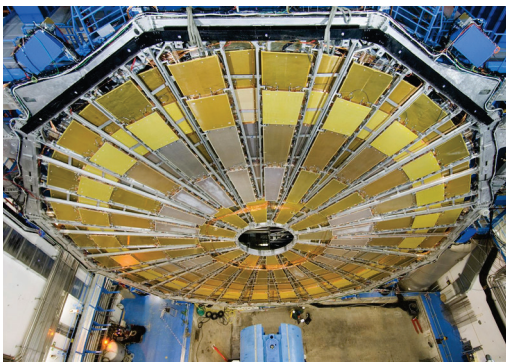
a) Sub-detector (diameter 27 m)   b) Photogrammetric analysis
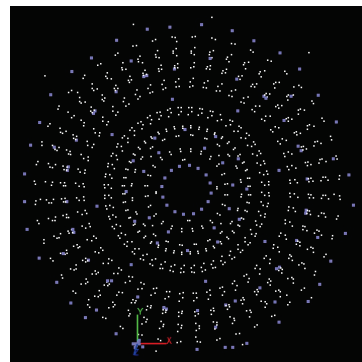
Fig. 8.64: Sub-detector of particle accelerator LHC (CERN).

The photogrammetric task encompasses the recording of around 1200 object points using approximately 1000 images (Nikon D2XS, f = 17 mm und 24 mm) which ensures coverage of areas which are difficult to access. In addition, measurements with a total station provide scale information to an accuracy of around 0.3 mm, and enable the

transformation of the photogrammetric data into the global coordinate system with an accuracy of around 0.5 mm. At average scales of 1:250–1:350, analysis with the AICON 3D Studio software provides an object coordinate precision of 0.022 mm in XY and 0.065 mm in Z (1-sigma RMS). Fig. 8.64b shows measured object points from part of the photogrammetric analysis. The absolute necessity for automatic outlier detection should be noted. Without this feature, a project like this, with some 89 000 observations and 9400 unknowns, could not be processed successfully.

8.5.2 Aerospace industry

Photogrammetric applications in the aerospace industry are distinguished by extremely high accuracy specifications for very large objects. Typical specifications for relative accuracies are between 1:100 000 and 1:250 000 or more. Example application areas of manual and automated industrial photogrammetry include:
– measurement of parabolic antennas;
– measurement of large tooling jigs and mechanical gauges;
– production control of large components and assembly interfaces;
– completeness checks of assembled parts;
– space simulations.

8.5.2.1 Inspection of tooling jigs

For the inspection of large tooling jigs in the aircraft industry, accuracy is typically specified to about 0.1 mm for object sizes up to 75 m (length of aircraft), i.e. a relative accuracy of up to 1:750 000 is required. This task can only be solved using highly redundant, multi-image networks or laser tracker measurements. Fig. 8.65a shows an older application of photogrammetric measurement using the Rollei LFC and GSI CRC-1 large-format analogue cameras (see section 1.4). Fig. 8.65b shows a similar task where the GSI INCA digital metric camera has been used.

Verifying the achieved accuracy is problematic in measuring tasks of these dimensions. Reference scale bars longer than 3 m are difficult to handle in practice and significant effort is required to provide reference coordinates (see section 7.2.3). If time permits, laser trackers commonly offer a solution for reference measurement. However, with these systems target measurement is sequential and this option is expensive in terms of system and manpower resources. There is a further need for targeting which is measurable by both systems in order to connect their measurements together.
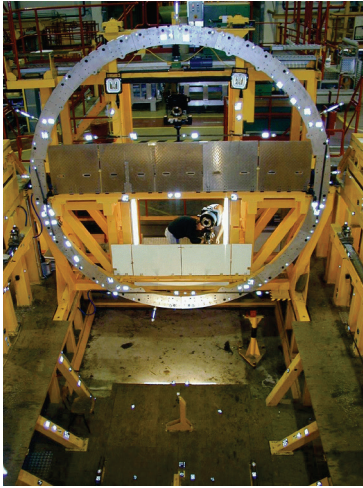
a) Periodic checking of a tooling jig using Rollei LFC and GSI CRC-1 film cameras
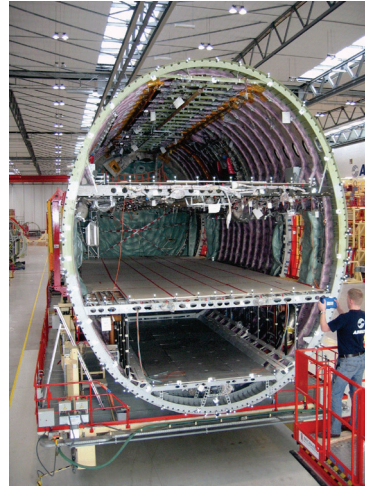
b) Measurement of A380 fuselage interface with GSI INCA digital metric camera

Fig. 8.65: Photogrammetric measurement of aircraft jigs and components (Airbus).

8.5.2.2 Production control

Fig. 6.34 shows the application of online photogrammetry to process control in aircraft manufacture. Here two cameras are mounted on an unstable platform, i.e. their relative orientation is not constant. However, the exterior orientation is continuously updated by measurement of coded targets in fixed locations around the aircraft door, which is the component being measured. Simultaneous measurement of a manually placed touch probe at critical points around the door provides the 3D coordinates of the touch point in the same coordinate system. The achievable measurement accuracy is 10 µm + 10 µm/m.

8.5.2.3 Visual checking

Fig. 8.66 shows an example of digitally supported visual checking in the aircraft industry. The CAD model of a part to be attached during assembly is projected onto the attachment location for a visual check that it is mounted in the right or wrong position, or warn if it is missing. A fringe projection system coupled with a projection device is used to make the correct geometric connection between the actual situation and the stored CAD models of parts to be checked (see also section 6.12.3).
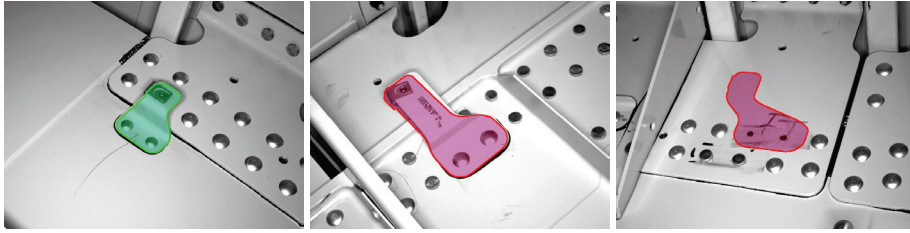
Fig. 8.66: Visual conformance check in the aircraft industry (Premium AEROTEC)
left: mounting ok; middle: erroneous position; right: missing part.

8.5.2.4 Antenna measurement

For the measurement of parabolic antennas and mirror telescopes, the shape of a hyperbolic surface must be checked. Antenna sizes range from 1 m to over 100 m. Applications cover a spectrum which includes size and shape control under varying thermal loads and the adjustment of large mirrors to the correct form. Typically, the objects can only be measured from one side and the imaging configuration is designed so that every object point can be measured from as many convergent images as possible (see Fig. 7.6). The configuration is chosen to provide a homogeneous accuracy along all three coordinate axes. Retro-reflective targets are used to mark the object surface (Fig. 8.67). The size of the object, and access restrictions, typically demand the use of hydraulic lifting platforms (Fig. 8.68) which prevents the use of alternative techniques such as laser tracker measurement.
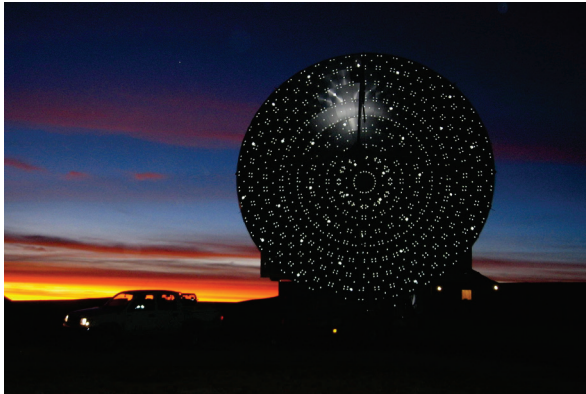
Fig. 8.67: Targeted parabolic reflector (GDV, Vertex Antennentechnik).

Fig. 8.68: Image acquisition from hydraulic platform (GDV, Vertex Antennentechnik).

The APEX telescope on the Chajnantor Plateau in the Chilean Andes (Fig. 8.67) provides a good example of the capabilities of photogrammetric measurement. Due to its position at 5100 m above sea level, only a limited time is available for measurement. The task is to measure the entire mirror surface (diameter 12 m) with a specified accuracy of 0.05 mm, in order to make adjustments to the component mirrors. Around 170 images are recorded using an INCA camera and evaluated using V-STARS (GSI). Measurement at approximately 1300 points determines the influence of gravity on the shape of the antenna in different positions. Measurement accuracy in object space is 10 μm (RMS 1-sigma).
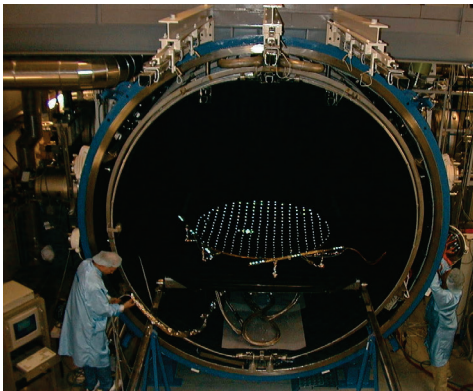
a) Imaging task in a thermal vacuum chamber   b) Results of shape analysis
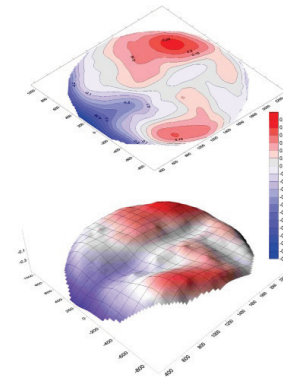
Fig. 8.69: Measurement of a parabolic antenna in a space simulation chamber (GDV, IABG).

If parabolic antennas are deployed in space, they are subject to extreme environmental conditions which can be re-created in a simulation chamber (Fig. 8.69a). Here antennas up to 4 m in diameter are exposed to temperatures between −120° C and +150° C. Either the camera or the antenna can be moved on a circular path in order to create the regular multi-image network. Deformation analyses can then be derived from the measured object coordinates. Measurement accuracy is in the range 10−20 μm (1-sigma RMS).

8.5.3 Car industry

Three-dimensional measurement technology has, for some time, been one of the most important tools for quality control in the car industry. Mechanically probing coordinate measuring machines (CMMs) are mainly used for the high-precision measurement (1–10 μm) of small components. In contrast, optical 3D measuring methods, with typical accuracies of 0.05 to 0.2 mm, are mostly applied where an

object cannot be transported to a stationary measuring system, cannot be probed mechanically or a very large number of object points must be measured within a short time. Here is a selection of areas where photogrammetry can be applied:

– production control (shape and size);
– installation of production cells and assembly facilities;
– robot control;
– surface measurement of design models (reverse engineering);
– car body measurement in a wind tunnel;
– deformation measurement in torsion and car safety tests;
– inspection of parts from third party suppliers (see windscreen measuring system, Fig. 6.46);
– driver assistance systems;
– control of production machines (e.g. brake pipes, Fig. 6.42).

Measuring systems in production environments today almost exclusively utilize digital cameras, on or off line, in order to handle the necessary high data flows. In addition to the actual optical 3D measurement, these systems generally also provide data interfaces to CAD or CAM systems.

8.5.3.1 Rapid prototyping and reverse engineering

Prototypes are designed and manufactured prior to series production. If the corresponding parts contain free-form contours or surfaces, then the conventional generation of production data, e.g. for milling machines, is a costly and normally iterative process. Rapid prototyping methods can lead to much faster manufacturing of prototypes. The measuring task is essentially the complete 3D acquisition of the model in order to derive machine control data (reverse engineering).

Both single point and surface measurement systems are employed. They are portable, flexible and can be taken directly to the object. Since it is costly and time-consuming to place targets, it is common here to use online systems with manual touch probing (target adapters and touch probes, see also section 6.5). For surface measurement, area-based systems are used. Where large surfaces must be covered, additional photogrammetrically measured reference points are used to register individual scans (see Fig. 6.55 and Fig. 6.87).

Fig. 8.70 shows the photogrammetric measurement of local reference points placed on a car body. These are used to connect together individual surface point clouds generated by a fringe projection system. The final 3D model can be used for a variety of test and design tasks.
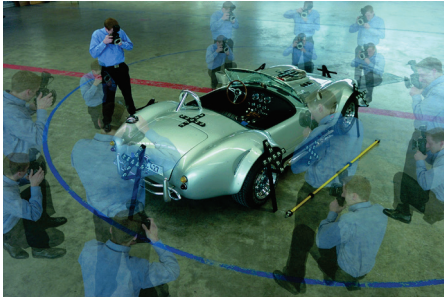
a) Photogrammetric reference point measurement

b) Surface scan by fringe projection
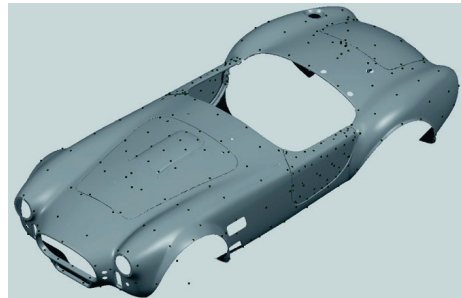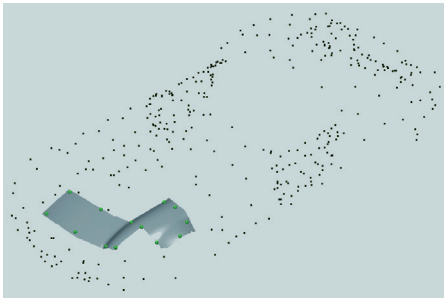
c) Reference points and individual surface scans d) Complete model from merged individual scans

Fig. 8.70: 3D modelling of a car body (Zeiss).

8.5.3.2 Car safety tests

In car safety tests (crash tests), photogrammetric evaluation of high-speed video is one of the tools used to the detailed displacements and changes in the car and dummies which represent the passengers (compare with section 3.5.3).

Typical investigations in car safety testing include:

– positioning of dummies;
– front and side crashes;
– protection of pedestrians;
– deformation analyses (engine compartment, roof and window structures, footwells).

Fig. 8.71 illustrates the Humanetics DPS (Dummy Positioning System), used for automated measurement of the pose (6DOF) at relevant points on the body of a crash test dummy. The system is based on AICON's MoveInspect system (section 3.5.3). Deviations from nominal positions are displayed in real-time. Further relevant features can be added subsequently using a manual touch probe.

Fig. 8.71: Real-time dummy positioning (Humanetics).

Here imaging configurations can be classified into two types (Fig. 8.72):
–   fixed camera(s) outside the car (A);
–   moving camera(s) installed inside the car (B).
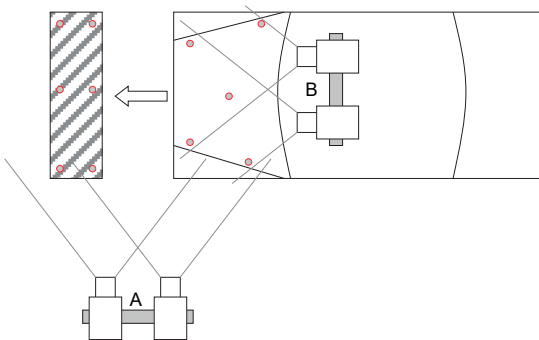
Fig. 8.72: Imaging configurations in car crash tests.

In the case of fixed cameras, interior and exterior orientations remain constant. The object coordinate system can be defined on the car or outside it, for example in one of the cameras. If the cameras are attached to the car, then they are exposed to high accelerations and their orientation parameters may change. These can then only be determined by the simultaneous measurement of fixed reference points.

To determine lateral displacements, image sequences are evaluated by 2D image analysis. To determine three-dimensional displacements and deformations, full photogrammetric techniques are employed. In crash tests, a typical accuracy is in the range 1–5 mm.

t = −15 ms

t = 0 ms

t = 10 ms

t = 33 ms
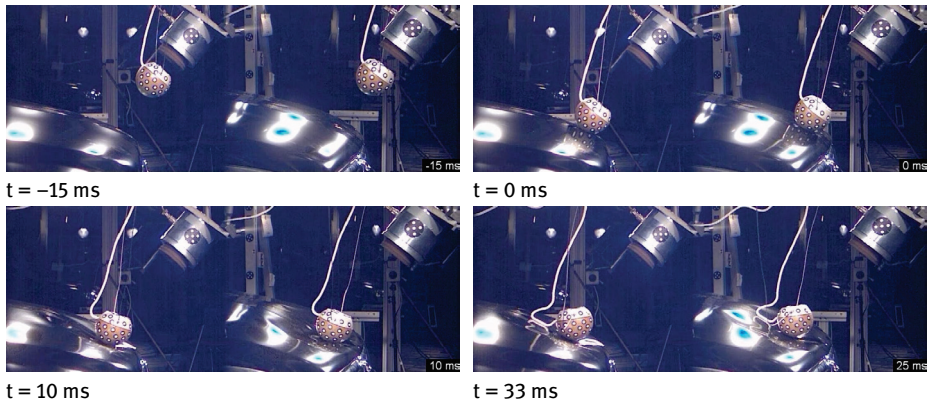
Fig. 8.73: Stereo image sequence of a head-impact test (Volkswagen).

t = −20 ms

t = 4 ms
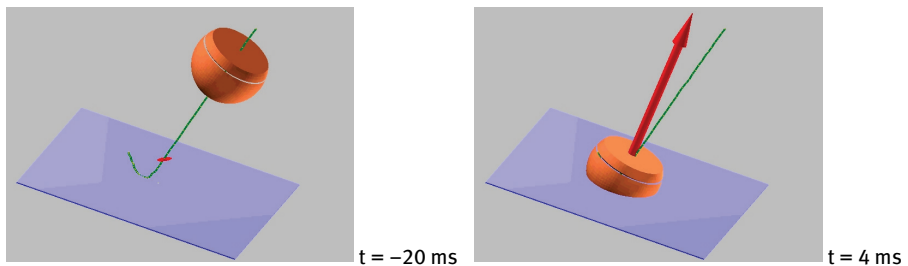
Fig. 8.74: Animated trajectory with overlay of acceleration vector (Volkswagen).

Fig. 8.73 shows extracts from an image sequence used for head-impact tests. They were made using a NAC HiDcam with a stereo mirror attachment (see section 3.5.3, Fig. 3.124). The objective is to measure the penetration depth of the head impact on the bonnet (hood). Trajectories and orientations of the impacting head can be determined in a global coordinate system from the sets of photogrammetrically measured 3D points, as illustrated in Fig. 8.74. By a combination of 3D image sequence analysis and electronic sensor technology, the forces and accelerations in pedestrian safety and frontal crash tests can be visualised.

8.5.3.3 Car body deformations

To evaluate the effects of dynamic and thermal loading, car parts or entire cars are measured photogrammetrically point-by-point. If the individual deformation states are quasi static and so remain unchanged over a short period of time, e.g. before and after testing in a climate chamber, then the target points can be measured by a multi-image network using the offline photogrammetric method.

Fig. 8.75 shows the measurement of a car door used for deformation analysis in a climate chamber. The surface is marked with targets whose movement is shown, in this example, at a temperature of 80° C. The door is attached to a stable fixture which is marked with coded targets to establish a stable reference coordinate system. The accuracy is reported to be around 0.02mm in object space.
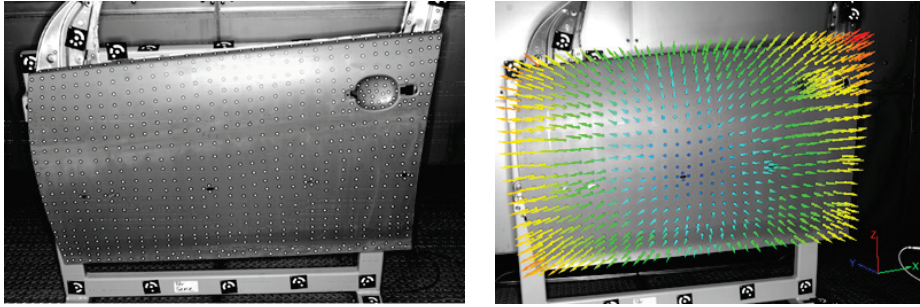
a) Metric image with targets   b) Measured deformations

Fig. 8.75: Photogrammetric part testing in a climate chamber (GOM).

8.5.4 Ship building industry

Photogrammetric tasks in shipbuilding include the following:
– measurement of steel plates and their orientation on metal cutting machines;
– measurement of ship sections;
– measurement of windows, hull and fittings;
– surface inspection of hulls.

Metrology applications in the ship building industry are characterized by:
– measurement of large objects (>10 m);
– restricted access;
– vibrations and disadvantageous environmental conditions.

An example of photogrammetric measurement in shipbuilding is the determination of shrinkage in the welding of large steel parts. For typical component sizes of 10–12 m, shrinkage during welding can amount to several mm and is taken into account in the construction plan. Fig. 8.76 illustrates the process.

Further typical applications are photogrammetric form measurement of hulls, and fixtures such as railings and windows. In addition to photogrammetric methods, geodetic techniques and laser tracker measurements are also applied.

a) Steel part

b) Targeting

c) Image acquisition from a crane

d) Imaging configuration
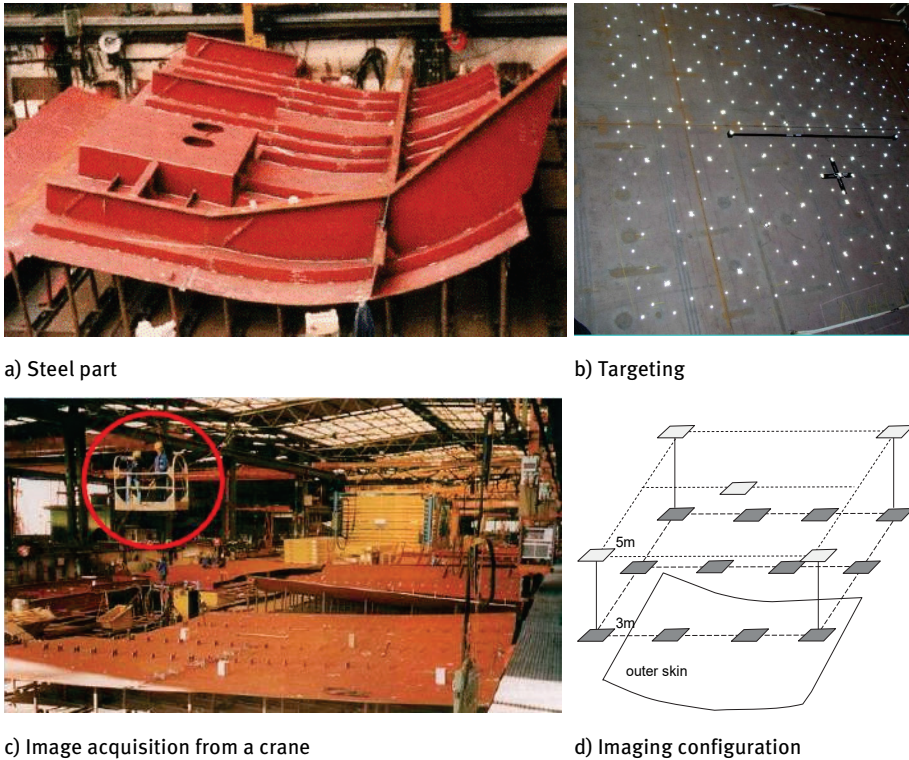
Fig. 8.76: Photogrammetric acquisition of welding shrinkage (Hexagon).

8.6 Underwater photogrammetry

The use of photogrammetric measurement methods in or through the medium of water entails a number of additional challenges:

– accounting for refraction, which is dependent on temperature and salinity, by suitable mathematical modelling of the optical path (see section 4.6);
– degradation of optical image quality due to turbidity and the associated transmission and absorption properties;
– any necessary artificial lighting;
– waterproof housing corresponding to the required diving depth (pressure);
– limited (wireless) data communication under water.

8.6.1 Measurement of water and other fluid surfaces

The measurement of water surfaces, especially waves, is an important precondition for determining the properties of river and tidal currents, ship resistance and other

hydro-mechanical parameters. Due to the reflective and transparent characteristics of water, a normal photogrammetric measurement is practical only if the water surface can be made visible with floating particles. However, the pollution this causes is normally undesirable, and the particles can also disturb the fluid properties. As an alternative, the position and normal vector at a particular location on the surface can be calculated from the distorted image of a three-dimensional point field under the water surface (example in Fig. 3.4). However, this solution only works for clear water and low wave amplitudes.
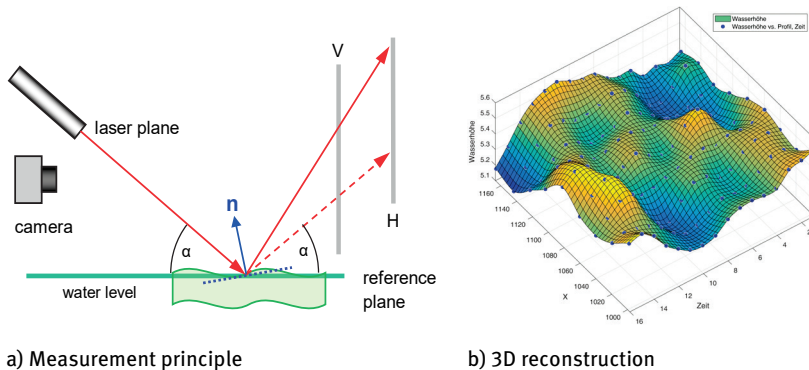
a) Measurement principle   b) 3D reconstruction

Fig. 8.77: Photogrammetric wave measurement (IPF TU Dresden).

Alternatively, Fig. 8.77a illustrates the side view of a photogrammetric set-up where a laser plane is projected at an angle onto the water surface where it is reflected onto two separated projection planes. The front plane (V), nearest to the laser projector, has vertical slits through which the laser plane is projected onto the back plane (H). A camera photographs the projected pattern which is detected in the image and measured to sub-pixel resolution.

Fig. 8.78 shows images for a calm water surface and a moving water surface. The intersections of the reflected laser plane with the slits on the front projection plane correspond to the laser points on the back plane. By connecting corresponding points from front and back plane, a vector is defined which can be intersected with the projected laser plane (see solid lines in Fig. 8.77a). If the geometric configuration of the components is known, including a reference water level (from calibration), the coordinate of points on the water surface can be calculated. If the law of reflection (illustrated by angle α on the reference surface in Fig. 8.77b) is taken into account, the normal vector n can also be calculated (blue vector). Fig. 8.77b shows the interpolated 3D water surface from an image sequence taken at 25 fps. An accuracy of around 1/10 mm was demonstrated empirically.
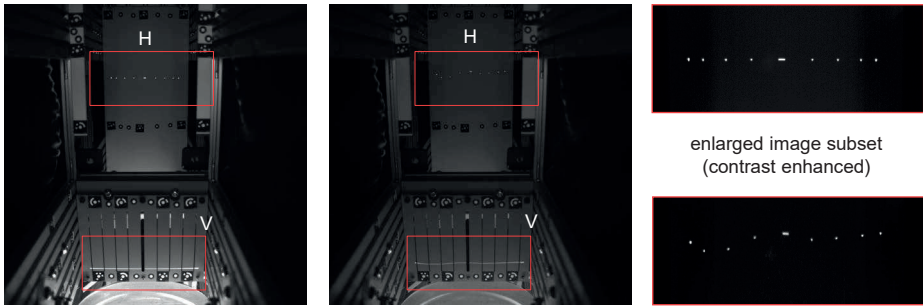
Fig. 8.78: Camera images of a calm surface (left) and with waves (right) (IPF TU Dresden).

8.6.2 Underwater measurement of welds

A continuing growth in offshore construction such as wind farms, as well as expanding port and harbour development, naturally leads to increasing demands for welding inspection and checking, particularly underwater. Current visual inspection by divers is increasingly supported by sensor technologies. Subjective visual inspection is being usefully supplemented by (autonomous) optical 3D measurement tools which will increasingly be adopted in the future. In addition to the general challenges described above, particular problems include:

– resolution of surface point clouds to less than 0.05 mm;
– local accuracy of the order of 0.05 mm;
– weld seam widths between 5 mm and 200 mm;
– small systems to measure complex object shapes;
– restricted depth of field with macro imagery;
– suitable lighting and reduction of reflective hot spots at inhomogeneous surfaces;

Fig. 8.79 shows a monocular system consisting of an industrial camera in an underwater housing with dome port, LED strip illumination and a stationary reference frame which is magnetically attached to the weld. By moving the camera, several camera positions in space can be oriented by bundle adjustment. Previously determined distances on the reference frame, and the previously calibrated parameters of the interior orientation of the optical system, are introduced as constraints.

a) Reference frame above weld
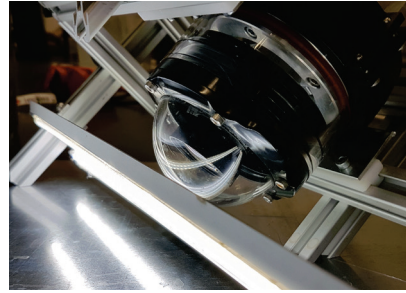
b) Prototype on linear slide

Fig. 8.79: Inspection system for weld seams (IAPG).

The point cloud of the weld seam surface can be calculated by spatial intersections from at least two corresponding image points, which can be determined beforehand by area-based image matching, e.g. LSM (section 5.5.3.5). Local point clouds of partial sections of the seam can then be registered globally using ICP procedures (section 6.9.6.2) to generate a complete surface model. The accuracy and completeness of the point cloud as well as the required computing time differ depending on the number and configuration of the available images, the surface texture and the available reference points.

Fig. 8.80b shows an example of a reconstructed point cloud and the deviations (cloud-to-cloud) against a reference measurement of a 3D fringe projection system. The currently achievable average accuracy, in clear water and laboratory conditions, is approx. 0.05 mm and decreases to approx. 0.11 mm in artificially created, heavily turbid water (approx. 140 mm visibility). The measured 3D data can then be processed in such a way that geometries relevant to analysis, such as the "throat" parameter, can be derived automatically.
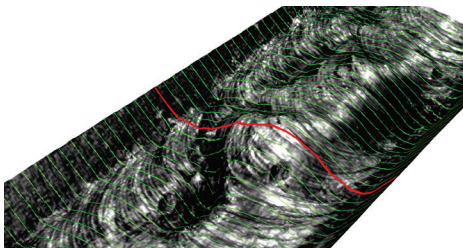
a) Textured model with extracted profile lines
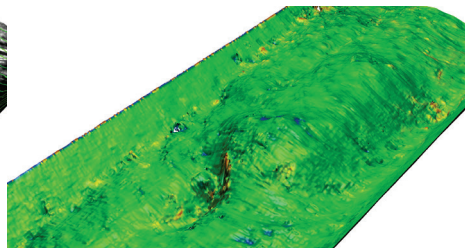
b) Deviations against reference (cloud-to-cloud)

Fig. 8.80: 3D reconstruction of a weld (IAPG).

8.6.3 Underwater photogrammetry of the Costa Concordia shipwreck

A combination of close-range photogrammetry techniques, under and above the water, was used to document and digitally reconstruct the 3D shape of the damaged hull of the Italian cruise ship, Costa Concordia. The vessel partially sank off the coast of Giglio Island in Italy in 2012, coming to rest on her starboard side at an inclination of about 70 degrees to the vertical (Fig. 8.81). The technical investigation of the accident required a special surveying procedure of the damaged part, partially submerged, in a seamless solution including both the underwater and above-water parts. The photogrammetric survey, carried out in June 2012, was designed to provide a metric documentation to sub-centimetre accuracy and resolution when surveying the main leak and smaller cracks in the hulls.

Fig. 8.81: The Costa Concordia resting on her starboard side at the main harbour of Giglio island, Italy (FBK Trento, Uniparthenope, Naples).

An area of damage was surveyed which was 60m long and extended about 6m and 4 m below the sea surface. A Nikon D300 with a 35 mm lens was used above water. for Fhe underwater survey, a camera with a 24 mm lens in a waterproof dome-port housing with flash was used. The recording configuration consisted of two separate blocks above and below water, each containing several strips of parallel and convergent images. Due to the shipwreck position, no control points could be measured. To define the scale, and to link the recorded measurement blocks above and below water, three 2 m scale bars and five 3 m rods were attached to the hull in such a way that they were partially submerged to link the two 3D models obtained from photogrammetry (SfM). Each rod consisted of three plates each with four coded targets whose 3D coordinates were accurately measured beforehand by photogrammetry (Fig. 8.82). In addition, about 500 circular magnetic targets were placed on the hull.
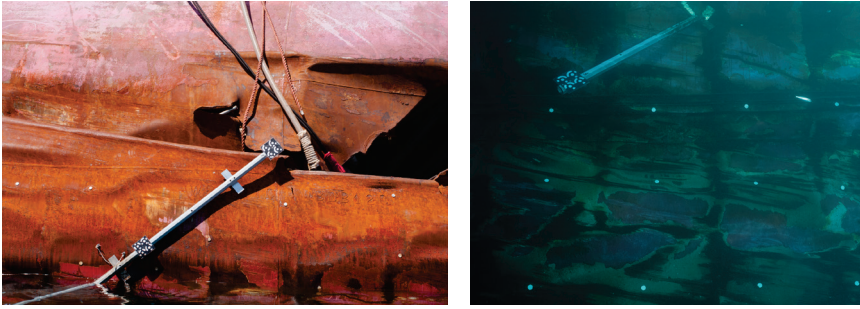
Fig. 8.82: A linking rod as seen in a picture taken above (left) and below (right) the water (FBK Trento, Uniparthenope, Naples).

Seven local 3D models were created (five plates plus two image bundles above and below water) and were registered via a best-fit similarity transformation. An RMS of the spatial residuals between the models of 1.4 mm, with a maximum absolute value of 3.7 mm, were achieved. The transformation parameters produced a seamless 3D model with a resolution better than 5 mm (Fig. 8.83).
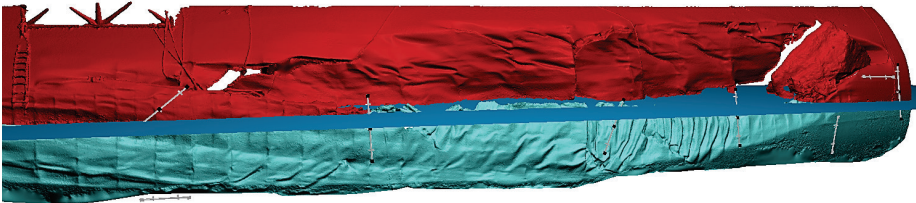
Fig. 8.83: 3D mesh models of the damaged hull of the Costa Concordia obtained using close-range and underwater photogrammetry techniques (FBK Trento, Uniparthenope, Naples).

8.7 Medicine

In medicine, photogrammetry is mainly used to measure parts of the body, for example:
– to prepare and carry out operations (surgical instrument navigation, inserting implants, cutting, drilling and operational access);
– to construct prostheses and implants;
– in plastic surgery;
– in the therapy for bone and spinal deformations;
– in motion studies;
– to monitor growth.

Photogrammetric solutions are applied in many areas of medicine such as orthopaedics, neurosurgery, dentistry and sports medicine.

8.7.1 Surface measurement

Medical surface measurements are mostly characterized by:
– the measurement of unstructured, soft, free-form surfaces;
– the recording of subjects which are moving or not static;
– the absence of permanent reference points on the subject.

Poorly textured, free-form surfaces must be given an artificial structure, e.g. by fringe or pattern projection (see section 6.7.3). Stereoscopic and multiple-image configurations can be chosen for surface measurement. In all cases image measurement must be synchronized to handle movement of the subject.

Fig. 8.84 shows the schematic measurement of a human head from five images. The Face Shape Maxi5 3D system (Polishape 3D) was used for photogrammetric processing. This has a reported accuracy around 0.4 mm, which is considered sufficient for the flexible surface of the human body. Such 3D models can be used to plan operations for plastic surgery, to establish databases of the human face or for biometric face recognition. A similar approach is taken for other body parts.

a) Schematic imaging configuration   b) Calculated 3D model
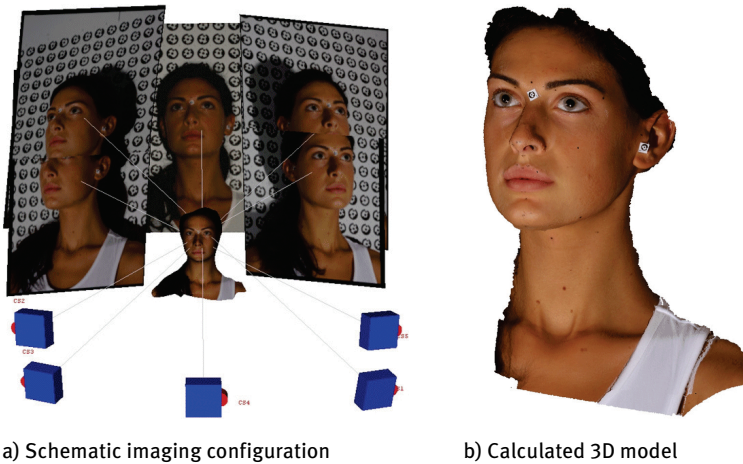
Fig. 8.84: Face reconstruction by multi-image photogrammetry (from Deli et al. 2013).

Features of the photogrammetric process are:
– non-contact measurement avoiding stress to the patient;
– short measurement time (tenths of a second);
– no radiation exposure;

– suitable also for infants and children;
– status documentation at a given time;
– subsequent measurement and diagnostic comparisons over time.

8.7.2 Online navigation systems

Medicine increasingly uses digital, photogrammetric, online measuring systems with contact probing. Typically, the imaging unit is two cameras fixed a mobile housing.

Fig. 8.85 shows examples of such dual camera, online systems with point-by-point tactile probing. These can not only use target adapters with probing tips (see section 6.2.2.2), but also the actual tools used in operations, if equipped with suitable targets (Fig. 8.86). They are a critical component in image-based planning and execution of operations (*image guided surgery*, IGS) where a spatial relationship between patient and surgical instrument must be established. In medicine this process is known as *navigation*.

a) Projection and measurement of a laser point (BrainLab)

b) Online system for use in operation hall (Aesculap)

Fig. 8.85: Dual-camera systems for medical applications.

The key problem for navigation is the unstable position of the patient during the operation. For this reason, local reference target arrays (locators) are attached to the patient whose potential motion can then be continuously monitored by the navigation system. The spatial position of surgical tools or predefined operation data, e.g. related computer tomograms (CT), can be transformed into the coordinate system of the locator. Fig. 8.86b shows an example where a locator fixed to a bone is measured, together with a moving locator mounted on a surgical robot. In this way it is possible to compensate for any motion of the head. Furthermore, there are reference points (landmarks) at the head that also allow spatial reference to a CT.
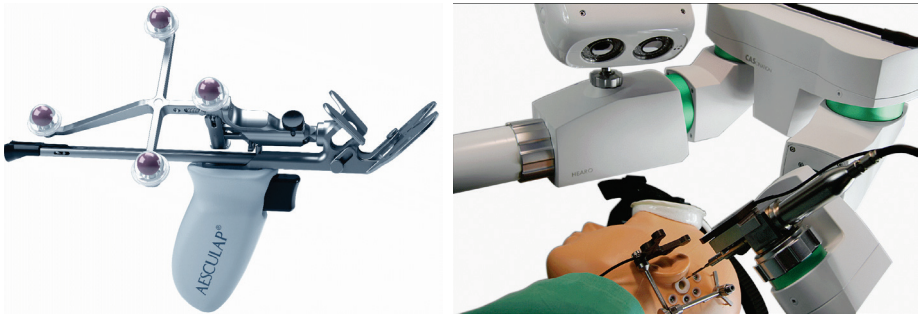
a) Multi-function locator (Aesculap, AXIOS 3D)   b) Surgical robot for cochlea implantation
(Cascination)

Fig. 8.86: Surgical tools with spatial reference points.

Accuracy requirements for navigation systems depend on application and range from around 0.05 mm for cochlea implantation, 0.2-0.5 mm for spinal surgery, around 0.5-1 mm for implanting knee and hip joints and up to 2 mm or so in brain surgery. Since the medical operation requires stereo-tracking cameras to be located some distance away, resulting in height-to-base ratios of 6:1 or greater, the accuracy requirements for the camera measurement technology are very high. These applications require measurement frequencies of 10 to 100 3D measurements per second, with up to 20 points or more simultaneously measured or tracked in 3D space. Motion analysis requires significantly lower accuracies but much higher measurement frequencies (section 6.10.3).

Fig. 8.87: Implementation plan for inserting a hip joint (Plus Orthopedics).

Photogrammetric navigation systems are also employed for pre-operative planning which is usually derived from CT images. Fig. 8.87 shows an example of the planned insertion of a CAD-designed hip joint in a CT image. Using anatomical landmarks, the transformation between CT and real bone can be generated during the operation.

8.8 Miscellaneous applications

8.8.1 Forensic applications

Forensic photogrammetry covers applications such as:
–   traffic accident recording;
–   recording and reconstruction of aircraft crashes;
–   scene-of-crime measurement;
–   estimating the height of criminals;
–   reconstructing bullet trajectories;
–   object reconstruction from amateur images;
–   body scanning of the living and the deceased;
–   detecting environmental pollution from aerial images.

8.8.1.1 Accident recording

Accident recordings are usually carried out by means of laser scanning, and only now in rare cases with photogrammetry. Laser scanning and photogrammetry, also with the use of UAVs, can also be combined here. Since control points cannot usually be provided, the results are given in a local coordinate system. Results can be output as scaled plans, CAD drawings, visibility studies or coordinates of the accident site. In many European countries, photogrammetric measurements are accepted as valid evidence in criminal trials and legal disputes.

A typical measurement situation at an accident site is shown in Fig. 8.88. The accident scene was recorded with a 7 Megapixel amateur camera, the Olympus C7070WZ. The area covered by the camera network and object points have an extent of 80 m and maximum height variation of only 2 m. Imaging restrictions on site lead to largely horizontal and narrow ray intersections which demand particularly robust image orientation methods. The iWitness software (Photometrix) was used to orient the network of 22 images and produce drawings and CAD models. Accuracy in object space was around 3 cm.

In addition to surveying the accident site, issues such as the driver's perspectives are also relevant, Analysing these may require other optical 3D measurement methods, such as fringe projection systems (section 6.7.3) and hand-held scanners (section 6.9.7). Fig. 8.89 shows visibility analyses for a truck driver. Based on the 3D documentation of the truck, as well as measurements relating to the driver, the visibility range and the visibility ratio of the driver before turning (e.g. right in EU, left in UK) can be calculated. In particular, the mirrors (plane or curved) must be modelled with high precision here in order to be able to simulate a correct rear view.

a) Sample measurement images
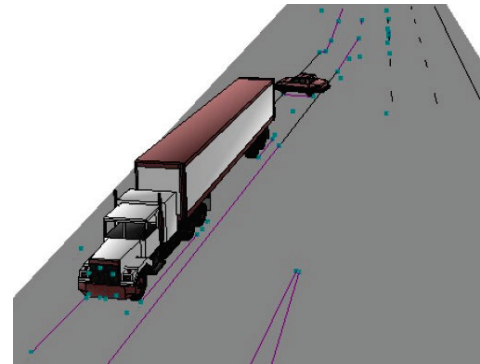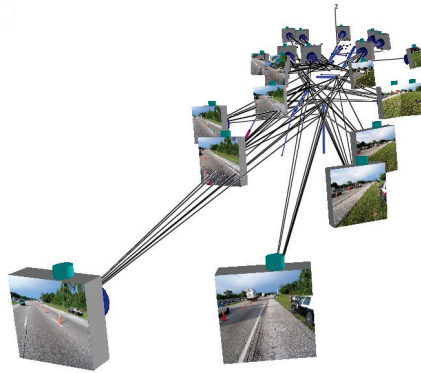
b) Imaging configuration   c) Derived CAD model

Fig. 8.88: Photogrammetric accident recording (Photometrix).
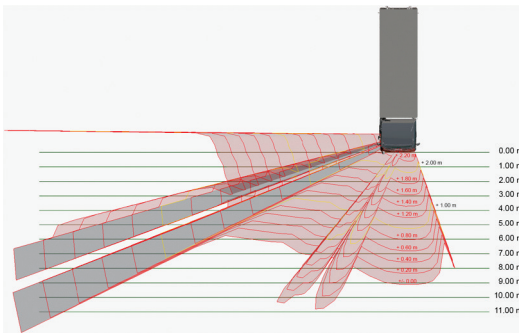
a) Visibility range of the truck driver   b) Visibility before turning right

Fig. 8.89: Visibility analyses (3D Center Zürich).

8.8.1.2 Scene-of-crime recording

The photogrammetric surveying of crime scenes is often now part of the process of securing evidence as it as it enables the scent to be recorded and documented to high accuracy without alterning the surrounding situation.

Fig. 8.90 shows the result of a scene-of-crime reconstruction with full modelling of the local scene (a), overlay of CAD model and photo detail (b), blood spray trajectories (c) and the possible locations of the individuals involved (d). The scene in this case was recorded using a Nikon D700 digital camera. Elcovision 10 (PMS) and AutoCAD (Autodesk) were used for photogrammetric analysis and modelling. Measurement accuracy was in the mm region.

a) Measurement image from crime scene
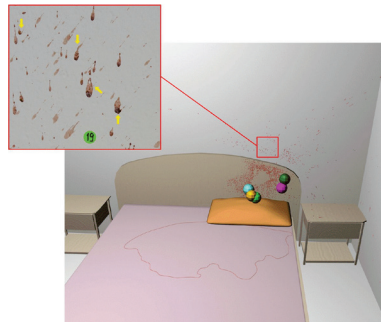
b) 3D model with image detail overlay
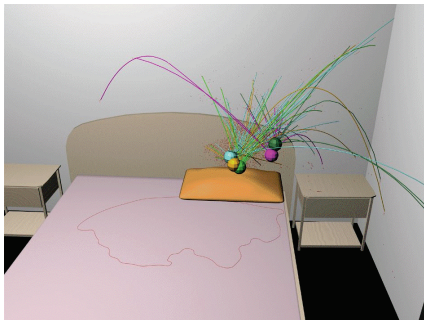
c) Reconstructed blood spray trajectories

d) Reconstructed positions of attacker and victim

Fig. 8.90: Photogrammetric crime record (Institut für Rechtsmedizin, University of Bern).

Fig. 8.91a shows a location surveyed with laser scanning, which was processed and meshed using various software packages. With the involvement of experts from other disciplines, such as ballistics and forensic pathology, the event can then be reconstructed (Fig. 8.91b). The achievable accuracy of approx. 3 mm is usually sufficient for forensic 3D reconstructions. In cases where higher accuracy is required, individual details can additionally be documented using more precise methods.

a) Laser scan of a crime scene   b) Reconstruction of a dynamic event

Fig. 8.91: Example of a crime reconstruction (3D Center Zürich).

8.8.1.3 Anthropological analyses

Photogrammetric 3D data is also used in forensic anthropology. One approach is facial or whole-body documentation using a multi-camera system. This data can be used for a biometrical comparison of body dimensions, especially height, against video surveillance footage of crime suspects. Fig. 8.92a shows the generation of a whole-body model using 70 simultaneously triggered cameras. Fig. 8.92b shows the comparison of the reconstructed 3D body dimensions with a distortion-free computed surveillance video using a 3D laser scan of the environment visible in the video. This procedure does not identify suspects but can exclude them. Facial measurements can be used to compare identifying anthropological features.
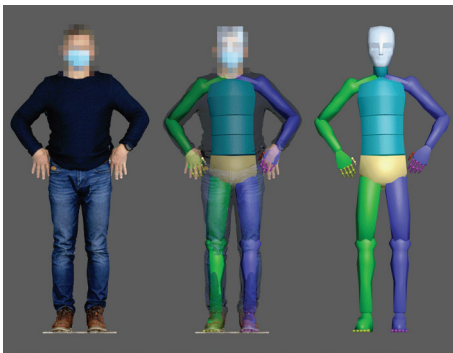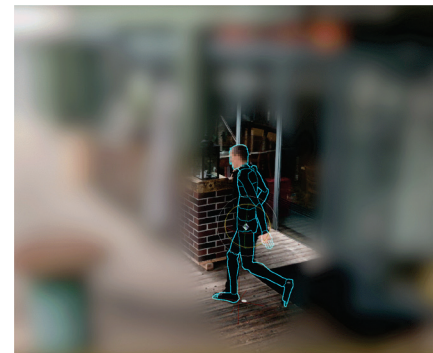
a) Full-body model   b) Comparison with surveillance video

Fig. 8.92: Perpetrator reconstruction (3D Center Zürich).

8.8.2 Scientific applications

Close-range photogrammetry can be advantageously applied in many scientific applications such as:

– recording physical phenomena;
– reconstructing biological processes, e.g. plant and animal growth and form;
– monitoring and modelling events for the earth sciences, e.g. landslides;
– static and dynamic measurement of glaciers;
– etc.

8.8.2.1 Monitoring glacier movements

Another example of using photogrammetric analyses of image sequences is in monitoring glacier movements. The Jakobshavn Isbrae Glacier on the west coast of Greenland had, for many years, a constant speed of 20 m per day, but in more recent years this has increased to 40 m per day, combined with a dramatic retreat of the glacier front. Photogrammetric measurement campaigns were carried out in 2004, 2007 and 2010 with the objective of determining the glacier's spatial and temporal movement patterns using a sequence of terrestrial images taken with a digital camera from a hill.

a) Camera in position   b) Example of a measurement image

Fig. 8.93: Photogrammetric recording of glacier movement (IPF TU Dresden).

The intention is to determine two-dimensional movement vectors showing the glacier movement in the direction of flow, as well as movements in a vertical direction which are induced by tidal effects. The recorded data was a monocular sequence of images made with a digital SLR camera. Images were taken at 30-minute intervals over a period of 12–36 hours. Scale was introduced with the aid of a hybrid geodetic and photogrammetric network. The application of correlation techniques enabled movement and trajectory vectors to be determined with a standard deviation of around 0.1–0.2 % of the trajectory length over a whole day. Accuracy was mainly limited by the difficult topography presented by the glacier surface and the effect of shadows.

Fig. 8.93 shows the camera in position and one of the images. The height changes over a 24-hour period at a point on the glacier near the glacier front, as calculated from the image sequence, are shown in Fig. 8.94. Here the glacier movement shows an almost perfect correlation with the tidal curve. This proves that the glacier front is floating on the fjord. The tidal influence fades away about 1 km behind the front. From this the position of the grounding line can be determined (Fig. 8.95). The results of this terrestrial photogrammetric measurement form the basis of glaciological studies of the Jakobshavn Isbrae Glacier.
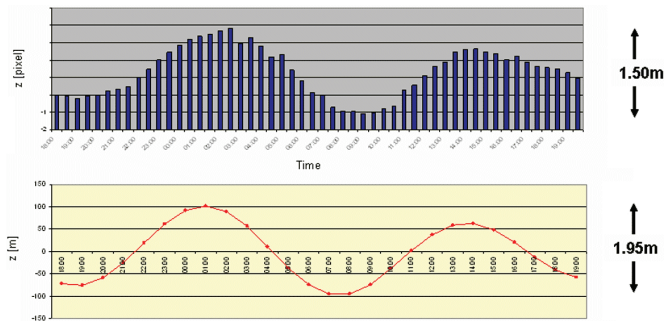
Fig. 8.94: Calculated height changes at a glacier location over 24 hours (upper diagram) and corresponding tidal curve (lower diagram) (IPF TU Dresden).
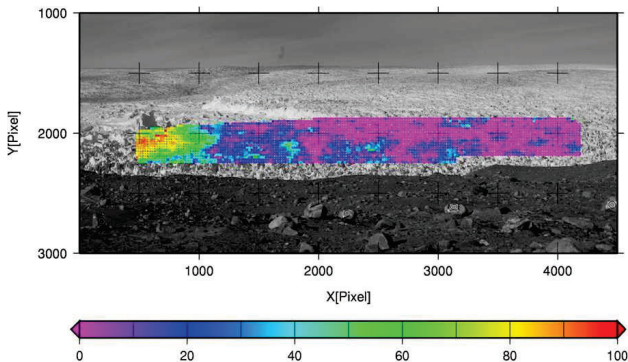
Fig. 8.95: Participation of the vertical glacier movement component with the tidal range (IPG/IPF TU Dresden).

8.8.2.2 Earth sciences

Recording and quantifying processes relevant to the earth sciences demands precise measurement methods. Digital photogrammetry is applied to a wide range of projects at both close and long ranges. Often surfaces must be modelled in order to compare

them at different epochs. This may require the comparison of parameters such as roughness of the ground surface or an absolute comparison in order to quantify erosion and rates of material loss.

For example, photogrammetric methods can be used to monitor changes in ground surface or the evolution of rill network development due to rain. This can either be natural rain or rain simulated in a laboratory. In lab tests, ground samples with an extent of 1×2 m^2 up to 3.7×14.4 m^2 are measured photogrammetrically on a grid with 3×3 mm^2 resolution. For erosion research, comparison of surfaces at different epochs enables conclusions to be made concerning the change in surface roughness and the development and structure of drainage channels. Digital height models can be made in GIS software. Modelling such large areas to millimetre resolution places high demands on the design of the measurement network.

a) Artificial rain generation and location of earth sample   b) Imaging configuration
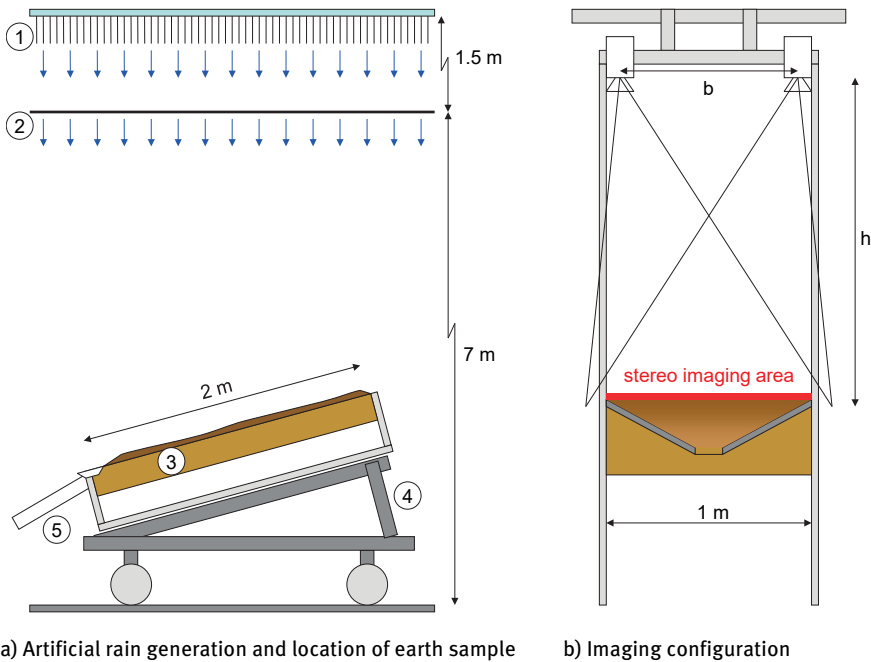
Fig. 8.96: Schematic arrangement of a rainfall test (University of Bern).

Fig. 8.96a shows the schematic arrangement of a rainfall test in the lab. Water drops (1) are formed on the end of capillary tubes from where they drop 1.5 m onto a net (2) in order to create a wider spectrum of drops. The drop from net to the upper edge of the sample box is approximately 6 m. This ensures that the drops reach a velocity corresponding to natural rainfall. The sample box (3) in this test has an area of 1×2 m^2 and is filled with a sample of earth or soil to be tested. Positioning struts (4) can be

used to alter the tilt of the box. The surface water runoff is drained away at the lower end of the box (5). Sediment and water samples are taken here. The surface of the sample is recorded from two camera positions (Fig. 8.96b). The principal points of the cameras are displaced using lenses with fixed shift (see section 3.4.3.6) in order to obtain as large an area of stereo coverage as possible.

Fig. 8.97 shows the surface of the test sample before artificial rain (a), after 20 minutes of simulated rainfall (b) and after 65 minutes (c). It can be see that after 20 minutes the surface runoff has created a drainage pattern which becomes more established with increasing length of rainfall. Natural interaction makes the drainage pattern more efficient with time.

a) Before raining   b) After 20 min   c) After 65 min
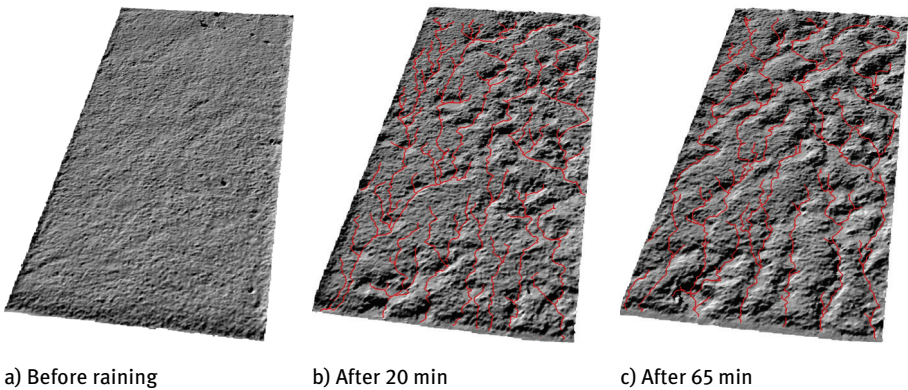
Fig. 8.97: Photogrammetrically generated height model and drainage pattern (University of Bern).

8.8.2.3 Entomology

Insects are particularly challenging objects for photogrammetry due to their small size, detailed structures and often "optically uncooperative" surfaces which might be reflective, semi-transparent or covered in fine hairs. On the other hand, insects often provide dense surface textures, are easy to handle and can be viewed from all sides if suitably mounted. Structure-from-motion can then be used for largely automated modelling (SfM, section 5.5.2.2).

A specific problem to be solved when examining insects is the small depth of field available in the camera's macro range. *Focus stacking* offers a solution in which the camera is moved towards or away from the sample to acquire one focussed plane after the other. Subsequently, image regions of best sharpness are combined to a final image with extended depth of field.
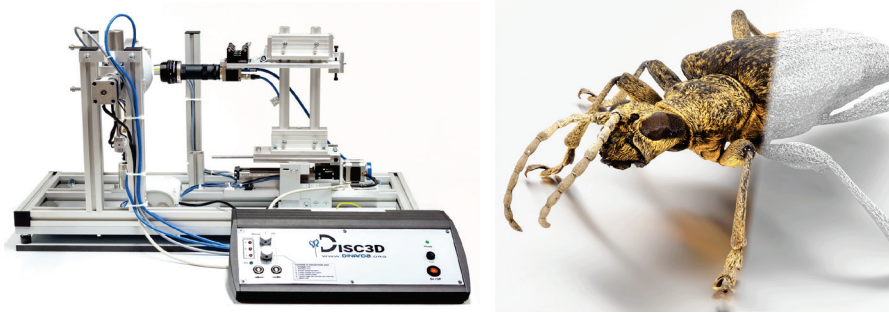
a) Set-up   b) 3D model of an insect

Fig. 8.98: The Darmstadt Insect Scanner (Hochschule Darmstadt / TU Darmstadt).

An important application of 3D modelling by photogrammetry is the exact geometric recording of insect dimensions. Compared with conventional measurement by microscope, this has advantages in easier and more secure handling of the specimens and the digital archiving of results. In addition, physiologically important shape parameters such as volume and surface form can be determined reliably only from 3D models.

Fig. 8.98a shows the "Darmstadt Insect Scanner DISC3D" (Hochschule Darmstadt, TU Darmstadt). The insect is placed on a rotating and tilting mount within two spherical half-shells that provide uniform illumination according to the principle of the integrating sphere (Ulbricht sphere which creates diffuse light). A total of approx. 400 poses are imaged, which provides all-around recording of the insect. Fig. 8.98b shows the 3D model of a scanned insect (black-spotted pincer beetle, Rhagium mordax), which was created using Agisoft Metashape software. The pixel resolution on the object is approx. 5-10 μm.

8.8.2.4 Measurement of a soap bubble

For a record attempt to create the largest free-flying soap bubble in the world, it was required to measure its volume by photogrammetry. Three PCO DIMAX high-speed cameras (Fig. 3.121a) recorded three orthogonal views of the object in synchronized image sequences (example in Fig. 8.99a). Calibration and orientation were calculated on site by means of a field of reference points measured in advance by offline photogrammetry. The images taken at the instant of maximum bubble size were rectified onto best-fit planes which were defined by contour points around the bubble (Fig. 8.99b,c). Within these rectifications, sections separated by small increments were defined and corresponding elliptical areas computed for them. In this way. the bubble's total volume of 13.9 m³ was calculated by integration of the individual ellipse areas (Fig. 8.99d). By comparative measurement of reference bodies of known volume, the accuracy was estimated to about 10 % of the bubble's volume.

a) Original front view
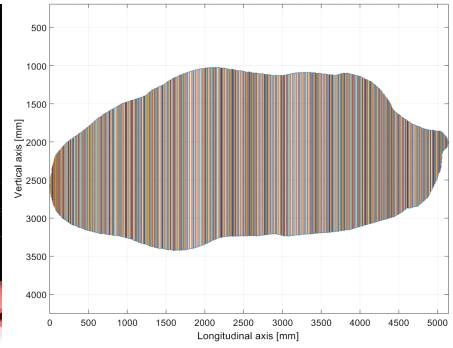
d) Sections from side view
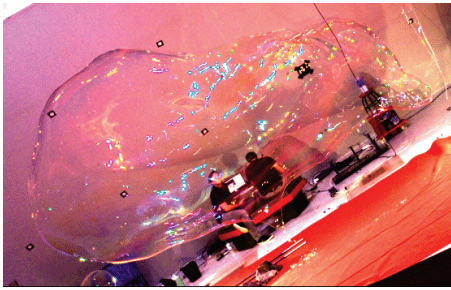
b) Rectified side view
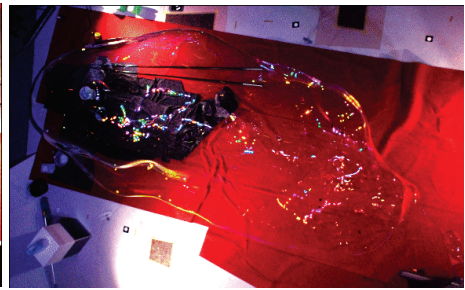
c) Rectified top view

Fig. 8.99: 3D measurement of a soap bubble (IAPG).

9 Literature

This chapter provides an overview of relevant literature on photogrammetry and its neighbouring fields. It does not claim to be complete and serves as a rough orientation in this complex field. Internet sources are not given here. Individual references to citations or illustrations in the text are also listed here.

9.0 Textbooks

9.0.1 Photogrammetry

Atkinson, K.B. (ed.) (2001): Close Range Photogrammetry and Machine Vision. Whittles Publishing, Caithness, UK, 371 p.

Förstner, W., Wrobel, B.P. (2016): Photogrammetric Computer Vision – Statistics, Geometry, Orientation and Reconstruction. Springer International Publishing, Cham, Schweiz, 816 p.

Kasser, M., Egels, Y. (2002): Digital Photogrammetry. Taylor & Francis, London, UK.

Kraus, K. (2007): Photogrammetry – Geometry from Images and Laserscans. 2nd ed., W. de Gruyter, Berlin, 459 p.

Lemmens, M. (2022): Introduction to Pointcloudmetry. Whittles Publishing, Dunbeath, Caithness, 352 p.

Luhmann, T., Robson, S., Kyle, S., Boehm, J. (2020): Close Range Photogrammetry and 3D Imaging.3rd., W. de Gruyter Verlag, Berlin, 822 p.

Luhmann, T. (2023): Nahbereichsphotogrammetrie. 5th ed., Wichmann, Heidelberg, 796 p.

McGlone, J.C. (ed.) (2013): Manual of Photogrammetry. 6th ed., American Society for Photogrammetry and Remote Sensing, 1318 p.

Mikhail, E.M., Bethel, J.S., McGlone, J.C. (2001): Introduction to Modern Photogrammetry. John Wiley & Sons, New York, 479 p.

Wiggenhagen, M., Steensen, T. (2021): Guide for Photogrammetry and Remote Sensing. 6th ed., Wichmann, Heidelberg, 360 p.

9.0.2 Optic, camera and imaging techniques

Arnold, C.R., Rolls, P.J., Stewart, J.C.J. (1971): Applied Photography, Focal Press, London, 510 p.

Born, M., Wolf, E., Bhatia, A.B. (1999): Principles of Optics. Cambridge University Press, 986 p.

Hecht, E. (2003): Optics. Addison Wesley, Longman, 698 p.

Langford, M.J. (1982): Basic Photography. 4th ed., Focal Press, London, 397 p.

Langford, M.J. (1983): Advanced Photography. 4th ed., Focal Press, London, 355 p.

Reinhard, E., Ward, G., Pattanaik, S., Debevec, P. (2006): High Dynamic Range Imaging. Elsevier, Amsterdam, 502 p.

Ray, S.F. (1994): Applied Photographic Optics. 2nd ed., Focal Press, Oxford, 586 p.

Rowlands, A. (2017): Physics of Digital Photography. IOP Publishing, 330 p.

Teubner, U., Brückner, H. J. (2019): Optical Imaging and Photography. De Gruyter, Berlin, 592 p.

9.0.3 Digital image processing, computer vision and pattern recognition

Haralick, R.M., Shapiro, L.G. (1992): Computer and Robot Vision. Vol. I+II, Addison-Wesley, Reading, Ma., USA.

Hartley, R., Zisserman, A. (2013): Multiple View Geometry in computer vision. 10th ed., Cambridge University Press, Cambridge, UK, 655p.

Jähne, B., Haußecker, H., Geißler, P. (1999): Handbook of Computer Vision and Applications. Vol. 1 Sensors and Imaging; Vol. 2 Signal Processing and Pattern Recognition; Vol. 3 Systems and Applications. Academic Press.

Koller, D., Friedman, N. (2009): Probabilistic Graphical Models: Principles and Techniques. MIT Press, Massachusetts.

Osten, W., Reingand, N. (2012): Optical Imaging and Metrology – Advanced Technologies. Wiley-VCH, Berlin, 482 p.

Parker, J.R. (1996): Algorithms for Image Processing and Computer Vision. John Wiley & Sons, 432 p.

Russ, J.C. (2002): The Image Processing Handbook. CRC Press, 752 p.

Szeliski, R. (2022): Computer Vision – Algorithms and Applications. 2nd ed., Springer, Berlin, 925 p.

Wöhler, C. (2013): 3D Computer Vision: Efficient Methods and Applications. 2nd ed., Springer, London, 382 p.

9.0.4 Mathematics and 3D computer graphics

Ahn, S. (2004): Least Squares Orthogonal Distance Fitting of Curves and Surfaces in Space. Springer-Verlag, Berlin.

Brüderlin, B., Meier, A. (2001): Computergrafik und Geometrisches Modellieren. Teubner Verlag, Stuttgart, 312 p.

Foley, J.D., VanDam, A., Feiner, S.K. (1995): Computer Graphics. Addison-Wesley Professional, 1200 p.

Kuipers, J.B. (2002): Quaternions and Rotation Sequences. Princeton University Press, 371 p.

Laga, H., Guo, Y., Tabia, H., Fisher, R. B., Bennamoun, M. (2019): 3D Shape Analysis: Fundamentals, Theory, and Applications. John Wiley & Sons, 368 p.

Shirley, P. (2002): Fundamentals of Computer Graphics. A K Peters, 500 p.

Stroud, K.A. (2001): Engineering Mathematics. 5th ed., Industrial Press, 1236 p.

9.0.5 Least-squares adjustment and statistics

Mikhail, E.M. (1976): Observations and Least Squares. IEP, New York, 497 p.

Mikhail, E.M, Gracie, G. (1981): Analysis & Adjustment of Survey Measurements. Van Nostrand Reinhold Company, 368 p.

Niemeier, W. (2002): Ausgleichungsrechnung. W. de Gruyter Verlag, Berlin, 407 S.

Robert, C. P., Casella, G. (2002): Monte Carlo Statistical Methods. Springer, Berlin.

Wolf, P.R., Ghilani, C.D. (1997): Adjustment Computations: Statistics and Least Squares in Surveying and GIS (Surveying & Boundary Control). John Wiley & Sons.

9.0.6 Industrial, optical 3D metrology, quality control

Brinkmann, B. (2012): International Vocabulary of Metrology, German-English version, 4th ed., Beuth, Berlin, 76 p.

Confalone, G.C., Smits, J., Kinnare, T. (2023): 3D Scanning for Advanced Manufacturing, Design, and Construction. John Wiley & Sons, Hoboken, New Jersey, 224 p.

Gruen, A., Kahmen, H. (eds.) (1989–2009): Optical 3D Measurement Techniques. Wichmann, Heidelberg, since 2001 published by ETH Zurich and TU Vienna.

Luhmann, T., Schumacher, C. (2002–2022) (eds.): Photogrammetrie-Laserscanning-Optische 3D-Messtechnik – Beiträge der Oldenburger 3D-Tage, Wichmann, Heidelberg.

Malacara, D. (ed.) (2007): Optical Shop Testing. John Wiley & Sons, Hoboken, New Jersey, 862 p.

Nüchter, A. (2009): 3D Robotic Mapping. Springer Tracts in Advanced Robotics, Springer, Berlin.

Vosselman, G., Maas, H.-G. (eds.) (2010): Airborne and Terrestrial Laser Scanning. Whittles Publishing, Caithness, UK, 336 p.

9.0.7 Applications

Bentkowska-Kafel, A., MacDonald, L. (eds.) (2017): Digital Techniques for Documenting and Preserving Cultural Heritage. ARC Humanities Press, ITHAKA, New York, 369 p.

Eltner, A., Hoffmeister, D., Kaiser, A., Karrasch, P., Klingbeil, L., Stöcker, C., Rovere, A. (eds.) (2022): UAVs for the environmental sciences – Methods and applications. Wbg Academic, Darmstadt.

Fryer, J.G., Mitchell, H.L., Chandler, J.H. (eds.) (2007): Applications of 3D Measurement from Images. Whittles Publishing, Caithness, UK, 304 p.

Stylianidis, E., Remondino, F. (2016): 3D Recording, Documentation and Management of Cultural Heritage. Whittles Publishing, Caithness, UK, 388 p.

Thali, M., Dirnhofer, R., Vock, P. (eds.). (2009): The Virtopsy Approach: 3D Optical and Radiological Scanning and Reconstruction in Forensic Medicine, CRC Press.

9.1 Chapter 1 – Introduction and history

Ackermann, F., Ebner, H., Klein, H. (1970): Ein Rechenprogramm für die Streifentriangulation mit unabhängigen Modellen. Bildmessung und Luftbildwesen, Heft 4, pp. 206–217.

Adams, L.P. (2001): Fourcade: The centenary of a stereoscopic method of photogrammetric surveying. Photogrammetric Record, 17 (99), pp. 225–242.

Albertz, J. (2009): 100 Jahre Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V. Photogrammetrie-Fernerkundung-Geoinformation, Heft 6, pp. 487–560.

Arthur, D.W.G. (1960): An automatic recording stereocomparator. Photogrammetric Record, 3(16), pp. 298–319.

Atkinson, K.B. (1980): Vivian Thompson (1880–1917): not only an officer in the Royal Engineers. Photogrammetric Record, 10 (55), pp. 5–38.

Atkinson, K.B. (2002): Fourcade: The Centenary – Response to Professor H.-K. Meier. Correspondence, Photogrammetric Record, 17 (99), pp. 555–556.

Brown, D.C. (1958): A solution to the general problem of multiple station analytical stereotriangulation. RCA Data Reduction Technical Report No. 43, Aberdeen.

Brown, D.C. (1976): The bundle adjustment – progress and prospectives. International Archives of Photogrammetry, 21 (3), ISP Congress, Helsinki, pp. 1–33.

Burchardi, B., Meydenbauer, A. (1868): Ueber die Verwendbarkeit der Photographie für Terrain- und Architektur-Aufnahmen. Archiv für die Offiziere der Königlich Preußischen Artillerie- und Ingenieur-Korps 32(1868), 63. Bd., pp. 189–210.

Deville, E. (1895): Photographic Surveying. Government Printing Bureau, Ottawa. 232 p.

Deville, E. (1902): On the use of the Wheatstone Stereoscope in Photographic Surveying. Transactions of the Royal Society of Canada, Ottawa, 8, pp. 63–69.

Förstner, W. (1982): On the geometric precision of digital correlation. International Archives for Photogrammetry and Remote Sensing, Vol. 26/3, pp. 176–189.

Fourcade, H.G. (1903): On a stereoscopic method of photographic surveying. Transactions of the South African Philosophical Society, 14 (1), pp. 28–35.

Fraser, C.S., Brown, D.C. (1986): Industrial photogrammetry – new developments and recent applications. Photogrammetric Record, 12 (68), pp. 256–281.

Fraser, C.S. (1993): A resume of some industrial applications of photogrammetry. ISPRS Journal of Photogrammetry & Remote Sensing, 48(3), pp. 12–23.

Fraser, C.S. (1997): Innovations in automation for vision metrology systems. Photogrammetric Record, 15(90), pp. 901–911.

Grimm, A: (2007): The Origin of the Term Photogrammetry. In Fritsch (ed.): Photogrammetric Week 07, Herbert Wichmann Verlag, Heidelberg, pp. 53–60.

Grimm, A. (2021): Albrecht Meydenbauer: Bauingenieur – Fotograf – Photogrammeter. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science, Vol. 89, Nr. 5, pp. 371–389.

Gruber, O. von, (ed.), McCaw, G.T., Cazalet, F.A., (trans) (1932): Photogrammetry, Collected Lectures and Essays. Chapman & Hall, London.

Gruen, A. (1985): Adaptive least squares correlation – a powerful image matching technique. South African Journal of Photogrammetry, Remote Sensing and Cartography, 14 (3), pp. 175–187.

Gruen, A. (1994): Digital close-range photogrammetry: progress through automation. International Archives of Photogrammetry & Remote Sensing, Melbourne, 30(5), pp. 122–135.

Gruner, H. (1977): Photogrammetry: 1776 – 1976. Photogrammetric Engineering & Remote Sensing, Vol. 43, No. 5, pp. 569–574.

Haggrén, H. (1987): Real-time photogrammetry as used for machine vision applications. Canadian Surveyor, 41 (2), pp. 210–208.

Harley, I.A. (1967): The non-topographical uses of photogrammetry. The Australian Surveyor, 21 (7), pp. 237–263.

Heckes, J., Mauelshagen, L. (eds.) (1987): Luftaufnahmen aus geringer Flughöhe. Veröffentlichungen aus dem Deutschen Bergbaumuseum, Nr. 41, 135 p.

Helava, U.V. (1957): New principle for analytical plotters. Photogrammetria, 14, pp. 89–96.

Hinsken, L. (1989): CAP: Ein Programm zur kombinierten Bündelausgleichung auf Personalcomputern. Zeitschrift für Photogrammetrie und Fernerkundung, Heft 3, pp. 92–95.

Kruck, E. (1983): Lösung großer Gleichungssysteme für photogrammetrische Blockausgleichungen mit erweitertem funktionalem Modell. Dissertation, Wiss. Arbeiten der Fachrichtung Vermessungswesen der Universität Hannover, Nr. 128.

Laussedat, A. (1899): La Métrophotographie. Gauthier-Villars, Paris.

Loser, R., Luhmann, T. (1992): The programmable optical 3-D measuring system POM – applications and performance. International Archives of Photogrammetry and Remote Sensing, Vol. 29, B5, Washington, pp. 533–540.

Luhmann, T., Wester-Ebbinghaus, W. (1986): Rolleimetric RS – A new system for digital image processing. Symposium ISPRS Commission II, Baltimore.

Masry, S.E., Faig, W. (1977): The analytical plotter in close-range applications. Photogrammetric Engineering and Remote Sensing, January 1977.

Meydenbauer, A. (1912): Handbuch der Messbildkunst. Knapp Verlag, Halle, 245 p.

Poivilliers, G. (1961): Address delivered at the opening of the Historical Exhibition., International Archives of Photogrammetry, XIII/1, London.

Pulfrich, C. (1902): Über neuere Anwendungen der Stereoskopie und über einen hierfür bestimmten Stereo-Komparator. Zeitschrift für Instrumentenkunde, 22 (3), pp. 65–81.

Sander, W. (1932): The development of photogrammetry in the light of invention, with special reference to plotting from two photographs. In von Gruber, O. (ed.): McCaw, G. T. & Cazalet, F. A., (Trans.): Photogrammetry, Collected Lectures and Essays. Chapman & Hall, London, pp. 148–246.

Schmid, H. (1956–57): An analytical treatment of the problem of triangulation by stereo-photogrammetry. Photogrammetria, XIII, Nr. 2 and 3.

Thompson, E.H. (1962): Photogrammetry. The Royal Engineers Journal, 76 (4), pp. 432–444.

Thompson, V.F. (1908): Stereo-photo-surveying. The Geographical Journal, 31, pp. 534ff.

Torlegård, K. (1967): On the determination of interior orientation of close-up cameras under operational conditions using three-dimensional test objects. Doctoral Thesis, Kungl, Tekniska Höskolan.

Wester-Ebbinghaus, W. (1981): Zur Verfahrensentwicklung in der Nahbereichsphotogrammetrie. Dissertation, Universität Bonn.

Wheatstone, C. (1838): Contribution to the physiology of vision – Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. Philosophical Transactions of the Royal Society of London for the year MDCCCXXXVIII, Part II, pp. 371–94.

9.2 Chapter 2 – Mathematical fundamentals

9.2.1 Transformations and geometry

Dam, E.B., Koch, M., Lillholm, M. (1998): Quaternions, Interpolation and Animation. Technical Report DIKU-TR-98/5, Department of Computer Science, University of Copenhagen. → 2.2.2

Forbes, A.B. (1989): Least-squares best-fit geometric elements. National Physical Laboratory, Report DITC 140/89, Teddington, United Kingdom. → 2.2.4.3

Förstner, W. (2013): Graphical models in geodesy and photogrammetry. Photogrammetrie–Fernerkundung–Geoinformation, Heft 4, pp. 255–267. → 2

Förstner, W., Wrobel, B. (2013): Mathematical concepts in photogrammetry. In McGlone (ed.): Manual of Photogrammetry, 6th ed., pp. 63–233. → 2.2.3

Hinsken, L. (1987): Algorithmen zur Beschaffung von Näherungswerten für die Orientierung von beliebig im Raum angeordneten Strahlenbündeln. Deutsche Geodätische Kommission, Reihe C, Nr. 333. → 2.2.2, → 2.4.5

Horn, B.K. (1987): Closed-form solution of absolute orientation using unit quaternions. Journal of the Optical Society of America A, 4 (4), pp. 629–642. → 2.2.2.2

Lösler, M., Nitschke, M. (2010): Bestimmung der Parameter einer Regressionsellipse in allgemeiner Raumlage. Allgemeine Vermessungs-Nachrichten, Heft 3, pp. 113–117. → 2.3.1.3, → 2.3.2.3

9.2.2 Adjustment techniques

Ackermann, F., Förstner, W., Klein, H., Schrot, R., van Mierlo, J. (1980): Grobe Datenfehler und die Zuverlässigkeit der photogrammetrischen Punktbestimmung. Seminar, Universität Stuttgart. → 2.4.4

Baarda, W. (1968): A testing procedure for use in geodetic networks. Niederländische Geodätische Kommission, New Series 2, Nr. 5, Delft. → 2.4.4

Förstner, W., Wrobel, B. (2013): Mathematical concepts in photogrammetry. In McGlone (ed.): Manual of Photogrammetry, 6th ed., pp. 63–233. → 2.4

Förstner, W., Wrobel, B., Paderes, F., Fraser, C.S., Dolloff, J., Mikhail, E.M., Rujikietgumjorn, W. (2013): Analytical photogrammetric operations. In McGlone (ed.): Manual of Photogrammetry, 6th ed., pp. 785–955. → 2.4

Förstner, W., Wrobel, B.P. (2016): Photogrammetric Computer Vision – Statistics, Geometry, Orientation and Reconstruction. Springer International Publishing, Cham, Schweiz, 816 p. → 2.4

Holland, P.W., Welsch, R.E. (1977): Robust regression using iteratively reweighted least-squares. Communications in Statistics – Theory and Methods, 6(9), pp. 813–827. → 2.4

Huber, P.J. (1973): Robust regression: asymptotics, conjectures, and Monte Carlo. The Annals of Statistics, Vol. 1, No. 5, pp. 799–821. → 2.4, → 7.1.4

Hinsken, L. (1989): CAP: Ein Programm zur kombinierten Bündelausgleichung auf Personalcomputern. Zeitschrift für Photogrammetrie und Fernerkundung, Heft 3, pp. 92–95. → 2.4

Jorge, J.M. (1977): The Levenberg-Marquardt algorithm: implementation and theory. In Watson (ed.): Numerical Analysis, Lecture Notes Math, 630, pp. 105–116. → 2.4.2.3

Klein, H. (1984): Automatische Elimination grober Datenfehler im erweiterten Blockausgleichungsprogramm PAT-M. Bildmessung und Luftbildwesen, Heft 6, pp. 273–280. → 2.4.4

Kruck, E. (1983): Lösung großer Gleichungssysteme für photogrammetrische Blockausgleichungen mit erweitertem funktionalem Modell. Dissertation, Wiss. Arbeiten der Fachrichtung Vermessungswesen der Universität Hannover, Nr. 128. → 2.4, → 2.4.4, → 4.4

Kruck, E. (1995): Balanced least squares adjustment for relative orientation. In Gruen/Kahmen (ed.): Optical 3-D Measurement Techniques III, Wichmann, Heidelberg, pp. 486–495. → 2.4

Levenberg, K. (1944): A method for the solution of certain problems in least squares. The Quarterly of Applied Mathematics, 2, pp. 164–168. → 2.4.2.3

Markovsky, I., Van Huffel, S. (2007): Overview of total least-squares methods. Signal processing, 87(10), pp. 2283–2302. → 2.4

Marquardt, D. (1963): An algorithm for least-squares estimation of nonlinear parameters. SIAM Journal on Applied Mathematics, 11, pp. 431–441. → 2.4.2.3

Niemeier, W. (2002): Ausgleichungsrechnung. W. de Gruyter Verlag, Berlin, 407 p. → 2.4

9.3 Chapter 3 – Imaging technology

9.3.1 Optics and sampling theory

Ernst, D. (2017): Polarisationsbildgebung. In Sackewitz (ed.): Handbuch zur Industriellen Bildverarbeitung – Qualitätssicherung in der Praxis, 3. Aufl., Fraunhofer Verlag, Stuttgart, pp. 100–104. → 3.1.1.5

Schmidt, F. (1993): 2D-Industriemeßtechnik mit CCD-Sensoren. Zeitschrift für Photogrammetrie und Fernerkundung, Heft 2, pp. 62–70. → 3.4.3.7

9.3.2 Camera modelling and calibration

Barazzetti, L., Mussio, L., Remondino, F., Scaioni, M. (2011): Targetless camera calibration. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38 (5/W16), on CD-ROM. → 3.3.3

Beyer, H. (1992): Geometric and radiometric analysis of a CCD-camera based photogrammetric close-range system. Dissertation, Mitteilungen Nr. 51, Institut für Geodäsie und Photogrammetrie, ETH Zürich. → 3.3.1, → 3.4.1

Brown, D.C. (1971): Close-range camera calibration. Photogrammetric Engineering, 37(8), pp. 855–866. → 3.3.3

Clarke, T.A., Fryer, J.G. (1998): The development of camera calibration methods and models. Photogrammetric Record, 16(91), pp. 51–66. → 3.3.3

Conrady, A. (1919): Decentered lens systems. Royal Astronomical Society, Monthly Notices, Vol. 79, pp. 384–390. → 3.3.3.2

Cronk, S., Fraser, C.S. Hanley, H.B. (2006): Automatic calibration of colour digital cameras. Photogrammetric Record, 21(116), pp. 355–372. → 3.3.3, → 3.3.4.2, → 3.3.4.6, → 3.4.1.6

Ebner, H. (1976): Self calibrating block adjustment. Bildmessung und Luftbildwesen, 44, pp. 128–139. → 3.3.4.2

Fraser, C.S. (1980): Multiple focal setting self-calibration of close-range metric cameras. Photogrammetric Engineering and Remote Sensing, 46(9), pp. 1161–1171. → 3.3.4

Fraser, C.S., Shortis, M. (1992): Variation of distortion within the photographic field. Photogrammetric Engineering and Remote Sensing, 58 (6), pp. 851–855. → 3.3.3, → 3.3.4

Fraser, C.S. (1997): Digital camera self-calibration. ISPRS International Journal of Photogrammetry & Remote Sensing, Vol. 52, pp. 149–159. → 3.3.3

Fraser, C.S., Cronk, S., Stamatopoulos, C. (2012): Implementation of zoom-dependent camera calibration in close-range photogrammetry. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 39 (B5), pp. 15–19. → 3.3.4

Fraser, C.S. (2013): Automatic camera calibration in close-range photogrammetry. Photogrammetric Engineering & Remote Sensing, 79(4), pp. 381–388. → 3.3.3

Fryer, J.G., Brown, D.C. (1986): Lens distortion for close-range photogrammetry. Photogrammetric Engineering & Remote Sensing, 52 (1), pp. 51–58. → 3.3.3, → 3.3.4

Fryer, J.G. (1996): Camera calibration. In Atkinson (ed.): Close Range Photogrammetry and Machine Vision, Whittles Publishing, Caithness, UK, pp. 156–179. → 3.3.2, → 3.3.3, → 3.3.4

Grün, A. (1978): Experiences with self-calibrating bundle adjustment. ACSM-ASP Annual Meeting, Washington D.C. → 3.3.2.4

Gruen, A., Huang, T.S. (2001): Calibration and Orientation of Cameras in Computer Vision. Springer, Berlin/Heidelberg. → 3.3.3

Kotowski, R. Weber, B. (1984): A procedure for on-line correction of systematic errors. International Archives of Photogrammetry and Remote Sensing, Vol. 25(3a), pp. 553–560. → 3.3.4.5

Luhmann, T. (2005): Zum photogrammetrischen Einsatz von Einzelkameras mit optischer Stereostrahlteilung. Photogrammetrie-Fernerkundung-Geoinformation, Heft 2, pp. 101–110. → 3.2.2.2, → 3.4.3.9, → 3.5.3

Luhmann, T., Fraser, C., Maas, H.-G. (2015): Sensor modelling and camera calibration for close-range photogrammetry. ISPRS Journal of Photogrammetry and Remote Sensing, November 2015, pp. 37–46. → 3.3

Luhmann, T., Piechel, J., Roelfs, T. (2013): Geometric calibration of thermographic cameras. In Kuenzer/Dech (eds.), Thermal Infrared Remote Sensing: Sensors, Methods, Applications, Springer-Verlag, Berlin, pp. 27–42. → 3.5.8

Maas, H.-G. (1999): Ein Ansatz zur Selbstkalibrierung von Kameras mit instabiler innerer Orientierung. Publikationen der DGPF, Band 7, München 1998. → 3.3.3.4

Piechel, J., Jansen, D., Luhmann, T. (2010): Kalibrierung von Zoom- und Shift-/Tilt-Objektiven. In Luhmann/Müller (ed.): Photogrammetrie-Laserscanning-Optische 3D-Messtechnik, Oldenburger 3D-Tage 2010, Wichmann, Heidelberg, pp. 284–291. → 3.4.3.6

Pollefeys, M., Koch, R., Van Gool, L. (1999): Self-calibration and metric reconstruction inspite of varying and unknown internal camera parameters. International Journal of Computer Vision, 32 (1), pp. 7–25. → 3.3.4.4

Remondino, F., Fraser, C.S. (2006): Digital camera calibration methods: considerations and comparisons. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36/5, pp. 266–272. → 3.3

Richter, K., Mader, D., Seidl, K., Maas, H.-G. (2013): Development of a geometric model for an all-reflective camera system. ISPRS Journal of Photogrammetry and Remote Sensing, 86, pp. 41–51. → 3.4.3.8

Robson, S., Shortis, M.R., Ray, S.F. (1999): Vision metrology with super wide angle and fisheye optics. Videometrics VI, SPIE Volume 3641, pp. 199–206. → 3.3.3, → 3.3.4, → 3.3.7, → 3.5.2

Scaramuzza, D. (2012): Omnidirectional camera. In Ikeuchi (ed.): Encyclopedia of Computer Vision, Springer-Verlag, Berlin. → 3.3.7

Schneider, D. (2008): Geometrische und stochastische Modelle für die integrierte Auswertung terrestrischer Laserscannerdaten und photogrammetrischer Bilddaten. Dissertation, Technische Universität Dresden. → 3.3.4, → 3.4.3.4, → 3.5.6

Schneider, D., Maas, H.-G. (2003): Geometric modelling and calibration of a high resolution panoramic camera. In Grün/Kahmen (eds.): Optical 3D Measurement Techniques VI. Vol. II, Wichmann, Heidelberg, pp. 122–129. → 3.3.4, → 3.4.3.4, → 3.5.6

Schwalbe, E. (2005): Geometric modelling and calibration of fisheye lens camera systems. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36, Part 5/W8. → 3.3.7, → 3.4.3.3

Shortis, M.R., Bellman, C.J., Robson, S., Johnston, G.J., Johnson, G.W. (2006): Stability of zoom and fixed lenses used with digital SLR cameras. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36/5, pp. 285–290. → 3.3.3, → 3.3.4, → 3.5.2

Stamatopoulos, C., Fraser, C.S. (2011): Calibration of long focal length cameras in close-range photogrammetry. Photogrammetric Record, 26(135), pp. 339–360. → 3.3.4

Tecklenburg, W., Luhmann, T., Hastedt, H. (2001): Camera modelling with image-variant parameters and finite elements. In Grün/Kahmen (eds.): Optical 3D Measurement Techniques V, Wichmann, Heidelberg. → 3.3.3, → 3.3.4, → 3.5.2

Tsai, R.Y. (1986): An efficient and accurate camera calibration technique for 3-D machine vision. Proc. International Conference on Computer Vision and Pattern Recognition, Miami Beach, USA, pp. 364–374. → 3.3.4.1

Wester-Ebbinghaus, W. (1981): Zur Verfahrensentwicklung in der Nahbereichsphotogrammetrie. Dissertation, Universität Bonn. → 3.3., → 6.2

Wester-Ebbinghaus, W. (1985): Verfahren zur Feldkalibrierung von photogrammetrischen Aufnahmekammern im Nahbereich. In Kupfer/Wester-Ebbinghaus (eds.): Kammerkalibrierung in der photogrammetrischen Praxis, Deutsche Geodätische Kommission, Reihe B, Heft Nr. 275, pp. 106–114. → 3.3

Zhang, Z. (2000): A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 11, pp. 1330–1334. → 3.3.4

9.3.3 Sensors and cameras

Beyer, H. (1992): Geometric and radiometric analysis of a CCD-camera based photogrammetric close-range system. Dissertation, Mitteilungen Nr. 51, Institut für Geodäsie und Photogrammetrie, ETH Zürich. → 3.3.1, → 3.3.2, → 3.4.1

Brown, J., Dold, J. (1995): V-STARS – A system for digital industrial photogrammetry. In Gruen/Kahmen (eds.): Optical 3D Measurement Techniques III, Wichmann, Heidelberg, pp. 12–21. → 3.2.1, → 3.5.2

Dold, J. (1997): Ein hybrides photogrammetrisches Industriemeßsystem höchster Genauigkeit und seine Überprüfung. Dissertation, Heft 54, Schriftenreihe Studiengang Vermessungswesen, Universität der Bundeswehr, München. → 3.2.1.1, → 3.3.1, → 3.3.3, → 3.3.4.3, → 6.2

El-Hakim, S. (1986): Real-time image metrology with CCD cameras. Photogrammetric Engineering & Remote Sensing, 52 (11), pp. 1757–1766. → 3.3.2, → 3.3.3, → 3.4.1, → 3.5.1

Hauschild, R. (1999): Integrierte CMOS-Kamerasysteme für die zweidimensionale Bildsensorik. Dissertation, Universität Duisburg. → 3.4.1.3

Holst, G. (2014): Scientific CMOS camera technology: A breeding ground for new microscopy techniques. Microscopy and Analysis, January 2014, Digital Cameras and Image Analysis Supplement, pp. 4–10. → 3.4.1, → 3.4.1.3, → 3.5.3

Jähne, B. (2017): Der EMVA 1288 Standard: Mehr als objektive Kameracharakterisierung. Jahrestagung der Initiative Bildverarbeitung e.V. → 3.4.1.9

Läbe, T., Förstner, W. (2004): Geometric stability of low-cost digital consumer cameras. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 35 (1), pp. 528–535. → 3.3.2, → 3.5.2

Lenz, R., Fritsch, D. (1990): On the accuracy of videometry with CCD-sensors. International Journal for Photogrammetry and Remote Sensing (IJPRS), 45, pp. 90–110. → 3.1.5, → 3.1.6, → 3.3.4.1, → 3.4.1, → 3.5.5.1

Lenz, R., Beutlhauser, R., Lenz, U. (1994): A microscan/macroscan 3x12 bit digital color CCD camera with programmable resolution up to 20992 x 20480 picture elements. International Archives of Photogrammetry and Remote Sensing, Vol. 30/5, Melbourne, pp. 225–230. → 3.5.5.1

Luhmann, T., Tecklenburg, W. (2002): Bundle orientation and 3D object reconstruction from multiple-station panoramic imagery. ISPRS Symposium Comm. V, Korfu. → 3.5.6

Richter, K., Mader, D., Seidl, K., Maas, H.-G. (2013): Development of a geometric model for an all-reflective camera system. ISPRS Journal of Photogrammetry and Remote Sensing, 86, pp. 41–51. → 3.4.3.8

Robson, S., Shortis, M.R., Ray, S.F. (1999): Vision metrology with super wide angle and fisheye optics. Videometrics VI, SPIE Vol. 3641, pp. 199–206. → 3.4.3.4

Robson, S., Shortis, M.R. (1998): Practical influences of geometric and radiometric image quality provided by different digital camera systems. Photogrammetric Record, 16(92), pp. 225–248. → 3.5.2

Schmidt, F. (1993): 2D-Industriemeßtechnik mit CCD-Sensoren. Zeitschrift für Photogrammetrie und Fernerkundung, Heft 2, pp. 62–70. → 3.4.1, → 3.4.3.7

Wester-Ebbinghaus, W. (1989): Das Réseau im photogrammetrischen Bildraum. Zeitschrift für Photogrammetrie und Fernerkundung, Heft 3, pp. 3–10. → 3.3.2

9.3.4 Target materials and illumination

Coules, H.E., Orrock, P.J., Seow, C.E. (2019): Reflectance transformation imaging as a tool for engineering failure analysis. Engineering Failure Analysis 105(2019), pp. 1006–1017. → 3.6.3.5

Frank, E., Heath, S., Stein, C. (2021): Integration of photogrammetry, reflectance transformation imaging (RTI), and multiband imaging (MBI) for visualization, documentation, and analysis of archaeological and related materials. ISAW Papers 21. → 3.6.3.5

Heist, S., Mann, A., Kühmstedt, P., Schreiber, P., Notni, G. (2014): Array projection of aperiodic sinusoidal fringes for high-speed three-dimensional shape measurement. Opt. Eng., 53 (11), 112208. → 3.6.3.2

Mulsow, C. (2007): Ein photogrammetrisches Verfahren zur Kalibrierung eines Beamers. In Luhmann/Müller (ed.): Photogrammetrie-Laserscanning-Optische 3D-Messtechnik, Oldenburger 3D-Tage 2007, Wichmann, Heidelberg, pp. 58–67. → 3.6.3.2

Phong, B.T. (1975): Illumination for computer generated pictures. Comm. ACM, 18 (8), pp. 311–317. → 3.6.1

Robson, S., MacDonald, L., Sargeant, B. (2020): Metrology enabled reflection transformation imaging to reconstruct local detail in manufactured surfaces. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B2-2020, pp. 797–804. → 3.6.3.5

Wissmann, P., Forster, F., Schmitt, R. (2011): Fast and low-cost structured light pattern sequence projection, Opt. Express, 19 (24), pp. 24657–24671. → 3.6.3.2

Woodham, R.J., (1980): Photometric method for determining surface orientation from multiple images. Optical Engineering, 19(1), pp. 191139. → 3.6

Zheng, Q., Kumar, A., Shi, B., Pan, G. (2019): Numerical reflectance compensation for non-lambertian photometric stereo. IEEE Transactions on Image Processing, 28(7), pp. 3177–3191. → 3.6

9.4 Chapter 4 – Analytical methods

9.4.1 Analytical photogrammetry

Abdel-Aziz, Y., Karara, H.M. (1971): Direct linear transformation from comparator coordinates into object space coordinates in close range photogrammetry. ASP Symposium on Close-Range Photogrammetry. → 4.2.4.1

Andresen, K. (1991): Ermittlung von Raumelementen aus Kanten im Bild. Zeitschrift für Photogrammetrie und Fernerkundung, Nr. 6, pp. 212–220. → 4.4.7.2

Beder, C., Förstner, W. (2006): Direct solutions for computing cylinders from minimal sets of 3D points. In Leonardis et al. (eds.): Proceedings of the European Conference on Computer Vision, Springer, pp. 135–146. → 4.4.7.2

Brown, D.C. (1958): A solution to the general problem of multiple station analytical stereotriangulation. RCA Data Reduction Technical Report No. 43, Aberdeen. → 4.4

Förstner, W., Wrobel, B. (2013): Mathematical concepts in photogrammetry. In McGlone (ed.): Manual of Photogrammetry, 6th ed., pp. 63–233. → 4.2.4, → 4.3.4

Förstner, W., Wrobel, B., Paderes, F., Fraser, C.S., Dolloff, J., Mikhail, E.M., Rujikietgumjorn, W. (2013): Analytical photogrammetric operations. In McGlone (ed.): Manual of Photogrammetry, 6th ed., pp. 785–955. → 4.2, → 4.3, → 4.4

Förstner, W., Wrobel, B.P. (2016): Photogrammetric Computer Vision – Statistics, Geometry, Orientation and Reconstruction. Springer International Publishing, Cham, Schweiz, 816 p. → 4.2.4, → 4.3.4

Fraser, C.S. (2006) Evolution of network orientation procedures. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Dresden, 35(5), pp. 114–120. → 4.4

Haralick, R.M., Lee, C., Ottenberg, K., Nolle, M. (1994): Review and analysis of solutions of the three point perspective pose estimation problem. International Journal of Computer Vision, 13 (3), pp. 331–356. → 4.2.3, → 4.2.4

Hartley, R.I. (1997): In defense of the eight-point algorithm. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 6, pp. 580–593. → 4.3.3

Horn, B.K. (1987): Closed-form solution of absolute orientation using unit quaternions. Journal of the Optical Society of America A, 4(4), pp. 629–642. → 4.2.3

Kyle, S.A. (1990): A modification to the space resection. Allgemeine Vermessungs-Nachrichten, International Edition, Heft 7, pp. 17–25. → 4.2.3

Läbe, T., Förstner, W. (2005): Erfahrungen mit einem neuen vollautomatischen Verfahren zur Orientierung digitaler Bilder. Publikationen der DGPF, Band 14, pp. 271–278. → 4.2.4, → 4.3.4

Loop, C., Zhang, Z. (1999): Computing rectifying homographies for stereo vision. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Fort Collins, CO, Vol. 1, pp. 125–131. → 4.3.3

Luhmann, T. (2009): Precision potential of photogrammetric 6 DOF pose estimation with a single camera. ISPRS Journal of Photogrammetry and Remote Sensing, 64(3), pp. 275–284. → 4.2.5

Mayer, H. (2007): Automatische Orientierung mit und ohne Messmarken – Das Mögliche und das Unmögliche. Publikationen der DGPF, Band 16, pp. 457–464. → 4.4

Nister, D. (2004): An efficient solution to the five-point relative pose problem. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, No. 6, pp. 756–770. → 4.3.3

Sanso, F. (1973): An exact solution of the roto-translation problem. Photogrammetria, 29(6), pp. 203–216. → 4.2.3

Schwermann, R., Effkemann, C. (2002): Kombiniertes Monoplotting in Laserscanner- und Bilddaten mit PHIDIAS. In Luhmann (ed.): Photogrammetrie und Laserscanning, Wichmann Verlag, Heidelberg, pp. 57–70. → 4.2.7

Thompson, E.H. (1959): A rational algebraic formulation of the problem of relative orientation. The Photogrammetric Record, 3 (14), pp. 152–159. → 4.3.3

Thompson, E.H. (1966): Space resection: failure cases. The Photogrammetric Record, 5 (27), pp. 201–204. → 4.3.3

van den Heuvel, F.A. (1997): Exterior orientation using coplanar parallel lines. Proc. 10th Scandinavian Conference on Image Analysis, Lappeenranta, pp. 71–78. → 4.2.3

Wrobel, B. (1999): Minimum solutions for orientation. In Huang/Gruen (eds.): Calibration and Orientation of Cameras in Computer Vision, Washington D.C., Springer, Heidelberg. → 4.2.3

9.4.2 Bundle adjustment

Agarwal, S., Snavely, N., Seitz, S. M., Szeliski, R. (2011): Bundle adjustment in the large. European Conference on Computer Vision, Springer, pp. 29–42. → 4.4

Börlin, N., Murtiyoso, A., Grussenmeyer, P. (2020): Efficient computation of posterior covariance in bundle adjustment in DBAT for projects with large number of object points. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B2-2020, pp. 737–744. → 4.4

Brown, D.C. (1976): The bundle adjustment – progress and prospectives. International Archives of Photogrammetry, 21 (3), ISP Congress, Helsinki, pp. 1–33. → 4.4

El-Hakim, S.F., Faig, W. (1981): A combined adjustment of geodetic and photogrammetric observations. Photogrammetric Engineering and Remote Sensing, 47(1), pp. 93–99. → 4.4.2.3

Fraser, C.S. (1982): Optimization of precision in close-range photogrammetry. Photogrammetric Engineering and Remote Sensing, 48(4), pp. 561–570. → 4.4.5

Fraser, C.S. (1996): Network design. In Atkinson (ed.): Close Range Photogrammetry and Machine Vision, Whittles Publishing, Caithness, UK, pp. 256–281. → 4.4.6

Hastedt, H. (2004): Monte-Carlo simulation in close-range photogrammetry. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 35, Part B5, pp. 18–23. → 4.4.6.1

Hinsken, L. (1987): Algorithmen zur Beschaffung von Näherungswerten für die Orientierung von beliebig im Raum angeordneten Strahlenbündeln. Dissertation, Deutsche Geodätische Kommission, Reihe C, Nr. 333. → 4.4.4

Hunt, R.A. (1984): Estimation of initial values before bundle adjustment of close range data. International Archives of Photogrammetry and Remote Sensing, 25 (5), pp. 419–428. → 4.4.4

Jorge, J.M. (1977): The Levenberg-Marquardt algorithm: implementation and theory. In Watson (ed.): Numerical Analysis, Lecture Notes Math. 630, pp. 105–116. → 2.4.2.3, → 4.4

Kruck, E. (1983): Lösung großer Gleichungssysteme für photogrammetrische Blockausgleichungen mit erweitertem funktionalen Modell. Dissertation, Universität Hannover. → 4.4

Levenberg, K. (1944): A method for the solution of certain problems in least squares. The Quarterly of Applied Mathematics, 2, pp. 164–168. → 2.4.2.3, → 4.4

Lourakis, M.I., Argyros, A.A. (2009): SBA: A software package for generic sparse bundle adjustment. ACM Transactions on Mathematical Software (TOMS), 36(1), pp. 1–30. → 4.4

Marquardt, D (1963): An algorithm for least-squares estimation of nonlinear parameters. SIAM Journal on Applied Mathematics, 11, pp. 431–441. → 2.4.2.3, → 4.4

Tecklenburg, W., Luhmann, T., Hastedt, H. (2001): Camera modelling with image-variant parameters and finite elements. In Gruen/Kahmen (eds.): Optical 3-D Measurement Techniques V, Wichmann Verlag, Heidelberg. → 4.4.2.4

Triggs, B., McLauchlan, P., Hartley, R., Fitzgibbon, A. (1999): Bundle adjustment – a modern synthesis. ICCV '99, International Workshop on Vision Algorithms, Springer, Berlin, pp. 298–372. → 4.4

Wester-Ebbinghaus, W. (1985): Bündeltriangulation mit gemeinsamer Ausgleichung photogrammetrischer und geodätischer Beobachtungen. Zeitschrift für Vermessungswesen, 110 (3), pp. 101–111. → 4.4.2.3

Zinndorf, S. (1985): Freies Netz – Anwendungen in der Nahbereichsphotogrammetrie. Bildmessung und Luftbildwesen, Heft 4, pp. 109–114. → 4.4.3

9.4.3 Multi-media photogrammetry

Elnashef, B., Filin, S. (2019): Direct linear and refraction-invariant pose estimation and calibration model for underwater imaging. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 154, pp. 259–271. → 4.6

Fryer, J.G., Fraser, C.S. (1986): On the calibration of underwater cameras. The Photogrammetric Record, 12 (67), pp. 73–85. → 4.6.2

Höhle, J. (1971): Zur Theorie und Praxis der Unterwasser-Photogrammetrie. Dissertation, Deutsche Geodätische Kommission, Reihe C, Heft 163. → 4.6

Jordt-Sedlazeck, A., Koch, R. (2013): Refractive structure-from-motion on underwater images. Proceedings of the IEEE international Conference on Computer Vision, pp. 57–64. → 4.6

Kahmen, O., Rofallski, R., Luhmann, T. (2020): Impact of stereo camera calibration to object accuracy in multimedia photogrammetry. Remote Sensing, 12, 2057. → 4.6

Kotowski, R. (1988): Phototriangulation in multi-media photogrammetry. International Archives of Photogrammetry and Remote Sensing, Vol. 27, Kyoto. → 4.6

Maas, H.-G. (1995): New developments in multimedia photogrammetry. In Gruen/Kahmen (eds.): Optical 3D Measurement Techniques III, Wichmann, Heidelberg, pp. 91–97. → 4.6

Maas, H.-G. (2015): On the accuracy potential in underwater/multimedia photogrammetry. Sensors, 15(8), pp. 18140–18152. → 4.6

Menna, F., Remondino, F., Maas, H.-G. (2016): Sensors and techniques for 3D object modelling in underwater environments. MDPI Sensors, 368 p. → 4.6

Menna, F., Nocerino, E., Remondino, F. (2018): Photogrammetric modelling of submerged structures: influence of underwater environment and lens ports on three-dimensional (3D) measurements. In Remondino/Georgopolous/Gonzales-Aguilera/Agrafitois (eds.): Latest Developments in Reality-Based 3D Surveying and Modelling, MDPI Books. → 4.6

Mulsow, C. (2010): A flexible multi-media bundle approach. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38, Part 5, pp. 472–477. → 4.6

Mulsow, C., Maas, H.-G. (2014): A universal approach for geometric modelling in underwater stereo image processing. ICPR Workshop on Computer Vision for Analysis of Underwater Imagery, Stockholm, Sweden, pp. 49–56. → 4.6

Nocerino, E., Menna, F. Fassi, F., Remondino, F. (2016): Underwater calibration of dome port pressure housings. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XL-3/W4, Part 5, pp. 127–134. → 4.6.1.3

Putze, T. (2008): Erweiterte Verfahren zur Mehrmedienphotogrammetrie komplexer Körper. In Luhmann/Müller (eds.): Photogrammetrie-Laserscanning-Optische 3D-Messtechnik, Oldenburger 3D-Tage 2008, Wichmann, Heidelberg, pp. 202–209. → 4.6

Rofallski, R., Luhmann, T. (2022): An efficient solution to ray tracing problems in multimedia photogrammetry for flat refractive interfaces. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science, Vol. 90, Issue 1, pp. 37–54. → 4.6.1.2

Sedlazeck, A., Koch, R. (2012): Perspective and non-perspective camera models in underwater imaging – overview and error analysis. Outdoor and large-scale real-world scene analysis, 7474, pp. 212–242. → 4.6

Shortis, M.R., Harvey, E.S., Abdo, D.A. (2009): A review of underwater stereo-image measurement for marine biology and ecology applications. In Gibson et al. (eds.): Oceanography and Marine Biology: An Annual Review, Vol. 47, CRC Press, Boca Raton FL, USA, 342 p. → 4.6

Shortis, M. (2019): Camera calibration techniques for accurate measurement underwater. In Benjamin/van Duivenvoorde (eds.): 3D recording and interpretation for maritime archaeology, Springer International Publishing, Cham, pp. 11–27. → 4.6

Treibitz, T., Schechner, Y., Kunz, C., Singh, H. (2012): Flat refractive geometry. IEEE Trans Pattern Anal Mach Intell., 34(1), pp. 51–65. → 4.6.1.2

Westfeld, P., Maas, H.-G. (2011): Photogrammetric techniques for voxel-based flow velocimetry field measurement photogrammetry. The Photogrammetric Record, 26(136), pp. 422–438. → 4.6, → 6.9.2.4

9.4.4 Panoramic photogrammetry

Chapman, D., Kotowski, R. (2000): Methodology for the construction of large image archives of complex industrial structures. Publikationen der DGPF, Band 8, Essen 1999. → 4.5

Luhmann, T., Tecklenburg, W. (2004): 3-D object reconstruction from multiple-station panorama imagery. ISPRS Panoramic Photogrammetry Workshop, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 34, Part 5/W16, pp. 39–46. → 4.5.2

Luhmann, T. (2010): Panorama photogrammetry for architectural applications. Mapping, № 139, pp. 40–45. → 4.5

Schneider, D. (2008): Geometrische und stochastische Modelle für die integrierte Auswertung terrestrischer Laserscannerdaten und photogrammetrischer Bilddaten. Dissertation, Technische Universität Dresden. → 4.5

Schneider, D., Maas, H.-G. (2006): A geometric model for linear-array-based terrestrial panoramic cameras. The Photogrammetric Record 21(115), Blackwell Publishing Ltd., Oxford, UK, pp. 198–210. → 4.5.1

9.5 Chapter 5 – Digital image analysis

9.5.1 Fundamentals

Canny, J. (1986): A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8, Vol. 6, pp. 679–698. → 5.2.4.6

Graps, A. (1995): An introduction to wavelets. IEEE Computational Science and Engineering, 2 (2), Los Alamitos, USA. → 5.1.3.3

Pennebaker, W.B., Mitchell, J.L. (1993): JPEG still image data compression standard. Van Nostrad Reinhold, New York. → 5.1.3.3

Tabatabai, A.J., Mitchell, O.R. (1984): Edge location to subpixel values in digital imagery. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-6, No. 2. → 5.1.3.3

Weisensee, M. (1997): Wechselwirkungen zwischen Bildanalyse und Bildkompression. Publikationen der DGPF, Band 5, Oldenburg 1996, pp. 37–45. → 5.1.3.3

9.5.2 Pattern recognition and image matching

Ackermann, F. (1983): High precision digital image correlation. 39. Photogrammetric Week, Stuttgart. → 5.4.2.4

Baltsavias, E. (1991): Multiphoto geometrically constrained matching. Dissertation, ETH Zürich, Institut für Geodäsie und Photogrammetrie, Nr. 49. → 5.4.2.4, → 5.5.4.1

Bay, H., Ess, A., Tuytelaars, T., Van Gool, L. (2008): Speeded-up robust features (SURF). Computer vision and image understanding, 110(3), pp. 346–359. → 5.4.3.6

Bethmann, F., Luhmann, T. (2010): Least-squares matching with advanced geometric transformation models. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38, Part 5, pp. 86–91. → 5.4.2.4, → 5.5.3.5

Bethmann, F., Luhmann, T. (2017): Object-based semi-global multi-image matching. Photogrammetrie–Fernerkundung–Geoinformation, Heft 6, 2017. → 5.5.6.3

Bleyer, M., Rhemann, C., Rother, C. (2011): PatchMatch Stereo – Stereo Matching with Slanted Support Windows. British Machine Vision Conference 2011, pp. 1–11. → 5.5.6.4

Bulatov, D., Wernerus, P., Heipke, C. (2011): Multi-view dense matching supported by triangular meshes. ISPRS Journal of Photogrammetry and Remote Sensing, 66, pp. 907–918. → 5.5.6.4

Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., Leonard, J.J. (2016): Past, present, and future of simultaneous localization and mapping: Toward the robust perception age. IEEE Transactions on Robotics, 32 (6). → 5.5.7.5

Cavegn, S., Haala, N., Nebiker, S., Rothermel, M., Zwölfer, T. (2015): Evaluation of matching strategies for image-based mobile mapping. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. II-3/W, pp. 361–368. → 5.5.7

Chen, L., Rottensteiner, F., Heipke, C. (2020): Feature detection and description for image matching: from hand-crafted design to deep learning. Geo-spatial Information Science, Vol. 24/1, pp. 58–74. → 5.4.3

Conen, N, Jepping, C., Luhmann, T., Maas, H.-G. (2016): Rectification and robust matching using oriented image triplets for minimally invasive surgery. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume III-3, pp. 27–34. → 5.5.5.2

Engel, J. (2017): Large-Scale Direct SLAM and 3D Reconstruction in Real-Time. Dissertation, TU München, Fakultät für Informatik. → 5.5.7.5

Förstner, W. (1982): On the geometric precision of digital correlation. International Archives of Photogrammetry and Remote Sensing, Vol. 26/3, pp. 150–166. → 5.4.2.4

Förstner, W. (1985): Prinzip und Leistungsfähigkeit der Korrelation und Zuordnung digitaler Bilder. 40. Photogrammetric Week, Stuttgart. → 5.4.2.4

Förstner, W., Dickscheid, T., Schindler, F. (2009): Detecting interpretable and accurate scale-invariant keypoints. IEEE 12th International Conference on Computer Vision, pp. 2256–2263. → 5.4.3

Förstner, W., Gülch, E. (1987): A fast operator for detection and precise location of distinct points, corners and centres of circular features. ISPRS Intercommission Workshop on „Fast Processing of Photogrammetric Data", Interlaken, pp. 281–305. → 5.4.3.2

Fuentes-Pacheco, J., Ruiz-Ascencio, J., Rendón-Mancha, J.M. (2015): Visual simultaneous localization and mapping: a survey. Artif. Intell. Rev., 43, pp. 55–81. → 5.5.7.5

Furukawa, Y., Ponce, J. (2010): Accurate, dense and robust multi-view stereopsis. IEEE Trans. PAMI, Vol. 32, pp. 1362–1376. → 5.4.3

Galliani, S., Lasinger, K., Schindler, K. (2016): Gipuma: Massively parallel multi-view stereo reconstruction. Publikationen der DGPF, Band 25. → 5.5.5.1, → 5.5.6.4

Gallup, D., Frahm, J.-M., Mordohai, P., Yang, Q., Pollefeys, M. (2007): Real-time plane-sweeping stereo with multiple sweeping directions. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR'07. → 5.5.6.4

Gruen, A. (1996): Least squares matching: a fundamental measurement algorithm. In Atkinson (ed.): Close Range Photogrammetry and Machine Vision, Whittles Publishing, Caithness, UK, pp. 217–255. → 5.4.2.4, → 5.5.5.1

Gruen, A., Baltsavias, E. (1988): Geometrically constrained multiphoto matching. Photogrammetric Engineering and Remote Sensing, 54 (5), pp. 633–641. → 5.4.2.4, → 5.5.5.1

Haala, N., Cavegn, S. (2017): High densitiy aerial image matching: State-of-the-art and future prospects. EuroSDR contributions to ISPRS Congress XXIII, Special Sessions, No. 66, pp. 28–43. → 5.5.4

Haala, N., Rothermel, M. (2012): Dense multi-stereo matching for high quality digital elevation models. Photogrammetrie-Fernerkundung-Geoinformation, Vol. 4, pp. 331–343. → 5.5.4

Heipke C. (1992): A global approach for least squares image matching and surface reconstruction in object space. Photogrammetric Engineering & Remote Sensing (58) 3, pp. 317–323. → 5.5.6

Hirschmuller, H. (2005): Accurate and efficient stereo processing by semi-global matching and mutual information. Computer Vision and Pattern Recognition, Vol. 2, pp. 807–814. → 5.5.4.2

Hirschmüller, H. (2008): Stereo processing by semi-global matching and mutual information. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30 (2), pp. 328–341. → 5.5.4.2

Hirschmüller, H., Scharstein, D. (2009): Evaluation of stereo matching costs on images with radiometric differences. Proceedings of the Institute of Electrical and Electronics Engineers, Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 31 (9), pp. 1582–1599. → 5.5.4.2

Hosseininaveh, A., Robson, S., Boehm, J., Shortis, M., Wenzel, K. (2013): A comparison of dense matching algorithms for scaled surface reconstruction using stereo camera rigs. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 78, pp. 157–167. → 5.5.4

Jazayeri, I., Fraser, C.S. (2008): Interest operators in close-range object reconstruction. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 37, Part B5, pp. 69–74. → 5.4.3.1

Kersten, T., Lindstaedt, M. (2012): Automatic 3D object reconstruction from multiple images for architectural, cultural heritage and archaeological applications using open-source software and web services. Photogrammetrie–Fernerkundung–Geoinformation, Heft 6, pp. 727–740. → 5.5.2.2

Lowe, D.G. (2004): Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60 (2), pp. 91–110. → 5.4.3.1

Luhmann, T. (1986a): Automatic point determination in a réseau scanning system. Symposium ISPRS Commission V, Ottawa. → 5.4.2.5

Luhmann, T. (1986b): Ein Verfahren zur rotationsinvarianten Punktbestimmung. Bildmessung und Luftbildwesen, Heft 4, pp. 147–154. → 5.4.2.5

Maas, H.-G., Grün, A., Papantoniou, D. (1993): Particle tracking in threedimensional turbulent flows - Part I: Photogrammetric determination of particle coordinates. Experiments in Fluids Vol. 15, pp. 133–146. → 5.5.5.2

Mecca, R., Rodola, E., Cremers, D. (2015): Realistic photometric stereo using partial differential irradiance equation ratios. Computers and Graphics, Vol. 51, pp. 8–16. → 5.5.6.4

Morel, J. M., Yu, G. (2009): ASIFT: A new framework for fully affine invariant image comparison. SIAM Journal on Imaging Sciences, 2(2), pp. 438–469. → 5.4.3.6

Muja, M., Lowe, G.D. (2009): Fast approximate nearest neighbors with automatic algorithm configuration. International Conference on Computer Vision Theory and Applications. → 5.5.2

Mur-Artal, R., Tardós, J.D. (2017): ORB-SLAM2: an open source SLAM system for monocular, stereo and RGB-D cameras. IEEE Transactions on Robotics, 33(5), pp. 1255–1262. → 5.4.3.7, → 5.5.7.5

Nistér, D., Stewénius, H. (2006): Scalable recognition with a vocabulary tree. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, pp. 2161–2168. → 5.5

Pons, J.-P., Keriven, R., Faugeras, O. (2007): Multi-view stereo reconstruction and scene flow estimation with a global image-based matching score. International Journal of Computer Vision, Vol. 72(2), pp. 179–193. → 5.5.5.1

Roth, L., Kuhn, A., Mayer, H. (2017): Wide-baseline image matching with projective view synthesis and calibrated geometric verification. PFG Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 85 (2), S. 85–95. → 5.5.4.2

Rothermel, M., Wenzel, K., Fritsch, D., Haala, N. (2012): SURE: Photogrammetric surface reconstruction from imagery. Proceedings LowCost3D Workshop 2012, December 4–5, Berlin. → 5.5.4.2

Scharstein, D., Szeliski, R. (2002): A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision, Vol. 47 (1/2/3), pp. 7–42. → 5.5.4

Schneider, C.-T. (1991): Objektgestützte Mehrbildzuordnung. Dissertation, Deutsche Geodätische Kommission, Reihe C, Nr. 375. → 5.5.6

Schönberger, J. (2018): Robust Methods for Accurate and Efficient 3D Modeling from Unstructured Imagery. Dissertation, ETH Zürich, Institut für Visual Computing. Nr. 25370. → 5.5.2.2

Seitz, S.M., Curless, B., Diebel, J., Scharstein, D., Szeliski, R. (2006): A comparison and evaluation of multi-view stereo reconstruction algorithms. Computer Vision and Pattern Recognition, Vol. 1, pp. 519–526. → 5.5.4

Shortis, M.R., Seager, J.W., Robson, S., Harvey, E.S (2003): Automatic recognition of coded targets based on a Hough transform and segment matching. Videometrics VII, SPIE Vol. 5013, pp. 202–208. → 5.4.2

Snavely, N., Seitz, S.M., Szeliski, R. (2008): Modeling the world from Internet photo collections. International Journal of Computer Vision, 80(2), pp. 189–210. → 5.5.2.2

Stühmer, J., Gumhold, S., Cremers, D. (2010): Real-time dense geometry from a handheld camera. In Goesele et al. (eds.): DAGM 2010, LNCS 6376, Springer, Berlin, pp. 11–20. → 5.5.6.4

Sturm, P., Triggs, B. (1999): A factorization based algorithm for multi-image projective structure and motion. European Conference on Computer Vision, Springer, pp. 709–720. → 5.5.2.2

Triggs, B. (2004): Detecting keypoints with stable position, orientation, and scale under illumination changes. Computer Vision-ECCV, Springer, Berlin, pp. 100–113. → 5.4.3

Viola, P., Jones, M. (2001): Rapid object detection using a boosted cascade of simple features. Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, Vol. 1, pp. I/511-I/518. → 5.4.3

Weisensee, M. (1992): Modelle und Algorithmen für das Facetten-Stereosehen. Dissertation, Deutsche Geodätische Kommission, Reihe C, Nr. 374. → 5.5.6

Wenzel, K., Rothermel, M., Fritsch, D., Haala, N. (2013): Image acquisition and model selection for multi-view stereo. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XL-5/W1, pp. 251–258. → 5.5.4.2

Wrobel, B, Weisensee, M. (1987): Implementation aspects of facets stereo vision with some applications. Fast Processing of Photogrammetric Data, Interlaken, pp. 259–272. → 5.5.6

Wrobel, B. (1987): Facets stereo vision (FAST vision) – A new approach to computer stereo vision and to digital photogrammetry. Fast Processing of Photogrammetric Data, Interlaken, pp. 231–258. → 5.5.6

Yang, B., Gu, F., Niu, X. (2006): Block mean value based image perceptual hashing. International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP'06). → 5.5.3.3

Zabih, R., Woodfill, J. (1994): Non-parametric local transforms for computing visual correspondence. Third European Conference of Computer Vision, Stockholm, pp. 151–158. → 5.5.3.2

Zhou, G. (1986): Accurate determination of ellipse centers in digital imagery. ASPRS Annual Convention, Vol. 4, March 1986, pp. 256–264. → 5.4.2.5

9.6 Chapter 6 – Measurement tasks and systems

9.6.1 Target and reference systems

Ahn, S.J., Schultes, M. (1997): A new circular coded target for the automation of photogrammetric 3D surface measurements. In Gruen/Kahmen (eds.): Optical 3D Measurement Techniques IV. Wichmann, Heidelberg, pp. 225–234. → 6.2.1.1

Dold, J. (2004): Neue Laser-Technologien für die Industrievermessung. Photogrammetrie-Fernerkundung-Geoinformation, Heft 1, pp. 39–46. → 6.3.2.2

Fraser, C.S., Cronk, S. (2009): A hybrid measurement approach for close-range photogrammetry. ISPRS Journal of Photogrammetry & Remote Sensing, 64(3), pp. 328–333. → 6.2.1

Hattori, S., Akimoto, K., Fraser, C.S., Imoto, H. (2002): Automated procedures with coded targets in industrial vision metrology. Photogrammetric Engineering & Remote Sensing, 68(5), pp. 441–446. → 6.2.1.4

Hennes, M. (2009): Freiformflächenerfassung mit Lasertrackern – eine ergonomische Softwarelösung zur Reflektoroffsetkorrektur. Allgemeine Vermessungs-Nachrichten, Heft 5, pp. 188–194. → 6.3.2.2

Loser, R., Kyle, S. (2003): Concepts and components of a novel 6DOF tracking system for 3D metrology. In Gruen/Kahmen (eds.): Optical 3D Measurement Techniques VI, Vol. 2, pp. 55–62. → 6.3.2.2

Luhmann, T. (2014): Eccentricity in images of circular and spherical targets and its impact on spatial intersection. The Photogrammetric Record, 29(148), pp. 417–433. → 6.2.1.1, → 6.2.1.2

Muralikrishnan, B., Phillips, S., Sawyer, D. (2015): Laser trackers for large scale dimensional metrology: A review. Precision Engineering. → 6.3.2.2

Otepka, J.O., Fraser, C.S. (2004): Accuracy enhancement of vision metrology through automatic target plane determination. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 35(B5), pp. 873–879. → 6.2.1.1

Wasmeier, P. (2009): Videotachymetrie – Sensorfusion mit Potenzial. Allgemeine Vermessungs-Nachrichten, Heft 7, pp. 261–267. → 6.3.2.1

Wendt, K., Schwenke, H., Bösemann, W., Dauke, M. (2003): Inspection of large CMMs and machine tools by sequential multilateration using a single laser tracker. Laser Metrology and Machine Performance, Vol. 5, pp. 121–130. → 6.3.2.2

Zschiesche, K., Fitzke, M., Schlüter, M. (2022): Self-calibration and crosshair tracking with modular digital imaging total station. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science, Vol. 90, pp. 543–557. → 6.3.2.1

9.6.2 Interactive measuring and processing systems

Benning, W., Schwermann, R. (1997): PHIDIAS-MS – Eine digitale Photogrammetrieapplikation unter MicroStation für Nahbereichsanwendungen. Allgemeine Vermessungs-Nachrichten, Heft 1, pp. 16–25. → 6.4.1

González-Aguilera, D., López-Fernández, L., Rodriguez-Gonzalvez, P., Guerrero, D., Hernandez-Lopez, D., Remondino, F., Menna, F., Nocerino, E., Toschi, I., Ballabeni, A., Gaiani, M. (2016): Development of an all-purpose free photogrammetric tool. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XLI-B6, pp. 31–38. → 6.4.2

Kersten, T., Lindstaedt, M. (2012): Generierung von 3D-Punktwolken durch kamera-basierte low-cost Systeme – Workflow und praktische Beispiele. Terrestrisches Laserscanning 2012, Schriftenreihe des DVW, Band 69, Wißner-Verlag, Augsburg, pp. 25–46. → 6.4.2

Luhmann, T. (2010): Close-range photogrammetry for industrial applications. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 64/3, pp. 558–569. → 6.4.3

Luhmann, T. (2016): Learning Photogrammetry with Interactive Software Tool PhoX. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B6, pp. 39–44. → 6.4.4

Luhmann, T., Chizhova, M., Gorkovchuk, D. (2022): Developing a virtual laser scanner for training and research. GIM International, Issue 2, pp. 18–20. → 6.4.4

Remondino, F., Del Pizzo, S., Kersten, T., Troisi, S. (2012): Low-cost and open-source solutions for automated image orientation – a critical overview. In Ioannides/Fritsch/Leissner/Davies/Remondino/Caffo (eds.): Progress in Cultural Heritage Preservation, EuroMed 2012, Lecture Notes in Computer Science, Vol. 7616, Springer, Berlin, pp. 40–54. → 6.4.2

Zinndorf, S. (2004): Photogrammetrische Low-Cost-Systeme. Photogrammetrie–Fernerkundung–Geoinformation, Heft 1, pp. 47–52. → 6.4.3

9.6.3 Measurement of points and contours

Martin, O.C., Muelaner, J., Robson, S. (2016): Comparative performance between two photogrammetric systems and a reference laser tracker network for large-volume in-industrial measurement. The Photogrammetric Record, 31 (155), pp. 348–360. → 6.3.2.2, → 6.6.1.1

Sinnreich, K., Bösemann, W. (1999): Der mobile 3D-Meßtaster von AICON – ein neues System für die digitale Industrie-Photogrammetrie. Publikationen der DGPF, Band 7, München 1998, pp. 175–181. → 6.6.2.1

9.6.4 Laser-scanning systems

Beraldin, J.-A., Blais, F., Lohr, U. (2010): Laser scanning technology. In Vosselman/Maas (eds.): Airborne and Terrestrial Laser Scanning, Whittles Publishing, Caithness, UK, pp. 1–42. → 6.8

Gaisecker, T., Fowler, A., Pfennigbauer, M., Ullrich, A. (2016): Vom Messgerät zum Punktwolken-generator – Laserscanningsystem für statische und kinematische Datenerfassung mit Echtzeitdatenverarbeitung und Cloud-Anbindung. In Luhmann/Schumacher (eds.): Photogrammetrie–Laserscanning–Optische 3D-Messtechnik, Oldenburger 3D-Tage 2016, Wichmann, Berlin/Offenbach, pp. 20–31. → 6.8.1

Kersten, T., Mechelke, K., Lindstaedt, M., Sternberg, H. (2009): Methods for geometric accuracy investigations of terrestrial laser scanning systems. Photogrammetrie-Fernerkundung-Geoinformation, Heft 4, pp. 301–316. → 6.8.1

Kersten, T., Stallmann, D., Tschirschwitz, F. (2015): Entwicklung eines neuen Low-Cost-Innenraum-Aufnahmesystems – Systemaufbau, Systemkalibrierung und erste Ergebnisse. In Luhmann/Müller (eds.): Photogrammetrie–Laserscanning–Optische 3D-Messtechnik, Oldenburger 3D-Tage 2015, Wichmann, Berlin/Offenbach, pp. 131–144. → 6.8.3

Mechelke, K., Kersten, T., Lindstaedt, M. (2007): Comparative investigations into the accuracy behaviour of the new generation of terrestrial laser scanning systems. In Gruen/Kahmen (eds.): Optical 3D Measurement Techniques VIII, Zurich, Vol. I, pp. 319–327. → 6.8.1

Runne, H., Niemeier, W., Kern, F. (2001): Application of laser scanners to determine the geometry of buildings. In Grün/Kahmen (eds.): Optical 3D Measurement Techniques V, Wichmann, Heidelberg, pp. 41–48. → 6.8.1

Schwarte, R. (1997): Überblick und Vergleich aktueller Verfahren der optischen Formerfassung. GMA-Bericht 30, Optische Formerfassung, Langen, pp. 1–12. → 6.9.3

Schwermann, R., Effkemann, C. (2002): Kombiniertes Monoplotting in Laserscanner- und Bilddaten mit PHIDIAS. In Luhmann (ed.): Photogrammetrie und Laserscanning, Wichmann Verlag, Heidelberg, pp. 57–70. → 6.8.1

9.6.5 Measurement of surfaces

Berger, M., Tagliasacchi, A., Seversky, L.M., Alliez, P., Levine, J.A., Sharf, A., Silva, C.T. (2014): State of the art in surface reconstruction from point clouds. The Eurographics Association, Eurographics 2014 – State of the Art Reports. → 6.7

Gühring, J. (2002): 3D-Erfassung und Objektrekonstruktion mittels Streifenprojektion. Dissertation, Universität Stuttgart, Fakultät für Bauingenieur- und Vermessungswesen. → 6.7.3

Heist, S., Mann, A., Kühmstedt, P., Schreiber, P., Notni, G. (2014): Array projection of aperiodic sinusoidal fringes for high-speed three-dimensional shape measurement. Opt. Eng., 53, Nr. 11, 112208. → 6.7.3.4

Heist, S., Lutzke, P., Schmidt, I., Dietrich, P., Kühmstedt, P., Notni, G. (2016): High-speed three-dimensional shape measurement using GOBO projection. Opt. Lasers Eng., 87, pp. 90–96. → 6.7.3.4

Imkamp, D., Schuster, E., Basel, M., Daxauer, H., Fuchs, A., Mayer, T. (2016): Systeme zur Messung komplexer Oberflächen und ihre Leistungsbewertung. VDI-Berichte / VDI-Tagungsbände 2285, Optische Messung von Funktionsflächen, VDI-Verlag. → 6.8

Niini, I. (2003): 3-D glass shape measuring system. In Gruen/Kahmen (eds.): Optical 3D Measurement Techniques VI, Vol. II, ETH Zürich, pp. 176–181. → 6.6.2.3

Pfeifer, N., Karel, W. (2008): Aufnahme von 3D-Punktwolken mit hoher zeitlicher Auflösung mittels aktiver Sensoren. In Luhmann/Müller (eds.): Photogrammetrie-Laserscanning-Optische 3D-Messtechnik, Oldenburger 3D-Tage 2008, Wichmann, Heidelberg, pp. 2–13. → 6.7.1

Ritter, R. (1995): Optische Feldmeßmethoden. In Schwarz (ed.): Vermessungsverfahren im Maschinen- und Anlagenbau, Wittwer, Stuttgart, pp. 217–234. → 6.7.4.1

Schewe, H. (1988): Automatische photogrammetrische Karosserievermessung. Bildmessung und Luftbildwesen, Heft 1, pp. 16–24. → 6.7.5.1

Schreiber W., Notni G. (2000): Theory and arrangements of self-calibrating whole-body three-dimensional measurement systems using fringe projection technique. Optical Engineering 39, pp. 159–169. → 6.7.3

Stahs, T., Wahl, F.M. (1990): Oberflächenmessung mit einem 3D-Robotersensor. Zeitschrift für Photogrammetrie und Fernerkundung, Heft 6, pp. 190–202. → 6.7.3

Thesing, J., Behring, D., Haig, J. (2007): Freiformflächenmessung mit photogrammetrischer Streifenprojektion. VDI Berichte Nr. 1996, Optische Messung technischer Oberflächen in der Praxis. → 6.7.3

VDI (2016): Optische Messung von Funktionsflächen. VDI-Fachtagung Form- und Konturmesstechnik, VDI-Berichte 2285, VDI-Verlag, Düsseldorf, 300 S. → 6.7

Wahl, F.M. (1986): A coded light approach for depth map acquisition. In Hartmann (ed.): Mustererkennung 1986, Springer, Berlin, pp. 12–17. → 6.7.3

9.6.6 Range and 3D cameras

Böhm, J., Pattinson, T. (2010): Accuracy of exterior orientation for a range camera. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38, Part 5, pp. 103–108. → 6.7.6

Karel, W. (2008): Integrated range camera calibration using image sequences from hand-held operation. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 37, Part 5, pp. 945–951. → 6.7.6

Kersten, T., Przybilla, H.-J., Lindstaedt, M. (2016): Investigations of the geometrical accuracy of handheld 3D scanning systems. Photogrammetrie–Fernerkundung–Geoinformation, Heft 5–6, pp. 271–283. → 6.9.7

Khoshelham, K., Elberink, S.O. (2012): Accuracy and resolution of Kinect depth data for indoor mapping applications. Sensors, 12, pp. 1437–1454. → 6.7.6.1

Luhmann, T., Jepping, C., Herd, B. (2014): Untersuchung zum messtechnischen Genauigkeitspotenzial einer Lichtfeldkamera. Publikationen der DGPF, Band 23, Hamburg. → 6.7.6.3

Newcombe, R.A., Davison, A.J., Izadi, S., Kohli, P., Hilliges, O., Shotton, J., Fitzgibbon, A. (2011): KinectFusion: Real-time dense surface mapping and tracking. 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 127–136. → 6.7.6.1

Omelanowsky, D., Kersten, T., Lindstaedt, M. (2013): Untersuchungen von Low-Cost-Systemen zur 3D-Rekonstruktion kleiner Objekte. In Luhmann/Müller (eds.): Photogrammetrie–Laserscanning–Optische 3D-Messtechnik, Oldenburger 3D-Tage 2013, Wichmann, Berlin/Offenbach, pp. 217–229. → 6.9.7

Remondino, F., Stoppa, D. (eds.) (2013): TOF Range-Imaging Cameras. Springer. → 6.7.6

9.6.7 Registration and orientation

Akça, D., Gruen, A. (2005). Recent advances in least squares 3D surface matching. In Gruen/Kahmen (eds.): Optical 3D Measurement Techniques VII, Vol. II, TU Vienna, pp. 197–206. → 6.9.6

Akça, D., Gruen, A. (2007): Generalized least squares multiple 3D surface matching. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36, Part 3/W52. → 6.9.6

Besl, P.J. (1988). Surfaces in range image understanding. Vol. 27, Springer, New York. → 6.9.6.2

Besl, P.J., McKay, N.D. (1992): A method for registration of 3D-shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14 (2), pp. 239–256. → 6.9.6.2

Böhm, J., Becker, S. (2007): Automatic marker-free registration of terrestrial laser scans using reflectance features. In Grün/Kahmen (eds.): Optical 3D Measurement Techniques, pp. 338–344. → 6.9.6.1

Dold, C., Ripperda, N., Brenner, C. (2007): Vergleich verschiedener Methoden zur automatischen Registrierung von terrestrischen Laserscandaten. In Luhmann/Müller (eds.): Photogrammetrie-Laserscanning-Optische 3D-Messtechnik, Oldenburger 3D-Tage 2007, Wichmann, Heidelberg, pp. 196–205. → 6.8.1, → 6.9.6

Lichti, D., Skaloud, J. (2010): Registration and calibration. In Vosselman/Maas (eds.): Airborne and Terrestrial Laser Scanning, Whittles Publishing, Caithness, UK, pp. 83–133. → 6.8.1, → 6.9

Staiger, R., Weber, M. (2007): Die passpunktlose Verknüpfung von Punktwolken – ein Erfahrungsbericht. Terrestrisches Laserscanning, Schriftenreihe des DVW, Band 53, pp. 91–107. → 6.9.6.2

9.6.8 Dynamic and mobile systems

D'Appuzo, N. (2009): Recent advances in 3D full body scanning with applications to fashion and apparel. In Grün/Kahmen (eds.): Optical 3-D Measurement Techniques IX, Vol. II, TU Wien, pp. 243–253. → 6.5, → 6.6.3

Boeder, V., Kersten, T., Hesse, C., Thies, T., Sauer, A. (2010): Initial experience with the integration of a terrestrial laser scanner into the mobile hydrographic multi-sensor system on a ship. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38, part 1/W17. → 6.11.1

Colomina, I., Blázquez, M., Molina, P., Parés, M.E., Wis, M. (2008): Towards a new paradigm for high-resolution low-cost photogrammetry and remote sensing. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37(B1), pp. 1201–1206. → 6.11.2

Eisenbeiss, H. (2009): UAV Photogrammetry. Dissertation, ETH Zurich, Switzerland, Mitteilungen Nr. 105, 235 p. → 6.11.2

Filion, A., Joubair, A., Tahan, A.S., Boneva, I.A. (2018): Robot calibration using a portable photogrammetry system. Robotics and Computer-Integrated Manufacturing, Vol. 49, pp. 77–87. → 6.10.2.1

Kutterer, H. (2010): Mobile mapping. In Vosselman/Maas (eds.): Airborne and Terrestrial Laser Scanning, Whittles Publishing, Caithness, UK, pp. 293–311. → 6.11.1

Maas, H.-G., Grün, A., Papantoniou, D. (1993): Particle tracking in threedimensional turbulent flows - Part I: Photogrammetric determination of particle coordinates. Experiments in Fluids, Vol. 15, pp. 133–146. → 6.10.2.4

Maas, H.-G. (2005): Werkzeuge und Anwendungen der photogrammetrischen 3D-Bewegungs-analyse. In Luhmann (ed.): Photogrammetrie-Laserscanning-Optische 3D-Messtechnik, Oldenburger 3D-Tage 2005, Wichmann, Heidelberg, pp. 207–213. → 6.10.3

Nex, F., Remondino, F. (2014): UAV for 3D mapping applications: a review. Applied Geomatics, 6 (1), pp. 1–15. → 6.11.2

Perfetti, L., Fassi, F. (2022): Handheld fisheye multicamera system: surveying meandering architectonic spaces in open loop mode – accuracy assessment. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVI-2/W1-2022, pp. 435–442. → 6.11.1

Przybilla, H-J., Kraft, T., Geßner, M., Zurhorst, A. (2017): Untersuchungen und erste Ergebnisse zur geometrischen Qualität marktgängiger Kameras für den UAV-Einsatz. In Luhmann/Schumacher (eds.): Photogrammetrie–Laserscanning–Optische 3D-Messtechnik, Oldenburger 3D-Tage 2017, Wichmann Verlag, Berlin/Offenbach, pp. 235–247. → 6.11.2

Raguse, K., Heipke, C. (2005): Photogrammetrische Auswertung asynchroner Bildsequenzen. In Luhmann (ed.): Photogrammetrie-Laserscanning-Optische 3D-Messtechnik, Oldenburger 3D-Tage 2005, Wichmann, Heidelberg, pp. 14–21. → 6.10.1

Remondino, F., Barazzetti, L., Nex, F., Scaioni, M., Sarazzi, D. (2011): UAV photogrammetry for mapping and 3D modelling – current status and future perspectives. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 38(1). → 6.11.2

Westfield, P., Maas, H.-G. (2011): Photogrammetric techniques for voxel-based flow velocity field measurement. The Photogrammetric Record 26(136), pp. 422–438. → 6.10.2.4

9.7 Chapter 7 – Quality issues and optimization

9.7.1 Project planning and simulation

Abdullah, Q., Bethel, J., Hussain, M., Munjy, R. (2013): Photogrammetric project and mission planning. In McGlone (ed.): Manual of Photogrammetry, 6th ed., American Society of Photogrammetry & Remote Sensing, Bethesda, Maryland, pp. 1187–1219. → 7.1

Alsadik, B.S., Gerke, M., Vosselman, G. (2012): Automated camera network design for 3D modeling of cultural heritage objects. Journal of Cultural Heritage. → 7.1.5.1

Efron, B., Tibshirani, R.J. (1993): An Introduction to the Bootstrap. Springer-Verlag, 456 p. → 7.1.4.4

Fraser, C.S. (1984): Network design considerations for non-topographic photogrammetry. Photogrammetric Engineering and Remote Sensing. 50(8), pp. 1115–1126. → 7.1.5.1

Fraser, C.S. (1987): Limiting error propagation in network design. Photogrammetric Engineering and Remote Sensing, 53(5), pp. 487–493. → 7.1.5.1

Fraser, C.S. (1996): Network design. In Atkinson (ed.): Close Range Photogrammetry and Machine Vision, Whittles Publishing, Caithness, UK, pp. 256–281. → 7.1.5.1

Fraser, C. S., Woods, A., Brizzi, D. (2005): Hyper redundancy for accuracy enhancement in automated close range photogrammetry. The Photogrammetric Record, 20 (111), pp. 205–217. → 7.1.2, → 7.2.1

Hastedt, H. (2004): Monte Carlo simulation in close-range photogrammetry. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 35, part B5, pp. 18–23. → 7.1.4

Huber, P.J. (1973): Robust regression: asymptotics, conjectures, and Monte Carlo. The Annals of Statistics, Vol. 1, No. 5, pp. 799–821. → 7.1.4

Reznicek, J., Ekkel, T., Hastedt, H., Luhmann, T., Kahmen, O., Jepping, C. (2016): Zum Einfluss von Maßstäben in photogrammetrischen Projekten großer Volumina. In Luhmann/Schumacher (eds.): Photogrammetrie−Laserscanning−Optische 3D-Messtechnik − Beiträge der 15. Oldenburger 3D-Tage, Wichmann Verlag, Berlin/Offenbach, pp. 286–295. → 7.1.5.2

Schmitt, R., Fritz, P., Jatzkowski, P., Lose, J., Koerfer, F., Wendt, K. (2008): Abschätzung der Messunsicherheit komplexer Messsysteme mittels statistischer Simulation durch den Hersteller. 4. Fachtagung Messunsicherheit, VDI-Wissensforum, Düsseldorf. → 7.1.4

Scott, W.R., Roth, G., Riverst, J.-F. (2003): View planning for automated three-dimensional object reconstruction and inspection. ACM Computing Surveys, Vol. 35(1), pp. 64–96. → 7.1.5

Trummer, M., Munkelt, C., Denzler, J. (2010): Online next-best-view planning for accuracy optimization using an extended criterion. IEEE International Conference on Pattern Recognition (ICPR'10), pp. 1642–1645. → 7.1.5

Zinndorf, S. (1986): Optimierung der photogrammetrischen Aufnahmeanordnung. Dissertation, Deutsche Geodätische Kommission, Reihe C, Heft Nr. 323. → 7.1.5

9.7.2 Quality

Hennes, M. (2007): Konkurrierende Genauigkeitsmaße – Potential und Schwächen aus der Sicht des Anwenders. Allgemeine Vermessungs-Nachrichten, Heft 4, pp. 136–146. → 7.2.1, → 7.2.2

Luhmann, T., Bethmann, F., Herd. B., Ohm, J. (2010): Experiences with 3D reference bodies for quality assessment of free-form surface measurements. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38, Part 5, pp. 405–410. → 7.2.3.5

Luhmann, T. (2011): 3D imaging – how to achieve highest accuracy. In Remondino/Shortis (eds.): Videometrics, Range Imaging, and Applications XI, Proc. SPIE 8085, 808502. → 7.1.2, → 7.2

Martin, O.C., Muelaner, J., Robson, S. (2016): Comparative performance between two photogrammetric systems and a reference laser tracker network for large-volume industrial measurement. The Photogrammetric Record, 31 (155), pp. 348–360. → 7.2

Muralikrishnan, B., Sawyer, D., Blackburn, C., Phillips, S., Borchardt, B., Estler, W.T. (2009): ASME B89.4.19 Performance evaluation tests and geometric misalignments in laser trackers. Journal of Research of the National Institute of Standards and Technology, Vol. 114, No. 1, pp. 21–35. → 7.2

Puerto, P., Heißelmann, D., Müller, S., Mendikute, A. (2022): Methodology to evaluate the performance of portable photogrammetry for large-volume metrology. Metrology 2022, 2, pp. 320–334. → 7.2

Rautenberg, U., Wiggenhagen, M. (2002): Abnahme und Überwachung photogrammetrischer Messsysteme nach VDI 2634, Blatt 1. Photogrammetrie-Fernerkundung-Geoinformation, Heft 2, pp. 117–124. → 7.2.3

Staiger, R. (2004): Was ist eigentlich Metrologie? In Luhmann (ed.): Photogrammetrie-Laserscanning-Optische 3D-Messtechnik, Oldenburger 3D-Tage 2004, Wichmann, Heidelberg, pp. 2–11. → 7.2.1, → 7.2.2

Wiggenhagen, M., Raguse, K. (2003): Entwicklung von Kenngrößen zur Qualitätsbeurteilung optischer Prozessketten. Photogrammetrie-Fernerkundung-Geoinformation, Heft 2, pp. 125–134. → 7.2.1, → 7.2.2, → 7.2.3

9.7.3 Camera calibration

Abraham, S., Förstner, W. (1997): Zur automatischen Modellwahl bei der Kalibrierung von CCD-Kameras. In Paulus/Wahl (eds.): Mustererkennung 1997, 19. DAGM-Symposium, Springer-Verlag, Berlin, pp. 147–155. → 7.3.4

Brown, D.C. (1971): Close-range camera calibration. Photogrammetric Engineering, 37(8), pp. 855–866. → 7.3

Cronk, S., Fraser, C.S., Hanley, H. (2006): Automated metric calibration of colour digital cameras. The Photogrammetric Record, 21(116), pp. 355–372. → 7.3.1

Fraser, C.S., Shortis, M. (1992): Variation of distortion within the photographic field. Photogrammetric Engineering and Remote Sensing, 58(6), pp. 851–855. → 7.3.1

Fraser, C.S., Cronk, S., Stamatopoulos, C. (2012): Implementation of zoom-dependent camera calibration in close-range photogrammetry. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 39(B5), pp. 15–19. → 7.3.1

Fraser, C.S. (2013): Automatic camera calibration in close-range photogrammetry. Photogrammetric Engineering & Remote Sensing, 79(4), pp. 381–388. → 7.3.1

Fryer, J.G. (1996): Camera calibration. In Atkinson (ed.): Close Range Photogrammetry and Machine Vision, Whittles Publishing, Caithness, UK, pp. 156–179. → 7.3

Godding, R. (1999): Geometric calibration and orientation of digital imaging systems. In Jähne et al. (eds.): Handbook of Computer Vision and Applications, Vol. 1, Academic Press, pp. 441–461. → 7.3.1, → 7.3.2

Gruen, A., Huang, T.S. (2001): Calibration and Orientation of Cameras in Computer Vision. Springer, Berlin/Heidelberg. → 7.3

Luhmann, T. (2010): Erweiterte Verfahren zur geometrischen Kamerakalibrierung in der Nahbereichsphotogrammetrie. Habilitationsschrift, Deutsche Geodätische Kommission, Reihe C, Nr. 645. → 7.1.4, → 7.3

Luhmann, T., Godding, R. (2004): Messgenauigkeit und Kameramodellierung – Kernfragen der Industriephotogrammetrie. Photogrammetrie-Fernerkundung-Geoinformation, Heft 1, pp. 13–21. → 7.2.1, → 7.2.2, → 7.4.1, → 7.4.2

Maas, H.-G. (1999): Image sequence based automatic multi-camera system calibration techniques. ISPRS Journal of Photogrammetry and Remote Sensing, 54 (5–6), pp. 352–359. → 7.3.1, → 7.3.2.3

Remondino, F., Fraser, C.S. (2006): Digital camera calibration methods: considerations and comparisons. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36/5, pp. 266–272. → 7.3

Stamatopoulos, C., Fraser, C.S. (2013): Target-free automated image orientation and camera calibration in close-range photogrammetry. ASPRS Annual Conference, Baltimore, Maryland, March 24–28, 8 p, (on CD-ROM). → 7.3.1

Wester-Ebbinghaus, W. (1981): Zur Verfahrensentwicklung in der Nahbereichsphotogrammetrie. Dissertation, Universität Bonn. → 7.3.1, → 7.3.2

Wester-Ebbinghaus, W. (1983): Einzelstandpunkt-Selbstkalibrierung – ein Beitrag zur Feldkalibrierung von Aufnahmekammern. Deutsche Geodätische Kommission, Reihe C, Nr. 289. → 7.3.1, → 7.3.2

Wester-Ebbinghaus, W. (1985): Verfahren zur Feldkalibrierung von photogrammetrischen Aufnahmekammern im Nahbereich. In Kupfer/Wester-Ebbinghaus (eds.): Kammerkalibrierung in der photogrammetrischen Praxis, Deutsche Geodätische Kommission, Reihe B, Heft Nr. 275, pp. 106–114. → 7.3.1, → 7.3.2

9.8 Chapter 8 – Applications

9.8.1 Architecture, archaeology, city models

Acevedo Pardo, C., Brenner, J., Willemsen, T., Sternberg, H., Schramm, T., Zobel, K., Temme, M. (2013): 3D-Modellierung aus einer Laserscanneraufnahme kombiniert mit Luftbildern aus einem Oktokopter. In Luhmann/Müller (eds.): Photogrammetrie–Laserscanning–Optische 3D-Messtechnik, Oldenburger 3D-Tage 2013, Wichmann Verlag, Berlin/Offenbach, pp. 68–78. → 8.1.2.1

Akça, D., Grün, A., Breuckmann, B., Lahanier, C. (2007): High-definition 3D-scanning of arts objects and paintings. In Grün/Kahmen (eds.): Optical 3D Measurement Techniques VIII, Vol. II, ETH Zürich, pp. 50–58. → 8.2.1.1

Albertz, J. (ed.) (2002): Surveying and documentation of historic buildings – monuments – sites. International Archives of Photogrammetry and Remote Sensing, Vol. 34, Part 5/C7. → 8.1

Chizhova, M., Luhmann, T., Gorkovchuk, D., Hastedt, H., Chachava, N., Lekveishvili, N. (2019): Combination of terrestrial laserscanning, UAV and close-range photogrammetry for 3D reconstruction of complex churches in Georgia. GEORES 2019, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XLII-2/W11. → 8.1.2.3

El-Hakim, S.F., Beraldin, J.-A. (2007): Sensor integration and visualisation. In Fryer et al. (eds.): Applications of 3D Measurement from Images, Whittles Publishing, Caithness, UK, pp. 259–298. → 8.1.1, → 8.3

Fassi, F., Achille, C., Fregonese, L. (2011): Surveying and modelling the main spire of Milan Cathedral using multiple data sources. The Photogrammetric Record, 26, pp. 462–487. → 8.1.2.2

Fellbaum, M., Hau, T. (1996): Photogrammetrische Aufnahme und Darstellung des Doms von Siena. Zeitschrift für Photogrammetrie und Fernerkundung, Heft 2, pp. 61–67. → 8.1.1.1

Greiwe, A., Gehrke, R., Spreckels, V., Schlienkamp, A. (2014): Ableitung eines Geländemodells (Haldenkörper) aus multispektralen UAS-Bilddaten. Publikationen der DGPF, Band 23. → 8.3.2

Guidi, G., Remondino, F., Russo, M., Menna, F., Rizzi, A., Ercoli, S. (2009): A multi-resolution methodology for the 3D modeling of large and complex archaeological areas. International Journal of Architectural Computing, 7 (1), pp. 40–55. → 8.3

Haala, N., Kada, M. (2010): An update on automatic 3D building reconstruction. ISPRS Journal of Photogrammetry and Remote Sensing, 65, pp. 570–580. → 8.3

Hess, M., Colson, A., Hindmarch, J. (2018): Capacity building and knowledge exchange of digital technologies in cultural heritage institutions. Museums in a Digital World 70, pp. 48–61. → 8.1, → 8.2

Hastedt, H., Luhmann, T., Colson, A. (2019): Large-volume photogrammetric deformation monitoring of the Bremen Cog. 4th Joint International Symposium on Deformation Monitoring, Athens, Greece. → 8.2.2.4

Kalinowski, P., Füldner, K., Hindmarch, J., Mittmann, M., Schierbaum, A., Luhmann, T. (2022): High accuracy 3D digitisation of the Goethe Elephant skull using hand-held 3D scanning systems and structure-from-motion – a comparative case study. 3D Arch, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVI-2/W1-2022, pp. 275–281. → 8.2.1.4

Kersten, T. (2006): Kombination und Vergleich von digitaler Photogrammetrie und terrestrischem Laserscanning für Architekturanwendungen. Publikationen der DGPF, Band 15, pp. 247–256. → 8.1

Kersten, T., Sternberg, H., Mechelke, K. (2009): Geometrical building inspection by terrestrial laser scanning. Civil Engineering Surveyor – The Journal of the Chartered Institution of Civil Engineering Surveyors, pp. 26–31. → 8.4.1.2

Menna, F., Nocerino, E., Remondino, F., Dellepiane, M., Callieri, M., Scopigno, R. (2016): 3D digitization of a heritage masterpiece – a critical analysis on quality assessment. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLI-B5, pp. 675–683. → 8.2.1.3

Nocerino, E., Menna, F., Remondino, F. (2014): Accuracy of typical photogrammetric networks in cultural heritage 3D modeling projects, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XL-5, pp. 465–472. → 8.2.1.3

Patias, P. (2007): Cultural heritage documentation. In Fryer et al. (eds.): Applications of 3D Measurement from Images, Whittles Publishing, Caithness, UK, pp. 225–257. → 8.1

Perfetti, L., Fassi, F., Gulsan, H. (2019): Generation of gigapixel orthophoto for the maintenance of complex buildings – challenges and lessons learnt. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-2/W9, pp. 605–614. → 8.1.2.2

Reich, T., Schumann, R. (2016): 3D survey and documentation of an early iron age burial from Otzing, near Deggendorf in Lower Bavaria. COSCH e-Bulletin, 3-2016. → 8.2.2.3

Remondino, F. (2007): Image-based detailed 3D geometric reconstruction of heritage objects. Publikationen der DGPF, Band 16, pp. 483–492. → 8.1

Remondino, F., El-Hakim, S. (2006): Image-based 3D modelling: a review. The Photogrammetric Record, 21(115), pp. 269–291. → 8.1

Remondino, F. (2011): Heritage recording and 3D modeling with photogrammetry and 3D scanning. Remote Sensing, 3(6), pp. 1104–1138. → 8.1

Stylianidis, E., Remondino, F. (2016): 3D Recording, Documentation and Management of Cultural Heritage. Whittles Publishing, Caithness, UK, 388 p. → 8.1

Wiedemann, A., Tauch, R. (2005): Mosaikbildung in der Architekturphotogrammetrie. In Luhmann (ed.): Photogrammetrie-Laserscanning-Optische 3D-Messtechnik, Oldenburger 3D-Tage 2005, Wichmann, Heidelberg, pp. 116–121. → 8.1.1

Wiggenhagen, M., Elmhorst, A., Wissmann, U. (2004): Digitale Nahbereichsphotogrammetrie zur Objektrekonstruktion der Bremer Hanse Kogge. Publikationen der DGPF, Band 13, Halle, pp. 383–390. → 8.2.2.4

9.8.2 Engineering and industrial applications

Bösemann, W. (2016): Industrial photogrammetry – accepted metrology tool for exotic niche. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XLI-B5, pp. 15–24. → 8.5

Burnard, D. (2021): Mobile metrology within the aerospace industry. The Journal of the CMSC, Spring 2021, pp. 44–48. → 8.5.2

Chapman, D. (2013): As-built modelling of large and complex industrial facilities. In McGlone (ed.): Manual of Photogrammetry, 6th ed., American Society of Photogrammetry & Remote Sensing, Bethesda, Maryland, pp. 1082–1093. → 8.4.1.1

Flynn, R., McRae, N., Thompson, S. (2021): Case study of a fuselage join automation. The Journal of the CMSC, Spring 2021, pp. 18–25. → 8.5.2

Fraser, C.S. (1986): Microwave antenna measurement. Photogrammetric Engineering & Remote Sensing, 52(10), pp. 1627–1635. → 8.5.2.4

Fraser, C.S., Mallison, J.A. (1992): Dimensional characterization of a large aircraft structure by photogrammetry. Photogrammetric Engineering and Remote Sensing, 58(5), pp. 539–543. → 8.5.2

Fraser, C.S., Brizzi, D. (2002): Deformation monitoring of reinforced concrete bridge beams. Proceedings, 2nd International Symposium on Geodesy for Geotechnical and Structural Engineering, Berlin, 21–24 May, pp. 338–343. → 8.4.2

Fraser, C.S. (2007): Structural monitoring. In Fryer/Mitchell/Chandler (eds.): Applications of 3D Measurement from Images, Whittles Publishing, Caithness, UK, pp. 37–64. → 8.4, → 8.5

Fraser, C.S. (2013): Industrial and engineering measurement. In McGlone (ed.): Manual of Photogrammetry, 6th ed., American Society of Photogrammetry & Remote Sensing, Bethesda, Maryland, pp. 1075–1082. → 8.5

Göring, M., Luhmann, T. (2020): Development of a fan-shaped distance measuring system for measuring rotor blades – concept, photogrammetric orientation and first results. Annals of the Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., Volume V-2-2020, pp. 711–718. → 6.8.2, → 8.4.2.2

Hampel, U., Maas, H.-G. (2009): Cascaded image analysis for dynamic crack detection in material testing. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 64, pp. 345–350. → 8.4.3.2

Lindqvist, R. P., Strand, D., Nilsson, M., Collins, V., Torstensson, J., Kressin, J., Spensieri, D., Archenti, A. (2023): 3D model-based large-volume metrology supporting smart manufacturing and digital twin concepts. Metrology 2023, 3, pp. 29–64. → 8.5

Luhmann, T. (2010): Close range photogrammetry for industrial applications. ISPRS Journal for Photogrammetry and Remote Sensing, Vol. 64/3, pp. 558–569. → 8.5

Luhmann, T., Voigt, A. (2010): Automatische projektive Bildentzerrung am Beispiel der bildgestützten Planung von Solardachanlagen. Publikationen der DGPF, Band 19, pp. 469–476. → 8.4.4

Luhmann, T., Göring, M., Jepping, C. (2016): Optische Messung der Verformung von Rotorblättern unter Windlast. 5. VDI-Fachtagung Optische Messung von Funktionsflächen, VDI-Berichte 2285, VDE-Verlag, Düsseldorf, pp. 33–45. → 8.4.2.2

Luhmann, T., Maas, H.-G. (2017): Industriephotogrammetrie. In Freeden/Rummel (eds.): Handbuch der Geodäsie, Band „Photogrammetrie und Fernerkundung", (C. Heipke, ed.), Springer-Verlag, Berlin, pp. 1–5. → 8.5

Menna, F., Nocerino, E., Remondino, F., Saladino, L., Berri, L. (2020): Towards online UAS-based photogrammetric measurements for 3D metrology inspection. The Photogrammetric Record, 35(172), pp. 467–486. → 6.11, → 8.6

Nietiedt, S., Wester, T.T.B., Langidis, A., Kröger, L., Rofallski, R., Göring, M., Kühn, M., Gülker, G., Luhmann, T. (2022): A wind tunnel setup for fluid structure interaction measurements using optical methods. Sensors 2022, 22, 5014. → 8.5.1.1

Przybilla, H.-J., Woytowicz, D. (2004): Dokumentation industrieller Anlagen: Vom 2D-Bestandsplan über das GIS zur virtuellen Realität – eine Standortbestimmung. Photogrammetrie-Fernerkundung-Geoinformation, Heft 1, pp. 53–58. → 8.4.1.1

Robson, S., Shortis, M. (2007): Engineering and manufacturing. In Fryer et al. (eds.): Applications of 3D Measurement from Images, Whittles Publishing, Caithness, UK, pp. 65–101. → 8.5

Schmidt Paulsen, U., Erne, O., Möller, T., Sanow, G., Schmidt, T. (2009): Wind turbine operational and emergency stop measurements using point tracking videogrammetry. Conference proceedings Bethel (US), SEM. → 8.4.2.2

Sternberg, H., Kersten, T., Jahn, I., Kinzel, R. (2004): Terrestrial 3D laser scanning – data acquisition and object modelling for industrial as-built documentation and architectural applications. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 35, Part B2, pp. 942–947. → 8.4.1.1

Tommasi, C., Achille, C., Fassi, F. (2016) From point cloud to BIM: a modelling challenge in the cultural heritage field. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI (B5), pp. 429–436. → 8.1.1.2, → 8.4.1.2

Trabasso, L.G. (2020): Light automation for aircraft fuselage assembly. The Aeronautical Journal, 124(1272), pp. 216–236. → 8.5.2

Tuominen, V. (2016): The measurement-aided welding cell—giving sight to the blind. International Journal of Advanced Manufacturing Technology 86, pp. 371–386. → 8.5.3

Winstroth, J., Seume, J.R. (2015): Error assessment of blade deformation measurements on a multi-megawatt wind turbine based on digital image correlation. Proceedings of the ASME Turbo Expo 2015: Turbine Technical Conference and Exposition, Montréal, Canada. → 8.4.2.2

9.8.3 Underwater photogrammetry

Bräuer-Burchardt, C., Heinze, M, Schmidt, I., Kühmstedt, P., Notni, G. (2016): Underwater 3D surface measurement using fringe projection based scanning devices. Sensors. 16(1), pp. 13. → 8.6

Kahmen, O., Luhmann, T. (2022): Monocular photogrammetric system for 3D reconstruction of welds in turbid water. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science, Vol. 90, Issue 1, pp. 19–35. → 8.6.2

Menna, F., Nocerino, E., Troisi, S., Remondino, F. (2013): A photogrammetric approach to survey floating and semi-submerged objects. Videometrics, Range Imaging, and Applications XII; and Automated Visual Inspection, Vol. 8791, SPIE, pp. 117–131. → 8.6.3

Nocerino, E., Menna, F. (2020): Photogrammetry: linking the world across the water surface. Journal of Marine Science and Engineering, 8(2), p. 128. → 8.6.3

Rofallski, R., Tholen, C., Helmholz, P., Parnum, I., Luhmann, T. (2020): Measuring artificial reefs using a multi-camera-system for unmanned underwater vehicles. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B2-2020, pp. 999–1008. → 8.6

Shortis, M.R., Harvey, E.S., Abdo, D.A. (2009): A review of underwater stereo-image measurement for marine biology and ecology applications. In Gibson et al. (eds.): Oceanography and Marine Biology: An Annual Review, Vol. 47, CRC Press, Boca Raton FL, USA, 342 p. → 8.4

9.8.4 Medicine, forensics, earth sciences

Breitbeck, R., Ptacek, W., Ebert, L., Fürst, M., Kronreif, G., Thali. M. (2013): Virtobot – A robot system for optical 3D scanning in forensic medicine. 4th International Conference on 3D Body Scanning Technologies, Long Beach CA, USA, pp. 19–20. → 8.8.1

Buck, U., Naether, S., Räss, B., Jackowski, C., Thali, M.J. (2013): Accident or homicide – Virtual crime scene reconstruction using 3D methods. Forensic Science International, Vol. 225, Issues 1–3, pp. 75–84. → 8.8.1

D'Appuzo, N. (2009): Recent advances in 3D full body scanning with applications to fashion and apparel. In Grün/Kahmen (eds.): Optical 3-D Measurement Techniques IX, Vol. II, TU Vienna, pp. 243–253. → 8.7.1

Deli, R., Galantucci, L.M., Laino, A., D'Alessio, R., Di Gioia, E., Savastano, C., Lavecchia, F., Percoco, G. (2013): Three-dimensional methodology for photogrammetric acquisition of the soft tissues of the face: a new clinical-instrumental protocol. Progress in Orthodontics, pp. 14–32. → 8.7.1

Dietrich, R., Maas, H.-G., Baessler, M., Rülke, A., Richter, A., Schwalbe, E., Westfeld, P. (2007): Jakobshavn Isbrae, West Greenland: Flow velocities and tidal interaction of the front area from 2004 field observations. Journal of Geophysical Research, 112 (f3): F03S21. → 8.8.2.1

Fraser, C.S., Hanley, H.B., Cronk, S. (2005): Close-range photogrammetry for accident reconstruction. In Grün/Kahmen (eds.): Optical 3D Measurement Techniques VII, Vol. II, TU Vienna, pp. 115–123. → 8.8.1.1

Fryer, J.G. (2007): Forensic photogrammetry. In Fryer et al. (eds.): Applications of 3D Measurement from Images, Whittles Publishing, Caithness, UK, pp. 103–137. → 8.8.1

Krüger, T. (2008): Optische 3D-Messtechnik in der dentalen Implantologie. In Luhmann/Müller (eds.): Photogrammetrie-Laserscanning-Optische 3D-Messtechnik, Oldenburger 3D-Tage 2008, Wichmann, Heidelberg, pp. 31–42. → 8.7.2

Maas, H.-G., Schwalbe, E., Dietrich, R., Bässler, M., Ewert, H. (2008): Determination of spatio-temporal velocity fields on glaciers in West-Greenland by terrestrial image sequence analysis. International Archives of Photogrammetry, Remote Sensing and Spatial Information Science, Vol. 37, Part B8, pp. 1419–1424. → 8.8.2.1

Mitchell, H. (2007): Medicine and sport: measurement of humans. In Fryer et al. (eds.): Applications of 3D Measurement from Images, Whittles Publishing, Caithness, UK, pp. 171–199. → 8.7

Mugnier, C.J., Gillen, L., Lucas, S.P., Walford, A. (2013): Forensic photogrammetry. In McGlone (ed.): Manual of Photogrammetry, 6th ed., American Society of Photogrammetry & Remote Sensing, Bethesda, Maryland, pp. 1108–1130. → 8.8.1

Mulsow, C., Maas, H.-G., Hentschel, B. (2016): Optical triangulation on instationary water surfaces. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B5, pp. 85–91. → 8.6.1

Näther, S. (2010): 3-D-Vermessungstechniken im Einsatz bei Polizei und Rechtsmedizin in der Schweiz. Forum, Zeitschrift des Bundes für Öffentlich bestellte Vermessungsingenieure, Heft 1, pp. 192–201. → 8.8.1.2

Nguyen, C., Adcock, M., Anderson, S., Lovell, D., Fisher, N., La Salle, J. (2017): Towards high-throughput 3D insect capture for species discovery and diagnostics. arXiv preprint arXiv:1709.02033. → 8.8.2.3

Rieke-Zapp, D., Nearing, M.A. (2005): Digital close range photogrammetry for measurement of soil erosion. Photogrammetric Record, 20 (109), pp. 69–87. → 8.8.2.2

Ströbel, B., Justice, K., Heethoff, M. (2017): Mikrophotogrammetrie präparierter Insekten. In Luhmann/Müller (eds.): Photogrammetrie–Laserscanning–Optische 3D-Messtechnik, Wichmann, Berlin/Offenbach, pp. 433–356. → 8.8.2.3

9.9   Other sources of information

9.9.1   Standards and guidelines

ASME B89.4.19 (2006): Performance Evaluation of Laser-Based Spherical Coordinate Measurement Systems. www.asme.org.

DIN 1319: Fundamentals of metrology. Beuth, Berlin.

DIN 18716: Photogrammetry and remote sensing. Beuth, Berlin.

DIN 18709: Concepts, abbreviations and symbols in surveying. Beuth, Berlin.

DIN 18710: Ingenieurvermessung. Teil 1–4. Beuth, Berlin.

DIN V ENV 13005 (1999): Leitfaden zur Angabe der Unsicherheit beim Messen. ENV 13005, Beuth, Berlin.

GUM (1995): ISO Guide to the Expression of Uncertainty in Measurement (GUM). International Bureau of Weights and Measures, ISBN 92 67 10188 9.

ISO 10360: Geometrical Product Specifications (GPS) - Acceptance and reverification tests for coordinate measuring machines (CMM), Beuth, Berlin.

JCGM 100 (2008): Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM 1995 with minor corrections). Joint Committee for Guides in Metrology.

VDA 5 (2021): Measurement and Inspection Processes. Capability, Planning and Management. 3rd revised edition, Verband der Automobilindustrie e.V.

VDI/VDE 2617: Accuracy of coordinate measuring machines; characteristics and their checking. VDI/VDE guide line, Part 1–9, Beuth, Berlin.

VDI/VDE 2634: Optical 3D measuring systems. VDI/VDE guide line, part 1–3, Beuth, Berlin.

VIM (2010): International dictionary of metrology. 3rd ed., Beuth, Berlin.

9.9.2   Working groups and conferences

3DMC (3D Metrology Conference):
    Conferences
    www.3dmc.events

ASPRS (The American Society for Photogrammetry and Remote Sensing, USA):
    Publications: Photogrammetric Engineering and Remote Sensing
    www.asprs.org

CIPA (Comité International de Photogrammétrie Architecturale):
    Publications and conference proceedings
    http://cipa.icomos.org/

CMSC (Coordinate Metrology Society Conference):
    Publications and conference proceedings
    www.cmsc.org

DAGM (Deutsche Arbeitsgemeinschaft für Mustererkennung):
    Publications: Mustererkennung, Springer (proceedings of annual conferences)
    www.dagm.de

DGPF (Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation):
    Publications: Bildmessung und Luftbildwesen (until 1989), Zeitschrift für Photo-grammetrie und Fernerkundung (until 1997), Photogrammetrie-Fernerkundung-Geoinformation (since 1997); Publikationen der DGPF (proceedings of annual conferences)
    www.dgpf.de

DGZfP (Deutsche Gesellschaft für zerstörungsfreie Prüfung):
>   Publications: Technical guidelines, conference proceedings
>   www.dgzfp.de

ISPRS (International Society for Photogrammetry and Remote Sensing):
>   Commission I: Sensor Systems
>   Commission II: Photogrammetry
>   Publications: International Archives/Annals of the Photogrammetry, Remote Sensing and
>   Spatial Information Sciences; ISPRS Journal of Photogrammetry and Remote Sensing
>   www.isprs.org

SPIE (The International Society for Optical Engineering):
>   Publications and conference proceedings
>   www.spie.org

The Remote Sensing and Photogrammetry Society (UK):
>   Publications: The Photogrammetric Record
>   www.rspsoc.org

VDI/VDE-GMA (VDI/VDE-Gesellschaft für Mess- und Automatisierungstechnik):
>   Publications: guidelines, conference proceedings.
>   www.vdi.de

Abbreviations

ADC	analogue-to-digital converter
AEC	architecture, engineering and construction
AGC	automatic gain control
AI	artificial intelligence
ASIFT	affine invariant SIFT
ASCII	American Standard Code for Information Interchange
AVI	audio video interleaved
BI (BSI)	back (side) illumination
BIM	building information modelling
BRDF	bidirectional reflection distribution function
BVLOS	beyond visual line of sight
CAAD	computer-aided architectural design
CAD	computer-aided design
CAM	computer-aided manufacturing
CCD	charge-coupled device
CCIR	International Radio Consultative Committee
CID	charge injection device
CIE	Commission Internationale de l'Éclairage (International Commission on Illumination)
CIPA	Comité International de Photogrammétrie Architecturale
CIR	colour infrared
CMM	coordinate measuring machine
CMOS	complementary metal oxide semi-conductor
CT	computed tomography
CTF	contrast transfer function
DCT	discrete cosine transform
DGPF	Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V.
DGPS	differential GPS
DIC	digital image correlation
DLP	digital light processing
DLT	direct linear transformation
DMD	digital mirror device
DOF	degree(s) of freedom
DoG	difference of Gaussian
DRAM	dynamic random access memory
DSLR	digital single lens reflex (camera)
DSM	digital surface model
DVD	digital versatile (video) disk
DXF	Drawing Exchange Format (AutoCAD)
EP	entrance pupil

EPS	encapsulated postscript
FAST	features from accelerated segment test
FFD	flange focal distance
FFT	full frame transfer or fast Fourier transform
FMC	forward motion compensation
FOV	field of view
FPGA	field programmable gate array
fps	frames per second
FT	frame transfer
GIF	graphics interchange format
GIS	geo- or geographic information system
GNSS	global navigation satellite system
GPS	global positioning system
GUM	Guide to the Expression of Uncertainty in Measurement
HDR	high dynamic range (imaging)
HDTV	high definition television
HEVC	high efficient video coding
HSL	hue, saturation, lightness
ICP	iterative closest point
IEEE	Institute of Electrical and Electronics Engineers
IFOV	instantaneous field of view
IGS	image guided surgery
IHS	intensity, hue, saturation
IL	interline transfer
IMU	inertial measurement unit
INS	inertial navigation system
IR	infrared
ISPRS	International Society for Photogrammetry and Remote Sensing
JCGM	Joint Committee for Guides in Metrology
JPEG	Joint Photographic Expert Group
LCD	liquid crystal display
LCOS	liquid crystal on silicon
LED	light emitting diode
LMA	Levenberg-Marquardt algorithm
LME	length measurement error
LoG	Laplacian of Gaussian
LSM	least squares matching
LUT	lookup table
LZW	Lempel-Ziv-Welch (compression)
MDL	minimum description length
MoCap	motion capture
MOS	metal oxide semiconductor

MPE	maximum permitted error
MPEG	Motion Picture Expert Group
MR	magnetic resonance
MR, MRT	magnetic resonance tomography
MTF	modulation transfer function
NC	numerically controlled (machine)
NCC	normalised cross correlation
NIR	near infrared
OSGM	object-based semi-global matching
PLL	phase-locked loop
PLY	polygon file format
PMD	photonic mixing device
PNG	portable network graphics
PSF	point spread function
PTX	file format for point clouds
RANSAC	random sample consensus
RGB	red, green, blue
RGBD	red, green, blue, distance
RMS	root mean square
RMSE	root mean square error
RP	resolving power
RPAS	remotely piloted aircraft system
RPV	remotely piloted vehicle
RTI	reflectance transformation imaging
SAD	sum of absolute differences
sCMOS	scientific CMOS
SfM	structure from motion
SGM	semi-global matching
SIFT	scale invariant feature transform
SLR	single lens reflex (camera)
SMR	spherically mounted retro-reflector
SNR	signal-to-noise ratio
SPIE	The International Society for Optics and Photonics; originally founded as the Society of Photographic Instrumentation Engineers
SSD	sum of squared differences
STL	stereo lithography
SURF	speed-up robust features
SUSAN	smallest univalue segment assimilating nucleus
SURF	speed-up robust features
SUSAN	smallest univalue segment assimilating nucleus
SVD	single value decomposition
SWIR	short-wave infrared

TIFF	tagged image file format
TLS	terrestrial laser scanning
TOF	time-of-flight
UAS	unmanned aerial system
UAV	unmanned aerial vehicle (also: unpiloted aerial vehicle)
USB	universal serial bus
UTM	universal transverse Mercator (projection)
UV	ultraviolet
VDE	Verband der Elektrotechnik, Elektronik, Informationstechnik e.V.
VDI	Verband Deutscher Ingenieure e.V.
VLL	vertical line locus
VLOS	visual line of sight
VR	virtual reality
VRML	virtual reality modelling language
WFD	wave form digitizer
WiFi	wireless fidelity
WLAN	wireless local area network
WMV	Windows media video

Image sources

3D Center Zurich, Switzerland: 8.89, 8.91, 8.92

Aeroscout GmbH, Hochdorf, Switzerland: 6.104

Aesculap AG, Tuttlingen, Germany: 8.85b, 8.86a

Agisoft LLC, St. Petersburg, Russia: 1.48, 6.22

Airbus Operations GmbH, Hamburg, Germany: 8.65

Apple Corp., Cupertino, CA, USA: 6.68a

Artec Europe, Luxembourg: 6.91a

Atesos, Switzerland: 6.5a

Austrian Archaeological Institute (Österreichisches Archäologisches Institut), Vienna: 8.23

Automated Precisions Inc. (API), Rockville, MD, USA: 6.8, 6.16, 6.19, 6.86b

AWAIBA GmbH, Nürnberg, Germany: 3.112b

AXIOS 3D Services GmbH, Oldenburg, Germany: 3.127a, 3.156, 4.14, 6.5b, 6.40, 6.45, 7.18, 8.86a

Bayerische Staatsbibliothek München, Germany: 1.25, ZDB 131403-8; urn:nbn:de:bvb:12-bsb10256151-0

Boredom Projects: 6.39

Boston Dynamics, Waltham, USA: 6.103d

BrainLAB AG, Heimstetten, Germany: 8.85a

Carl Zeiss AG (Zeiss, ZI Imaging, Oberkochen, Jena; Steinbichler, GOM), Germany: 1.29, 1.30, 1.33, 1.35, 1.41, 3.47, 6.37, 6.55, 6.60, 6.65, 6.83, 6.84, 6.86a, 6.87, 8.50, 8.70, 8.75

Cascination AG, Bern, Switzerland: 8.86b

CERN, Switzerland: 8.64 (photo: Claudia Marcelloni)

Cognex Corp., Natick, USA: 6.61

Creaform, Canada; AMETEK GmbH, Leinfelden-Echterdingen, Germany: 6.36, 6.82, 6.85

CycloMedia Technology B.V., Waardenburg, Netherlands: 3.139, 3.140

Dantec Dynamics GmbH, Ulm, Germany: 6.97

DhochN-Jade Digital Engineering GmbH, Oldenburg, Germany: 8.44

Dr. Clauß Bild- und Datentechnik GmbH, Zwönitz, Germany: 6.78

EOS Systems Inc., Vancouver, Canada: 1.47

Excelitas PCO GmbH, Kehlheim, Germany: 3.91, 3.121a, 3.122

Faro GmbH, Krefeld, Germany: 6.18

Fokus GmbH, Leipzig, Germany: 4.29

Fondazione Bruno Kessler (FBK) Trento, Italy: 8.15, 8.16, 8.17, 8.20, 8.21, 8.22, 8.81, 8.82, 8.83

ForWind, Universität Hannover, Germany: 8.51, 8.63

Frankfurt University of Applied Sciences (UAS), Germany: 8.38, 8.39, 8.40, 8.41

Fraunhofer-Einrichtung für Großstrukturen in der Produktionstechnik (FhG IGP), Rostock, Germany: 6.118b

Fraunhofer-Institut für Angewandte Optik und Feinmechanik (FhG IOF), Jena, Germany: 3.157d, 6.52, 6.56, 8.60

Fraunhofer-Institut für Graphische Datenverarbeitung (FhG IGD), Darmstadt, Germany: 6.81

Fraunhofer-Institut für Windenergiesysteme (FhG IWES), Bremerhaven, Germany: 8.60 (photo: Frank Bauer)

Fujifilm Europe GmbH, Düsseldorf, Germany: 3.114a

GDV Systems GmbH, Bad Schwartau, Germany: 1.16, 8.67, 8.68, 8.69

GEXCEL srl, Brescia, Italy: 6.102b

Giroptic SAS, Lille, France: 3.141a

GSI Geodetic Services Inc., Melbourne, FL, USA: 1.38, 1.45, 3.119, 3.157a b, 4.76, 5.2a, 5.34, 6.9, 6.27b, 6.34, 6.35, 8.65

HafenCity University (HCU) Hamburg, Geomatics Lab, Germany: 6.79, 6.115, 8.5

Hasselblad Svenska AB, Göteborg, Sweden: 3.118

Hexagon AB (Leica Geosystems, Wild, Kern, AICON, Breuckmann), Sweden: 1.31, 1.32, 1.34, 1.42, 3.120, 3.121b, 3.127b, 3.157c, 3.161, 4.4, 6.12, 6.15, 6.17, 6.19, 6.23, 6.42, 6.43, 6.44, 6.49, 6.58, 6.88, 6.94, 6.102a, 6.103c-d, 6.118a, 8.12, 8.13, 8.23, 8.24, 8.47, 8.61, 8.62, 8.76

Hochschule Darmstadt / TU Darmstadt, Germany: 8.98

Hochschule München, Germany: 3.145b

High Speed Vision GmbH, Ettlingen, Germany: 3.124

Hubbs Machine & Manufacturing Inc., Fenton, USA: 6.8b

Humanetics Europe GmbH, Heidelberg, Germany: 6.30, 6.32, 8.71

i3mainz, Hochschule Mainz, Germany: 8.14, 8.25, 8.26

IES Elektronikentwicklung, Braunschweig, Germany: 3.123

Imetric S.A., Porrentruy, Switzerland: 2.5

Industrieanlagen-Betriebsgesellschaft mbH (IABG), Ottobrunn, Germany: 8.69

InfraTec GmbH – Infrarotsensorik und Messtechnik, Dresden, Germany: 3.143

iNovitas AG, Baden-Dättwil, Switzerland: 6.101

Institut für Geodäsie und Photogrammetrie (IGP), ETH Zürich, Switzerland: 1.22

Institut für Geologie, Universität Bern, Switzerland: 8.96, 8.97

Institut für Photogrammetrie, TU Berlin, Germany: 1.26

Institut für Photogrammetrie, Universität Bonn, Germany: 1.18

Institut für Photogrammetrie und Fernerkundung (IPF), TU Dresden, Germany: 8.57, 8.77, 8.78, 8.93, 8.94, 8.95

Institut für Geodäsie und Photogrammetrie (IGP), TU Braunschweig, Germany: 8.46

Institut für Photogrammetrie und GeoInformation (IPI), Leibniz Universität Hannover, Germany: 8.27, 8.28

Institut für Rechtsmedizin, Universität Bern, Switzerland: 8.90

Intel Corp., Santa Clara, CA, USA: 6.68b-d, 6.69

INVERS – Industrievermessung und -systeme, Essen, Germany: 1.21

International Society for Photogrammetry and Remote Sensing (ISPRS): 6.24

Jenoptik Laser-Optik-Systeme GmbH, Jena, Germany: 3.131

Kodak AG, Stuttgart, Germany: 1.44

KONTRON S&T AG, Augsburg, Germany: 3.129

Landesmuseum Natur und Mensch, Oldenburg, Germany: 1.12, 1.13, 8.31

Leica Camera AG, Solms, Germany: 3.48, 3.116, 3.117

LIMESS Messtechnik und Software GmbH, Krefeld, Germany: 6.66

Lytro Inc., Mountain View, CA, USA: 6.71b

Mantis Vision Ltd, Petach Tikva, Israel: 6.91b

Mapvision Ltd, Espoo, Finland: 1.40, 6.46

Messbildstelle GmbH, Dresden, Germany: 8.1

Rotbucher Systeme, Bad Reichenhall, Germany: 6.8c

Metronor AS, Nesbru, Norway: 6.27a, 6.29

MicaSense, AgEagle, USA: 3.145a

Microsoft Corp., Redmond, USA: 6.67, 6.116a

National Physics Laboratory (NPL), London, UK: 7.26

NavVis GmbH, Munich, Germany: 6.103a

Nikon Corp,, Minato, Japan: 3.115

Nikon Metrology, Metris, Leuven, Belgium: 3.128b, 6.93a

Otto-Friedrich Universität Bamberg, Lehrstuhl für Digitale Denkmaltechnologien: 8.2 (student work by Cornelia Müller), Germany: 8.10, 8.11

Panono Professional360 GmbH, Berlin, Germany: 3.141b

Pentacon GmbH, Dresden, Germany: 3.132

Phocad Ingenieurgesellschaft mbH, Aachen, Germany: 1.46, 4.23, 6.20

Photogauge Inc., Alamo, CA, USA: 6.80

Photometrix, Kew, Australia: 1.20, 6.21, 8.88

Pix4D S.A., Lausanne, Switzerland: 8.46

Plus Orthopedics, Precision Instruments, Smith & Nephews, Aarau, Switzerland: 8.87

Photron, Tokyo, Japan: 6.13b

Politecnico di Milano, Italy: 6.102, 8.6, 8.7, 8.8, 8.9, 8.20, 8.21, 8.22

Porsche AG, Stuttgart, Germany: 6.95

Premium AEROTEC GmbH, Nordenham, Germany: 8.66

Propellor Aerobotics Pty Ltd, Surry Hills, NSW, Australia: 6.110

Qualisys AB, Göteborg, Sweden: 1.19, 6.98

Raytrix GmbH, Kiel, Germany: 6.71a

Riegl Laser Measurement Systems GmbH, Horn, Austria: 4.23, 6.73c, 6.105

Rollei Fototechnic GmbH, Rolleimetric GmbH, Braunschweig, Germany: 1.36, 1.37, 1.39, 1.43, 3.63, 3.93, 8.65

Römisch-Germanisches Zentralmuseum (RZGM), Mainz, Germany: 8.14

Schneider Digital, Josef J. Schneider e. K., Miesbach, Germany: 6.114a

Schölly Fiberoptik GmbH, Denzlingen, Germany: 3.142

SenseFly Ltd, Cheseaux-Lausanne, Switzerland: 6.107

SICK AG, Waldkirch, Germany: 6.76

Sony Deutschland GmbH, Köln, Germany: 3.112c, 3.114b

SpheronVR AG, Zweibrücken, Germany: 3.135

SPM3D, Kiev, Ukraine: 8.4, 8.42, 8.43, 8.45

ThyssenKrupp Accessibility BV, Krimpen aan den IJssel, Netherlands: 8.47

Technical University Munich, Computer Vision Group, Germany: 5.107, 5.113

Teledyne Technologies Inc, USA: 3.75a

Topcon Deutschland Positioning GmbH, Hamburg, Germany: 6.14

Trimble GmbH, Raunheim, Germany: 6.100

University College London (UCL), UK: 5.90, 6.89

University of Naples, Parthenope, Italy: 8.81, 8.82, 8.83

Vertex Antennentechnik GmbH, Duisburg, Germany: 1.16, 8.67, 8.68

Vialux GmbH, Chemnitz, Germany: 6.53, 6.59

Volkswagen AG, Wolfsburg, Germany: 1.15, 2.5, 3.126, 7.12, 8.73, 8.74

Vuzix Corp., West Henrietta, NY, USA: 6.114b

Winterthur, Switzerland: 6.113

X-Rite Inc., Grand Rapids, MI, USA: 5.28b

Zoller & Fröhlich GmbH, Wangen im Allgäu, Germany: 6.73a, 6.109

All unspecified images were created either by the author or by the Institute for Applied Photogrammetry and Geoinformatics (IAPG), Jade University in Oldenburg. All information was given to the best of our knowledge.

Index

Leica Geosystems BLK

LEICA BLK3D

HANDHELD 3D IMAGER

Combining measurement sensors, software, and on-device edge data processing capabilities, the BLK3D c in-picture measurements with professional-grade accuracy in real-time possible. Every image captured is a complete and precise 3D measurement record.

Leica Geosystems AG
BLK3D.com
realitycapture.geo@leica-geosystems.com

Measure 3D coordinates with TRITOP:
Compact, mobile and highly accurate

TRITOP is the solution for measuring large parts under most demanding environmental conditions. It delivers highly precise measuring data with a maximum ease of use even for difficult-to-access parts on ships, wind turbines or industrial plants.

The photogrammetry system captures three-dimensional coordinates of singular measurement points and provides the basis for the inspection of static objects and dynamic deformations.

zeiss.ly/eipn

Retroreflective
Targets

OrthoPilot ® Elite

Hearo® Robotic Surgery

Our solutions

- Ready-to-use and long-term stable optical measuring systems.

- Camera calibrations with a wide range of imaging and distortion models from 25° to 210° field of view angle with several million applications per year.

- Software modules for camera calibration, single shot camera calibration, tool calibration, image measurement, positioning, position monitoring and bundle adjustment.

Let's talk about your project!
info@axios3d.de
www.axios3d.de
+49-441-217-47-00