Vincent G. Duffy (Ed.)

Digital Human Modeling

First International Conference on
Digital Human Modeling, ICDHM 2007
Held as Part of HCI International 2007
Beijing, China, July 22-27, 2007
Proceedings

Springer

Volume Editor

Vincent G. Duffy
Purdue University
Regenstrief Center for Healthcare Engineering Falculty
Schools of Industrial Engineering and Agricultural & Biological Engineering
Grissom Hall, West Lafayette
IN 47907, USA
E-mail: duffy@purdue.edu

Library of Congress Control Number: 2007929546

CR Subject Classification (1998): H.5, H.1, H.3, H.4.2, I.2-6, J.3

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

ISSN 0302-9743
ISBN-10 3-540-73318-3 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-73318-8 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12082759 06/3180 5 4 3 2 1 0

Foreword

The 12th International Conference on Human-Computer Interaction, HCI International 2007, was held in Beijing, P.R. China, 22-27 July 2007, jointly with the Symposium on Human Interface (Japan) 2007, the 7th International Conference on Engineering Psychology and Cognitive Ergonomics, the 4th International Conference on Universal Access in Human-Computer Interaction, the 2nd International Conference on Virtual Reality, the 2nd International Conference on Usability and Internationalization, the 2nd International Conference on Online Communities and Social Computing, the 3rd International Conference on Augmented Cognition, and the 1st International Conference on Digital Human Modeling.

A total of 3403 individuals from academia, research institutes, industry and governmental agencies from 76 countries submitted contributions, and 1681 papers, judged to be of high scientific quality, were included in the program. These papers address the latest research and development efforts and highlight the human aspects of design and use of computing systems. The papers accepted for presentation thoroughly cover the entire field of Human-Computer Interaction, addressing major advances in knowledge and effective use of computers in a variety of application areas.

This volume, edited by Vincent G. Duffy, contains papers in the thematic area of Digital Human Modeling, addressing the following major topics:

- Shape and Movement Modeling and Anthropometry
- Building and Applying Virtual Humans
- Medical and Rehabilitation Applications
- Industrial and Ergonomic Applications

The remaining volumes of the HCI International 2007 proceedings are:

- Volume 1, LNCS 4550, Interaction Design and Usability, edited by Julie A. Jacko
- Volume 2, LNCS 4551, Interaction Platforms and Techniques, edited by Julie A. Jacko
- Volume 3, LNCS 4552, HCI Intelligent Multimodal Interaction Environments, edited by Julie A. Jacko
- Volume 4, LNCS 4553, HCI Applications and Services, edited by Julie A. Jacko
- Volume 5, LNCS 4554, Coping with Diversity in Universal Access, edited by Constantine Stephanidis
- Volume 6, LNCS 4555, Universal Access to Ambient Interaction, edited by Constantine Stephanidis
- Volume 7, LNCS 4556, Universal Access to Applications and Services, edited by Constantine Stephanidis
- Volume 8, LNCS 4557, Methods, Techniques and Tools in Information Design, edited by Michael J. Smith and Gavriel Salvendy
- Volume 9, LNCS 4558, Interacting in Information Environments, edited by Michael J. Smith and Gavriel Salvendy

- Volume 10, LNCS 4559, HCI and Culture, edited by Nuray Aykin
- Volume 11, LNCS 4560, Global and Local User Interfaces, edited by Nuray Aykin
- Volume 13, LNAI 4562, Engineering Psychology and Cognitive Ergonomics, edited by Don Harris
- Volume 14, LNCS 4563, Virtual Reality, edited by Randall Shumaker
- Volume 15, LNCS 4564, Online Communities and Social Computing, edited by Douglas Schuler
- Volume 16, LNAI 4565, Foundations of Augmented Cognition 3rd Edition, edited by Dylan D. Schmorrow and Leah M. Reeves
- Volume 17, LNCS 4566, Ergonomics and Health Aspects of Work with Computers, edited by Marvin J. Dainoff

I would like to thank the Program Chairs and the members of the Program Boards of all Thematic Areas, listed below, for their contribution to the highest scientific quality and the overall success of the HCI International 2007 Conference.

Ergonomics and Health Aspects of Work with Computers

Program Chair: Marvin J. Dainoff

Arne Aaras, Norway
Pascale Carayon, USA
Barbara G.F. Cohen, USA
Wolfgang Friesdorf, Germany
Martin Helander, Singapore
Ben-Tzion Karsh, USA
Waldemar Karwowski, USA
Peter Kern, Germany
Danuta Koradecka, Poland
Kari Lindstrom, Finland
Holger Luczak, Germany
Aura C. Matias, Philippines
Kyung (Ken) Park, Korea
Michelle Robertson, USA
Steven L. Sauter, USA
Dominique L. Scapin, France
Michael J. Smith, USA
Naomi Swanson, USA
Peter Vink, The Netherlands
John Wilson, UK

Human Interface and the Management of Information

Program Chair: Michael J. Smith

Lajos Balint, Hungary
Gunilla Bradley, Sweden
Hans-Jörg Bullinger, Germany
Alan H.S. Chan, Hong Kong
Klaus-Peter Fähnrich, Germany
Michitaka Hirose, Japan
Yoshinori Horie, Japan
Richard Koubek, USA
Yasufumi Kume, Japan
Mark Lehto, USA
Robert Proctor, USA
Youngho Rhee, Korea
Anxo Cereijo Roibás, UK
Francois Sainfort, USA
Katsunori Shimohara, Japan
Tsutomu Tabe, Japan
Alvaro Taveira, USA
Kim-Phuong L. Vu, USA
Tomio Watanabe, Japan
Sakae Yamamoto, Japan

Jiye Mao, P.R. China
Fiona Nah, USA
Shogo Nishida, Japan
Leszek Pacholski, Poland
Hidekazu Yoshikawa, Japan
Li Zheng, P.R. China
Bernhard Zimolong, Germany

Human-Computer Interaction

Program Chair: Julie A. Jacko

Sebastiano Bagnara, Italy
Jianming Dong, USA
John Eklund, Australia
Xiaowen Fang, USA
Sheue-Ling Hwang, Taiwan
Yong Gu Ji, Korea
Steven J. Landry, USA
Jonathan Lazar, USA
V. Kathlene Leonard, USA
Chang S. Nam, USA
Anthony F. Norcio, USA
Celestine A. Ntuen, USA
P.L. Patrick Rau, P.R. China
Andrew Sears, USA
Holly Vitense, USA
Wenli Zhu, P.R. China

Engineering Psychology and Cognitive Ergonomics

Program Chair: Don Harris

Kenneth R. Boff, USA
Guy Boy, France
Pietro Carlo Cacciabue, Italy
Judy Edworthy, UK
Erik Hollnagel, Sweden
Kenji Itoh, Japan
Peter G.A.M. Jorna, The Netherlands
Kenneth R. Laughery, USA
Nicolas Marmaras, Greece
David Morrison, Australia
Sundaram Narayanan, USA
Eduardo Salas, USA
Dirk Schaefer, France
Axel Schulte, Germany
Neville A. Stanton, UK
Andrew Thatcher, South Africa

Universal Access in Human-Computer Interaction

Program Chair: Constantine Stephanidis

Julio Abascal, Spain
Ray Adams, UK
Elizabeth Andre, Germany
Margherita Antona, Greece
Chieko Asakawa, Japan
Christian Bühler, Germany
Noelle Carbonell, France
Jerzy Charytonowicz, Poland
Pier Luigi Emiliani, Italy
Michael Fairhurst, UK
Zhengjie Liu, P.R. China
Klaus Miesenberger, Austria
John Mylopoulos, Canada
Michael Pieper, Germany
Angel Puerta, USA
Anthony Savidis, Greece
Andrew Sears, USA
Ben Shneiderman, USA
Christian Stary, Austria
Hirotada Ueda, Japan

Gerhard Fischer, USA
Jon Gunderson, USA
Andreas Holzinger, Austria
Arthur Karshmer, USA
Simeon Keates, USA
George Kouroupetroglou, Greece
Jonathan Lazar, USA
Seongil Lee, Korea
Jean Vanderdonckt, Belgium
Gregg Vanderheiden, USA
Gerhard Weber, Germany
Harald Weber, Germany
Toshiki Yamaoka, Japan
Mary Zajicek, UK
Panayiotis Zaphiris, UK

Virtual Reality

Program Chair: Randall Shumaker

Terry Allard, USA
Pat Banerjee, USA
Robert S. Kennedy, USA
Heidi Kroemker, Germany
Ben Lawson, USA
Ming Lin, USA
Bowen Loftin, USA
Holger Luczak, Germany
Annie Luciani, France
Gordon Mair, UK
Ulrich Neumann, USA
Albert "Skip" Rizzo, USA
Lawrence Rosenblum, USA
Dylan Schmorrow, USA
Kay Stanney, USA
Susumu Tachi, Japan
John Wilson, UK
Wei Zhang, P.R. China
Michael Zyda, USA

Usability and Internationalization

Program Chair: Nuray Aykin

Genevieve Bell, USA
Alan Chan, Hong Kong
Apala Lahiri Chavan, India
Jori Clarke, USA
Pierre-Henri Dejean, France
Susan Dray, USA
Paul Fu, USA
Emilie Gould, Canada
Sung H. Han, South Korea
Veikko Ikonen, Finland
Richard Ishida, UK
Esin Kiris, USA
Tobias Komischke, Germany
Masaaki Kurosu, Japan
James R. Lewis, USA
Rungtai Lin, Taiwan
Aaron Marcus, USA
Allen E. Milewski, USA
Patrick O'Sullivan, Ireland
Girish V. Prabhu, India
Kerstin Röse, Germany
Eunice Ratna Sari, Indonesia
Supriya Singh, Australia
Serengul Smith, UK
Denise Spacinsky, USA
Christian Sturm, Mexico
Adi B. Tedjasaputra, Singapore
Myung Hwan Yun, South Korea
Chen Zhao, P.R. China

Online Communities and Social Computing

Program Chair: Douglas Schuler

Chadia Abras, USA
Lecia Barker, USA
Amy Bruckman, USA
Peter van den Besselaar, The Netherlands
Peter Day, UK
Fiorella De Cindio, Italy
John Fung, P.R. China
Michael Gurstein, USA
Tom Horan, USA
Piet Kommers, The Netherlands
Jonathan Lazar, USA
Stefanie Lindstaedt, Austria
Diane Maloney-Krichmar, USA
Isaac Mao, P.R. China
Hideyuki Nakanishi, Japan
A. Ant Ozok, USA
Jennifer Preece, USA
Partha Pratim Sarker, Bangladesh
Gilson Schwartz, Brazil
Sergei Stafeev, Russia
F.F. Tusubira, Uganda
Cheng-Yen Wang, Taiwan

Augmented Cognition

Program Chair: Dylan D. Schmorrow

Kenneth Boff, USA
Joseph Cohn, USA
Blair Dickson, UK
Henry Girolamo, USA
Gerald Edelman, USA
Eric Horvitz, USA
Wilhelm Kincses, Germany
Amy Kruse, USA
Lee Kollmorgen, USA
Dennis McBride, USA
Jeffrey Morrison, USA
Denise Nicholson, USA
Dennis Proffitt, USA
Harry Shum, P.R. China
Kay Stanney, USA
Roy Stripling, USA
Michael Swetnam, USA
Robert Taylor, UK
John Wagner, USA

Digital Human Modeling

Program Chair: Vincent G. Duffy

Norm Badler, USA
Heiner Bubb, Germany
Don Chaffin, USA
Kathryn Cormican, Ireland
Andris Freivalds, USA
Ravindra Goonetilleke, Hong Kong
Anand Gramopadhye, USA
Sung H. Han, South Korea
Pheng Ann Heng, Hong Kong
Dewen Jin, P.R. China
Kang Li, USA
Zhizhong Li, P.R. China
Lizhuang Ma, P.R. China
Timo Maatta, Finland
J. Mark Porter, UK
Jim Potvin, Canada
Jean-Pierre Verriest, France
Zhaoqi Wang, P.R. China
Xiugan Yuan, P.R. China
Shao-Xiang Zhang, P.R. China
Xudong Zhang, USA

In addition to the members of the Program Boards above, I also wish to thank the following volunteer external reviewers: Kelly Hale, David Kobus, Amy Kruse, Cali Fidopiastis and Karl Van Orden from the USA, Mark Neerincx and Marc Grootjen from the Netherlands, Wilhelm Kincses from Germany, Ganesh Bhutkar and Mathura Prasad from India, Frederick Li from the UK, and Dimitris Grammenos, Angeliki Kastrinaki, Iosif Klironomos, Alexandros Mourouzis, and Stavroula Ntoa from Greece.

This conference could not have been possible without the continuous support and advise of the Conference Scientific Advisor, Prof. Gavriel Salvendy, as well as the dedicated work and outstanding efforts of the Communications Chair and Editor of HCI International News, Abbas Moallem, and of the members of the Organizational Board from P.R. China, Patrick Rau (Chair), Bo Chen, Xiaolan Fu, Zhibin Jiang, Congdong Li, Zhenjie Liu, Mowei Shen, Yuanchun Shi, Hui Su, Linyang Sun, Ming Po Tham, Ben Tsiang, Jian Wang, Guangyou Xu, Winnie Wanli Yang, Shuping Yi, Kan Zhang, and Wei Zho.

I would also like to thank for their contribution towards the organization of the HCI International 2007 Conference the members of the Human Computer Interaction Laboratory of ICS-FORTH, and in particular Margherita Antona, Maria Pitsoulaki, George Paparoulis, Maria Bouhli, Stavroula Ntoa and George Margetis.

Constantine Stephanidis
General Chair, HCI International 2007

HCI International 2009

The 13th International Conference on Human-Computer Interaction, HCI International 2009, will be held jointly with the affiliated Conferences in San Diego, California, USA, in the Town and Country Resort & Convention Center, 19-24 July 2009. It will cover a broad spectrum of themes related to Human Computer Interaction, including theoretical issues, methods, tools, processes and case studies in HCI design, as well as novel interaction techniques, interfaces and applications. The proceedings will be published by Springer. For more information, please visit the Conference website: http://www.hcii2009.org/

General Chair
Professor Constantine Stephanidis
ICS-FORTH and University of Crete
Heraklion, Crete, Greece
Email: program@hcii2009.org

Table of Contents

Part I: Shape and Movement Modeling and Anthropometry

Simulation of Complex Human Movement Through the Modulation of Observed Motor Tasks 3
Giuseppe Andreoni, Marco Rabuffetti, and Antonio Pedotti

Simulation of Digital Human Hand Postures of Car Controls Using a Data Based Approach 13
Georges Beurier, Nicolas Chevalot, Gilles Monnier, Jules Trasbot, and Xuguang Wang

Human Articulation Efforts Estimation in the Automobile Vehicle Accessibility Movement – A Pilot Study 23
Jean-François Debril, Philippe Pudlo, Mohand Ouidir Ait El Menceur, Philippe Gorce, and François Xavier Lepoutre

Study on the Appraisal Methods of Hand Fatigue 33
Li Ding, Feng Yang, Chunxin Yang, Xiugan Yuan, and Yang Li

Experimental Research on Human Body Motion Simulation Based on the Motion Capture Technology 42
Dayong Dong, Lijing Wang, and Xiugan Yuan

Modeling of Human's Pointing Movement on the Effect of Target Position 48
Junmin Du, Haiwen Shi, and Xiugan Yuan

A Modified Particle Swarm Optimizer Using an Adaptive Dynamic Weight Scheme 56
Shu-Kai S. Fan and Ju-Ming Chang

Prediction of Discomfort During Arm Movements 66
Florian Fritzsche and Heiner Bubb

A Motion Compensated De-interlacing Algorithm for Motive Object Capture 74
Lei Gao, Chao Li, Chengjun Zhu, and Zhang Xiong

Color 3D Digital Human Modeling and Its Applications to Animation and Anthropometry 82
Bao-zhen Ge, Qing-guo Tian, K. David Young, and Yu-chen Sun

Advanced Human Body and Head Shape Representation and Analysis ... 92
Afzal Godil

Hand Grasping Motion Simulation for Astronauts Training ... 101
Qiang Guo and Yuqing Liu

Predefined Manikins to Support Consideration of Anthropometric Diversity by Product Designers ... 110
Dan Högberg and Keith Case

Comparison of Human and Machine Recognition of Everyday Human Actions ... 120
Trevor D. Jones, Shaun W. Lawson, David Benyon, and Alistair Armitage

A New Decoding Algorithm in MIMO-ZP-OFDM Systems ... 130
Rui Kang, ChaoXiong Xu, Hao Chen, and HongBo Xu

Computer Graphic Modeling and Simulation of Human Musculoskeletal System for Biomechanical Research ... 136
Yoon Hyuk Kim

The Strength Factor in Digital Human Modeling and Simulation: A Case for a New Framework ... 144
Kang Li and Xudong Zhang

The Effects of the False Vocal Fold Gaps in a Model of the Larynx on Pressures Distributions and Flows ... 147
Sheng Li, MingXi Wan, and SuPin Wang

A Robust Algorithm for a System Identification Approach to Digital Human Modeling: An Application to Multi-fingered Hand Movement ... 157
Kang Li, Sang-Wook Lee, and Xudong Zhang

Mathematical Methods for Shape Analysis and form Comparison in 3D Anthropometry: A Literature Review ... 161
Jianwei Niu, Zhizhong Li, and Gavriel Salvendy

A Case Study of Multi-resolution Representation of Heads ... 171
Jianwei Niu, Zhizhong Li, and Gavriel Salvendy

The Application of Kane Equation in the Impact Prediction of Human Motion ... 179
Mu Qiao, Chunxin Yang, and Xiu-gan Yuan

A Fast Motion Estimation Algorithm for H.264 ... 189
Jianbin Song, Bo Li, and Qinglei Meng

Validating Optical Motion Capture Assessments of the Dynamic Aspects of Work 197
Jackie Sutherland and Vincent Duffy

Modeling Human Bipedal Navigation in a Dynamic Three Dimensional Virtual Environment 205
Mark.D. Thomas, Daniel W. Carruth, Bryan Robbins, John A. McGinley, and Alex Morais

A Data-Based Modeling Approach of Reach Capacity and Discomfort for Digital Human Models 215
Xuguang Wang, Elodie Chateauroux, and Nicolas Chevalot

Motion Retrieval Based on Temporal-Spatial Features by Decision Tree 224
Jian Xiang and HongLi Zhu

Motion Retrieval Based on an Efficient Index Method for Large-Scale Mocap Database 234
Jian Xiang and Hongli Zhu

Modeling of Layered Fuzzy Facial Expression Generation 243
Yu-Li Xue, Xia Mao, Zheng Li, and Wei-He Diao

An Inverse Dynamical Model for Slip Gait 253
Jiankun Yang, Dewen Jin, Linhong Ji, Jichuan Zhang, Rencheng Wang, Xin Fang, and Dawei Zhou

Redundant Muscular Force Analysis of Human Lower Limbs During Rising from a Squat 259
Yiyong Yang, Rencheng Wang, Ming Zhang, Dewen Jin, and Fangfang Wu

Optimal Control and Synergic Pattern Analysis of Upper Limb Reaching-Grasping Movements 268
Yiyong Yang, Rencheng Wang, Ming Zhang, Dewen Jin, and Fangfang Wu

Low Cost 3D Shape Acquisition System Using Strip Shifting Pattern 276
Li Yao, Lizhuang Ma, Di Wu

Automatic Joints Extraction of Scanned Human Body 286
Yong Yu, Zhaoqi Wang, Shihong Xia, and Tianlu Mao

Wavelet Transform and Singular Value Decomposition of EEG Signal for Pattern Recognition of Complicated Hand Activities 294
Xiaodong Zhang, Weifeng Diao, and Zhiqiang Cheng

Capturing 3D Human Motion from Monocular Images Using Orthogonal Locality Preserving Projection 304
Xu Zhao and Yuncai Liu

Human Motion Simulation and Action Corpus 314
Gang Zheng, Wanqing Li, Philip Ogunbona, Liju Dong, and Igor Kharitonenko

Part II: Building and Applying Virtual Humans

User Experience Quality: A Conceptual Framework for Goal Setting and Measurement 325
Russell Beauregard and Philip Corriveau

Integrating Perception, Cognition and Action for Digital Human Modeling 333
Daniel W. Carruth, Mark D. Thomas, Bryan Robbins, and Alex Morais

The Usability of Metaphors with Different Degree of Abstract in Interface Design 343
Ming-Chuen Chuang and Inglen Lo

Applications of the Visible Korean Human 353
Min Suk Chung and Jin Seo Park

Role of Humans in Complexity of a System-of-Systems 363
Daniel DeLaurentis

Digital Human Modeling for Product Lifecycle Management 372
H. Onan Demirel and Vincent G. Duffy

Bridging the Gap: Exploring Interactions Between Digital Human Models and Cognitive Models 382
Robert G. Feyen

Translating User Experience to Requirements 392
Ji Gao, Glen Anderson, Brian W. Bramlett, Ryan Palmer, and Delbert Marsh II

Involving Engineers in User Research and User Experience Design of ICT for China 399
Chaoyu Huang and Huogao He

Shape Analysis of Human Brain with Cognitive Disorders 409
Tianzi Jiang, Feng Shi, Wanlin Zhu, Shuyu Li, and Xiaobo Li

Finding Origin Points for New Coordinate System Suitable for Sign Animation 415
Tomohiro Kuroda, Kazuya Okamoto, Ryo Kitauchi, Tadamasa Takemura, Naoki Ohboshi, and Hiroyuki Yoshihara

User Experience Modeling and Enhancement for Virtual Environments That Employ Wide-Field Displays 423
James J.W. Lin and Donald E. Parker

AIPlayer: A Platform of Intelligent Simulation of Virtual Human in Virtual Environment 434
JianGuo Liu, YanSheng Lu, and JiuYun Chen

Understanding RUTH: Creating Believable Behaviors for a Virtual Human Under Uncertainty 443
Insuk Oh and Matthew Stone

Automatic, Body Measurements Based Generation of Individual Avatars Using Highly Adjustable Linear Transformation 453
Andreas Volz, Rainer Blum, Sascha Häberling, and Karim Khakzar

A Composite Measure for the Evaluation of Mental Workload 460
Lee-ming Wang, Vincent G. Duffy, and Yingzi Du

Brain-Computer Interfaces Based on Attention and Complex Mental Tasks 467
Jue Wang, Nan Yan, Hailong Liu, Mingyu Liu, and Changfeng Tai

Fractal Modeling of Human Psychomotor Skills Acquisition Process 474
James Wanliss, Dahai Liu, Vadim Uritsky, and Michael Wuergler

Effect of Noise-Enhanced on the Balance Control Ability in Older Adults 483
Fangfang Wu, Rencheng Wang, Dewen Jin, Xiao Hu, Yiyong Yang, JiChuang Zhang, and Noboru Youshimura

Development of the Virtual-Human Santos^{TM} 490
Karim Abdel-Malek, Jingzhou Yang, Joo H. Kim, Timothy Marler, Steve Beck, Colby Swan, Laura Frey-Law, Anith Mathai, Chris Murphy, Salam Rahmatallah, and Jasbir Arora

Validation of Predicted Posture for the Virtual Human Santos^{TM} 500
Jingzhou Yang, Salam Rahmatalla, Tim Marler, Karim Abdel-Malek, and Chad Harrison

Modeling of Multi-organization Performance for Emergency Response 511
Yufei Shu and Kazuo Furuta

Estimating Mental Fatigue Based on Multichannel Linear Descriptors Analysis 521
Chong Zhang, Chongxun Zheng, Xiaomei Pei, and Xiaolin Yu

Chinese Visible Human Data Sets and Their Applications 530
Shao-Xiang Zhang, Pheng Ann Heng, Zheng-Jin Liu, Li-Wen Tan, Ming-Guo Qiu, Qi-Yu Li, Rong-Xia Liao, Kai Li, Gao-Yu Cui, Yan-Li Guo, and Yong-Ming Xie

Visualizing User Experience Through "Perceptual Maps": Concurrent Assessment of Perceived Usability and Subjective Appearance in Car Infotainment Systems 536
Xianjun Sam Zheng, James J.W. Lin, Salome Zapf, and Claus Knapheide

Part III: Medical and Rehabilitation Applications

Digital Human Modelling: A Global Vision and a European Perspective 549
Gordon J. Clapworthy, Peter Kohl, Hans Gregerson, S.R. Thomas, Marco Viceconti, D.R. Hose, D. Pinney, John Fenner, K. McCormack, P. Lawford, S. Van Sint Jan, S. Waters, and P. Coveney

ICT Methodologies to Model and Simulate Parts of Human Body for Prosthesis Design 559
Giorgio Colombo, Stefano Filippi, Paolo Rissone, and Caterina Rizzi

Simulating Cancer Radiotherapy on a Multi-level Basis: Biology, Oncology and Image Processing 569
Dimitra D. Dionysiou, Georgios S. Stamatakos, and Kostas Marias

Applied User Performance Modeling in Industry – A Case Study from Medical Imaging 576
Marcela Esteves, Tobias Komischke, Salome Zapf, and Antje Weiss

Computer Aided Lumbar Support Design and Application 586
Susanne Frohriep and Jan Petzel

Facilitating Pronunciation Skills for Children with Phonological Disorders Using Human Modelling 595
Jennifer George and Paul Gnanayutham

FPF-SB: A Scalable Algorithm for Microarray Gene Expression Data Clustering 606
Filippo Geraci, Mauro Leoncini, Manuela Montangero, Marco Pellegrini, and M. Elena Renda

A Finite Element 3D Model of in Vivo Human Knee Joint Based on MRI for the Tibiofemoral Joint Contact Analysis 616
Zhixiu Hao, Dewen Jin, Yu Zhang, and Jichuan Zhang

Advances in Visible Human Based Virtual Medicine 623
Pheng Ann Heng

Finite Element Analysis of a Six-Component Force Sensor for the Trans-Femoral Prosthesis 633
Xiao Hu, Rencheng Wang, Fangfang Wu, Dewen Jin, Xiaohong Jia, Jichuan Zhang, Fuwen Cai, and Shuangxi Zheng

The Influence of Shoe-Heel Height on Knee Muscle Activity of Transtibial Amputees During Standing 640
Xiaohong Jia, Jichuan Zhang, Rencheng Wang, Lidan Fang, Dewen Jin, and Ming Zhang

3D Modeling of the Vessels from X-Ray Angiography 646
Na-Young Lee, Gye-Young Kim, and Hyung-Il Choi

Finite Element Modeling to Aid in Refining the Rehabilitation of Amputees Using Osseointegrated Prostheses 655
Winson Lee, Laurent Frossard, Nicola Cairns, Rickard Branemark, John Evans, Clayton Adam, and Mark Pearcy

Actions of an External Electrical Shock on Human Atrial Excitation – A Computer Model Study 659
Jihong Liu, A.V. Holden, and Henggui Zhang

Study and Application of Medical Image Visualization Technology 668
Jihong Liu, Weina Ma, Fei Liu, Ying Hu, Jinzhu Yang, and Xinhe Xu

A Method for Gene Identification by Dynamic Feature Choosing 678
Jia-Wei Luo, Li Yang, and Xi-Zhen Zhang

Latest Development of an Interventional Radiology Training Simulation System: *NeuroCath* 684
Xin Ma

Methodologies to Evaluate Simulations of Cardiac Tissue Abnormalities at a Cellular Level 694
Nicos Maglaveras and Ioanna Chouvarda

Multi-level Analysis and Information Extraction Considerations for Validating 4D Models of Human Function 703
Kostas Marias, Dimitra D. Dionysiou, Georgios S. Stamatakos, Fotini Zacharopoulou, Eleni Georgiadi, Thanasis Margaritis, Thomas G. Maris, and Ioannis G. Tollis

Clinical Patient Safety—Achieving High Reliability in a Complex System ... 710
Kathryn Rapala and Julie Cowan Novak

Novel Methods for Human-Computer Interaction in Multimodal and Multidimensional Noninvasive Medical Imaging ... 717
Tomasz Soltysinski

A Hybrid AB-RBF Classifier for Surface Electromyography Classification ... 727
Rencheng Wang, Yiyong Yang, Xiao Hu, Fangfang Wu, Dewen Jin, Xiaohong Jia, Fang Li, and Jichuan Zhang

An Epileptic Seizure Prediction Algorithm from Scalp EEG Based on Morphological Filter and Kolmogorov Complexity ... 736
Guanghua Xu, Jing Wang, Qing Zhang, and Junming Zhu

A New Virtual Dynamic Dentomaxillofacial System for Analyzing Mandibular Movement, Occlusal Contact, and TMJ Condition ... 747
Chi Zhang, Lei Chen, Fengjun Zhang, Hao Zhang, Hailan Feng, and Guozhong Dai

Mechanism of Bifurcation-Dependent Coherence Resonance of Excitable Neuron Model ... 757
Guang-Jun Zhang, Jue Wang, Jian-Xue Xu, Xiang-Bo Wang, and Hong Yao

An Integrated Approach for Reconstructing Surface Models of the Proximal Femur from Sparse Input Data for Surgical Navigation ... 767
Guoyan Zheng and Miguel A. González Ballester

Part IV: Industrial and Ergonomic Applications

Future Applications of DHM in Ergonomic Design ... 779
Heiner Bubb

The Design and Exploitation of Visual Feedback System for Rowing ... 794
Chunmei Cao, Chuncai Wang, Linhong Ji, Zixi Wang, and Xiaoping Chen

A Method for Garment Pattern Generation by Flattening 3D Body Scan Data ... 803
Young Lim Choi, Yunja Nam, Kueng Mi Choi, and Ming Hai Cui

Human Body Modeling for Riding Comfort Simulation ... 813
Hyung Yun Choi, Kyung Min Kim, Jiwon Han, Sungjin Sah, Seok-Hwan Kim, Su-Hwan Hwang, Kwang No Lee, Jong-Kweon Pyun, Nicole Montmayeur, Christian Marca, Eberhard Haug, and Inhyeok Lee

Applications of Digital Human Modeling in Industry 824
H. Onan Demirel and Vincent G. Duffy

A Computer-Aided Ergonomic Assessment and Product Design System Using Digital Hands 833
Yui Endo, Satoshi Kanai, Takeshi Kishinami, Natsuki Miyata, Makiko Kouchi, and Masaaki Mochimaru

Human Age and Vehicle Speeds Affect on Vehicle Ingress Motion Pattern 843
Hanson Lars, Yong Li, and Falkmer Torbjörn

Digital Human Model Based Participatory Design Method to Improve Work Tasks and Workplaces 847
Kaj Helin, Juhani Viitaniemi, Jari Montonen, Susanna Aromaa, and Timo Määttä

Investigation on Ergonomics Characteristics of Protective Clothing Based on Capture of Three-Dimensional Body Movements 856
Huimin Hu, Li Ding, Chunxin Yang, and Xiugan Yuan

Strategy to Operate Cylindrical Interface-Operation Difference According to the Dimension of the Cylinder and That of the Hand 865
Ohki Kanezashi, Natsuki Miyata, and Jun Ota

Modeling of Human Head for Custom Wig Production 874
Youngjun Kim, Jungbum Cho, Bohyun Kim, and Kunwoo Lee

Evaluation of Navy Shipboard Habitability for a Warship Design Using Human Model 884
Hongtae Kim, Jin H. Park, Hojin Hwang, and Chang-Min Lee

Two Vibration Modes of a Human Body Sitting on a Car Seat- The Relationship Between Riding Discomfort Affected by the Material Properties of the Seat Cushion and the Two Vibration Modes 894
Mitsunori Kubo, Fumio Terauchi, and Hiroyuki Aoki

The Effects of Human Interaction on Biometric System Performance ... 904
Eric P. Kukula, Stephen J. Elliott, and Vincent G. Duffy

Design and Realization of Synthesis Assessment System for Cockpit Ergonomics 915
Yinxia Li, Kaiying La, and Xiugan Yuan

Development of a Test-Bed for Synthetical Ergonomics Evaluation of Pilot 923
Wei Liu, Xiu-gan Yuan, Zhong-qi Liu, Rui Ma, and Wei-yong Kang

Experimental Research of Evaluation of Temperature Ergonomics of EVA Spacesuit Glove 929
Han Longzhu and Xiu-gan Yuan

Digital Humans for Virtual Assembly Evaluation 939
Dimitris Mavrikios, Menelaos Pappas, Marios Kotsonis, Vassiliki Karabatsou, and George Chryssolouris

Foot Digitalization for Last Design and Individual Awareness of Personal Foot Characteristics 949
Paolo Olivato, Manuela Morricone, Enrica Fubini, and Alessandra Re

Can We Use Technology to Train Inspectors to Be More Systematic? ... 959
Sajay Sadasivan and Anand K. Gramopadhye

Enhancing Sense of Reality by Efficient and Precise Collision Detection in Virtual Environments 969
Chuan-Jun Su

Dynamic Generation of Human-Populated VR Models for Workspace Ergonomic Evaluation 979
Tien-Lung Sun, Wen-Yang Feng, and Chin-Jung Chao

Effecting Validity of Ergonomics Analysis During Virtual Interactive Design 988
Renran Tian, Vincent G. Duffy, and John McGinley

Advanced Technology Training for Operating the Microlithography Panel Printer 998
Bharat Upadrasta, Mohammad T. Khasawneh, and Sarah S. Lam

Ramsis – The Leading Cad Tool for Ergonomic Analysis of Vehicles 1008
Peter van der Meulen and Andreas Seidl

Using Multimodal Technologies to Enhance Aviation Maintenance Inspection Training 1018
Carl Washburn, Paris Stringfellow, and Anand Gramopadhye

Colored Petri Net Based Formal Airport Control Model for Simulation and Analysis of Airport Control Processes 1027
Bernd Werther, Christoph Moehlenbrink, and Michael Rudolph

Research on Modeling of Complicate Traffic Simulation System 1037
Jiankun Wu, Linpeng Huang, Jian Cao, Minglu Li, Dejun Wang, and Mingwen Wang

Design and Implementation of Ergonomics Evaluation System of 3D Airplane Cockpit 1047
Libo Zhang, Xiugan Yuan, Lijing Wang, and Dayong Dong

A Novel Method for Cloth-Body Collision Detection 1056
Xiaolong Zhu, Shihong Xia, Yong Yu, and Tianlu Mao

Author Index 1065

Part I

Shape and Movement Modeling and Anthropometry

Simulation of Complex Human Movement Through the Modulation of Observed Motor Tasks

Giuseppe Andreoni[1], Marco Rabuffetti[2], and Antonio Pedotti[1]

[1] Politecnico di Milano, Bioengineering Dept., P.zza L. Da Vinci. 32,
20133 Milan, Italy
giuseppe.andreoni@polimi.it
[2] Fond. Do C. Gnocchi IRCCS, Centro di Bioingegneria., via Capecelatro. 66,
20148 Milan, Italy

Abstract. A method for the simulation of human movements driven by real data and correlated with modification of constraints in the external environmental is presented. It was applied to the simulation of the car ingress changing the configuration of the doorway to check early on in the design the man-machine-interface requirements for choosing the best ergonomic solution among different alternative solutions without the physical construction of prototypes. The method for the simulation of the movement is based on the modulation of a real measured performance recorded through an opto-electronic system for motion analysis. The algorithm implements a multifactorial target function to solve the redundancy problem. The reliability of the method was tested through the comparison of simulated and real data showing promising developments in ergonomics.

Keywords: movement, simulation, pattern modulation, ergonomics, virtual prototyping.

1 Introduction

The Central Nervous System (CNS) controls any human motor performance by setting sequences of muscular activation in respect of the several constraints (environmental as well as biomechanical) acting on the human [1], [2], [3]. Controlled by the CNS, the biomechanical system is characterised by a very high degree of complexity so that the mechanical structure is redundant thus allowing for different motor strategies in realising the same motor task [4], [5]. These strategies may have completely different joint motion patterns, or may represents slightly different modulations of a recognisable unique motor strategy. The study of the neural control of human movement aims to identifying the criteria assumed by the CNS in the design of the muscle activation patterns [1], [5]. These studies have considered mostly simple movements (i.e. movements involving few anatomical segments, the joints connecting them and the muscles acting on that part), but the same concepts have been adopted in the explanation of complex movements of the whole body.

The quantitative approach in the measurement of human motion as well as the mechanical modeling of the human body lead to a mathematical interpretation of the

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 3–12, 2007.

CNS criteria [2], [6]: some "cost functions" have been defined and it is assumed that the CNS controls the several Degrees of Freedom (DoF) in the anatomical structure in order to minimise those costs through a sort of optimisation algorithm.

The contribution of the several authors consisted in explicitly introduce in those cost functions all the elements that are thought to be determinant in the definition of a motor strategy: the anthropometric features, the physiological/pathological conditions, the level of motor ability, a preliminary phase of specific training, the energetic cost, the respect of the implicit and explicit constraints, the feedback [3], [7], [8], [9], [10]. In particular, in synthetic generation of the movement the three main categories of ill-posed problems can be identified: the simulation of trajectories, the inverse kinematics and the inverse dynamics. For the system redundancy previously described, in human motion simulation, the main problem to face is that the solution to the problem is not unique. The simplest approach to solve and to find an unique solution is the introduction of a target function with a performance index to be optimized. The criterions for the optimization of this performance index are based on the characteristics of the motion or on the mechanical properties of the muscle-skeletal system implementing one of the followings functions [8], [9], [10], [11], [12], [13], [14]: a) minimum jerk; b) minimum spatial deviation; c) minimum change of the muscular tension; d) minimum torque change.

In the present paper a real data-driven method for the simulation of complex motor tasks (when many approaches fail) is presented through the application example of the analysis of the car ingress-egress movements.

2 Materials and Methods

The method requires the availability of an experimental measure of a reference similar movement: the simulation of the new movement according to virtual new constraints is performed by modulating the measured movement.

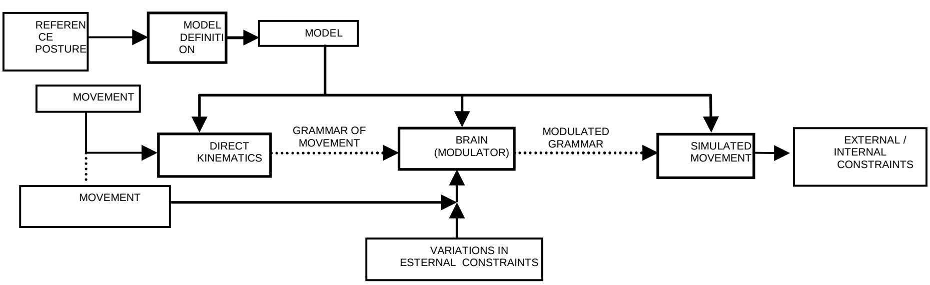


Fig. 1. The flow-chart diagram of the method for the virtual synthesis of the movement

The approach we propose here for the simulation consists in the modulation of the experimental grammar of movement, i.e. the set of vectors of all the kinematic variables describing the observed movement. This means that the simulation is the modification of a reference recorded movement according to new virtual constraints, through an algorithm that finds the solution of the redundancy problem by minimizing a target function. The purpose is to obtain a synthetic movement to be coherent with all the given intrinsic (i.e. the DoF of the model) and extrinsic (i.e. external obstacles) constraints, and indistinguishable from the correspondent movement performed in the reality for an external observer.

2.1 The Biomechanical Model

Human body was modeled with 16 anatomical segments: head, arms, forearms, hands, thighs legs, feet, thorax, abdomen and pelvis. Markers were positioned on the head, on the trunk, on the lower limbs while upper limbs were not considered and so they were thought in a rest position along the body. The 11 markers are: for the head, front, vertex and occiput; for the shoulders, right and left acromion; for the hips, the right and left great trochanter; for the knees, lateral condyles of femuri; ankle.

The experimental protocol was integrated by the measurement of some anthropometric parameters: weight (kg); height (cm); arm length; forearm length; sternum height; pelvis width; pelvis height in sitting position; upper iliac crest point height from sitting plane; knee diameter; ankle diameter.

Each segment was characterized by fixed anatomical axes and constant inertial geometric properties: the segment mass was concentrated in one fixed point. The mass of each body segment was estimated by means of regression equations [15]. The total body centre of mass (COM) position was computed as the weighted mean of each body segment centre of mass position.

2.2 The Experimental Protocol

Three-dimensional measurement of human movement is the detection of the trajectories of several points that identify the positions of the body segments in space through a motion capture optoelectronic system (ELITE by BTS, Milan, Italy) recording the position of spherical or hemispherical passive markers fixed onto the body. The vehicle was implemented by a car simulator provided by all the structural elements required to reproduce a real car.

Six healthy subjects (five male and one female), university students, took part in the study (age: 26 ÷ 31 years, height: 170 ÷ 183 cm and weight: 50 ÷ 94 kg).

The protocol consisted in ingress and egress with reference to left back seat. When entering, the subject started in the standing position near the car mock-up and ended when in the sitting position. In the egress, the movement went in the opposite way. The subject chose velocity and strategy of movement. During the experiments, the door height was progressively lowered with steps of 2 cm starting from higher quote of 154 cm into other 8 fixed positions. For every set-up the subject did a sequence of 10 trials, 5 ingresses and 5 egresses, so that in total he made 80 trials. Data of homogeneous trials (same subject, same vehicle set-up) were time normalized and averaged for the statistical analysis.

2.3 The Motor Task Model (Grammar of Movement)

In the case of car ingress, the main consideration is that any subject has to lower the COM while translating from out to the interior of the vehicle. A similar but opposite consideration can be done for egress. The events describing the motor task are the maximum (start) and the minimum (end) of the vertical co-ordinate of the COM; similarly, but in inverse order, for egress.

The grammar of movement is set of *n* vectors containing the values of the *n* DoF of the model. These DoF are the angles between connected segments - (n – 6) relative rotations – and the position and the orientation in the 3D space of the segment that is the root of the kinematic chain (6 DoF). A grammar is computed from an observed real movement considering the corresponding model through an optimization procedure minimizing frame by frame the differences between the given experimental points and the positions of the manikin (model).

2.4 The Simulation as a Modulation of an Observed Movement

The model and the grammar of movement converge in the algorithm that integrates the intrinsic limits of the model and the external environmental constraints and mathematically implements the criterions for the solution of the redundancy. A grammar of movement can be divided in its linear and non-linear components (Fig. 2.a and 2.b) defined according to the following equations:

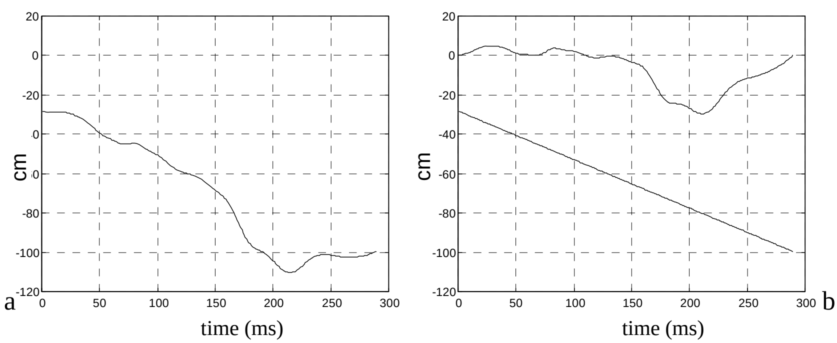


Fig. 2. The separation of a variable Y (a) in its the Linear Component YL and Non-Linear Component YNL (b)

Grammar of movement: $Y_i = YL_i + YNL_i \qquad \forall i \in [1, N]$ (1)

linear component: $YL_i = Y_1 + \frac{Y_N - Y_1}{N - 1}(i - 1)$ (2)

non linear component: $YNL_i = Y_i - YL_i$ where i is the time i = 1, ,T. (3)

The separation of the linear and non-linear components is consistent with the consideration that in general a movement does not modify the starting and final posture such as in the example of the ingress/egress to a vehicle in which a variation of the geometry of the doorway does not modify the initial (standing) and final (sitting) positions, but it influences the intermediate trajectory. In this sense, the informative content of the reference grammar is exclusively in its non-linear component because, during the simulation, the linear component is defined while the non-linear component is modulated. The only requirement for the reliability of the presented method is that the conditions of the reference grammar should be similar to those of the virtual conditions that are going to be verified.

The simulation of the movement is realised as follows: considering the general kinematic variable Z^s of the simulated movement, the linear component ZL^s_i is the same one of the observed data ZL^o_i: $ZL^s_i = ZL^o_i$

To simulate a movement modulating the non-linear component of its grammar of movement means to define a constant multiplying factor of the reference non-linear component YNL_i:

$$ZNL^s{}_i = k \cdot ZNL^o{}_i \qquad\qquad i = 1,..,N \tag{4}$$

and, finally it is possible to obtain the simulated vector of values of the variable Z through the sum of the two components:

$$Z_i = ZL^s{}_i + ZNL^s{}_i = ZL^o{}_i + k \cdot ZNL^o{}_i \tag{5}$$

Therefore, the three unknown parameters of the function are the extremes of the linear component, and the coefficient of the non-linear component to be computed considering the constraints along the trajectory. But, given the starting and final postures from the observed movement, the problem of the simulation of the movement is reduced to the calculation of the vector of the coefficients k_i K = [$k_1 \ldots k_i \ldots k_n$], i.e. of the multiplier factors for each non-linear component of all the DoF.

The algorithm for the virtual synthesis of the movement integrates information about the subject, that is the model and its intrinsic limits (DoF and RoM); the new geometry of the environment (the car in the proposed application); the grammars of movement, i.e. the experimentally observed motor patterns.

The target function consists in the product of simple sub-functions that singularly implement the following different aspects:

f_1: clearance, e.g. no interference between subject and environment;

f_2: constraints on the absolute positions of segments, e.g. respect of environmental constraints;

f_3: constraints on free coordinates, e.g. respect of RoM;

f_4: attraction towards a reference posture, e.g. the standing.

Therefore, the global target function F is:

$$F = \prod_{j=1}^{4} f_j \tag{6}$$

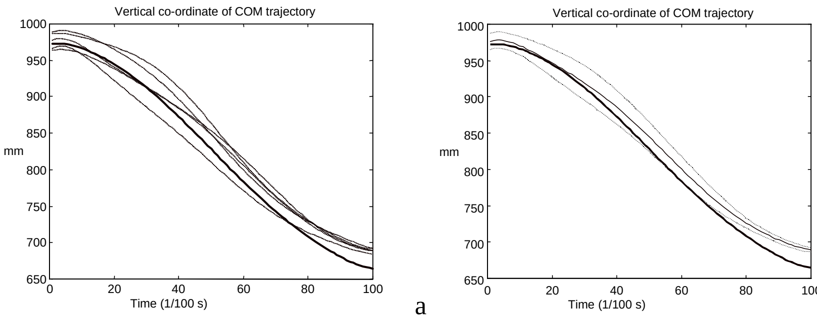
The algorithm finds the *n* values k_i that minimize the function F.

After the calculation of the coefficients k_i of the non linear component, the algorithm provides the vectors of the new DoFs (i.e. the virtual or simulated grammar) to be imposed to the biomechanical model to obtain the simulated movement.

3 Results

The validation of the procedure of simulation of the movement was realized using the central configuration (corresponding to the real condition in the commercial car model) as reference for the extraction of the grammar of movement. Subsequently tests of ingress and egress movements with the roof respectively lowered of 6 cm and raised up of 8 cm were used for comparison with the simulated movements. Then the kinematic data of the variables, and the trajectories of the COM were compared (Fig. 3.a and 3.b).

Also by visual inspection it is evident how the algorithm did not change the global motor strategy adopted by the subject when in presence of non-extreme constraints, as in the verified situations that represent general and common situations of typical cars' doorways of the market central segments (Fig. 4). Data analysis did not reveal differences between the two joint angular patterns (simulated and real ones).



a b

Fig. 3. a) vertical co-ordinate YCOM trajectories for a subject of 176 cm in 5 observed tests (dotted lines) and in the simulation (solid line) of the car ingress with the roof height lowered of 6 cm; b) the average and standard deviation of the vertical co-ordinate COM trajectory for the same subject in the 5 trials (solid and dotted lines) and in the simulation (solid wider line)

In the last part of the movement a lowering is observed for the vertical coordinate of the COM in the simulation in comparison with the observed tests: probably it happens because the modulation coefficient is constant along all the movement so that to pass through the doorway and under the superior region of the doorway, all the body is somehow lowered also in the intermediate positions and therefore the COM reaches lower final values. With a more detailed model of the environment with the seat and of the human body with the ‘muscular envelope’ a better results could be achieved. Also the 20 mm threshold for the clearance factor could be over-estimated as shown in the figure 4 where in correspondence of the instant of maximum head interference in the real movement the head is closest to the doorway than the set value.

To quantify the correspondence of the simulated movement to the real one we introduced a statistical index of coherence computed for the two main kinematic variables of the movement we analysed (vertical co-ordinate of COM - YCOM and Left Hip Flexion – LHF). Defined the range of natural variability of a generic variable X as the interval average ± 2 standard deviations computed on 5 observations of the same movement, the PIX index is computed as follows:

$$PI_i = \begin{cases} 1 & \text{if the simulated angle is in the range of the natural variability} \\ 0 & \text{if the simulated angle is not in the range of the natural variability} \end{cases} \qquad (7)$$

and finally $PI_X = \sum_{i=1}^{100} PI_i$ where *i* is the normalized time. (8)

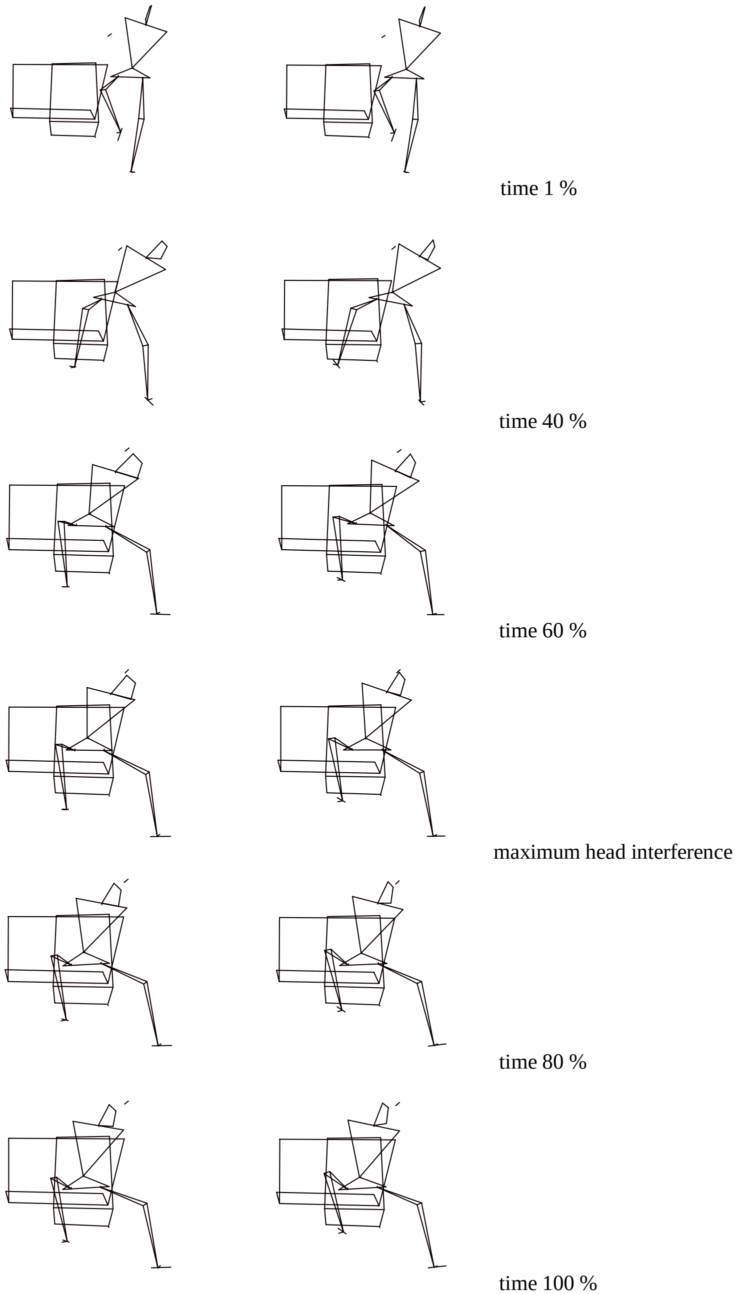


Fig. 4. Comparison of the simulated (on the left) with the observed movement (on the right)

Table 1. The coherence index of the simulation with respect to the actual movement ant its natural variability for the two main kinematic variables: YCOM = vertical co-ordinate of COM, LHF = Left Hip Flexion

Subject	PIYCOM (%)	PILHF (%)
1	76	89
2	80	93
3	79	91
4	81	88
5	87	94
6	83	90
m ± sd	81,0 ± 3.7	90.8 ± 2.3

The aim is to verify how much the simulation is in the range of the natural movement repetition of the same task. The results are shown in table 1.

However, the simulated data were in the natural variability of real movements for 91% of their execution concerning the anatomical angles and 81% of the vertical co-ordinate of COM trajectory: so we can say that the algorithm provides a correct and likely simulation of the motor task.

4 Discussion

The obtained movement is intrinsically similar to real movements already observed, thus looking like a real movement. In the spirit of the proposed methodology, the new grammar has not been already observed but, at the same time, is not strictly virtual. It can be concluded that a simulation realized starting from an observed and then modified grammar is virtual, or it has not been ever observed, but it has characteristics of likelihood and naturalness proper of the observable movements, or it could be observed in the future.

The algorithm was developed and validated for the ingress-egress movements to vehicle, but it is possible to calculate the simulation of any movement given a biomechanical model and the observed behaviour in experimental conditions similar to those to be simulated (to conserve the qualitative behaviour), as demonstrated by the application in the simulation of reaching tasks.

Because of the proper sizing of the movement with the anthropometry of the subject, the reference grammar of the movement must be used for simulations of movement of manikins or subjects with a biomechanical model that is anthropometrically similar to the original experimental subject. Therefore it is opportune to have available many different models of movement, each one from an experimental analysis on subjects of different heights (percentiles).

5 Conclusions

In the presented approach for human motion simulation seems to be sufficiently strong and reliable, and to allow an on-line verification of the human-car interaction early on in the design of virtual prototypes or new models.

The main field of application of the method developed in this research is the ergonomic design of environments and products. The simulation of the movement based on real motor strategies that are adapted to a new virtual space configuration, could assume a relevant importance for the on-line ergonomic evaluation during the design process of environments and objects [16], 17], [18], [19], [20]. In fact our approach proposes that the simulated movement has to be not only congruent with the biomechanical model and its constraints but that also with the qualitative human behaviour (i.e. the different motor strategies that can be adopted by different subjects or by the same subject in modified conditions, speed and accelerations, continuity of the movement). So this concept is implemented and mirrored by the built reliable and likely simulation of what happens in the reality, i.e. a faithful reproduction of the man-machine-environment interface.

References

1. Bernstein, N.: Some emergent problems of the regulation of motor acts. In: The Co-ordination and Regulation of Movements, Pergamon Press, New York, USA (1967)
2. Pedotti, A., Crenna, P.: Individual strategies of muscle recruitment in complex natural movements. In: Multiple Muscle Systems, Biomechanics and Movement Organization, Springer, New York, USA (1990)
3. Winter, D.A.: Biomechanics and motor control of human movement. John Wiley and Sons Inc., New York, USA (1990)
4. Crenna, P., Frigo, C., Massion, J., Pedotti, A.: Forward and backward axial synergies in man. Experimental Brain Research 65, 538–548 (1987)
5. Pedotti, A.: A study of motor coordination and neuromuscular activities in human locomotion. Biol. Cybern 26, 53–62 (1977)
6. Pedotti, A., Crenna, P., Deat, A., Frigo, C., Massion, J.: Postural synergies in axial movement: short and long term adaptation. Experimental Brain Research 74, 3–10 (1989)
7. Ito, K.: Dynamic control of the musculoskeletal system. In: Computational Biomechanics, Springer, Tokyo, Japan (1996)
8. Uno, Y., Kawato, M., Suzuki, R.: Formation and control of optimal trajectory in human multijoint arm movement: minimum torque-change model. Biol. Cybern 61, 89–101 (1989)
9. Kawato, M., Maeda, Y., Uno, Y., Suzuki, R.: Trajectory formation of arm movement by cascade neural network model based on minimum torque-change criterion. Biol. Cybern 62, 275–288 (1990)
10. Lee, S., Kil, R.M.: Redundant arm kinematic control with recurrent loop. Neural Networks 7(4), 643–659 (1994)
11. Jordan, M.I., Flash, T., Arnon, Y.: A model of the learning of arm trajectories from spatial deviations. Journal of Cognitive Neuroscience 6(4), 359–376 (1994)
12. Massone, L., Bizzi, E.: A neural network model for limb trajectory formation. Biol. Cybern 61, 417–425 (1989)
13. Kawato, M., Furukawa, K., Suzuki, R.: A hierarchical neural-network model for control and learning of voluntary movements. Biol. Cybern 57, 169–185 (1987)
14. Borghese, N.A., Arbib, M.A.: Generation of temporal sequences using local dynamic programming. Neural Networks 8(1), 39–54 (1995)

15. Zatsiorsky, V., Seluyanov, V.: The mass and inertia characteristics of the main segments of the human body. In: Biomechanics VIII-B, Human Kinetics, Champaign, IL, USA (1983)
16. Dan, M.P.: Using man modelling CAD systems and expert systems for ergonomic vehicle interior design. In: Proceedings of the 13th Triennial Congress of the International Ergonomics Association, Tampere, Finland, June 29-July 4 1997, vol. 2 (1997)
17. Das, B., Sengupta, A.K.: Computer-aided human modelling programs for workstation design. Ergonomics 38(9), 1958–1972 (1995)
18. Haslegrave, C., Holmes, K.: Integrating ergonomics and engineering in the technical design process. Applied Ergonomics 25(4), 211–220 (1994)
19. Miller, J.S.: 3D simulation: a virtual environment for Proactive Ergonomics. In: Proceedings of the 13th Triennial Congress Of the International Ergonomics Association, Tampere, Finland, June 29-July 4 1997, vol. 2 (1997)
20. Wilson, J.R.: Virtual environments and ergonomics: needs and opportunities. Ergonomics 40(10), 1057–1077 (1997)

Simulation of Digital Human Hand Postures of Car Controls Using a Data Based Approach

Georges Beurier[1], Nicolas Chevalot[1], Gilles Monnier[2], Jules Trasbot[3], and Xuguang Wang[1]

[1] INRETS - LBMH, 25 avenue François Mitterrand, Case 24, 69675 Bron, France
[2] Altran Technologies, 58, Bd. Gouvion St Cyr, 75017 Paris
[3] Ergonomie 64270 - 1 Avenue du Golf 78288 GUYANCOURT Cedex
Georges.beurier@inrets.fr

Abstract. This paper introduces a data-based approach to simulate with digital human models hand postures for grasping car control objects, which takes into account hand anthropometry, grasp type and object size. This paper presents more the experimental approach part than the simulation part itself. The simulation part is essentially done by the RPx software. This paper presents mainly a protocol in order to obtain data for the simulation.

Keywords: hand, data based approach, car controls, digital human.

1 Introduction

This study was motivated by the fact that only a limited number of hand postures are generally available in a digital human model. In addition, it is difficult to adapt a predefined hand posture for different hand and object size due to the complexity of the hand. In general, a hand is modeled with more than 20 degrees of freedom. It is not easy to define a desired hand posture by direct manipulation of each joint axis. This is particularly true for the users of digital human models who are most often design engineers and are not hand specialist.

The simulation of hand posture can take many forms. Often, only simulations are used, and the problem is how to be certain that the obtained posture is realist. The simulation can be only cinematic or dynamic like [1]. Here only the kinematical way is considered. In some work specific hand model are used for the hand posture simulation [2]. In our work, a constraint is to compute the posture for a specific commercial digital Human.

In order to constitute a data base of hand postures for car interior design, an experiment was carried out to capture hand postures when grasping different car control commands using an optoelectronic system (VICON). Five voluntary subjects with differently sized hand participated in the experiment. Ten imposed hand postures for each subject were captured when grasping steering wheel, interior mirror, parking brake, push button, two different rotating knobs, gearshift and glove box. These objects were chosen because they have a great diversity of grasping posture, which can be dividing in two main categories: force and precision posture. Prior to the

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 13–22, 2007.

experiment, the hand anthropometric dimensions were measured. In addition, several photographs of the hand with markers attached were taken in a calibrated space in order to determine the local coordinates of the markers in the hand model.

In this paper, we will at first present the method used for posture reconstruction.

In order to modify a pre-defined posture by an inverse kinematic algorithm, the geometric interaction between hand and object was investigated, especially the contact zones between the fingers and the object. From the reconstructed posture these contact zones are identified for each object, by taking the maximum depth of penetration of the fingers in the object. This is due to the fact that the used digital manikin has a rigid body model with regard to its environment. So, using some rules, for each object, it will be possible to determine the object contact zones by taking into account the hand anthropometry, and to determine the relative fingers contact zones. Postural adaptation to a differently sized hand and a differently sized object is ensured by modifying a captured posture so that hand contact points match with those of the object.

The emphasis of the present paper will be put on the whole process of posture data construction and its use for simulation. One example will be shown. Limitations and perspectives of the proposed approach will also be discussed, especially how to associate the hand grasping posture with the arm when simulating a whole body reach posture.

2 Experiments

2.1 Participants

The experiment was done with 5 voluntary subjects (2 women and 3 men) with no sign of neurological or joint impairment. They have been chosen for their anthropometric dimension. For this experimentation, subjects were chosen using the hand length criteria such that the subjects selection is broadest possible. The Fig. 1 shows the definition of our direct hand dimension measurement. For this experiment, and for this paper, only the right hand was considered.

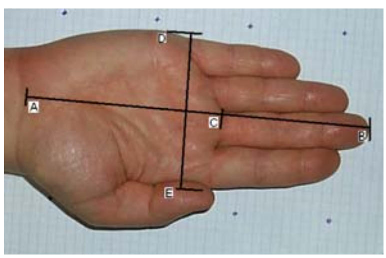


Fig. 1. Hand measurement

The direct hand measurement was:

- The hand length AB
- The palm length AC
- The palm width DE

The following table gives the hand dimensions of the selected subjects (for the first line M means Men and W means women).

Table 1. Subject hand dimension

Subject #	M00	M01	W02	M03	W04	Mean	Std dev
Hand length (mm)	174	198	185	180	159	179.2	14.3
Palm width (mm)	78	89	76	81	68	78.4	7.6
Palm length (mm)	94	109	102.5	101	86	98.5	8.8

2.2 Apparatus

This study take place in the car industry context, thus we have used a physical mock-up, already used in other experiments. This allows us to place the subjects in a driving posture, and to present the object to grasp in their reach spaces.

Fig. 2. Parametric physical mock-up

The objects used were selected from these which can be found in a car. They are fixed on a movable small beam, and they are located with small markers. The following Fig. 3 illustrates these objects.

2.3 Procedures

This part describes the experimental path followed by all subjects. The first step after the hand measurement done during the selection phases was to photocopy the both side of the hand in order to be able to determine the phalanges length.

The second step was the placement of the makers. The makers were located on the five metacarpo-phalangeal joints (MCP), on the ten inter-phalangeal joints, on the five nails and two on the wrist like in the following figure (Fig. 4). Some markers were placed on the arm but they weren't used.

In the third step, the subject takes place in the mock-up and places his right hand in a calibration tools (see Fig. 5) with a flat hand posture. During an acquisition with the optoelectronic system, some photographs were taken from orthogonal point of view.

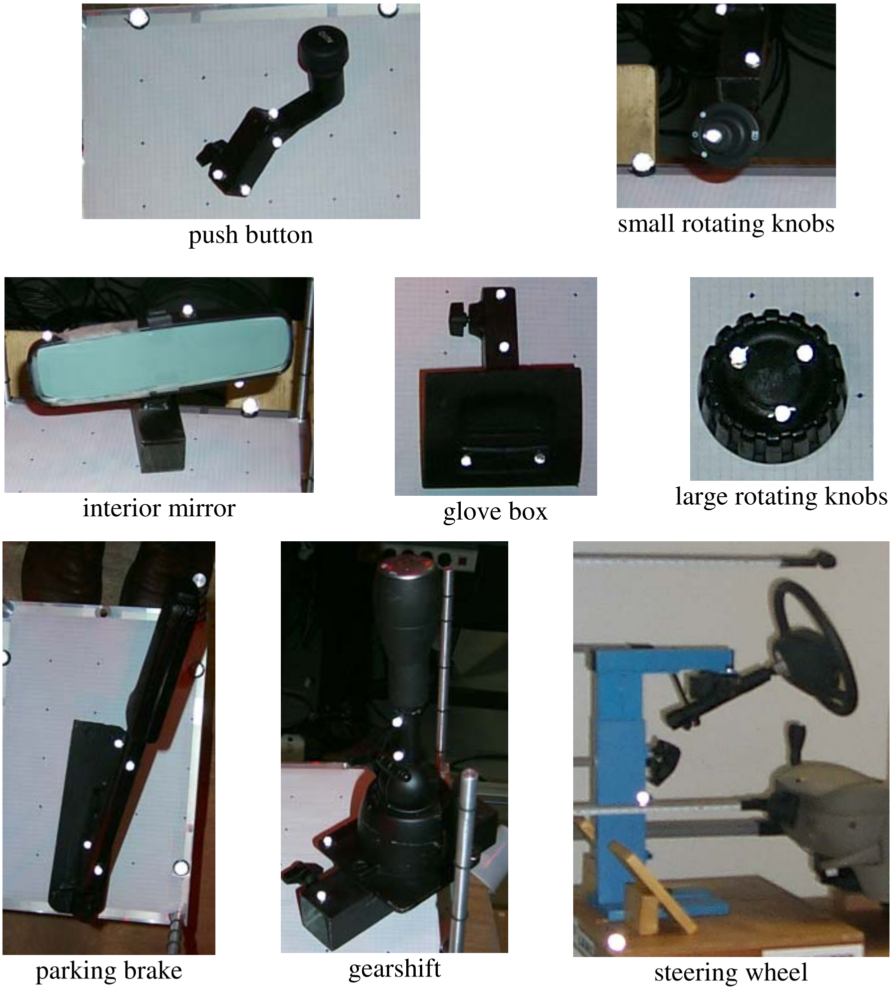


Fig. 3. Selected objects

Fig. 4. Marker placement on the hand

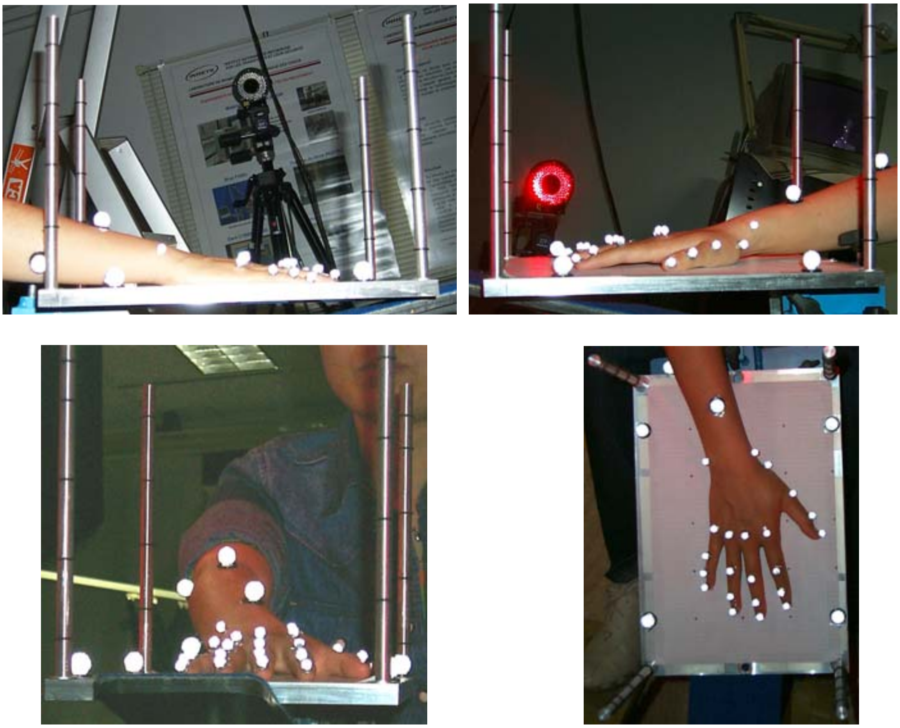


Fig. 5. Hand photographs and the calibration tool

The fourth step is the experiment itself, for each object each subject has to do three times the motion. An acquisition with the motion recorder was done for each motion. Generally, the motion consigns have three defined posture phases, the initial posture, the action postures (~10s), and finally the subject come back to the initial posture. The action postures are described bellow in italic for each asked motion. There can be several postures by objects.

- push button 1 finger
  10s with a contact without pushing + 10s with a maximum pushing contact force
- small rotating knobs 3 fingers
  10s contact + turning to the right + 10s contact
- interior mirror 5 fingers
  10s contact + pushing action + 10s contact
- steering wheel 5 fingers the thumb along the wheel
  10s contact + turning the wheel to the left + 10s contact
- steering wheel 5 fingers the thumb around the wheel
  10s contact + turning the wheel to the left + 10s contact
- parking brake 5 fingers the thumb along the parking brake
  10s contact + pulling the parking brake up + 10s contact

- gearshift 5 fingers the palm in contact with the top
  10s contact
- glove box 4 fingers
  10s contact + opening + 10s contact
- large rotating knobs 5 fingers the thumb around the knobs, the palm in contact
  10s contact (force posture)
- large rotating knobs 5 fingers all finger along the knobs
  10s contact (precision posture)

The 10s in contact means also without motion from the subjects.

2.4 Raw Data Processing – Motion Reconstruction

This first step is to dimension the numerical manikin hand. From the hand photocopies, the segment lengths of the hand are estimated. The photographs taken are calibrated with the DLT method [3] (Direct Linear Transformation) with the aim of making coincide the photographs and the numerical environment (Fig. 6). Then, in the second step, the joint angles of the digital manikin hand are modified in order to acquire the hand posture, using a superposition of its image and the DLT calibrated photographs. In the same, the hand dimensions are checked and eventually adjusted (Fig. 7). The photographs were taken during an acquisition with the motion recorder, so we have the coordinates of each marker in a laboratory frame in which the digital manikin is defined. Once the individual hand model is well defined, the second step is to 'attach' the markers on the hand model, i.e. attach each marker to the corresponding body of the hand, and to calculate the local coordinates of the markers in the relative frame.

The third step is to reconstruct each hand posture acquired by minimizing distance between the model based marker positions and the measured marker positions. Unfortunately, during the acquisition, some markers disappear for a more or less long time. So a period, in the marker coordinates data file, was selected as large as possible with no motion from the subject, and a mean was done for each marker coordinate in order to obtain data corresponding to a unique mean posture. From the three acquisition trial, the best one was selected. The reconstruction method used is already described in [4] which is based on [5]. Fig. 8 illustrates the reconstruction phase.

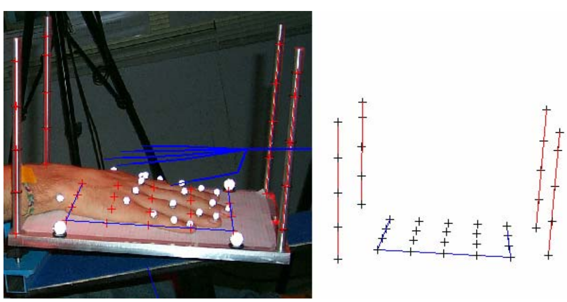


Fig. 6. DLT picture calibration

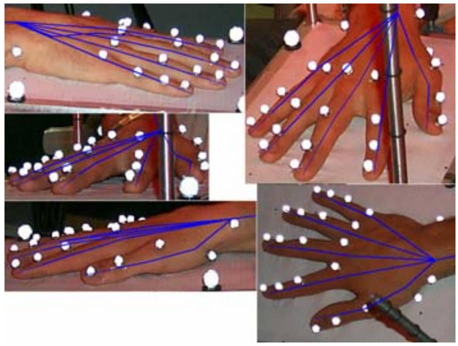


Fig. 7. Adjustment of the hand posture and the hand dimension

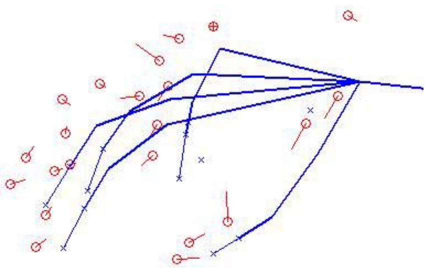


Fig. 8. Skeleton model of the hand with local and measured marker during the reconstruction step. The hand markers, defined in their manikin local frame, are represented by circles, the targets, from the measured data, are at the end of the segments.

3 Posture Simulations

Until here the procedure was identical for all the experiments, but for each object or grasping posture consign, the case is different. So in this part, only the large rotating knobs case with 5 fingers along the knobs (precision posture) will be considered. For each object or posture we have to consider their specific properties (i.e. for the kind of shape of the objects in the contact zone: cylindrical, spherical, planar …). In each case the contact zones have been reduced to one point, this allows to specify constraints to the simulation tool. But the constraints aren't always point to point, they can be point to line, point to plane …

In the used numerical manikin the envelope of the finger is defined by a succession of polygonal ring, which are used to draw the surface of the skin (Fig. 9). Each vertex of the polygon is called skin point. For the large rotating knobs case, the contact zones have been reduced to one point for each finger and for each subject. In this case the contact point was defined by the nearest skin point of the second ring of the distal phalange to the large rotating knobs axis. The following figure (Fig. 9) illustrates the definition of a contact point where the distal phalange axis is parallel to the knobs axis.

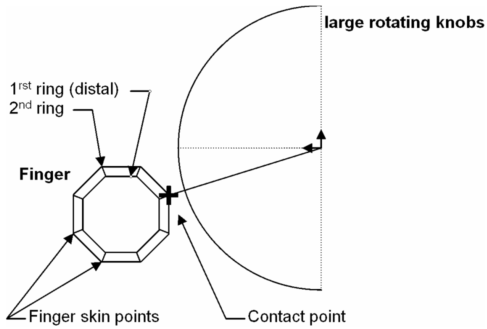


Fig. 9. Contact point definition

All subjects haven't grasped the knobs with the same manner. So for the simulation we have defined that the thumb contact point gives the zero angle for the orientation of the large rotating knobs towards the hand (see Fig. 10). Except for the middle finger, all contact zones of the other fingers are in a sphere of 1cm radius for each finger. For the simulation, with this fact, only the contact zones averages were considered (see Fig. 10), for the other objects the contact points can be dependent of the hand anthropometry.

For the simulation, with a digital manikin with any anthropometry, the algorithm will use only the defined contact zone as constraint and the angle joints will be initialized with the posture corresponding to the closer subject in term of hand size. The inverse kinematic solver is differential based as initially proposed by Liegois [5].

$$\Delta\theta = J^{+}\Delta X + \left(I - J^{+}J\right)\left(\theta - \theta_{ref}\right) \text{ with:}$$

ΔX the incremental marker displacement,

$\Delta\theta$ the incremental joint angle change

J called the Jacobian and relates linearly ΔX and $\Delta\theta$

J^{+} the pseudo inverse of J

θ the joint angle posture

θ_{ref} the referential posture

The first term of this equation tends to minimize the distance between the object contact points and the model-determined corresponding skin points and is given high priority. Whereas the second term minimizes the distance between the actual posture and the selected referential (and also the initial) posture and is given a low priority. So in terms of angle joint, the posture solution does not tend to move away from the reference posture.

For the posture simulation with similar object, with a different diameter, with a similar grasp, the chosen solution is to keep constant the distance between the contact points and the object surface, by taking into account a constant angular distribution of the contact points around the knob from the knob axis. The following figure shows an example of simulation (Fig. 11).

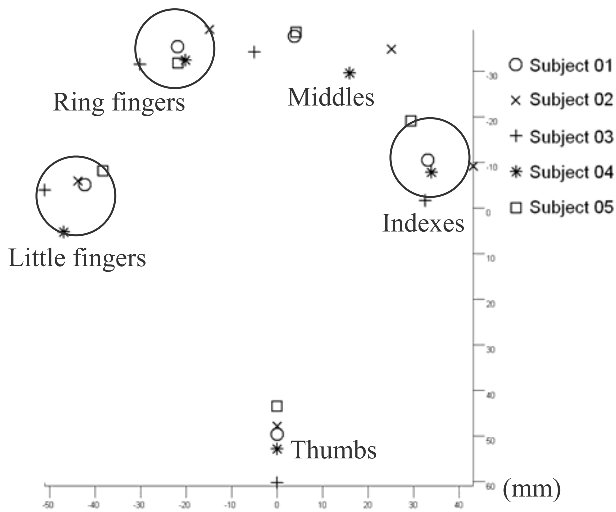


Fig. 10. Contact points projection distribution for the rotating knobs. Indexes are on the right.

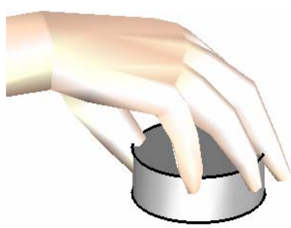


Fig. 11. Simulation result

4 Conclusions

This paper presents an experimental approach for the simulation of grasp posture. This approach is based on databases, and is adapted for a digital manikin. The simulation gives a realistic representation of the posture with a limited number of inputs for the users instead of a joint axis handling. One of the limitations can be the number of postures in the database; but in our car interior design context, only few postures are used. The hand manikin model definition can be another critical point. The digital manikin is defined with skin points which aren't dependent of the contact, so some fingers can shallowly penetrate in the object structure. This point often is not so important because the manikin is generally globally considered. For digital manikin used in classical CAD software, the hand model definition (commonly 250-500 skin points for one hand) is not enough accurate for studies which require a great precision.

Another point of difficulty is the manikin hand model kinematic definition. In this study the manikin hand model has some limitations: the palm is always flat, the joint angle limits aren't always well defined, the kinematic structure of the palm shows some symmetry… To get better results, a more complete model can be envisaged [6], but in this case, it will be not compatible with the chosen manikin. Another point of development, is how to associate the simulation of hand posture grasping with the reach of a global posture. In this case a first response element is to cumulate the constraints, but users have to take care of the constraints compatibility.

References

1. Rougeron, M., Le Garrec, J., Micaelli, A., Ouezdou, F.B.: A real tile human grasp simulation with deformable fingertips. In: proc. of the 17th IASTED International Conference Modeling and simulation May 24-26 Montreal QC, Canada, pp. 13–19 (2006)
2. Miyata, N., Kouchi, M., Mochimaru, M.: Posture Estimation for Design Alternative Screening by DhaibaHand - Cell Phone Operation. In: Proc. SAE 2006 Digital Human Modeling for Design and Engineering Conference, Lyon, pp. 2006–01–2327 (2006)
3. Abdel Aziz, Y.I., Karara, H.M.: Direct Linear Transformation from comparator coordinates into object space coordinate in close-range photogrammetry. In: proc. of the Symposium on Close-Range Photogrammetry, Falls Church, American Society of Photogrammetry, pp. 1–18 (1971)
4. Monnier, G., Wang, X., Beurier, G., Trasbot, J.: Coordination of spine degrees of freedom during a motion reconstruction process. In: proc. of the SAE Digital Human Modeling Conference, Univ. of Washington, Seattle, USA 07DHM-73 (in press) (2007)
5. Liegeois, A.: Automatic supervisory control of the configuration and behavior of multybody mechanisms. IEEE Transactions on systems, man and cybernetics 7(12), 868–871 (1977)
6. Savescu, A.V., Cheze, L., Wang, X., Beurier, G., Verriest, J.P.: A 24 Degrees of Freedom Hand Geometrical Model for Better Hand Attitude Simulation. In: SAE International conference and exposition of Digital Human Modeling for Design and Engineering, June 14-16, Oakland University, Rochester, Michigan, USA. SAE paper 2004-01-2196 (2004)

Human Articulation Efforts Estimation in the Automobile Vehicle Accessibility Movement – A Pilot Study

Jean-François Debril[1], Philippe Pudlo[1], Mohand Ouidir Ait El Menceur[1], Philippe Gorce[2], and François Xavier Lepoutre[1]

[1] LAMIH UMR CNRS 8530,
Université de Valenciennes et du Hainaut-Cambrésis, Le Mont Houy
59313 Valenciennes Cedex 9, France
{jean-francois.debril,philippe.pudlo,mohandouidir.aitelmenceur,
francois-xavier.lepoutre}@univ-valenciennes.fr
[2] HANDIBIO-ESP EA 31-62,
Université du Sud Toulon et du Var, Avenue de l'université,
83957 La Garde, France
philippe.gorce@univ-tln.fr

Abstract. To build a criterion of discomfort evaluation, based on the kinematics and the dynamics of the performed movement, is an expectation of the car manufacturers. There is no reference concerning the evaluation of the articulation efforts during the automobile vehicle accessibility movement. On this point, we propose to give a first reference. Our method and our computation were verified by a well known movement, the walking. The automobile vehicle accessibility movement of a young and healthy subject is studied. The articulation efforts presented are coherent with the produced movement's analysis.

Keywords: Automobile vehicle accessibility, Inverse dynamics, Articulation efforts.

1 Introduction

To enter or leave an automobile vehicle is a complex action requiring a precise coordination of the articular movements of the human body [7]. Felt at the time of its realization is an important factor for the customers mainly if this latter are elderly or having locomotor apparatus dysfunctions [14]. Traditionally, the car manufacturers realise ergonomic evaluation of the ingress/egress movements on real scale prototypes or mock-ups. This one is carried out with a significant number of subjects representative of the targeted customers. The subjective judgement given by the people taking part in the experiments is then correlated with dimensions of the vehicle [16], [1]. These approaches imply times and costs of significant treatment. If this approach is of habit, another possibility emerges and aims at evaluating discomfort only from the produced movement [4]. To go in this direction, our objective is to build a criterion (or criteria) of discomfort evaluation based on kinematics (e.g. articular angles) and the dynamics (e.g. articular torques) of the performed movement.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 23–32, 2007.

For the moment, only discomfort evaluation based on kinematics was approached [4]. Moreover, until now, the articular torques estimation of the automobile vehicle accessibility movement was never studied.

The objective of this paper aims at the calculation of the articulations' efforts developed in the lower limbs while performing vehicle accessibility movements. A first step is to validate our method with a well-known data of the literature, a walking data. The second step is to use our method for the accessibility movement and to verify its relevance with an ergonomic analysis.

In a first time we will present the methodology that we have followed in our study, where the biomechanical model and the mathematical tools used in the modelling will be detailed. The experiments allowing the determination of the different parameters of the model will be then presented. Some data processing will come after that. Then we will present the results. Finally our paper will be concluded and some perspectives will be presented as well.

2 Methodology

2.1 Theoretical Postulates

The body segments are considered solid and indeformable during the movement. For this reason, the positions of the centre of gravity, the mass and the moment of inertia in the centre of gravity are considered constant.

The articulations are considered perfect, without friction and slip. The movement generation is due only to the articulations' torques.

2.2 Kinematic Model

The vehicle accessibility movement is a three-dimensional movement [1]. A multi-chain/poly-articulated three-dimensional skeletal model is chosen [6].

The physiology of the human body breaks up the lower limb into 3 segments [8]: the thigh, the leg and the foot. These segments form a simple kinematic chain for each of the two lower limbs, right and left. These 2 kinematic chains are articulated on the pelvis, taken as body of reference, whose position and orientation are known at every moment.

The articulation of the different body segments is modelled by perfect connections of spherical type. The hip, the knee and ankle are defined by 3 DOF, this makes 9 DOF per chain or 18 DOF for the whole model.

The position and the orientation of the different segments the ones in opposite to the others and to the referential of reference R_0 are defined by homogeneous matrices ${}^{i}_{i+1}T$ (1):

$$
{}^{i}_{i+1}T = \begin{matrix} \vec{X}_i \\ \vec{Y}_i \\ \vec{Z}_i \\ \; \end{matrix}
\overset{\begin{matrix} \vec{X}_{i+1} & \vec{Y}_{i+1} & \vec{Z}_{i+1} & \overrightarrow{O_iO_{i+1}} \end{matrix}}{\begin{bmatrix} & \left[{}^{i}_{i+1}R\right] & & \\ 0 & 0 & 0 & 1 \end{bmatrix}} \tag{1}
$$

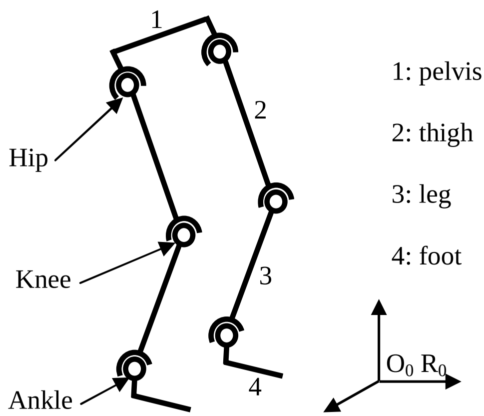


Fig. 1. The human body lower limbs' kinematic model

- $\overrightarrow{O_i O_{i+1}}$ represents the position of the medium of the pelvis or those of the articular centers into the referential R_i.
- $\left[{}^{i}_{i+1}R\right]$ represents the matrix of rotation of the referential $R_i \left(O_i, \vec{X}_i, \vec{Y}_i, \vec{Z}_i\right)$ in opposite to the referential $R_{i+1} \left(O_{i+1}, \vec{X}_{i+1}, \vec{Y}_{i+1}, \vec{Z}_{i+1}\right)$.

The local referentials R_i are built for each body segment according to the recommendations of the I.S.B. (International Society of Biomechanics). The matrix of rotation ${}^{i}_{i+1}R$, ortho-normalized, is then obtained by the z, x, y sequence Euler formalism [17].

The position of the end effector, the 2nd metatarsal, in the referential related to the pelvis R_1 is expressed according to the equation (2):

$$ {}^{1}_{4}T = {}^{1}_{2}T \times {}^{2}_{3}T \times {}^{3}_{4}T \tag{2} $$

The position of the end effector in the referential of reference R_0 becomes:

$$ {}^{0}_{4}T = {}^{0}_{1}T \times {}^{1}_{4}T \tag{3} $$

2.3 Angles Correction

During the kinematics movement reconstruction, an error appears between the ankle calculate position and ankle measured position. This error is due amongst others to the instrumental errors and the experimental errors. This can generated a collision of the foot with the vehicle's sill during the automobile vehicle accessibility movement. The coherence of the kinematic reconstruction with the measured movement is evaluated using a calculation of correlation coefficient and RMS. If a consequent skew is noted, the ankle's position is optimized by a correction of the articular angles [12], [13].

2.4 Anthropometric Model

The building of the lower limbs' anthropometric model requires, for each body segment, (the foot, the leg and the thigh) to determine its mass, geometrical and inertial properties. For this study, the regression equations [18] are used to obtain, starting from a series of the studied subject's anthropometric measurements:

- The mass of the segments,
- The position of the centre of gravity along the longitudinal axis of the segments,
- Moments of inertia of the segments in the centre of gravity around their principal axes.

2.5 Calculation of the Internal Efforts

The estimation of the articulation efforts concerns an inverse dynamics problem. For this purpose, the Newton-Euler dynamic equilibrium expression is used.

The Newton-Euler dynamic equilibrium is written by using the formalism of the homogeneous matrices [10], [11]. It facilitates the writing of the recursive algorithm of inverse dynamics [3]. Moreover, this formalism makes it possible to follow the recommendations of the I.S.B. (International Society of Biomechanics) [17].

3 Experiment

3.1 Device

The experimental device is similar for the walking experimentation and for the accessibility experimentation. The experimental device aims the capture of the movements. Three types of data are desired: anthropometric, kinematic and dynamic.

The kinematics data are obtained by an optoelectronic motion capture system VICON®. The eight cameras are laid out around each vehicle. The system records the Cartesian position of the reflective markers laid out on the vehicle and on the subject.

Two KISTLER® force plates are used to measure the ground reaction forces. The external efforts are applied at the end of the chain kinematics. Force plates are set in the medium of the way of walk or near to each vehicle in the case of ingress/egress movement.

Four vehicles of various dimensions, present in the trade, were used: a small car, a median car, a minivan and a small utilitarian. To allow the capture of the movements, the vehicles were stripped to the maximum. We present the small utilitarian's results.

3.2 Protocol

The base of the protocol is the same for the accessibility movement and for the walking movement.

Anthropometric measurements carried out are those necessary to the anthropometric model [18].

Twenty markers, 16 anatomical and 4 technical, are laid out on each subject : 1^{st} and 5^{th} tarsals, medial and lateral malleolus, medial and lateral tibial condyles, anterior superior iliac spines, pubis and vertebra L5.

For the two applications, a capture of the lower limbs' circumduction movement is carried out. This one allows the determination of the articulation centres of the hips by a using the sphere fitting method [9], [1<]. The subject in sitting position is captured for the estimation of the marker L5 when it is masked by the automobile seat [5].

On the one hand, a walking movement is captured and on the other hand, four ingress/egress movements on each of the 4 vehicles are captured. The repetition aims to remedy the occlusion or the fall of reflective markers.

3.3 Population

Forty one subjects carried out the vehicle accessibility movement's protocol. The population is varied: young and healthy subjects, elderly subjects and disabled subjects. This study is interested on a young and healthy subject: 27 years old, 1.75m and 74 kg.

For the walking movement, the presented subject is also a young and healthy subject: 25 years old, 1.87 m and 80 kg.

4 Data Processing

4.1 Evaluation of Articular Kinematics

Initially, the positions of the reflective markers measured by system VICON® are filtered. A forth order Butterworth filter without lag at 5 Hz is used. In the second time, the parameters of the kinematic model are calculated.

Determination of the articular centers. The center of the knee is estimated at the median position of the markers laid out on the internal and external condyles and the centre of ankle at the median position of the markers laid out on the internal and external malleoli [2]. The position of the hips centres is determined by the calculation of sphere fitting [9], [15].

Determination of the medium of the pelvis and segmentary lengths. The segmentary lengths of the thigh and the leg are defined by the average distance between the hip and the knee on the one hand and between the knee and ankle on the other hand. In the same way, the width of the pelvis is defined as the distance between the two articular centers of the hips. The medium of the pelvis is defined in the medium of this segment.

Evaluation of the angular data. If need be, an optimization of the foot's position by correction of the articular angles is carried out. The corrected angles are considered valid when the movement reconstruction presents no collision of the foot with the vehicle's floor.

4.2 Calculation of the External Efforts

The torques of the mechanical action issued from the force plates are expressed in the local referential R_k. This information being disturbed, a filtering of the components of these torques is carried out by a forth order Butterworth filter at 5 Hz.

The transfer of these torques in the referential of reference R_0 (10), makes it possible to meet the needs of the inverse dynamics algorithm.

$$[\Phi]_{R_0} = {}^0_k T \cdot [\Phi]_{R_k} \cdot \left({}^0_k T\right)^{-1} \tag{10}$$

- $[\Phi]_{R_k}$ represents the filtered torque issued from the force plate k expressed in R_k,
- $[\Phi]_{R_0}$ represents the filtered torque issued from the force plate k expressed in R_0,
- ${}^0_k T$ represents the matrix of passage from the referential R_k to the referential R_0, and $\left({}^0_k T\right)^{-1}$ its reverse.

5 Results

5.1 Walking Movement

Kinematics

The reconstruction of the walking movement has a good fidelity with the trajectory of the feet:

- All the correlation coefficients between the positions of ankle measured and reconstructed are at 1.00. (>0.95)
- The RMS is 0.42 cm. (< 1 cm)

This kinematics is used for the calculation of the articular efforts.

Articular Efforts

The results of this walking study seem to consolidate the experimental method, the kinematic model and the inverse dynamics algorithm (Fig. 2).

5.2 Automobile Vehicle Accessibility Movement

Kinematics

The reconstruction of ingress-egress movement presents difference between the trajectory of feet built and reconstructed (the RMS is 1.42 cm). So the articular angles are corrected. Following optimization, the RMS is 0.67 cm (< 1 cm) and the correlation coefficients are at 1.00 (> 0.95). This acceptable movement reconstruction is used for the calculation of the articular efforts.

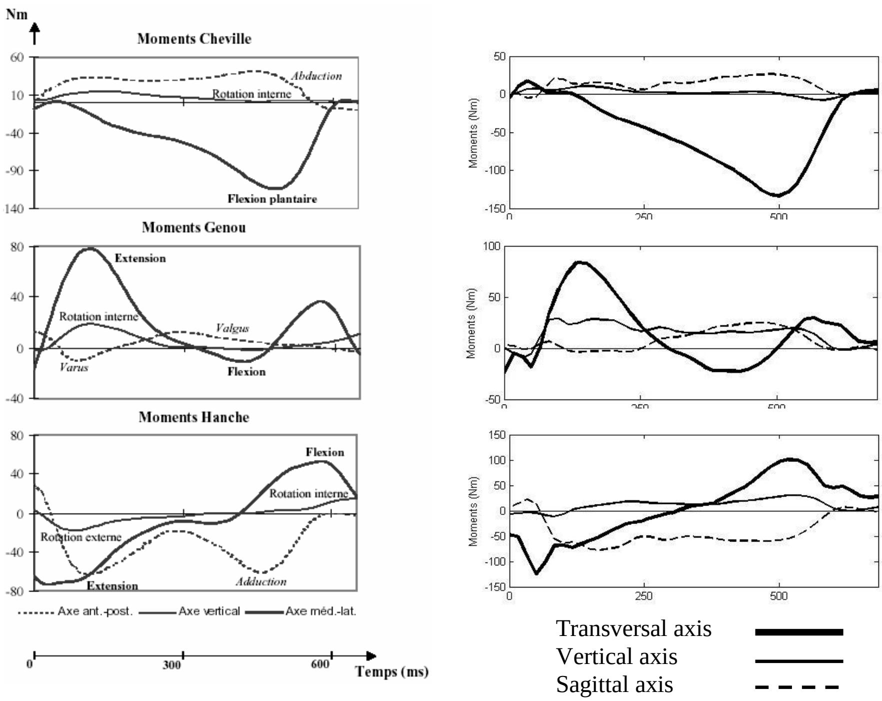


Fig. 2. Moments around the ankle (in top), the knee (in the medium) and the hip (in bottom), for the literature ([3] on the left) and for our subject (on the right) during the walking single stance

Abduction-adduction efforts of the left hip during the ingress phase

The figure 3 presents the comparison of the couples calculated around the sagittal axis and the articular angles associated. These results are at the end of the chain recursive algorithm.

Analysis of the three phases

Three phases are defined by the curves analysis (Fig. 3)

- **A** - concentric abduction phase
- **B** - concentric adduction phase
- **C** - excentric adduction phase

This analysis of the articular torques and kinematics corresponds to the ergonomic analysis. Phase **A** corresponds to the concentric opening of two lower members, which allows the subject to pose its right foot inside the vehicle. From this support, the phase **B** starts. The subject relocates its buttocks on the seat by tightening its legs. This is a concentric movement. Phase **C**, the subject begins to get down towards the seat. It is a light eccentric opening that it retains its fall.

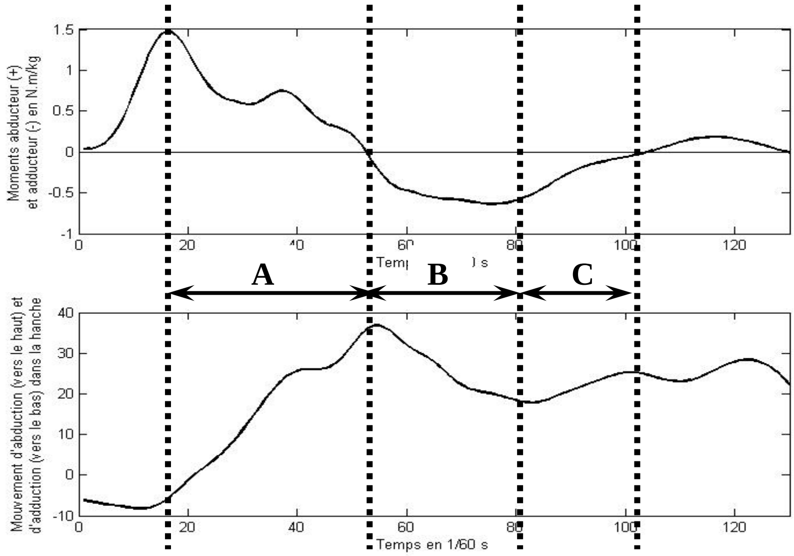


Fig. 3. Normalized torques by the weight and articular angles of abduction/adduction for the left hip during the phase of support unipode of entry in the small utilitarian (healthy subject of 1,75 m and 74 kg). A: Takeoff of the right foot to its installation on the vehicle's floor. B: Placement of the buttocks to the top of the seat. C: Movement of descent towards the seat.

6 Conclusions and Perspectives

Kinematics model and anthropometric model were proposed. The articular efforts were estimated by inverse dynamics for the walking manipulation. The calculated articular torques are in coherence with the literature.

The automobile vehicle accessibility movement reconstruction require an optimization of the foot's position by a correction of the articular angles. The analysis of the results obtained by the calculation of dynamics is concordant with ergonomic analysis.

Moreover, for the automobile vehicle ingress movement, around the sagittal axis of the left hip, three phases were recognized: a concentric abduction phase since the takeoff of the right foot to its installation on the vehicle's floor, a concentric adduction phase corresponding to the transfer of the buttocks inside the vehicle above the seat, an eccentric adduction phase slowing down the fall of the subject during the descent in the seat.

The articular efforts are calculated for the lower limbs' articulations. The results obtained are coherent with the produced movement. The instrumentation being limited, the articular efforts can be calculated only for the beginning of the movement. More pushed instrumentation of the vehicle would allow an increased knowledge of the articular efforts.

A forthcoming study is interested in the population as a whole (elderly, disabled). The various assumptions of modeling and calculation are discussed there.

Acknowledgments. Acknowledgments to the Ministry Delegated to the Higher Education and Research and to the IRRH and the IFRH for their support.

References

1. Andreoni, G., Rabuffetti, M.: New approaches to car ergonomics evaluation oriented to virtual prototyping. EURO-BME Course on Methods and Technologies for the Study of Human Activity and Behaviour, March 19-20 1997, Milano, Italy, pp. 1–13 (1997)
2. Cappozzo, A., Catani, F., Croce, U., Leardini, A.: Position and orientation in space of bones during movements: anatomical frame definition and determination. Clinical Biomechanics 10(4), 171–178 (1995)
3. Doriot, N.: Modélisation dynamique du membre inférieur pour l'estimation des forces articulaires et musculaires mises en jeu pendant la phase d'appui de la marche. PhD thesis, Université Claude Bernard, Lyon 1. N°259–2001
4. Dufour, F., Wang, X.: Discomfort assessment of car ingress/egress motions using the concept of neutral movement. In: SAE Digital Human Modeling For Design and Engineering Conference, Iowa City, Iowa, USA (June 15-17, 2005)
5. Gillet, C.: Analyse biomécanique de la marche et proposition de classes de marcheurs - application au portage de sacs à dos. PhD Thesis. Université de Valenciennes et du Hainaut-Cambrésis (2004)
6. Gorce, P., Villard, C., Fontaine, J.G.: Grasping, coordination and optimal force distribution in multifingered mechanisms. Robotica 12(2), 243–251 (1994)
7. Gransitzki, K.: Biomechanics of entering a car. Second International Congress on Physiological Anthropology, University of Kiel, Germany (September 12-16, 1994)
8. Kapandji, I.A.: Physiologie articulaire, fascicule II, membre inférieur, Librairie Maloine S.A., Paris, ISBN: 2-224-01052-4
9. Leardini, A., Cappozo, A., Catani, F., Toksvig-Larsen, S., Petitto, A., Sforza, V., Cassanelli, G., Giannini, S.: Validation of a functional method for the estimation of hip joint centre location. Journal of biomechanics 32(1), 99–103 (1999)
10. Legnani, G., Casolo, F., Righttini, P., Zappa, B.: A homogeneous matrix approach to 3D kinematic and dynamics-I Theory. Mech. Mach. Theory 31(5), 573–587 (1996)
11. Legnani, G., Casolo, F., Righttini, P., Zappa, B.: A homogeneous matrix approach to 3D kinematic and dynamics-II Applications to chains of rigid bodies and serial manipulators. Mech. Mach. Theory 31(5), 589–605 (1996)
12. Lempereur, M.: Simulation du mouvement d'entrée dans un véhicule automobile. PHD thesis, Université de Valenciennes et de Hainaut Cambrésis. 06/04 (2006)
13. Lempereur, M., Pudlo, P., Gorce, P., Lepoutre, F.-X.: Optimization approach for the simulation of car accessibility movement. In: IEEE International Conference on Systems, Man and Cybernetics, Washington D-C, USA (October 5-8, 2003)
14. Petzäll, J.: The design of entrances of taxis for elderly and disabled passengers An experimental study. Applaid Ergonomics 26, 343–352 (1995)
15. Piazza, S.J., Okitaj, N., Cavanagh, P.R.: Accuracy of the functional method of hip joint center location: effects of limited motion and varied implementation. Journal of Biomechanics 34(7), 967–973 (2001)
16. Sternini, D., Cerrone, M.: Setup of an experimental procedure for measuring the human body movements when entering and exiting the vehicle. In: Proceedings of the 3rd International Conference on Vehicle Comfort and Ergonomics, pp. 257–265 (1995)

17. Wu, G., Siegler, S., Allard, P., Kirtley, C., Leardini, A., Rosenbaum, D., Whittle, M., D'Lima, D.D., Cristofolini, L., Witte, H.: ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion–part I: ankle, hip, and spine. Journal of Biomechanics 35(4), 543–548 (2002)
18. Zatsiorsky, V., Seluyanov, V.: In vivo body segment inertial parameters determination using a gamma scanner method, Biomechanics of Human movement: applications in rehabilitation, sports and ergonomics, Berme, N., Cappozo, A. (eds.): Bertec Corporation, Worthington, Ohio, U.S.A., pp. 196–202 (1990)

Study on the Appraisal Methods of Hand Fatigue*

Li Ding[1], Feng Yang[2], Chunxin Yang[3], Xiugan Yuan[3], and Yang Li[1]

[1] Bioengineering Department, BeiHang University, Xueyuan Road 37# , Beijing 100083, China PR
[2] Department of Physical Therapy, University of Illinois at Chicago, IL, 60612, USA
[3] School of Aeronautics Science & Engineering, BeiHang University, Xueyuan Road 37# , Beijing 100083, China PR
{Li Ding,Feng Yang,Chunxin Yang,Xiugan Yuan,Yang Li, ding1971316}@buaa.edu.cn

Abstract. Objective: The evaluation methods of hand fatigue are built and the aim is to provide reference for the numerical human fatigue appraisal. Method 12 females and 14 males in two groups participated in the experiments of changeless force fatigue, which included the grip and pinch and screw, and took part in the nondimensional force fatigue. Result ①Hand fatigue could be full evaluated basing on grip and pinch and screw. ②Using stable force could test the fatigue of manual work simply, but it would be largely affected by the characteristic of sex. ③The nondimensional evaluation can avoid the impact results from the different forces of testees. Conclusion Hand fatigue could be full evaluated basing on grip and pinch and screw. The nondimensional evaluation is a feasible method to research numerical human manual fatigue.

Keywords: Hand performance, Fatigue, Evaluation, Nondimensiona.

1 Introduction

Falling under the influence of many factors such as figure, structure, pressure of spacesuit glove system and the low temperature environment, the cosmonaut's manual work are easily to be tired, which could lead to the biggish impact on space work[1]. But it is need to make much experimental evaluation for any modification of the lifesaving equipment system [1]. Experiments would be largely reduced if numerical human evaluation could be used in advance [1]. But fatigue is a complex physiological phenomenon to describe difficultly in building numerical human model. In especial, hand is mainly tools and it needs to simplify muscle, ligament and other tissues model in manual fatigue evaluation. A logical evaluation method is asked for to compare the consistency between human and the numerical human.

Generally, manual fatigue can be evaluated from three aspects as follows: physiological fatigue, subjective fatigue and action decline [2~9]. By all appearances, physiological fatigue and subjective fatigue can not applied to the numerical human, and only action decline can. Though there have many studies on the action decline

* **Sponsored project**: National youthful nature science fund (50406008).

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 33–41, 2007.

[2~9], but those often base on the academic study [2],[7],[8], and these studies still can not satisfy the demand generated from manufacture and evaluation of spacesuit glove system. Such demand is “few but extractive” [2], [3]. Therefore, this paper tries to develop the bare hands’ evaluative research of dividing the fatigue into changeless force fatigue and nondimensional force fatigue according to the theory that force can be divided into changeless force and nondimensional force. The purpose of this research is providing the feasible appraisal methods for performance fatigue of spacesuit human – glove system.

2 The Efficacy Analysis of Fatigue

2.1 Mechanism for the Occurrence of Work Fatigue

Physical fatigue is the especial representation after actions lasted for some time such as weariness and lack of force. When the physical fatigue is severe, the muscle aches. The main reason of this phenomenon is the decline of Adenosine Triphosphate (ATP), which could generate energy for human. In addition, during works, the accumulation of lactic acid results from the lack of oxygen can make the tired feeling intensive.

2.2 Fatigue Analysis of Manual Work

According to the mechanism of manual action and space work of cosmonauts, the main function of manual space work can be defined as grab, hold, grasp, turn and twist. According to the definition of force [3], this function can be simplified. It is three kinds of force: grip, pinch and screw. Therefore, the evaluation of fatigue of manual space work also includes fatigue of grip pinch and screw. The fatigue of grip is the tired instance of all fingers except for the thumb. The main measure work is testing the tired situation of muscle of fingers and palm. Some muscles of fingers are involved in the pinch measure. Because of that these muscles’ force is definitely larger than the thumb [3], the pinch fatigue is testing the fatigue extent of some curving muscles that are connected with thumb. Screw is a combination of two actions, which are pressing and screwing. The muscles for griping and pinching and some muscles of hands generate the action called screwing. The measure for screw is a synthetical test of hand muscle’s fatigue.

3 The Design of Experiment

The fatigue discussion of Nina、Shingo、Tommy and Edwatds is mainly about the physiological process of intramuscular fatigue [3~7]. Moreover, some experiments of cosmonaut’s fatigue done by O’Hara and Serge are not comprehensive [2], [8]. It is inevitable that in order to program the cosmonaut’s space work, it is necessary for us to realize the fatigue that can influence the accuracy of work. The fatigue not only includes fatigue of grip but also involves pinch and screw. Therefore, this text firstly studied the fatigue of these three kinds of stable force, and then tested the fatigue of nondimensional forces. In order to study completely, the research added in the body's subjective feedback to help fatigue appraisal.

3.1 The Fatigue of Stable Forces

When the cosmonaut is doing the space work, he must use such actions as grasp or grip[2], [6], [7], which will be influenced by the counterforce generated by resistance of gloves and inside stress of clothes. Same kinds of gloves could generate the uniformly stable stress, which will not change by different people. According to this, a stable force can be used to evaluate the fatigue of hands. The test of fatigue of grip was the same as "griping the bicycle's handles" (Fig. 1.1). The testee griped the two handles together, when the roots of handles were tangent, the indicator light on the display lighted, then one action finished. The testee must finish this action with the frequency of 1 time per second on average. After 20 times and a five minutes rest, the testee would be asked some interrelated questions in table 1 [2], [5].

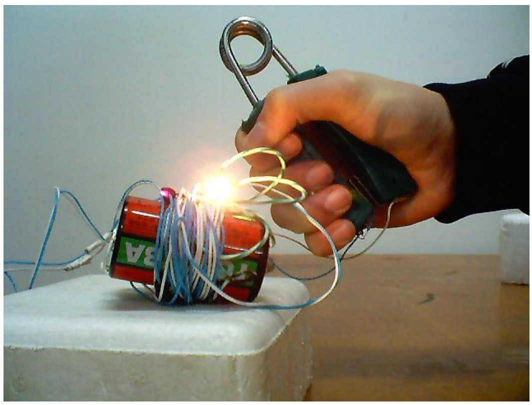


Fig. 1.1 Grip Fatigue

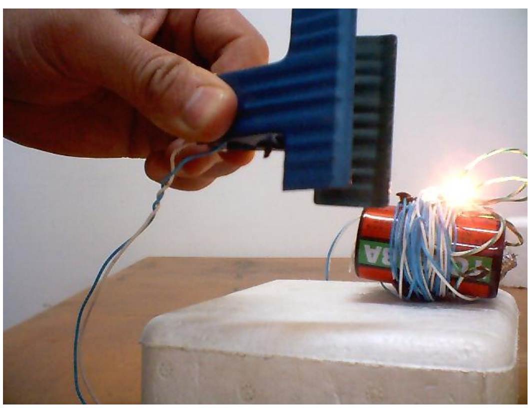


Fig. 1.2. Pinch Fatigue

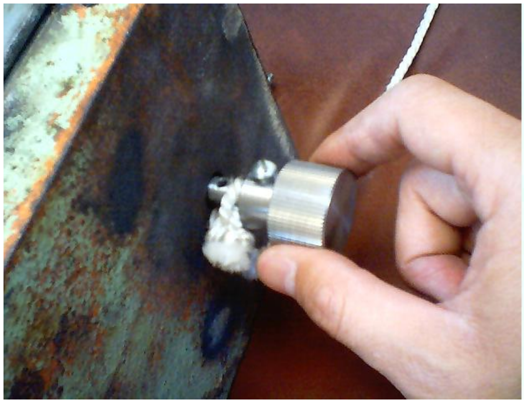


Fig. 1.3. Screw Fatigue

Table 1. The Fatigue Scale

Fatigue level	Fatigue scale
1	Feel the same as the beginning, no noticeable fatigue
2	Light fatigue — between1&3
3	Feel some resistance to squeezing the handle, moderate fatigue
4	Fatigue – between 3-5
5	Unable to effectively squeeze the handle, complete fatigue
6	Stop

After judging the fatigue extent of the testee, one circulation was end. The content of next circulation was the same as the first one (Fig. 1.2, 1.3). The test would not stop until the testee could not let the indicator light.

Tests of fatigue of pinch and screw were the same as grip. Because the forces of pinch of screw were little [9], the test time reduced to 10 times.

3.2 Nondimensional Fatigue

Because the differences among testees are remarkable, using the evaluation of stable force maybe can not let the testee with strong strength feel tired. Also, it maybe can not let the testee with small strength reach the real situation of fatigue. In order to deal with this question, the fatigue experiment of NASA used the nondimensional evaluation. However, when using this method in the test, we found that the percentage of the Maximal voluntary contraction (MVC) and the results of the test were questionable. Therefore, this text afresh designed the method of nondimensional evaluation through comparing experiments of grip fatigue. The experiment of fatigue included several circulations, which involved two states, static state and dynamic state (Fig.2). In the static state, testees were asked to keep 20% of their MVC for 10 seconds, while in the dynamic state, testees used the frequency of 15 times per minute to grasp and undo the hand dynamometer. Every numerical value of force must exceed 50% of the testee’s MVC. Moreover, testees were asked to repeat the experiment all long until they could not do it any more.

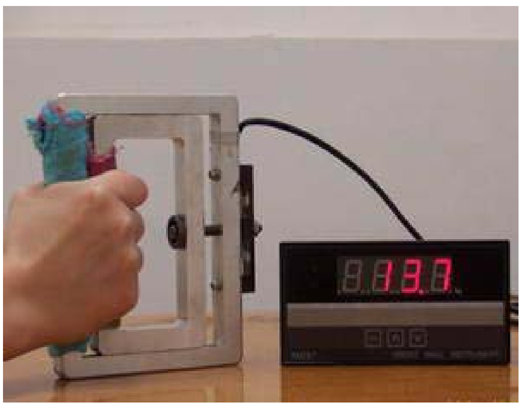


Fig. 2. Nondimensional Grip Fatigue

4 Method of Experiment

4.1 Personnel of Experiment

The experiment went along for two times. Every test involved 13 healthy voluntary students. (6 were female and 7 were male), all of them were right-handed. The first test (age: 25.5±4.1) measured the fatigue of stable force and nondimensional fatigue. The second test (age: 22.3±1.8) validated and standardized the data from the first one. The testees of every experiment were different.

4.2 The Equipment of Experiment

Fatigue device for stable force: the facility used for measuring fatigue adopts the spring for stable force (Fig. 1). According to the different contents of test, the testing forces are also different.

Fatigue device for nondimensional: The grip handle is connected with the display through cable so that it becomes easy to gather the data under the circumstance of experimental cabin (Fig. 2).

4.3 The Mode of Experiment

4.3.1 Fatigue of Stable Force

The testees were seated in front of table, griped the test device with the average frequency of 1 time per second. After 20 times, they could take a break of 5 seconds and answered some relevant questions of fatigue. Meanwhile, the tester measured off the fatigue grade according to the reactive extent of testees (Table 1). The test would not stop unless the testee could not let the indicator light.

Tests of fatigue of pinch and screw were the same as grip, the only difference was that the test time reduced to 10 times.

4.3.2 The Nondimensional Fatigue

The testee was seated in front of the table and griped the test device like figure 2, griping and undoing the device with the frequency of 15 times per second for 1 minute. Every numerical value of force must exceed 50% of the MVC. Then the testee was requested to hold 20% of his or her highest force for 10 seconds, meanwhile, the tester compartmentalized the grade (Table 1) of fatigue according to the reactive extent. The test would not stop unless they could not go on.

5 Result and Analysis

The experiment went on for two times. The first one was used to exploit methods and find out the basic conclusion, while the second one was used to validate and standardize the conclusion. Therefore, the data used in this text is from the second test.

Because of that the difference of force between male and female is comparatively great (male: female ≈ 8:5) [9], the relevant fatigue values of male is twice as high as female or higher (Table 2, $P<0.001$). Therefore, characteristic of sex must be considered in appraisal numerical human fatigue.

5.1 Fatigue of Grip

Compared with O’Hara’s method for measuring fatigue, the stable force fatigue adopted the alternation between work and break for 5 seconds, which is an effective and simple method for testing fatigue. When testees finished the experiment, their work efficiency would decline extremely ($P<0.05$), and they needed a long time to restore the fatigue ($t>1$ day). Meanwhile, the experiment adopted the work which includes breaks and was likely to attain the fatigue that accords with the various traits of space work. We also used the work without break to do the experiment, the result indicated that some testees did not really reach the fatigue and they restored immediately. Their work efficiency was not be influenced severely ($P>0.05$). However, when there were large differences among testees’ force, some testees with

Table 2. Statistical Data of Hand Fatigue (X±SD)

Efficiency index	Result	
	Female	Male
Grip fatigue/time	136±63	314±137
Pinch fatigue/time	83±31	184±60
Screw fatigue/time	189±78	400±251
Nondimensional grip fatigue /bout	4.0±0.14	9.14±1.7

large force did not feel tired or other testees' with less force could not really reach the fatigue situation. Therefore, this method was only suitable for testees with little differences of force, which had been fully proved in the second experiment. For the testees with large differences of force, we could divide the stable force into several grades and manage the experiment. But it would be hard to evaluate numerical human fatigue.

5.2 Fatigue of Pinch

Because the force of finger was small compared with grip (max grip : max pinch ≈ 4.1) [9], testees would feel tired easily in the test. At the same time, there was not a large difference of pinch among testees (max pinch – min pinch = 78N, max grip – min grip = 379N) [9]. Therefore, this kind of test did not include the phenomenon that strong testees could not feel tired and testees with small force could not pinch the device (Table 2). Two experiments indicated that after the test of pinch fatigue, the testees' work efficiency is influenced largely ($P<0.05$). Thereby, because this method does not have many limitations, the best method for testing pinch fatigue is designing two different devices of testing pinch fatigue for both males and females.

5.3 Fatigue of Screw

The coefficient of friction will decline after the hands perspire, and then the required force will increase. Therefore, compared with other two methods of test, this one could easily be influenced by the sudation of hands, which leads to differences among testees (Table 2, Fig. 3). Like the test of grip, this method is only appropriate to testees without many force differences.

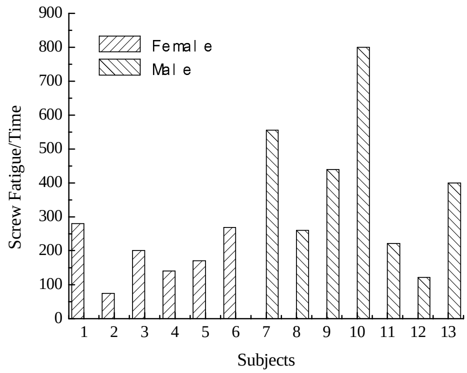


Fig. 3. The Compare of Screw Fatigue

5.4 Nondimensional Fatigue

In order to avoid the large differences of force among testees, this text designed the nondimensional evaluation of fatigue for grip. Experiment indicated that using nondimensional evaluation can effectively avoid the problem that strong testees and ones with small force can not really reach the state of fatigue. So it would be possible to enlarge the realm of selective testees (Table 2). Due to the influence from physiology and mentality, the testees' MVC must be tested many times in order to avoid the fact that the testees' real MVC can not be gained, which may lead to the distortion of the testees' result. Comparing the fatigue of stable force with the nondimensional force, it was found that the nondimensional could evaluate the condition of testees' bodies. The fact was that testees with good condition of body could remarkably complete many circulative times of fatigue test ($P<0.01$). Therefore, nondimensional method of evaluation could be used widely and it is a feasible method to research numerical human manual fatigue.

6 Conclusion

Fatigue is a special representation of muscular force's continuity. Because of differences of tired standard and dissimilar individuals, the experimental result will be influenced to some extent. Through the experimental study of hand's fatigue, we can draw the conclusions as bellows:

1. Fatigue of grip, pinch and screw could completely evaluate the fatigue of manual work.
2. Using stable force could test the fatigue of manual work simply, but there have some difficult in evaluating numerical human fatigue.
3. The nondimensional evaluation can avoid the impact results from the different forces of testees and evaluate the resilience of testees' muscle. It is a feasible method to research numerical human manual fatigue.

Some contents of these four conclusions are already applied in the experiment of wearing simulative spacesuit glove. The effect is remarkable[9].

References

1. Ding, L., Yang, F., Yang, C., et al.: Ergonomics Implications of Extravehicular Activity Spacesuit Glove. J. of Beijing University of Aeronautics and Astronautics 31, 416–420 (2005)
2. O'Hara., Brigann, J.M., Cleland, M., et al.: Extravehicular Activities Limitations Study Volume II Establishment of Physiological and Performance Criteria for EVA Goves-final Report. AS-EVALS-FR-8701. NASA: NAS-9-17702
3. Nina, K.V.: Measurement of Human Muscle Fatigue. Journal of Neuroscience Methods 74, 219–227 (1997)
4. Kim, S.: An Evaluation of Fatigue and Performance Changes During Intermittent Overhead Work, Virginia Polytechnic Institute and State University Master Thesis(2003)
5. Shingo, O., Noriyuki, K.: Neuromuscular fatigue during maximal concurrent hand grip and elbow flexion or extension. Journal of Electromyography and Kinesiology 11, 281–289 (2001)
6. Tommy, O.: Muscle fatigue and calibration of EMG measurements. J. Electromyography Kinesiology 5, 239–243 (1995)
7. Serge, H.R., John, M.: O'Hara: Evaluation of forearm fatigue during EVA pressure glove work. Work 8, 157–169 (1997)
8. Li, D., Longzhu, H., Cchunxin, Y.: An experimental study of effects of active-heating-system for extravehicular spacesuit gloves on working performance. Space Medicine & Medical Engineering 18, 47–51 (2005)
9. Chen, S., Ding, L., Yang, F., et al.: Study on Index optimization of Human mechanics based on basic ergonomic evaluation of spacesuit gloves. Space Medicine & Medical Engineering 19, 106–110 (2006)

Experimental Research on Human Body Motion Simulation Based on the Motion Capture Technology

Dayong Dong, Lijing Wang, and Xiugan Yuan

School of Aeronautic Science and Technique,
Beijing University of Aeronautics and Astronautics, Beijing 100083, China
{Dayong Dong}dy_dong@163.com

Abstract. Moveable human body model can be built based on human motion data, which is recorded by the VICON optical motion capture system and imported to the MOTION BUILDER software. In this experiment, two simulation projects were designed successfully, and satisfying results were obtained. The first test was an arm swinging test which simulates the pilot's action of manipulating dashboard. The other test involved arm swinging, walking and turning in standing mode, focused on testing the stability of the system in multi-degree-of-freedom complex motion situation

Keywords: motion capture, VICON, motion simulation, human body model, human factors.

1 Introduction

Human body motion simulation, which has been developed since the end of the 20th century, is a practical new & high technology integrating sensor, simulation, HMI and high parallel real-time calculate. In recent years, the combination of human body motion tracing and motion simulation technology has become a very promising field of research with its great potential development and pragmatic application

Since the 1970s, computer human simulation has been used in collision simulation. It was mainly used for simulating the motion of the human body and the degree of his injury, when the vehicle is being collided, and then the safety could be enhanced through the improvement of protecting equipments. In the field of man-machine engineering, motion simulation can be used to evaluate man's maneuverability and efficiency when he is in a limited space. In virtual designing, it can be used to investigate the feasibility of design in its early stage in order to enhance the design. In robot controlling, it can be used in designing controllers in simulating human balancing such as humanoid robot's the ability of homeostasis. Human body motion simulation can also be employed to analyze the speed, acceleration, force and torque of different parts of body when athletes, disabled people, and actors are moving. Interestingly, it can also be applied in computer animation. Computer animation has vast space of development and application in the fields of motion description, movie production, and games. In recent years, the concept of animated virtual human has been raised based motion control algorithm. Nowadays, the research on virtual human is in the ascendant, and virtual actor and

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 42–47, 2007.

digital agent have entered the stage of practicability research. With the introduction of artificial intelligence and multi-agent theory, virtual human is extending towards intellectualization and multitudinal.

2 Human Motion Tracing

There are many approaches on investigating human motion, and they can be summarized into two methods: the mathematics method and the experimental method. Both of these two methods aim to build modeling human body. The mathematical method is to establish mathematical models using the algorithm of Inverse Kinematics to drive different sections of the human body model to conduct simulation movements. The experimental method is to record the actual data of human body motion through employing human body motion tracing technology, then use these data to drive the human body model to conduct motion simulation. Comparing these two methods, the mathematical method is easier for mathematical models are easier to establish. But the human body motion is controlled totally by mathematical models which are different from natural human body movement; therefore, it can only be called "*animation simulation*". The experimental method can accurately simulate the actual movements of human body because it drives the human body model with real human motion data. But it requires a large quantity of experimental data, thus the job of data processing is fairly complicated. So we can choose a more suitable one from these two methods according to our needs. And this paper is mainly on the experimental method.

2.1 VICON Motion Capturing System

Vicon optical motion capturing system is developed by an English company, Vicon Motion Systems (VMS). It is used in human body motion capturing. The motion capturing system is composed of the following parts:

Sensor. It is the tracing equipment attached to certain positions of the moving object. It provides the Motion Capture system with information about the position of the moving object. In most cases, the number of sensors depends on the required particularity and the capturing equipment.

Signal Capturing Equipment. This kind of equipments varies according the types of Motion Capturing system. Their function is to capture signals of positions. In machinery system, it is a circuit board for capturing electronic signals; in optical system, it is a high-resolution IR camera.

Data Transfer Equipments. Motion Capture system, especially those that provide real-time data, need to transfer large quantities of data from signal capturing equipment to computer system accurately and rapidly. Data transfer system is designed to carry out this task.

Data Processing Equipment. Data captured by Motion Capture system need to be amended, processed and integrated into 3D model to fulfill the job of producing computer animation. This requires the use of data processing software or hardware.

The VICON system used in this test consists of 6 high-speed IR sensitive cameras. In the process of data collection, certain joints of the human body are attached with IR reflective markers. The high-speed IR sensitive cameras capture the motion trace of each marker, hence they can accurately capture the motion data of the joints. (See Fig. 1.)

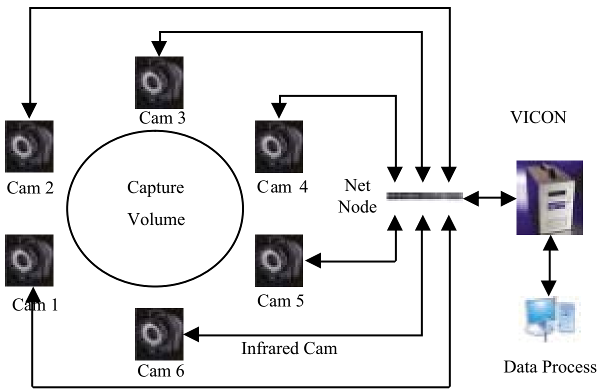


Fig. 1. The configuration of VICON motion tracing system

The tests firstly calibrate the system in static and dynamic mode to establish the 3D coordinate system. After that, test subject wearing markers enters the capture volume to conduct motion capturing.

2.2 Diagnosing and Mending the 3D Data

We can record human body motion through the system described above with sampling rate of 120Hz, and artificially validate the original data of the Marker points. Then we can conduct 3D reconstruction after finishing revision in VICON Workstation, and convert the captured motion trace in into 3D coordination data.

In the process of capturing, some useful data are missing or error data are recorded occasionally, because motions are sheltered or Marker points are lapped over. Accordingly, data amendment is to amend the mistakes in it, and to use some methods to revise the lost information in the process of data collecting.

When the data amendment has reached the anticipated value, we can use VICON Workstation system to export files in the following formats: *.eni, *.enf, *.tvd, *.mp, *sp, *.car, *.c3d. We can intercept the portion of frames exported to facilitate the following steps of the test. Then we choose several groups of well-recorded data from the data collected from the test and import them into Motion Builder to produce animation and save in *.c3d format.

3 Human Body Motion Simulation

We design two movements in this test. The first one is the motion of arm swinging in sitting posture to simulate the pilots' operating instruments. The second one is the motion of arm swinging in standing mode, of walking and of turning to test the performance and stability of the system in multi-dimensional complicated movements. To satisfy the need of capturing these two movements, we distribute the markers on the human body as shown in Fig. 2.

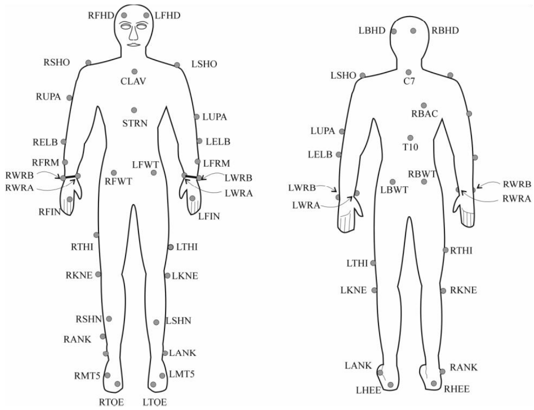


Fig. 2. Distribution of markers on human body

3.1 Motion Builder Platform

Motion Builder (MB for short), a product of Alias Company, is a human motion simulation software, which is widely used in field such as movie, animation and virtual compere. It is real-time and flexible to operate. What's more, MB provides a variety of data interfaces including *.c3d, which can seamlessly connect with the data obtained in motion capturing. In this test, we choose this software as our human body motion simulation platform.

3.2 Human Motion Simulation

We choose *c3d* files produced from VICON Workstation system, and import them into MB.A special attention should be paid to adjusting parameter, or disturbing markers will appear and severely affect the animation simulation. Before selecting "*impor*t", we shouldn't tick the "*Create Optical Segments*" option, or numerous unidentified disturbing makers will be caused. As soon as we import the data, the 3D data captured by motion tracing system are in the system immediately.

Create Actor. Actor is part of the human body model, and we select Actor from the asset browser and import it to the scene.

Bind 3D Data with Corresponding Section of the Actor. By doing this, the Actor can then move according to the captured 3D data, and be driven by the 3D data. (See Fig. 3.)

Amendments to Motion Simulation. After the binding, the human model can move according to the captured data, but there still exist some singularity points or crossover points, which will cause disorder in the motion. For this reason, it is required to examine the motion data frame by frame in MB to ensure the smoothly of the motion.

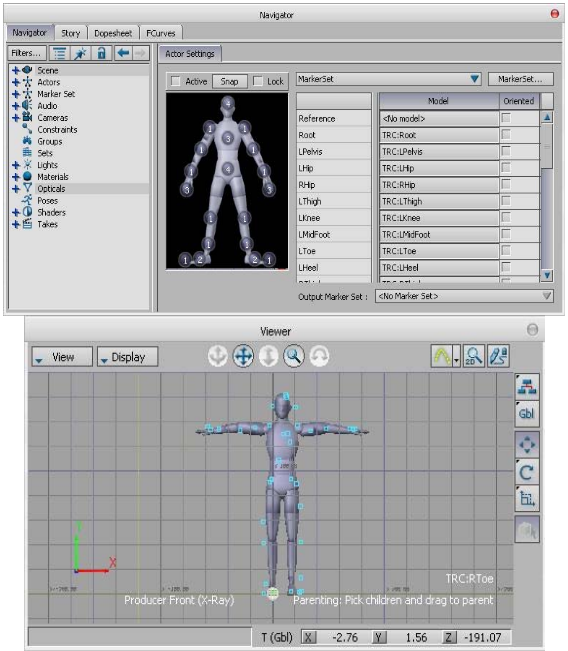


Fig. 3. Binding Actor and the 3D data

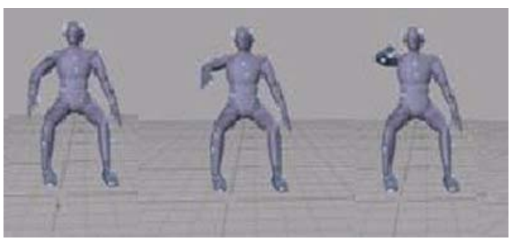


Fig. 4. Models of arm swinging movement

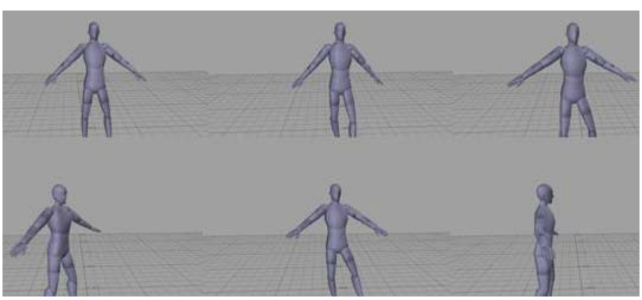


Fig. 5. Models in walking and turning status

Export to Video File. After the process of data binding and combination, we can then verify the human body motion in real-time. (See Fig. 4. and Fig.5.) The system can also export video files, and cooperate with other rendering software including MAYA, 3DS MAX or Lightwave etc. to produce videos of reality.

4 Conclusion

In this test, MB is chosen as the software platform for human body motion simulation. Through using this software, the author of this thesis find that the software itself can implement kinematics simulation fairly well, but it lacks support for kinetics. It will be insufficient as for more complicated tasks on ergonomics analysis. The whole simulation environment will more practical if we develop our own human body motion simulation platform in our future research, and add a variety of ergonomics evaluation tools about force, torque and operation gesture etc.

This investigation has actualized the technique of human motion capturing and motion simulation, which can be fairly useful in many fields such as ergonomics, human-machine engineering, sim-training, and biodynamics etc.

References

1. Wilhelms, J.: Using Dynamic Analysis for Animation of Articulated Bodies [J]. IEEE Computer Graphics & Applications 7(6), 12–27 (1987)
2. Ko, H., Badler, N.I.: Straight-line Walking Animation based on Kinematics Generalization that Preserves the Original Characteristics [A]. In: Proceedings of Graphics Interface’93 [C]. Los Angeles California, pp. 9–16 (1993)
3. Bingxiao, L., Heqin, Z., lei, W.: A Physical based human motion simulation combined with motion capture data (in Chinese). Computer Simulation 23(3), 187–189 (2006)
4. Rui, C., Guoyu, L., Guoying, Z., Hua, L.: 3D Human Motion Tracking Based on Sequential Monte Carlo Method [J]. Journal of computer-aided design & computer graphics 1, 85–92 (2005)
5. Yi, T., Yunjian, G., Jianju, J. et al.: Computer Simulation of Human Motion and Its Tendency [J]. Journal of System Simulation 16(5), 863–867 (2004)
6. Shanlin, S.: Research summarize on virtual human body model in ergonomics field [J] (in Chinese). Chinese Journal of Ergonomics 4(1), 52–57 (1998)
7. Changshui, Y., Zhaoji, W., wen, G., Yiqiang, C.: Skeleton Building of Individual Virtual Human Model [J] (in Chinese). Journal of computer-aided design & computer graphics 1, 67–72 (2004)
8. Software of Vicon Motion Systems. Vicon Motion Systems Ltd. United Kingdom (2002)

Modeling of Human's Pointing Movement on the Effect of Target Position

Junmin Du, Haiwen Shi, and Xiugan Yuan

Institute of Human-Machine-Environment Engineering,
Beijing University of Aeronautics and Astronautics, 100083 Beijing, China
{JunminDu,adujm}@126.com

Abstract. The purpose of this study is to construct a new pointing movement model considering the effect of target position. A pointing movement experiment was designed and carried out, and the pointing movement characteristics when human upper limb touched the targets on front board are studied. The result shows that the starting point position and target position greatly affect the movement time. A new pointing movement model is built, in which the effect of target position is introduced. The new model obtains higher contribution value and could describe the data better than the conventional models.

Keywords: Human upper limb, pointing movement, Mathematic model.

1 Introduction

The research of the pointing movement started in 1950's. According the information theory, Fitts[1] and others studied the relationship between target distance and movement time in one-dimensional movement, and brought the famous Fitts law first and foremost. After then, many researches extended the Fitts law from various aspects, and the research for pointing movement was also extended from one-dimensional to three-dimensional. As a result, some performance models of two-dimension and three-dimension pointing movement tasks were constructed [2-7]. However, these mathematic models were all built based on the immovable starting point. While the starting point usually has different position in practice. It is therefore the purpose of this study focuses on the experiment, and to construct a new model which considering the effect of the target position on movement time.

2 Methods

2.1 Subjects

21 male graduate students, who had neither musculoskeletal abnormalities nor optical disorders, participated in the experiment. The subjects were all right-handed. Their mean (SD) age, height and weight were 23.8 (2.0) years, 1.726 (0.021) m and 64.4 (5.3) kg respectively.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 48–55, 2007.

2.2 Design and Procedure

The whole pointing procedure is captured with Vicon motion tracking system. The sketch map of pointing experiment is shown as Fig.1. The target board is fixed on a table, and the board has 17° obliquity from a vertical plane. The subject is required to sit on a chair in front of the target board. The height and location of the chair is adjustable so that the subject could pose comfortable and meet the experiment requirements. The subject's body vector plane is superposed with the board perpendicularity line crossing the board centre point.

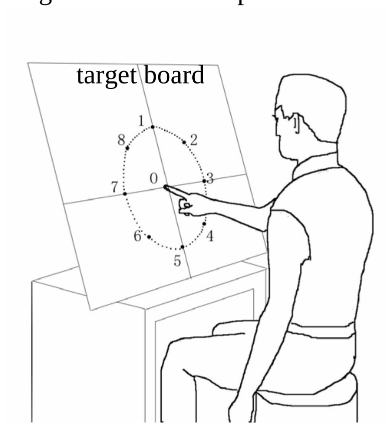


Fig. 1. The sketch map of pointing experiment

The location of nine target-circles is shown in Fig.2. Those target-circles' diameter are all 0.4m. The target-circles are arranged regularly for three rows and three columns. The O1~O9 signed in Fig.2 are the location of the center of target-circles. Eight targets scatter homogeneously on the circumference, as shown in Fig.3. In order to alleviating the vision interference, nine target-circles are drawn on six pictures according their location on the board, and there are only one or two target-circles on the board when the experiment is performed.

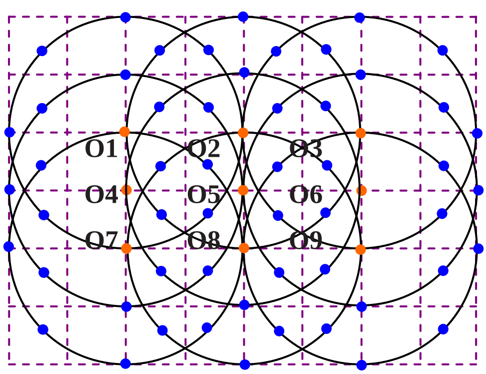


Fig. 2. The location of the target-circles on the board

Before the experiment start, each subject was aware of the experimental procedure. During the experiment, the subject must keep the shoulder steadily, if the subject's body posture is changed or shoulder moves greatly or touches wrong target, the experiment is abnegated and redo again. At the beginning of the experiment, the subject's index fingertip touches the center of a target-circle, which is the starting point. Subjects are instructed to carry out the movement as accurately and as quickly as possible. For the eight target points on one target-circle, the pointing order is clockwise, and each target point is touched three times by each subject.

The experimental factors include as the following:

(a) the distance from the starting point to the target point: 0.2m.

(b) the target shape and size: solid round with diameter of 0.013m.

(c) the approach direction from the starting point to the target point: take the line from target-circle center to 3rd target point as the datum line, define the 3rd angle as 0°. The angle of target points increase by anti-clockwise, shown as Fig.3.

(d) the direction of target-circle center: define the center line of middle column target-circles (O2-O5-O8) as datum line, the left forward direction set as negative, the right as positive.

(e) the distance between target-circles center: for row the distance is 0.1m, for column the distance is 0.2m.

(f) the position of target-circles: define the center line of middle column target-circles (O2-O5-O8) as datum line, the left forward direction set as negative, the right as positive.

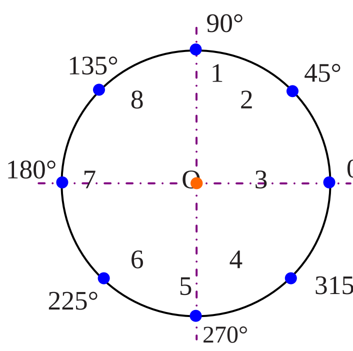


Fig. 3. The distribution of the targets on a single target-circle

3 Results

3.1 Fitting the Data to the Conventional Pointing Movement Model

Fitts law is a classical model for predicting one-dimensional pointing movement. MacKenzie[2] improved the Fitts model, and extended it to two-dimensional pointing movement, the formula is

$$t = a + b \bullet \log_2(2x/s + 1.0) \tag{1}$$

where x is the distance from the starting point to the target point, s is the target size, a and b are regression coefficients. The log term ($log_2(2x/s+1.0)$) in Eq.(1) indicates the index of difficulty(ID), viz. $ID=log_2(2x/s+1.0)$.

Based on Eq.(1), Iwase and Murata[5,6] built a three-dimensional pointing movement model of human body upper limb with considering the affect of approach direction to the target:

$$t = a + b \bullet \beta(\theta) \bullet \log_2(x(\theta)/s + 1) \qquad (2)$$

where $\beta(\theta)= log_e 2\cdot(x(\theta)+s)/(b\cdot v(\theta))$, $x(\theta)=x\cdot\{(1+ a\,(\sin\theta+1)\}$, a is a positive constant number, θ is the approach direction angle from starting point to target, which defined as shown in Figure3, $v(\theta)$ is mean velocity of pointing movement, a and b are regression coefficients. $ID=log_2(x(\theta)/s+1.0)$.

Compared with other conventional models, Eq.(1) and Eq.(2) could describe two-dimensional and three-dimensional pointing movement better[5,6]. So the two

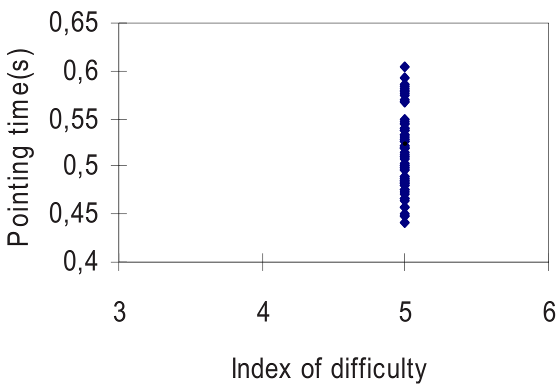


Fig. 4. Relationship between the index of difficulty and the pointing time using Eq.(1) (*R2= 4E-14*)

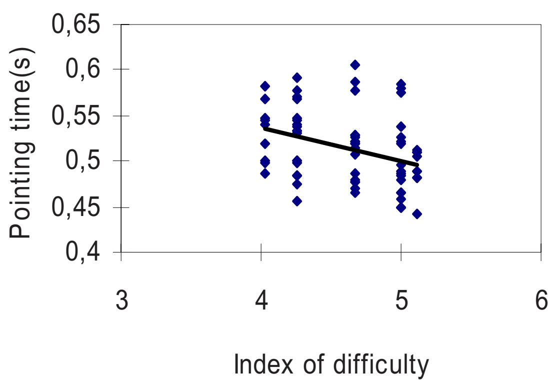


Fig. 5. Relationship between the index of difficulty and the pointing time using Eq.(2) (*R2= 0.1123*)

equations are used to do the linear regression of the experiment data. Fig.4 and Fig.5 are the relationship between the index of difficulty and the pointing time using Eq.(1) and Eq.(2). The contribution rates of the regression equation are very small, one is $R^2=4E\text{-}14$ and another is $R^2=0.1123$, which indicates that Eq.(1) and Eq.(2) cannot satisfy the demand of the experiment data. According to the experimental conditions, the distance from starting point to the target point *x* and the target size *s* are constant numbers, so *x/s* is constant. Although the approach direction is considered in Eq(2), the equation can't fit the data yet. The difference of target-circles position are not mentioned in Eq(1) and Eq(2), that's the main reason that the equations can't fit the experiment well.

3.2 Analyzing the Discipline of Pointing Movement Time

Fig.6 and Fig.7 reveal the pointing time of middle column target-circles and middle row target-circles respectively.

The significance test result reveals that the pointing time of eight targets on the same target-circle is significant ($p<0.05$, a=0.05). It means that wherever the target-circle

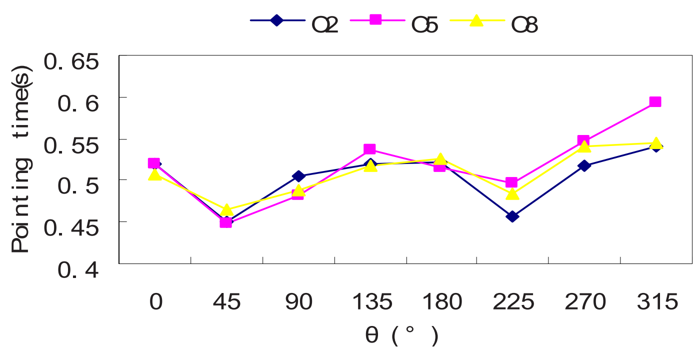


Fig. 6. The pointing time of middle column target-circles

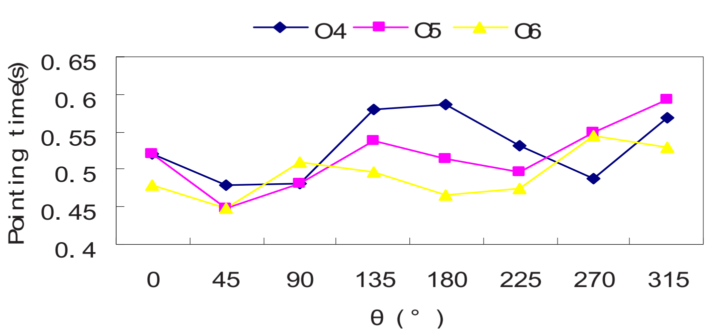


Fig. 7. The pointing time of middle row target-circles

is located, the pointing time of targets scattered on a same target-circle are different. This difference is mainly caused by the distinction of approach direction to the target. The significance test on the pointing time of the targets with same approach direction shows that the difference is insignificant when the target-circles in a column line, and the difference is significant when the target-circles in a row line. Fig.6 and Fig.7 could testify these results. In a column, the pointing time of the targets with same approach direction are very close and have no significant difference, so the pointing time could be regarded as equally. While in a row, the pointing time is strongly affected by the location of the target-circle. These analyses suggest that when structuring the new pointing movement model, it is necessary to consider the effect of the approach direction to the target as well as the target-circle location in the row, while the effect of the target-circle location in column could be ignored.

3.3 Constructing the New Pointing Movement Performance Model

First of all, introduce the approach direction from the starting point to the target into conventional Fitts model. The Fitts model is:

$$t = a + b \bullet \log_2(2x/s) . \quad (3)$$

Fig.6 is the pointing time of the target-circles in the middle column. Because the relationship between the approach direction and the pointing time is similar to the sin curve, we define the *sinθ* and *sin2θ* as the factors. Then the index of difficulty (ID) is revised as:

$$ID = [\log_2(x/s) - \sin(2\theta)][1 - k_1 \sin\theta] . \quad (4)$$

where K_1 is a constant positive number according to the data, θ is the approach direction angle.

Fig.2 reveals the relative position of the nine target-circles. Set the target-circles center line of middle column as the datum line, the target-circles in the left and the right column have different directions and have certain distance from the datum line. The different directions and distance contribute greatly to the pointing time difference. Define the direction of target-circle center and the distance between target-circles as the factors, and induct them into Eq.(4), then Eq.(4) could be improved to the following expression:

$$ID = [\log_2(x/s) - \sin(2\theta)][1 - k_1 \sin\theta] - k_2(\cos\theta + k_3) . \quad (5)$$

where K_2 is the target-circle center distance coefficient, which is defined by the relative distance of target-circles, K_3 is the target-circle center direction coefficient, which is defined as: $K_3 = -1$ when target-circle is in the left column, $K_3 = 0$ when target-circle is in the middle column, $K_3 = 1$ when target-circle is in the right column.

Fig.7 shows that, compared with the middle target-circle (O5), the pointing time of the left and the right target-circle (O4 and O6) have certain offsets. The offsets are caused mainly by the different locations of target-circles. Accordingly, a target-circle location modified item should be inducted in Eq.(5). So the Eq.(5) could be revised as:

$$ID = [\log_2(x/s) - \sin(2\theta + \alpha)][1 - k_1 \sin(\theta + \alpha)] - k_2(\cos\theta + k_3) \; . \quad (6)$$

where a is target-circle location modified coefficient, which has the same sign as K_3.

According to Eq.(6), a new performance model that considering the effect of starting point and target position could be obtained:

$$t = a + b \bullet [\log_2(x/s) - \sin(2\theta + \alpha)][1 - k_1 \sin(\theta + \alpha)] - k_2(\cos\theta + k_3) \; . \quad (7)$$

where a and b are regression coefficients, x is the distance from the starting point to the target, s is the target size, θ is the approach direction from the starting point to the target, K_1 is a constant positive number according to the data, K_2 is the target-circle center distance coefficient defined by the relative distance of target-circles center, K_3 is the target-circle center direction coefficient ($K_3 = -1$ when target-circle is in the left column, $K_3 = 0$ when target-circle is in the middle column, $K_3 = 1$ when target-circle is in the right column), a is target-circle location modified coefficient with the same sign as K_3.

Fig.8 shows the relationship between ID and pointing time using Eq.(7). The contribution rate of the regression equation (R^2=*0.751*) is much higher than that in Fig.4 and Fig.5. The new model could describe the data better than the conventional Fitts' models.

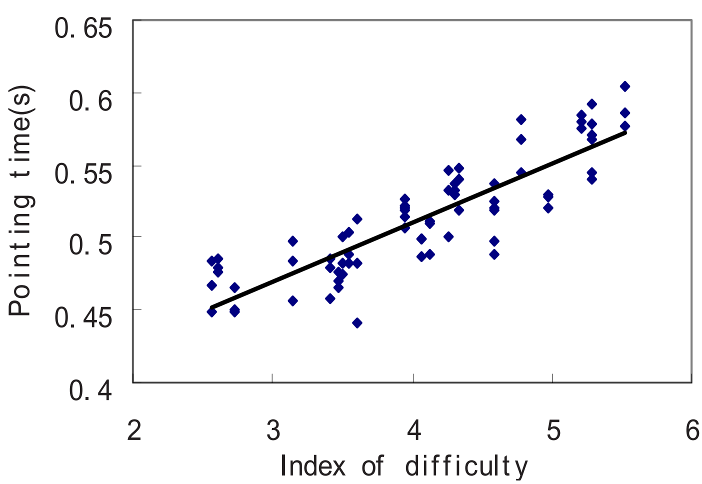


Fig. 8. Relationship between the index of difficulty and the pointing time using Eq.(7)(*R2= 0.7516*)

4 Conclusions

In this paper, a new performance model of pointing movement has been constructed, which is mainly based on the consideration of the effect coming from the starting point and the target position. Several factors have been introduced into the model, such as the approach direction from the starting point to the target, the target-circle center distance coefficient, the target-circle center direction coefficient, the target-circle location modified coefficient and etc. The new model could describe the data better than the conventional models.

References

1. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology 47(6), 381–391 (1954)
2. MacKenzie, I.S., Buxton, W.: Extending Fitts Law to Two-Dimensional Tasks. In: Proceedings of the CHI'92 conference on Human Factors in computing Systems. New York, pp. 219–226 (1992)
3. Mutata, A.: Empirical Evaluation of Performance Models of Pointing Accuracy and Speed With a PC Mouse. International Journal of Human-Computer Interaction 8(4), 457–469 (1996)
4. Yang, N.F., Wang, R.C., Jin, D.W.: Evaluation method of human upper limb movement function based on Fitts' law. Chinese journal of rehabilitation medicine 16(6), 336–339 (2001)
5. Iwase, H., Murata, A.: Modelling of human's three-dimensional movement-Extending Fitts' model to three-dimensional pointing task. In: IEEE international worksop on robot and human interactive communication, 594–599 (2001)
6. Murata, A., Iwase, H.: Extending Fitts' law to a three-dimensional pointing task. Human Movement Science 20, 791–805 (2001)
7. Zhai, S., Kong, J., Ren, X.: Speed-accuracy tradeoff in Fitts' law tasks – on the equivalency of actual and nominal pointing precision. International journal of human-computer studies 61, 823–856 (2004)

A Modified Particle Swarm Optimizer Using an Adaptive Dynamic Weight Scheme

Shu-Kai S. Fan and Ju-Ming Chang

Department of Industrial Engineering and Management
Yuan Ze University, Taoyuan County, Taiwan 320, Republic of China
simonfan@saturn.yzu.edu.tw

Abstract. Particle swarm optimization (PSO) is a stochastic, population-based optimization technique that is inspired by the emigrant behavior of a flock of birds searching for food. In this paper, a nonlinear function of decreasing inertia weight that adapts to current performance of PSO search is presented. Meanwhile, a dynamic mechanism to adjust decrease rates is also suggested. Through the experimental study, the new PSO algorithm with adaptive dynamic weight scheme is compared to the exiting models in terms of various benchmark functions. The computational experience shows some great promise.

Keywords: particle swarm optimization (PSO), dynamic inertia weight, population-based optimization technique.

1 Introduction

Many population-based search algorithms take inspiration from some natural system involving groups of interacting individuals, such as genetic algorithm (GA) [1], ant colony optimization (ACO) [2] and particle swarm optimization (PSO) [3]. Metaheuristics other than population-based approaches include simulated annealing (SA), tabu search (TS), variable neighborhood search (VNS), among others. The population-based evolutionary computation technique is operated through cooperation and competition among the potential solutions. One of prominent merits owned by the population-based techniques is the ability to anchor an approximate optimum solution quite effectively when applied to complicated optimization problems. PSO was first introduced by Eberhart and Kennedy [3] in 1995. It is noted for dealing with optimization in continuous, multimodal search spaces. Although PSO can produce remarkable performance for many practical problems of continuous optimization, the primitive PSO still suffers the consequences of slow convergence or convergence failure if the problem size is getting large [4]. Researchers from different research fields have been devoted to improving the performance of the original PSO. To date, the most well-known modification is the one incorporating an extra parameter called *inertia weight* into the original PSO algorithm. The objective is to balance the global and local exploration capability [5]. The inertia weight is deliberately designed to adjust the current velocity that resembles a local search behavior. This can be visualized by tracking the search locus of the best particle *en route*. By including the inertia weight together with sharing information among swarm, the particles tend to

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 56–65, 2007.

expand their search ability as much as they can; that is, they become more capable of exploring the new areas (namely, global search ability). Combining both the local and global search would definitely benefit solving highly nonlinear, multimodal optimization problems.

The inertia weight ω is employed as to control the impact of the previous history of velocities on the current velocity, therefore influencing the trade-off between global and local exploration of particle search. A large value of inertia weight facilitates global exploration, while a small inertia weight tends to facilitate local exploration that helps to fine-tune the current neighborhood search. A suitable selection of inertia weight can provide an appropriate balance between global and local exploration abilities, so less computation effort is anticipated for convergence. Shi and Eberhart [6] claimed that a reasonable choice of ω should decrease gradually while the swarm search progresses. As a general remark, they opined that a better performance would be expected if the inertia weight is chosen time-varying, linearly decreasing with iteration, rather than just using a constant value. Their argument is supported by examining a single case study. It was inferred that PSO should start with a high inertia weight for coarse global exploration and the inertia weight should linearly decrease to facilitate finer local explorations in later iterations. This should help PSO to approach the optimum efficiently.

Based on the hypothesis that dynamic adjustment of inertia weight is necessary for PSO, some improvements have been reported in the literature. Most recently, Chatterjee and Siarry [7] proposed a PSO model with dynamic adaptation that concentrates on the adjustability of the decreasing rate for the inertia weight ω. In their new approach, a set of exponential-like curves are used to create a complete set of free parameters for any given problem, saving the data from a tedious, trial-and-error-based approach to determine them for each specific problem. As evidenced by the experimental results, their approach can successfully solve several benchmark functions and attain better convergence than existing inertia weight PSO algorithms. Nonetheless, an obvious shortcoming is that the maximum number of iterations to define the decreasing inertia weight needs to be specified in advance.

2 The Particle Swarm Optimization (PSO)

PSO is a swarm intelligence algorithm that emulates a flock of birds searching over the solution landscape by sampling points and the swarm converges on the most promising regions. A particle moves through the solution space along a trajectory defined by its velocity, the draw to return to a previous promising search area, and an attraction towards the best area discovered by its neighbors. The force behind the convergence of a swarm is of social pressure, applied through the interplay between the particles in flight.

2.1 The Static PSO

Eberhart and Kennedy [3] developed a ground-breaking algorithm through simulating social behavior, which is called particle swarm optimization (PSO). In a particle swarm, each individual is influenced by its closest neighbors and each particle is a possible solution to the problem being optimized. PSO formulae define each particle as a potential solution in N-dimensional space, with each indicated as X_{id}. The update of the ith particle in the swarm of size s is accomplished according to the equations (1) and (2), as expressed by:

$$V_{id} = V_{id} + c_1 r_1 (P_{id} - X_{id}) + c_2 r_2 (P_{gd} - X_{id})\,; \tag{1}$$

$$X_{id} = X_{id} + V_{id}, \text{ for } d = 1,\ 2, \ldots,\ \text{N},\ i = 1,\ 2, \ldots,\ s\,. \tag{2}$$

Equation (1) calculates a new velocity for each particle based on its previous velocity V_{id}, and the previous best particle P_{id}, and the global best particle P_{gd}. Equation (2) updates each particle's position in the solution space. The two random numbers r_1 and r_2 are independently generated, and c_1 and c_2 are the learning factors determining the relative influence of the cognitive and social components, respectively. The value of each component in the velocity vector can be clamped to the range $[-v_{\max}, v_{\max}]$ so as to reduce the likelihood of particles leaving the search space. The value of $v_{\max}$ only limits the maximum distance that a particle can move in every iteration. Since the pioneering PSO work done by Kennedy and Eberhart [3], abundant research has been working on improving its performance in various ways, thereby publishing many interesting modified PSO versions. One of the most prominent PSO models [6] introduces a parameter called the inertia weight ω into the original PSO algorithm, and therefore the velocity update rule becomes:

$$V_{id} = \omega V_{id} + c_1 r_1 (P_{id} - X_{id}) + c_2 r_2 (P_{gd} - X_{id})\,. \tag{3}$$

The parameter of inertia weight is used to balance the global and local search abilities. A large inertia weight facilitates global search; a small inertia weight facilitates local search.

2.2 The Proposed Dynamic Adaptive PSO Algorithm

In an empirical study on PSO [6], Shi and Eberhart claimed that a linearly decreasing inertia weight could improve local search towards the end of a run, rather than using a constant value throughout. They supported their advocate with a single case study. Furthermore, it was also asserted that the swarm may fail to converge to a good solution by restricting the global search ability too much. It implies that a large value of the inertia weight could make the particles own more global search ability to explore new search areas; on the other hand, a small inertia weight produces less variation in velocity to perform a better local search. In this light, a decreasing function for the dynamic inertia weight can be devised in the following

$$\omega = \left(\text{iter}_{\max} - \text{iter}_{\text{cur}}\right)\left(\frac{\omega_{\text{initial}} - \omega_{\text{final}}}{\text{iter}_{\max}}\right) + \omega_{\text{final}}\,, \tag{4}$$

where ω_{initial} and ω_{final} represent the initial and final inertia weights respectively at the start of a given run, $\text{iter}_{\max}$ the maximum number of iterations in a offered run, and $iter_{cur}$ the current iteration number at the present time step. More recently, a different version of dynamic inertia weight was proposed by Chatterjee and Siarry [7]. They presented a different nonlinear function that modulates the inertia weight with time for improving the performance of the PSO algorithm. The main modification is the determination of the inertia weight as a nonlinear function of the current iteration number at each time step. The dynamic function of ω was modified as follows:

$$\omega = \left(\frac{iter_{max} - iter_{cur}}{iter_{max}} \right)^n (\omega_{inital} - \omega_{final}) + \omega_{final} . \tag{5}$$

Through the study of the nonlinear modulation parameter *n*, they also derived a reasonable set of choice for the free parameters. In addition, better results have been obtained for several benchmark problems while the nonlinear modulation n=1.2 was opted. The modified PSO is initially with a high inertia weight to explore new search areas. Later, with gradually decreasing inertia weights following different paths of different values of *n*, the final inertia weight is reached at the maximum number of iterations. In this paper, a new version of dynamic, adaptive inertia weight is proposed, which will be detailed in the next section.

2.3 Adaptive PSO Model with Dynamic Inertia Weight

With the decreasing inertia weight model, once the swarm finds the local area where the optimal solution resides, the particles would be attracted toward the neighborhood. The trajectory of each particle is pinned to its best ever visited solution. As new, better solutions are discovered, each particle updates its current solution for seeking out even better solutions. As such, the search spread is getting tight step by step. Consider a situation that the final destination is only a “local” optimal solution in that area. When converged, the decreasing inertia weight must turn out to be an extremely small value and the particles hardly escape from the local optimal area. A reasonable

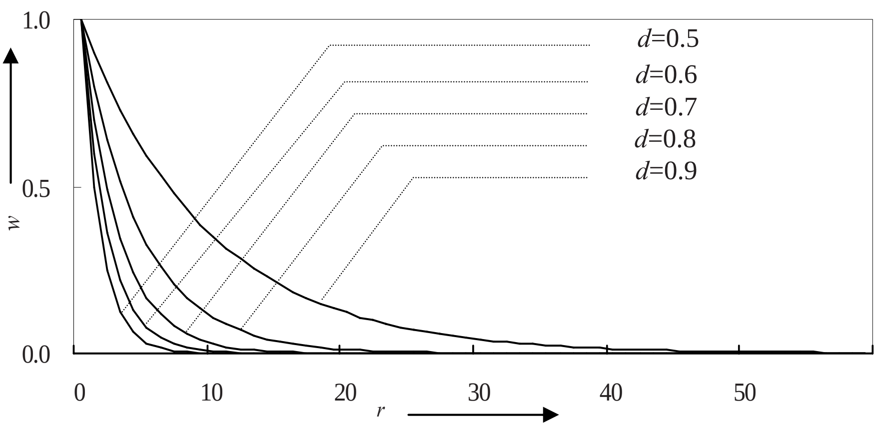


Fig. 1. Inertia weights with different values of the decrease rate *d*

strategy for improvement can be envisioned that if a particle finds a better solution then more energy (*i.e.*, weight) is given onto the current velocity to speed up exploitation, and *vice versa*. Under such circumstances, extended search and/or local refinement can be realized.

To achieve the foregoing concept, the present paper proposes a novel nonlinear function regulating inertia weight adaptation with a dynamic mechanism for enhancing the performance of PSO algorithms. The principal modification is the determination of the inertia weight through a nonlinear function at each time step. The nonlinear function is given by

$$\omega = (d)^r \omega_{initial} , \tag{6}$$

where d is the decrease rate ranging between 1.0 and 0.1, and r the dynamic adaptation parameter depending on the following rules for successive adjustment at each time step. For a minimization case, it follows

$$\text{if } f\left(P_{gd-new}\right) < f\left(P_{gd-old}\right) \text{ then } r \leftarrow r - 1 ; \tag{7}$$

$$\text{if } f\left(P_{gd-new}\right) \geq f\left(P_{gd-old}\right) \text{ then } r \leftarrow r + 1 , \tag{8}$$

where P_{gd-new} and P_{gd-old} denote the global best solutions at current and previous time steps, respectively. This mechanism wishes to make particles fly more quickly toward the potential optimal solution, and then through decreasing inertia weight to perform local refinement around the neighborhood of the optimal solution. Fig. 1 illustrates decreasing inertia weight plots with different decrease rates.

3 Behavior of Proposed Adaptive Dynamic PSO

To study the behavior of the proposed PSO model, the Sphere function was applied, which has a single minimum with adjustable problem size n, as defined by

$$\text{Sphere}_n(x) = \sum_{i=1}^{n} x_i^2 . \tag{9}$$

The Sphere function has the objective value 0 at its global minimum $x = (0, \ldots, 0)$. In this experiment of parameter analysis, the 10-dimensional Sphere

Table 1. Performance returned by the proposed PSO model for the Sphere function using various decrease rates (d)

Decrease rate (d)	Success rate (%)	Average function error	Iterations
0.5	100	9.9518e–7	519
0.6	100	9.8836e–7	362
0.7	100	8.6420e–7	281
0.8	100	7.4526e–7	197
0.9	100	7.1711e–7	137
0.95	100	9.3543e–7	153

model is used, the swarm size is $s = 20$, and the initial inertia weight is $\omega_{initial} = 1$. The search domain is restricted within $[-5.12,\ 5.12]$. The problem was optimized for 50 independent runs, and all particles in each dimension can have a maximum allowable velocity which is confined to the full range in a given dimension.

For evaluation purposes, it is necessary to give a formal definition of successful minimization, an instance to be considered as a success if the following inequality holds:

$$\left|\mathrm{FOBJ}_{\text{Proposed-PSO}} - \mathrm{FOBJ}_{\text{ANAL}}\right| < \varepsilon_{rel} \times \mathrm{FOBJ}_{\text{ANAL}} + \varepsilon_{abs}\,, \tag{10}$$

where the relative tolerance $\varepsilon_{rel} = 10^{-4}$ and the absolute tolerance $\varepsilon_{abs} = 10^{-6}$ are chosen. $\mathrm{FOBJ}_{\text{Proposed-PSO}}$ is referred to as the best function value achieved by the proposed algorithm, and $\mathrm{FOBJ}_{\text{ANAL}}$ is referred to as the exact analytical global objective. The optimization results are listed in Table 1, containing the average function error, the success rate over 50 runs, and the average iterations required. It can be observed from the table that the optimum decrease rate for the Sphere function is 0.9, requiring minimum iteration number 137 for convergence.

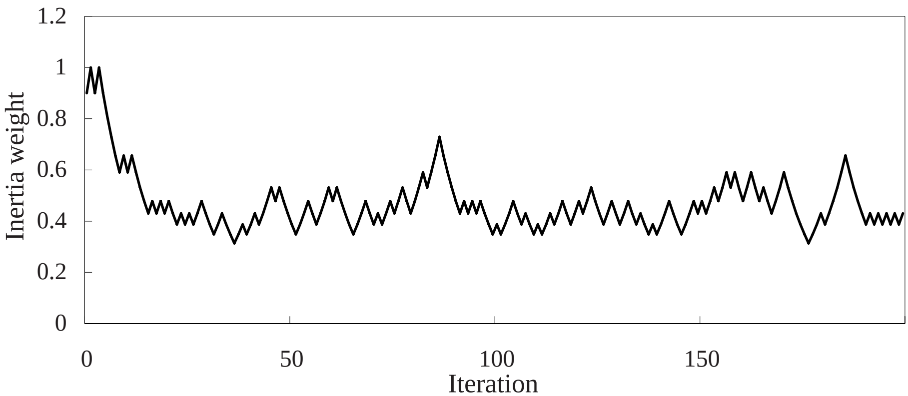


Fig. 2. Adaptive inertia weight profile using decrease rate *d* = 0.9 when solving the 10-D Sphere function

Fig. 2 pictorially shows the trend of adaptive inertia weights in a single optimization run of the 10-D Sphere function. Obviously, most of inertia weights fall from 0.4 to 0.7. This range of ω was also suggested in [8-9]. Hence, a choice of using d = 0.9 can be considered a rule of thumbs for general cases. In the figure, the spike indicates extended exploration conducted in the favorable search direction that has been found; a zigzagging that hovers around a specific inertia weight level indicates local refinement. The parameter setting recommended here will be used for subsequent experimental comparisons to be presented in section 4.

4 Performance of Proposed PSO Model

This section presents the performance of the proposed PSO model on 8 benchmark functions with 2 to 10 dimensions. They consist of 2-dimensional Rosenbrock, Quadric, Ackley, Rastrigin, Griewank, and Zakharov functions, and two higher dimensional problems, *i.e.*, 5- and 10-dimensional Zakharov functions. These benchmark functions have the global optimum objective of zero. Furthermore, the optimization results were also compared with the other two dynamic PSO models using different decreasing inertia weight functions. These two schemes are the nonlinear inertia weight PSO [7] and linear inertia weight PSO models [6]. The proposed PSO model uses the parameter setting: $\omega_{initial} = 1$ and $d = 0.9$. Note again that all experiments were run 50 times with 20 particles for 500 iterations. $v_{\max}$ is clamped to the full range in a given dimension for each benchmark function. The results reported are the averages calculated from all 50 runs. Table 2 shows the average objective function errors returned by using three dynamic inertia weight PSO versions. The proposed PSO model performed best (in solution precision) for all test problems except the Rastrigin function for which an equally good result was yielded by the other two schemes.

Table 2. Average objective function errors computed over 50 runs by three PSO models for 2 to 10 dimensional benchmark functions

Function	Decreasing Inertia Weight Method		
	Proposed Adaptive *w* PSO	Nonlinear *w* PSO [7]	Linear *w* PSO [6]
Rosenbrock	1.329985e−10	3.151920e−06	1.510120e−05
Quadric	7.420908e−101	2.534761e−98	2.756407e−91
Ackley	5.887218e−16	1.835229e−14	5.152140e−13
Rastrigin	0.000000e+00	0.000000e+00	0.000000e+00
Griewank	1.194600e−13	2.764728e−06	1.067861e−04
$Zakhavior_2$	4.413094e−114	3.030987e−88	1.588612e−80
$Zakhavior_5$	2.821841e−43	1.156531e−38	1.464352e−32
$Zakhavior_{10}$	1.520226e−12	1.347304e−07	5.346056e−04

To gain a better understanding of convergence property, Fig. 3 plots the logarithm of function error achieved by three PSO models versus iteration for the 10-D Zakhavior problem. The proposed PSO model exhibits an excellent, monotone convergence speed.

To further verify the proposed PSO model, 5 benchmark functions of higher dimension were tested in this experiment. The problem size of 30 was for every benchmark problems, and the performance was evaluated with 2000 iterations in each run. The optimization results were reported according to three different swarm sizes, 20, 40 and 60 particles. Table 3 presents the comparison results. Likewise, the proposed PSO model outperformed the other two models almost in every cases.

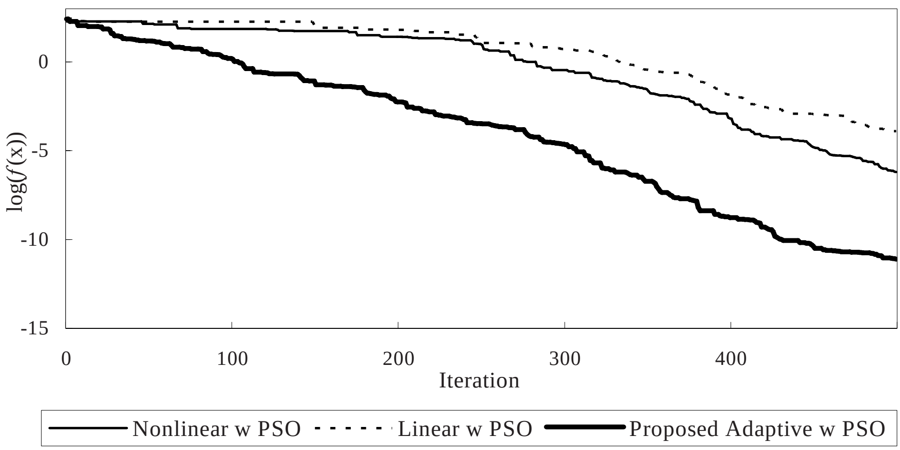


Fig. 3. Objective improvement of three PSO models for the 10-D Zakhavior function

Multimodal functions such as Ackley, Rastrigin and Griewank functions with many local minima positioned on a regular grid would pose a serious challenge for the algorithms. Yet, the proposed model still converged faster than the other two models, as can be seen from Fig. 4 for the Rastrigin function. The nonlinear and linear inertia weight PSO models appeared to get trapped in two different local minima.

Table 3. Average objective function errors returned by three PSO models for the 30-D problems using 3 different swarm sizes

Function	Pop.	Decreasing Inertia Weight Method		
		Proposed Adaptive *w* PSO	Nonlinear *w* PSO	Linear *w* PSO
Rosenbrock	20	8.386216e+01	1.361161e+02	1.538252e+02
	40	6.345287e+01	8.074967e+01	9.734088e+01
	60	3.785684e+01	6.664381e+01	8.009522e+01
Quadric	20	1.254913e+03	1.684360e+03	2.263458e+03
	40	1.297764e+02	6.945127e+02	8.299236e+02
	60	4.813780e+01	4.139014e+02	6.293952e+02
Ackley	20	2.450348e+00	2.775425e+00	3.825188e+00
	40	1.109738e+00	1.666886e+00	2.701517e+00
	60	9.706417e–01	1.103057e+00	1.275164e+00
Rastrigin	20	5.366067e+01	5.897047e+01	8.184283e+01
	40	3.676980e+01	3.634060e+01	7.364739e+01
	60	1.008213e+01	2.848330e+01	6.390452e+01
Griewank	20	8.849953e–02	1.344367e–01	9.558794e–01
	40	2.460268e–02	3.195703e–02	4.928395e–02
	60	9.857285e–03	2.956801e–02	9.358511e–02

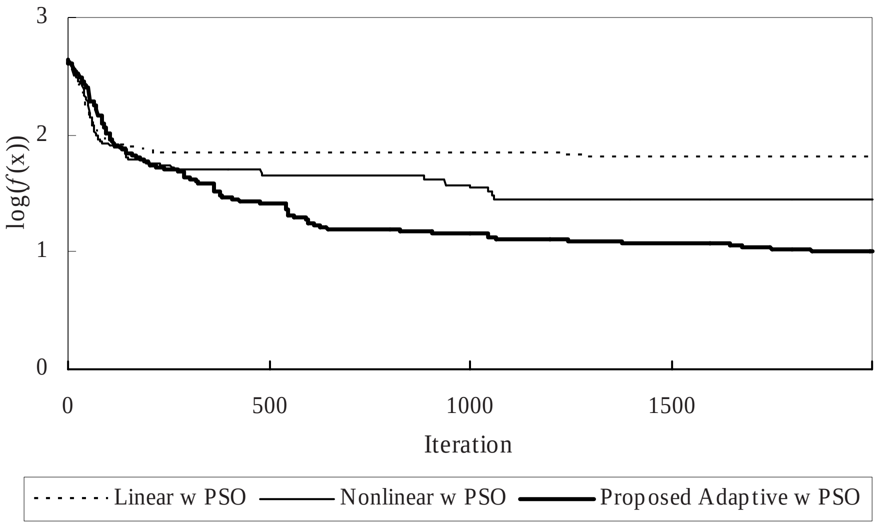


Fig. 4. Objective reduction versus iteration for the 30-D Rastrigin function

5 Conclusion

This paper presented a new adaptive scheme which employs a nonlinear function to provide dynamic inertia weights for PSO. This resulted in a significant improvement in performance, especially in terms of the solution quality and convergence speed. The success of the proposed model can be attributed mainly to an adequate selection of decrease rate. Meanwhile, a dynamic mechanism to adjust decrease rates is also suggested. A major difference from Chatterjee and Siarry's model is that the maximum number of iterations for designing the decreasing inertia weight need not be known beforehand. Through a series of experimental studies, the new PSO model was compared to two exiting models in terms of various benchmark functions. The computational experience shows some great promise.

References

1. Goldberg, D.E.: Genetic Algorithms in Search Optimization, and Machine Learning. Addison-Welsey, Reading MA (1989)
2. Colorni, A., Dorigo, M., Maniezzo, V.: Distributed optimization by ant colonies. In: Proceedings of CAL'91—European Conference on Artificial Life, pp. 134–142. Elsevier Publishing, Paris, France (1998)
3. Eberhart, R.C., Kennedy, J.: A new optimizer using particle swarm theory. In: Proceedings of the Sixth international Symposium on Micro Machine and Human Science. pp. 39–43, IEEE Service Center, Piscataway, NJ, Nagoya, Japan (1995)

4. Angeline, P.J.: Evolutionary optimization versus particle swarm optimization: philosophy and performance differences. In: Porto, V.W., Waagen, D. (eds.) Evolutionary Programming VII. LNCS, vol. 1447, pp. 601–610. Springer, Berlin (1998)
5. Shi, Y., Eberhart, R.C.: A modified particle swarm optimizer. In: Proceedings of the IEEE International Conference on Evolutionary Computation, pp. 69–73. IEEE Press, Piscataway NJ (1998)
6. Shi, Y., Eberhart, R.C.: Parameter selection in particle swarm optimization. In: Evolutionary Programming VII. Proc. EP98, pp. 591–600. Springer, New York (1998)
7. Chatterjee, A., Siarry, P.: Nonlinear inertia weight variation for dynamic adaptation in particle swarm optimization. Computers & Operations Research 33, 859–871 (2006)
8. Clerc, M.: Discrete Particle Swarm Optimization Illustrated by the Traveling Salesman Problem. Technical report (2000)
9. Parsopoulos, K.E., Vrahatis, M.N.: Initializing the particle swarm optimizer using the nonlinear simplex method. In: Grmela, A., Mastorakis, N.E. (eds.): Advances in Intelligent Systems, Fuzzy Systems, Evolutionary Computation, pp. 216–221, WSEAS Press (2002)

Prediction of Discomfort During Arm Movements

Florian Fritzsche and Heiner Bubb

Lehrstuhl für Ergonomie, Technische Universität München,
Boltzmannstr.15, 85747 Garching
fritzsche@tum.de

Abstract. On the way to a generalized discomfort model for movements steps are presented that calculate the determining parameters for the model. Discomfort is mainly dependant on posture and relative torque. A multi body system arm model is used to calculate the driving torques of a lifting task using inverse dynamics. A motion analysis of the movement was carried out and the corresponding angles were used to drive the arm model. In order to calculate relative torque a torque velocity relationship according to Hill was implemented in the arm model.

1 Introduction

In the field of ergonomic design the concept of discomfort is gaining more importance. Car manufacturers are aiming to predict the discomfort of customers performing typical tasks in and around the car, e.g. closing the door, pulling the hand brake or lifting a beverage crate out of the car boot. In the planning process of an assembly line the inner forces and moments of the worker are of interest in order to estimate if this work can be performed over a period of time. Recent studies investigated the ingress/egress movement. Relevant measures for the analysis of this movement are reactive joint forces, muscle effort [1] and discomfort [2]. Cherednichenko [3] examined the car ingress with the concept of leading body parts that control the movement. He used a force based method to simulate the dependant body parts which was applied to the human model RAMSIS. In order to analyze movements for the optimum ergonomic design process kinetic values like joint moments performed by a human during a specific task need to be calculated. This implies the use of inverse dynamic calculations.

Preliminary work on a strength based discomfort model for posture and movement was done by Zacher [4] at the Lehrstuhl für Ergonomie at the Technische Universität München. The aim is to develop a generalized discomfort model for movements of the whole body with the aim of predicting discomfort for different anthropometries and tasks. The results can be implemented in an existing digital human model like RAMSIS. Focusing first on the arm system knowledge about the static conditions of the model will be assigned to dynamic conditions.

This report will provide the necessary steps from movement analysis to the calculation of inner joint moments which are essential parameters in our approach of describing discomfort.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 66–73, 2007.

2 Experiment

The process was performed with a survey of an arm movement. The task of the subject (male, height: 188cm, weight: 102kg) was to lift a weight (10.5kg, 75% of the individual maximum weight) from a table to a platform. The table was adjusted to a height that the angle of the elbow was roughly 90°. The subject was instructed to keep the contact with the backrest and not to move forward with the trunk, so that the arm movement was not influenced by the movement of the spine. Two cameras were used to film the movement. Pictures with a frame rate of 25 HZ were saved to the hard disk. The motion was analyzed with the PCMAN [5], [6] measuring system. After the calibration of the system it is possible to superimpose the PCMAN model with the camera pictures (Fig. 1). First the segment lengths of the right clavicle, upper arm and fore arm were adjusted to the subject with the superimposition method. Afterwards the posture of the initial frame of the movement was adapted and than each successive frame was adapted by the motion tracking module. The algorithm tracks the complete movement automatically when the lifting height is 88mm and the movement is performed slowly as in this example (Fig. 2). In another configuration with a lifting height of 434mm the tracking has to be stopped when a drift occurs as the algorithm is not robust over a longer time. The next posture has to be adapted manually, so that the tracking turns into a semi automatic procedure.

Only the clavicle, shoulder and elbow joints were considered during the tracking. The angles of the wrist were only adapted for the initial frame and stayed constant afterwards. The algorithm was not accurate enough to detect axial rotations in the elbow or the small angle changes of the hand. The angles for each degree of freedom were saved in a text file.

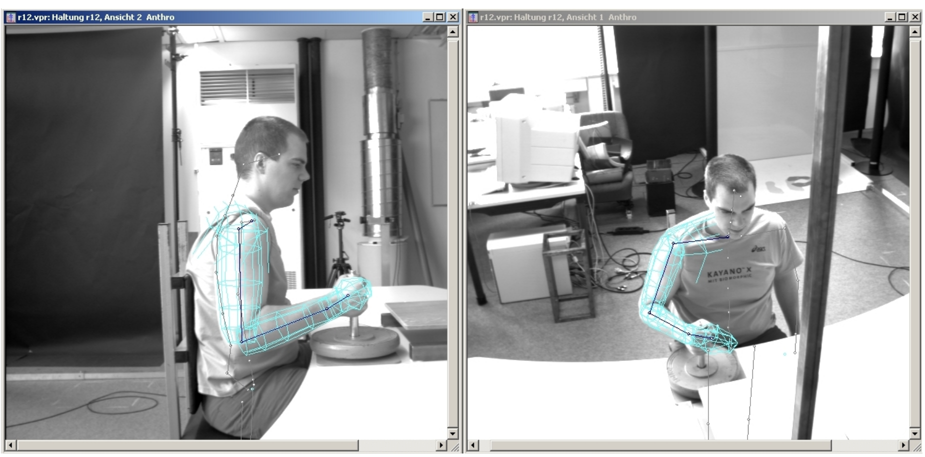


Fig. 1. Initial frame of the PCMAN motion tracking

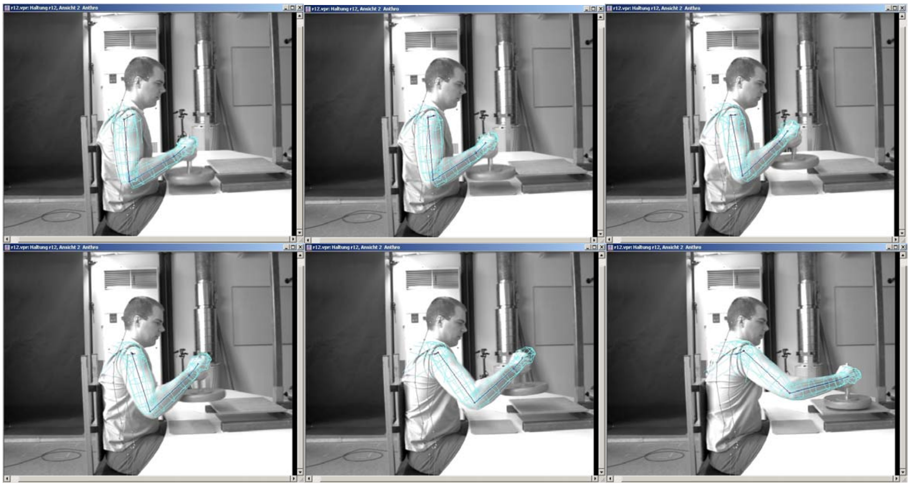


Fig. 2. Single images of a time sequence showing the lifting task

3 MBS Model DYNAMICUS

The software alaska is a toolbox for modeling and simulation of mechatronic systems developed by the Institut für Mechatronik at the Technische Universität Chemnitz. DYNAMICUS is a biomechanical multi body model [7]. It provides an interface with RAMSIS, so that posture and anthropometry of a RAMSIS model can be transferred to the MBS model. This means all joint angles, the mass of each body segment, the center of mass and the distance to the next joint in the kinematical chain. In the current version 5.0 of the software only the moments of inertia are calculated from the anthropometric model of Saziorski. Both digital human models are shown in Figure 2. The DYNAMICUS model used in this case consists of a constrained trunk, which is the basis of any model und serves as a linking to the arm system. Consequently, the human arm model consists of the clavicle, upper arm, fore arm and the hand. The DYNAMICUS Bibliography provides a variety of modeling components for each segment that differs in the room of movement (e.g. spherical and planar rotation, fixation). For the use of the RAMSIS interface special components are prepared that comply with the kinematical structure of the RAMSIS model. The hand joint is fixed as well as the axial rotation of the elbow as the motion tracking didn't detect these DOFs. Also the axial rotation of the clavicle joint is constrained as this motion is not physiological. A mass of 10,5kg was fixed to the hand at a distance that corresponds to the length of the handle of the weight that the subject lifted during the experiment.

Although both models have the same kinematical structure, the local coordinate systems and the representation of the rotations in the joints differ. So the matrix representation in each joint for the positions in each time step had to be calculated using

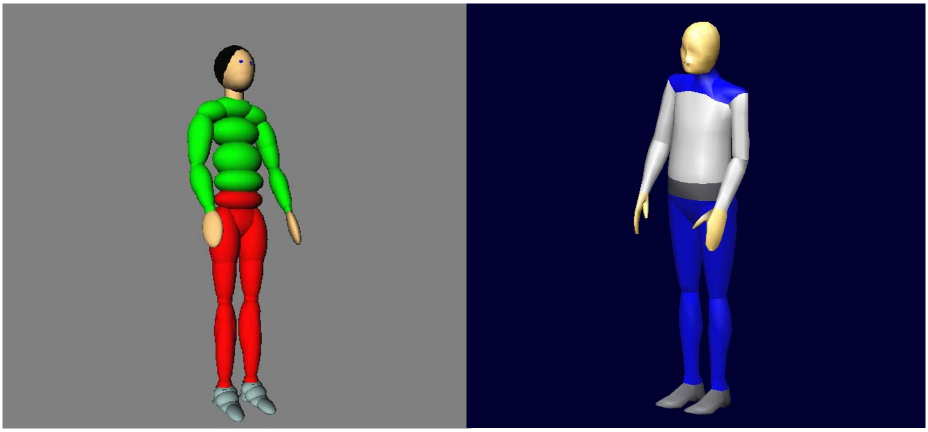


Fig. 3. The biomechanical multi body model DYNAMICUS (left) and the digital human model RAMSIS (right)

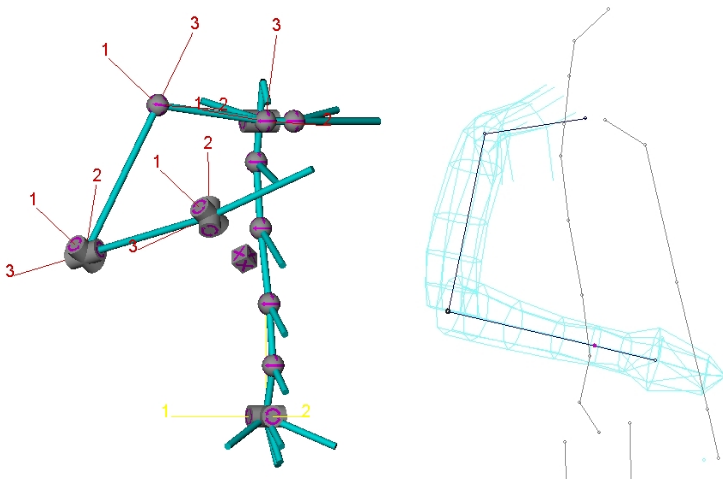


Fig. 4. Kinematic structure with connections between the joints of DYNAMICUS and RAMSIS

Matlab routines. The cardan angels and their derivatives were imported into the DYNAMICUS arm model. Thus the movement of the human arm system could be described providing the time dependant behavior of the joint angles. This temporal control is realized in the way that constraints are added to the model equations.

The torques resulting from the control of the joint angles that drive the simulation are important for further evaluation as well as the maximum torques that the subject is able to apply in each position. The maximum torques M_0 of the subject were measured by Zacher [4] under static conditions. The force velocity relationship of the muscles and consequently of the resulting torques are taken into account by implementing a Hill type function. The parameters for the hyperbolic function [8] describing the torque velocity relationship for concentric torques are taken from de Koning [9] who measured the angular velocity under different loads for the elbow flexors of untrained but sports minded males.

$$M = \frac{(M_0 + a) \cdot b}{(\omega + b)} - a \tag{1}$$

Using this equation the maximum concentric torque M of a DOF of a joint is calculated at the actual angular velocity ω of the movement. This is done for all joints and their respective DOF with the constants a=76 Nm and b=14 rad/s. In the eccentric phase the hyperbola was calculated according to Yeadon et al [10]. The composite function is continuous and there appears a ratio in the slopes of the eccentric and concentric function at ω=0. This slope factor is k=4.3.

4 Preliminary Results

First and intermediate results of the simulation of a lifting task are presented on torques, maximum torques and the relative torques (relation between both). This is an intermediate step to calculate discomfort during arm movements. Fig. 5 depicts the kinematical values joint angle and angular velocity of the shoulder elevation along the transversal axis. The upper arm is continuously lifted to the final height. The angular velocity shows a typical bell shaped form, resulting from the acceleration of the weight and the deceleration in the phase of placing the object to the final position. Both local maxima of the velocity curve result from a polynomial fit to the numerical derivative of the angles. Flexion angle and velocity of the elbow are shown in Fig. 7. First the elbow is slightly flexed in order to lift the weight in vertical direction and keep a security distance to the platform. Afterwards the elbow is extended to place the weight in the middle of the platform. During the extension the flexor muscles produce an eccentric flexion torque that is opposed to the direction of the velocity (negative velocity and positive torque). Fig. 6 depicts the maximum torque calculated from Hill's equation (Eq. 1) that is dependant on the velocity and the maximum isometric torque at the respective posture. The dotted curve is the isometric maximum torque from experiments with the subject. When ω=0 both curves are identical.

At positive velocities (shoulder: elevation, elbow: flexion) the muscles are able to produce a concentric torque that is less than the maximal isometric torque. At negative velocities like at the elbow in Fig. 7 the eccentric torque is higher than the isometric torque (Fig. 8). The step in the slope of the maximum torque is due to the slope factor of the composite Hill curve. The produced torques in shoulder and elbow (Fig. 6 and Fig. 8) are always positive as the weight is moved against gravity.

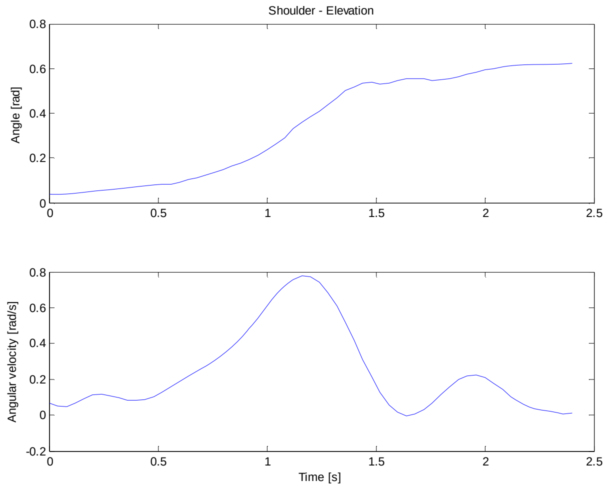


Fig. 5. Plots for the shoulder elevation along the transversal axis. Displayed are the angle and the angular velocity during the lifting task.

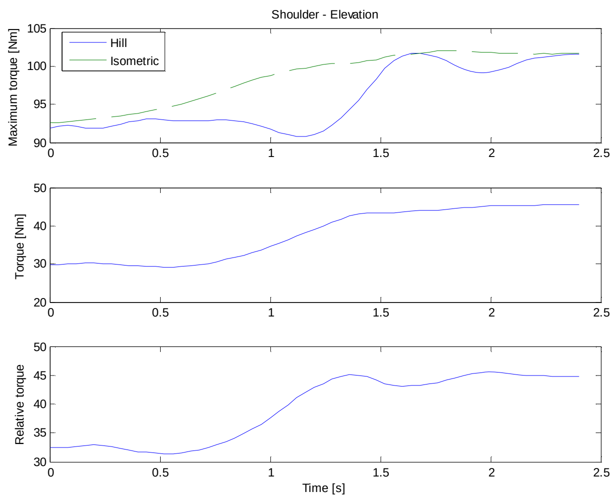


Fig. 6. Plots for the shoulder elevation along the transversal axis. The maximum torque at each time step/posture is given for the isometric condition (- - -) and for the concentric condition calculated with Hill's equation (---). Displayed are also the torque and the relative torque.

Relative torque is the ratio between maximum torque and actual torque in the respective direction of motion in the joint. This is a parameter that influences the discomfort.

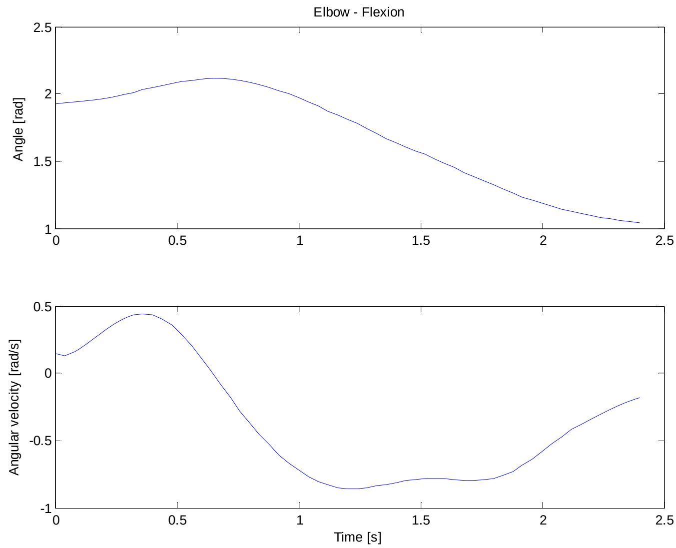


Fig. 7. Plots for the elbow flexion. Displayed are the angle and the angular velocity during the lifting task.

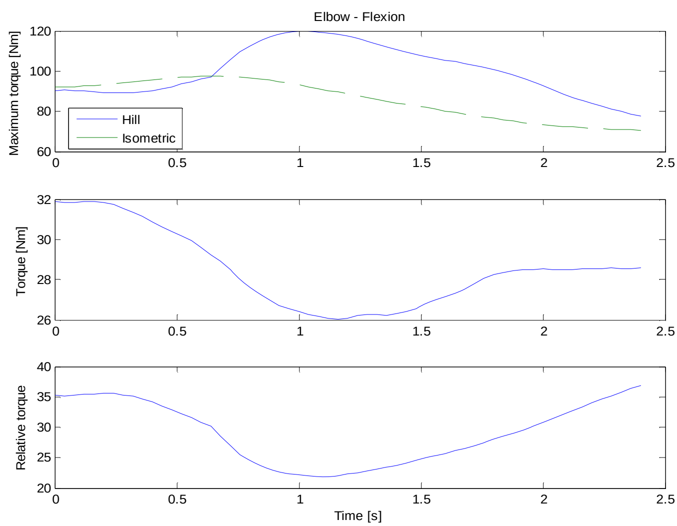


Fig. 8. Plots for the elbow flexion. The maximum torque at each time step/posture is given for the isometric condition (- - -) and for the concentric/eccentric condition calculated with Hill's equation (---). Displayed are also the torque and the relative torque.

5 Discussion and Conclusion

On the way to a generalized model to predict discomfort during movement intermediate results were generated. These results are parameters on which a discomfort model will be dependant. The discomfort is a subjective value, but as it can be evaluated with other factors that are connected with the concept of suffering it is possible to describe discomfort with physical values. These values act on the body or within the body like posture, force/torque and pressure from the contact with the surrounding. Thus we focus on two parameters that influence the local discomfort in a joint. These are joint angles and the relative joint torque. A posture and force dependant articular discomfort model for static conditions was presented by Zacher [4], Schäfer [11] and Wang [2]. If the local discomfort model is also applicable for movements needs to be shown by calculating the discomfort with the arm model and comparing the results with the discomfort ratings of the subject. It will also be interesting to find out if the time is another factor to influence the discomfort of movements as assumed by Wang. This would be similar to the physical demand of dynamic muscle work. Our aim is to give an objective evaluation of discomfort during movements.

References

1. Rasmussen, J., Christensen, S.T.: Musculoskeletal Modeling of Egress with the AnyBody Modeling System. In: Proceedings of the SAE DHMS 2005, Iowa, 2005-01-2721 (2005)
2. Dufour, F., Wang, X.: Discomfort Assessment of Car Ingress/Egress Motions using the Concept of Neutral Movement. In: Proceedings of the SAE DHMS 2005, Iowa, 2005-01-2706 (2005)
3. Cherednichenko, A., Assmann, E., Bubb, H.: Computational Approach for Entry Simulation. In: Proceedings of the SAE DHMC 2006, Iowa, 2006-01-2358 (2006)
4. Zacher, I., Bubb, H.: Strength Based Discomfort Model of Posture and Movement. In: Proceedings of the SAE DHMS 2004, Rochester, 04DHM-66 (2004)
5. Seitz, T., Balzulat, J., Bubb, H.: Anthropometrie and Measurement of Posture and Motion. International Journal of Industrial Ergonomics, 25, 447–453 (2000)
6. Seitz, T., Bubb, H.: Human-model based movement-capturing without markers for ergonomic studies. In: Proceedings of the SAE DHMC 2001, Arlington, 01DHM-7 (2001)
7. Benutzerhandbuch DYNAMICUS alaska 5.0. Institut für Mechatronik, Reichenhainerstr. 88, 09126 Chemnitz, Germany
8. Stern Jr, J.T.: Computer Modeling of Gross Muscle Dynamics. Journal of Biomechanics 7, 411–428 (1974)
9. de Koning, F.L., Binkhorst, R.A., Vos, J.A., van't Hof, M.A.: The force-velocity relationship of arm flexion in untrained males and females and arm-trained athletes. European Journal of Applied Physiology 34, 89–94 (1985)
10. Yeadon, M.R., King, M.A., Wilson, C.: Modelling the maximum voluntary joint torque/angular velocity relationship in human movement. Journal of Biomechanics 39, 476–482 (2006)
11. Schäfer, P., Zacher, I.: On the way to autonomously moving manikins – empowered by discomfort feelings. In: Proceedings of the XVI. World Congress on Ergonomics of the IEA, Maastricht (2006)

A Motion Compensated De-interlacing Algorithm for Motive Object Capture

Lei Gao, Chao Li, Chengjun Zhu, and Zhang Xiong

School of Computer Science and Engineering, Beihang University,
37th Xueyuan Road,Haidian District, Beijing, China, 10083
squirelive@gmail.com, licc@buaa.edu.cn, zhucj@cse.buaa.edu.cn, xiongz@buaa.edu.cn

Abstract. A motion compensated de-interlacing algorithm is proposed to recover the defects of interlaced video frame for capturing motion object. In this algorithm, two anti-noise background fields are formed by analyzing the temporal correlation of pixels between adjacent same parity fields. To each field, the subtraction with the corresponding background is used to detect motion object. To avoid the inaccurate detection caused by the difference between the spatial scanning positions of odd and even field, the motion objects are detected with same parity field and background field. Then motion estimation technology is used to measures the inter-field motion, find out the motion vector between the odd field and even field. Based on the motion vector, an interpolation filter is designed to shift the pixels of the motion object in the two temporally displaced fields to a common point in time. This de-interlacing algorithm maximizes the vertical resolution of the motion objects. Experimental results show that the proposed algorithm could achieve higher image quality on motion object, and the computational complexity is acceptable for consumer computer applications.

Keywords: de-interlacing, motion compensation, motion estimation, motion detect, motion object.

1 Introduction

The conventional television system based on interlaced video signal has been widely used for many decades. There are many systems transmit or record video in PAL or NTSC format. The interlaced video signal means that during the image capture process, the camera outputs the odd lines at one instant in time, and then several milliseconds later, outputs the even lines. Odd lines make up odd field and even lines make up even field. Interlacing reduces the bandwidth by half, is a way to display the nonmoving parts with full resolution and the moving parts with half resolution, but fluidly. It is a very clever way to cut bandwidth without sacrificing much quality [1]. But in some cases it is necessary to get one full frame from the two fields. A traffic surveillance and automobile license plate identification system, for example, is required to capture moving objects and to allow viewing of details within a moving image.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 74–81, 2007.

Fig. 1. An example of interlaced effect on moving object

Because of a time delay between two half-frames which is approximately 16ms for NTSC standard and approximately 20 ms for PAL and SECAM standards [2], a temporal shift occurs between the odd and even lines of the image. After combing two adjoining half-frames an interlaced effect appears on moving object’s edges. This problem is illustrated in Figure 1.

Therefore, the de-interlacing process attempts to remove interlace effect by creating a clean frame from the two fields. Many de-interlacing algorithms have been proposed [3-7]. These methods can be classified into two basic categories: motion compensated and non-motion compensated. When restoring a full frame from two half-frames, most of non-motion compensated methods base on one of the half-frames, and calculate moving pixels which need to be corrected from the basic one using an interpolation method. These mathematical methods lead inevitably to information loss right up to losing all the information of one of the half-frames [6], especially in regions with significant motion. In contrast with non-motion compensated method, motion compensated de-interlacing using maximum information of both half-frames. It is the most advanced method of de-interlacing [7], but it requires significant compute resources to implement.

In this paper, we present a low computation requirement de-interlacing algorithm based on motion compensation. It adapts to divide quick moving object from background and interpolate moving pixels based on motion vector. Comparing with conventional methods, it produces higher image quality on motive object.

The organization of this paper is as follows. The motion compensated de-interlacing is reviewed in section 2. The proposed algorithm is presented in section 3. Experiment results are presented in section 4. Finally, our conclusion is given in section 5.

2 Motion Compensated De-interlacing

Since the position of a moving object will be different in the two half-frames, if one measure the inter-filed motion vector and then align data between the two half-frames, more inter-field information can be used. Motion compensated de-interlacing is the technique attempts to interpolate pixels base on motion vectors.

Motion compensated de-interlacing can be symbolized as:

$$F(\vec{x},n)=\begin{cases} f(\vec{x},n), & y \bmod 2 = n \bmod 2 \\ f(\vec{x}-\vec{d}(\vec{x},n),n-1), & otherwise \end{cases} \tag{1}$$

where $f(\vec{x},n)$ is the brightness value of the pixel in field n, at position $\vec{x}=[x,y]^t$, t for transpose, $f(\vec{x},n)$ the input field and $F(\vec{x},n)$ the de-interlaced output frame, $\vec{d}(\vec{x},n)=(d_x(\vec{x},n),d_y(\vec{x},n))^t$ for the motion vector at spatial position $\vec{x}$ in field n.

Obviously, the accuracy of motion vector $\vec{d}(\vec{x},n)$ is key factor to de-interlace frame correctly. In general, finding out the motion vector is typically implemented as a block-matching motion estimation process. Typical block sizes range from 4x4 to 8x8. However, the complexity of real-time block matching is significantly higher. The extra hardware is required to perform the complex motion estimation and compensation process. For example, real-time motion estimation for CIF (352 x 288) 30 frames per second (fps) video with [-16, +15] search range requires 9.3 Giga-operations per second (GOPS). If the frame size is enlarged to D1 (720 x 480) 30fps with [-32, +32] search range, 127 GOPS is required [8]. To reduce computation in real-time system, an improved motion compensated de-interlacing algorithm for motive object capture is proposed.

3 The Proposed Motion Compensated De-interlacing Algorithm

The proposed algorithm consists of three stages, that is, motive object detection, motion estimation and motion compensated de-interlacing filtering. We will discuss them in detail in the following.

3.1 Motive Object Detection

Motive object detection is the essential pre-stage of motion estimation. In order to extract motive objects, background subtraction method is adopted. Widely used in surveillance system, background subtraction is an efficient method to discriminate moving objects from the still background [9]. The idea of background subtraction is to subtract the current image from the still background, which is acquired before the objects move in.

To avoid the inaccurate detection caused by the difference between the spatial scanning positions of odd and even field, the proposed method form two anti-noise background fields by capturing the statistics from the first N background frames:

$$B_0(\vec{x}) = \frac{1}{N}\sum_{k=0}^{N-1} f(\vec{x},2k)$$
$$B_1(\vec{x}) = \frac{1}{N}\sum_{k=0}^{N-1} f(\vec{x},2k+1) \tag{2}$$

where $B_0(\vec{x})$ is the brightness value of the pixel in odd background field, at position $\vec{x} = [x, y]^t$ and $B_1(\vec{x})$ is the even one. Then the motion objects located displaced in adjoining fields are detected by subtracting the current field from background field with same parity. The brightness difference value $D_0(\vec{x})$ and $D_1(\vec{x})$ are defined as:

$$\begin{aligned} D_0(\vec{x}) &= \left| f(\vec{x},n) - B_0(\vec{x}) \right| \quad n \bmod 2 = 0 \\ D_1(\vec{x}) &= \left| f(\vec{x},n) - B_1(\vec{x}) \right| \quad otherwise \end{aligned} \tag{3}$$

If $D_0(\vec{x})$ or $D_1(\vec{x})$ is larger than a given threshold, current pixel $\vec{x}$ is classified to the motive object. Otherwise, the pixel $\vec{x}$ belongs to the background. Then an assumable motion vector can be calculated by comparing the centroids of the motion object located displaced in the two half-frame.

3.2 Motion Estimation

As described in previous section, the accuracy of motion vector $\vec{d}(\vec{x}, n)$ is key factor to de-interlace frame correctly. The goal of the motion estimation is to get an accurate motion vector of inter-field object moving. To achieve this goal, an improved block-matching method is adopted.

Considering low computation requirement, only one block within the area of motive object is chosen in the even field, and the motion of the block in the even field is estimated by matching to candidate blocks in the previous odd field. The size and position of the block depend on the size and the user most interested part of motive object. In order to reduce search positions, the search positions are calculated base on the assumable motion vector which was found in motive object detection stage. Then a search of lowest brightness distortion block among the entire candidate blocks is performed.

The proposed block-matching can be expressed as:

$$SAD(x,y)=\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left| f([i,j]^t,n+1)-f([i+x,j+y]^t,n)\right|,\quad n \bmod 2=0$$
$$x_a-p\le x\le x_a+p-1$$
$$y_a-p\le y\le y_a+p-1 \tag{4}$$

$$\vec{d}(\vec{x},n+1)=\min(SAD(x,y)) \tag{5}$$

where $SAD(x,y)$ is the sum of absolute differences of the candidate block at search position (x,y), the block size is $M\times N$, $\{f([i,j]^t,n+1)|0\le i\le M-1, 0\le j\le N-1\}$ means current block data in even field $n+1$, $\{f([i+x,j+y]^t,n)|0\le i\le M-1, 0\le j\le N-1\}$ means candidate block data at search position (x,y) in odd filed n, $[-p,p-1]$ is the search range, $[x_a,y_a]$ is the assumable motion vector which was found in motive object detection stage, and $\vec{d}(\vec{x},n+1)$ is the motion vector of current block with minimum *SAD* among $(2p)^2$ search positions.

3.3 Motion Compensated De-interlacing Filter

In the previous two sections, the inter-field motion vector of motive object has been calculated by the proposed method. Based on the motion vector, an interpolation filter is designed and the missing pixels in motive object could be interpolated by applying median filter on the spatial neighbors in the vertical direction and the predict position in the next even field.

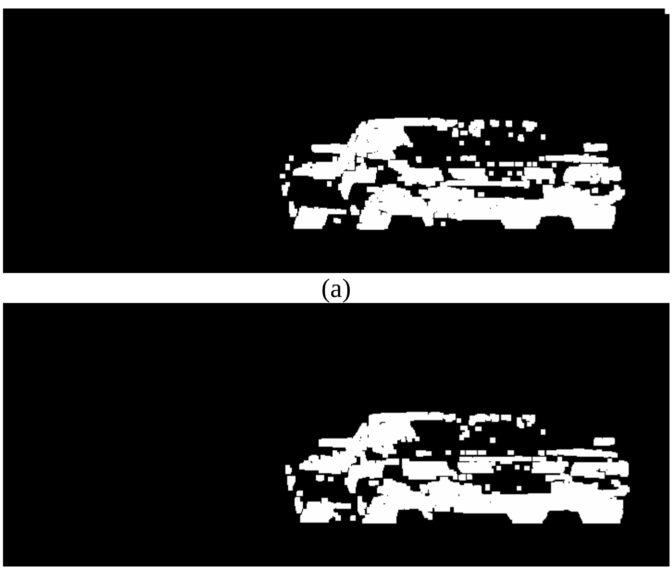
The final output can be defined as:

$$F(\vec{x},n)=\begin{cases} f(\vec{x},n), & y \bmod 2=0, n \bmod 2=0\\ median\begin{pmatrix} f(\vec{x}-\begin{bmatrix}0\\1\end{bmatrix},n)\\ f(\vec{x}+\begin{bmatrix}0\\1\end{bmatrix},n)\\ f(\vec{x}-\vec{d}(\vec{x},n+1),n+1)\end{pmatrix}, & y \bmod 2=1, n \bmod 2=0\end{cases} \tag{6}$$

4 Experimental Results

The proposed algorithm was simulated on a Celeron CPU 2.53GHz PC for performance evaluation. We used interlaced video sequences of spatial size 720 x 576

Fig. 2. An interlaced frame of the test sequences including a fast moving object



(a)

(b)

Fig. 3. Motive object detection results: (a) odd field, (b) even field

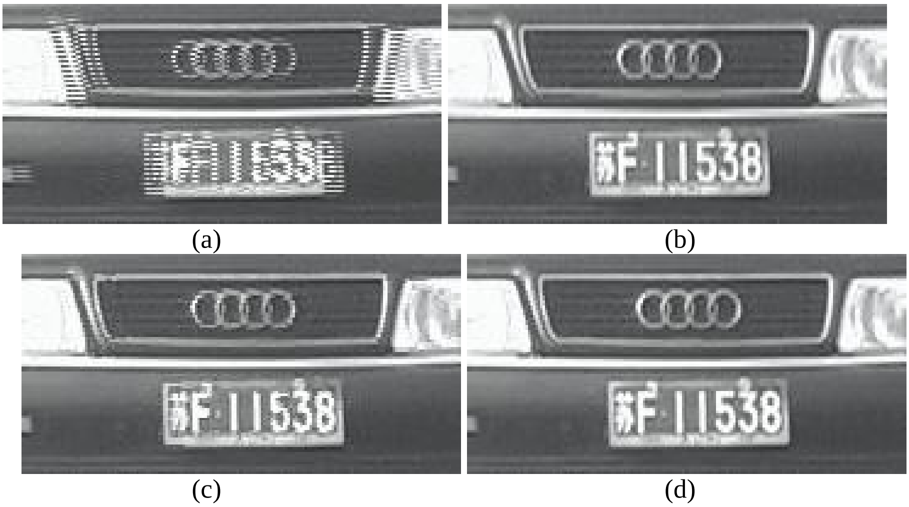


Fig. 4. Partial image of motive object: (a) original, (b) line average, (c) median filtering, (d) proposed method

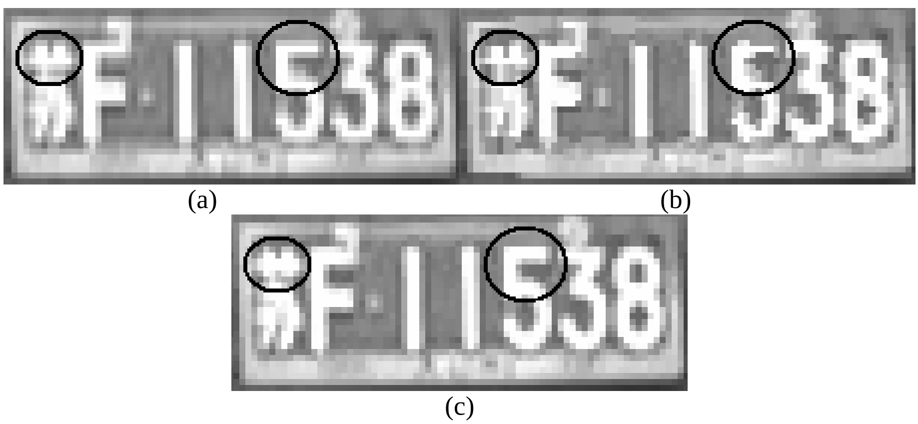


Fig. 5. Enlarged image of license plate areas: a) line average, (b) median filtering, (c) proposed method

which are captured on road for testing. The average processing speed is over 25 frames per second. For comparison, tow simple de-interlacing methods, line average and median filtering, were applied to same sequences.

Figure 2 is an interlaced frame of the test sequences including a fast moving object. The motive object detection results are shown in Figure 3. The de-interlacing results on this moving object are shown in Figure 4. Figure 5 are enlarged images of license plate areas. It can be seen that the image quality of the proposed method outperforms the line average and median filtering method.

5 Conclusion

The proposed motion compensated de-interlacing algorithm employed maximum information of both half-frames to interpolate the missing pixels of the motive object, maximized the vertical resolution of the motion objects. It could achieve higher image quality on motion object, and the computational complexity is acceptable for consumer computer applications.

References

1. What is Deinterlacing? Facts, solutions, examples http://www.100fps.com/
2. Wang, Y., Ostermann, J., Zhang, Y.-Q.: Video processing and communications. Prentice Hall, Upper Saddle River, NJ (2002)
3. Kuo, C.J., Liao, C., Lin, C.C.: Adaptive Interpolation Technique for Scanning Rate Conversion. IEEE Transactions on Circuits and System for Video Technology 6(3), 317–321 (1996)
4. Lin, S.F., Chang, Y.L., Chen, L.G.: Motion Adaptive Interpolation with Horizontal Motion Detection for De-interlacing. IEEE Transactions on Consumer Electronics 49(4), 1256–1264 (2003)
5. Tai, S.C., Yu, C.S., Chang, F.J.: A Motion and Edge Adaptive Deinterlacing Algorithm. In: IEEE International Conference on Multimedia and Expo (ICME), 27-30 June 2004, vol. 1, pp. 659–662 (2004)
6. Khudeev, R., Osipov, S., Afanasyev, V.: De-Interlacing Using Area Based Motion Compensation. In: Confernence on Control and Communications SIBCON, 2005, 21-22 October 2005, pp. 154–160 (2005)
7. Deame, J.: Motion Compensated De-Interlacing: The Key to Digital Video Transition. In: SMPTE 141st Technical Conference in NY (November 19-22, 1999)
8. Huang, Y.-W., Chen, C.-Y., Tsai, C.-H., Shen, C.-F., Chen, L.-G.: Survey on Block Matching Motion Estimation Algorithms and Architectures with New Results. Journal of VLSI Signal Processing 42, 297–320 (2006)
9. Pan, J., Lin, C.-W., Gu, C., Sun, M.-T.: A robust video object segmentation scheme with pre-stored background information. In: IEEE International Symposium on Circuits and Systems (ISCAS), 2002, 26-29 May 2002, vol. 3, pp. 803–806 (2002)

Color 3D Digital Human Modeling and Its Applications to Animation and Anthropometry

Bao-zhen Ge[1], Qing-guo Tian[1], K. David Young[2], and Yu-chen Sun[1]

[1] College of Precision Instruments and Opto-electronics Engineering and
Key Ministry of Education Laboratory of
Opto-electronics Information and Technical Science
Tianjin University Tianjin 300072. China
gebz@tju.edu.cn

[2] Embedded Systems Institute and Department of Electronic and Computer Engineering
The Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong
kkdyoung@ust.hk

Abstract. With the rapid advancement in laser technology, computer vision, and embedded computing, the application of laser scanning to the digitization of three dimensional physical realities has become increasingly widespread. In this paper, we focus on research results embodied in a 3D human body color digitization system developed at Tianjin University, and in collaboration with the Hong Kong University of Science and Technology. In digital human modeling, the first step involves the acquisition of the 3D human body data. We have over the years developed laser scanning technological know-how from first principles to support our research activities on building the first 3D digital human database for ethnic Chinese. The disadvantage of the conventional laser scanning is that surface color information is not contained in the point cloud data. By adding color imaging sensors to the developed multi-axis laser scanner, both the 3D human body coordinate data and the body surface color mapping are acquired. Our latest development is focused on skeleton extraction which is the key step towards human body animation, and applications to dynamic anthropometry. For dynamic anthropometric measurements, we first use an animation algorithm to adjust the 3D digital human to the required standard posture for measurement, and then fix the feature points and feature planes based on human body geometric characteristics. Utilizing the feature points, feature planes, and the extracted human body skeleton, we have measured 40 key sizes for the stand posture, and the squat posture. These experimental results will be given, and the factors that affect the measurement precision are analyzed through qualitative and quantitative analyses.

1 Introduction

Three dimensional (3D) digitization encompasses the applications of 3D information extraction methods to extract the 3D surface data of an object, real time 3D data

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 82–91, 2007.

pre-processing and generation of visual 3D computer graphic models of the object. The 3D information extraction methods are subdivided into two classes: contact and non-contact. A typical representative of the contact type is the coordinate measurement machine (CMM) [1] whose precision can be of microns, however the measurement speed is rather slow and furthermore inapplicable to soft objects. Non-contact methods primarily apply optical techniques which can be further subdivided into passive and active types. Passive optical measurement methods include monocular imaging and binocular stereo imaging approaches [2]. Active optical measurement methods include: (1) phase measurement approaches such as phase profilometry [3], Moire contouring [4], color-coded projection grating method [5] whose main advantage is fast extraction speed, but suffer from lower resolution, and greater limitation of the measurement volume by the optical aperture; (2) laser ranging techniques which are based on the principle of direct measurement of the time of flight difference between a target optical beam and a reference beam, and its conversion into distance which yields resolution typically about 1 mm [6]; (3) laser scanning measurement techniques which employ structured laser light source to illuminate the target objects and detect the reflected light distorted by the object's surface with CCD arrays, triangulation principle is applied to determine the 3D point locations in a reference coordinate system. Different methods have been developed depending on the types of structured laser light sources: point source structured light method [7], line source structured light method [8], and double line source structured light method [9]. Optical methods are widely use in many areas due to their inherit advantages of allowing non-contact, high resolution, and high speed measurements. In particular, structured light triangulated measurement based 3D color digitalization has become widely adopted, The approach taken by Yu [10] and the work at HKUST [11] both assume the target object which is a human body is stationary, and the structured laser line and sensing CCD arrays move vertically to obtain the 3D point cloud data of the human; color information however was not collected. Xu [12] developed 3D color measurement modeling method from monochrome 3D measurement theory, and applied to obtain 3D color information for small size objects. Hu [13] utilized color CCD arrays and structured laser line source, and with a constant speed turning platform and high speed optoelectronic shutters, obtained the target object's 3D point cloud and the corresponding color information. Pulli [14] deployed four NTSC format color CCD cameras and a white light projector affixed to a turning platform to form a scanning system.

This paper is divided into two halves: the first half introduces a color 3D human digitalization system which is based on the use of structured laser line source and multi-axis synchronized scanning. Basic principle of operation of the system will be given, with more in depth discussions on the method used for the calibration of the 3D sensors and color sensors. The second half covers the human surface mesh modelling and processing aspects, which includes skeleton extraction algorithms based on radial basis functions and steepest descent, and an animation computation algorithm which is based an improved skinning method. Finally, the developed animation algorithm is applied to human body size measurements for anthropometry and garment design.

2 Color 3D Human Digitization System Design

2.1 System Structure

After a thorough examination of the advantages and disadvantages of different approaches in laser scanning based 3D color digitalization, we present in this paper an approach which exploits independent yet correlated 3D sensor and color sensor development in order to improve the system flexibility. The system is pictorially represented in Figure 1. The laser scanning support structure is consists of a rectangular base on which four vertical columns are attached; on each of the vertical columns, a 3D sensor is mounted. Interleaving with the 3D sensor columns are four additional vertical supports on which four color sensors are mounted. The maximum volume covered in one complete scan is cylindrical: 1000mm diameter $\times$ 2000 in height. On the surface of a cylinder whose center is at the center of the rectangular loading base and at a radius of 500mm, the typical resolution is 1mm in the horizontal direction, 2mm in the vertical direction. One complete scan takes 16.7 seconds at a vertical spacing of 2mm.

The 3D sensor applies the principle of triangulation and utilizes a laser line structured light source. In order to minimize the obstruction of light by the target object, a dual CCD camera structure is employed to receive the reflected light beam which is modulated by the target object being measured. The single axis 3D sensor structure is schematically represented as Figure 2. On the vertical column, a mechanical translation platform which is constructed from the combination of a servo motor drive and a ball screw and slide is responsible for moving the sensor in the up down direction. On the moving platform, two gray scale CCD cameras are mounted at identical inclined

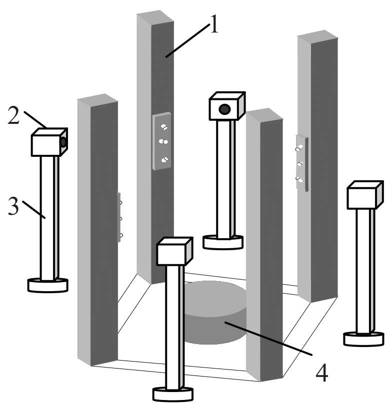


Fig. 1. 3D Laser Scanner based Color Digitalization System Schematic Diagram: 1. 3D Sensors 2. Color CCD Camera 3. 3D Camera Support 4. Load target platform.

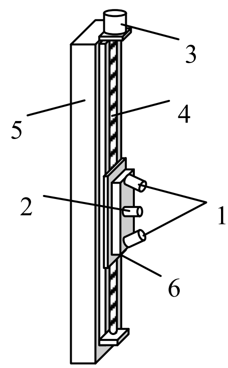


Fig. 2. Single axis 3D Sensor schematic diagram: 1 Gray Scale CCD Camera 2 Laser Line source 3 Servomotor 4 Ball Screw 5 Support Column 6 Sensor base

angles and equal distance from the centrally mounted laser line source. The key issue in this 3D color scanning system is calibration.

2.2 3D Sensor Calibration

The calibration of the 3D sensor is basically the calibration of the gray scale cameras. Yuan [15] applied a perspective transformation matrix method to calibrate the cameras, however he ignored nonlinear effects such as lens distortions. Although this method attains faster calibration speed, accuracy is lower. Faig [16] applied nonlinear optimization method, which took into account a variety of parameters to calibrate the camera system, and achieved higher accuracy, however the calibration outcome depends on the choice of initial values, making the calibration process rather unreliable. Reimar and Roger [17] combined the perspective transformation matrix approach with nonlinear optimization to perform calibration, although it can achieve high accuracy, this method requires solving for system parameters from nonlinear equations, it is also very time consuming. For comprise between speed and accuracy, we propose an approach which is based on linearized transformation matrix for partitioned regions to calibrate our 3D sensors.

Using a perspective transformation model, a point in space $P(X_w,Y_w,Z_w)$ a mapped onto the corresponding point $P_f(X_f,Y_f)$ on the pixel map through rotation and resizing transformations. Since Z_w within the same light band of an object is identical, this relationship can be represented using the following matrix equation:

$$\begin{bmatrix} X_f \\ Y_f \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \bullet \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} = M \bullet \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} \quad (1)$$

The purpose of calibration is to determine the values of the elements in the matrix M so that the correspondence between the point (X_f,Y_f) and (X_w,Y_w) can be defined. If we let $m_{33} = 1$, then theoretically, taking four calibration point pairs will be sufficient to determine the remaining eight unknown matrix coefficients. However, in order to obtain better accuracies, we sample more calibration point pairs. After we compute the matrix M, the 3D coordinates of the corresponding points on the object can be determined from the pixel map.

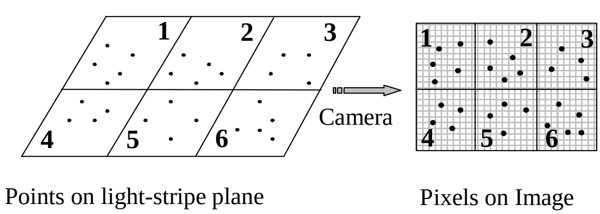


Fig. 3. Partitioned regions for calibration

In order to use linear transformation method while taking in account of lens distortion and other nonlineareffects, we adopt a partitioned region approachas indicated in

Figure 3 where points on the light stripe plane and the corresponding pixels on the imaging plane are subdivided into different zones. For each of the different zones, a matrix M as given by (1) is derived; yielding a number of different calibration matrix M_i, i=1, 2, …, N. In general, N=6 is sufficient.

2.3 Color Sensor Calibration

Color sensor captures 2D color images of the different side views of an object. Information of each pixel can be represented by a set of five variables (X_f,Y_f,R,G,B). Therefore the calibration of a color sensor is the determination of the projective relationship between a spatial point with color information (X_w,Y_w,Z_w,R,G,B) and a point (X_f,Y_f,R,G,B) on the color bitmap; and the imposition of the color value of the point on the bitmap on the corresponding point on the object. Although it is possible to use methods based on linearization with respect to partitioned regions to calibrate color sensors, however, due to the many-to-one mapping between the spatial point and the corresponding color bitmap pixel, a color sensor calibration method which is based on NP neural network is adopted.

A three layer BP neural network structure is utilized. There are three neurons at the input layer corresponding to the object's three spatial coordinates X_w, Y_w and Z_w . The output layer has two neurons which correspond to the color bitmap's pixel coordinates X_f and Y_f . According to the Kolmogorov theorem, only one hidden layer is needed. Thus there are a total of $2 \times 3 + 1 = 7$ neurons.

The calibration of color sensor is realized by a recursive learning process through which the weighting coefficients with respect to the neurons are determined. The learning process of a BP neural network is divided into one in the forward propagation path, and another one in the reverse propagation path. In the forward path, using the object's coordinates, the corresponding pixel coordinate estimates are computed; the coordinate deviations are also computed by comparing with the true pixel coordinates. In the reverse path, the coordinate deviations are propagated toward the input nodes, with the weighting parameters of the neurons updated as well. This process is repeated until the deviations are reduced below certain threshold value; and a set of BP neural network weighting parameters reaches their final values.

Figure 4(a) is an image showing the light strip bitmap of a mannequin's head sampled by a gray scale CCD camera. Figure 4(b) is the point cloud image of the same mannequin's head constructed by processing the light strip bitmap image with the 3D sensor calibration matrix M. There are a total of 35629 points in the point cloud data. Figure 4(c) depicts the color bitmap image of the mannequin's head captured by one of the color CCD camera; and Figure 4(d) is a 3D point cloud image embedded with strip color information which is derived from computing the estimated spatial coordinates using the converged BP neural network weighting parameters as gain values and the point cloud 3D coordinates as inputs. As shown, our system provides fairly accurate 3D color information.

Fig. 4. (a) Light strip bitmap

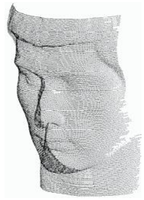


Fig. 4. (b) 3D point cloud

Fig. 4. (c) Color bitmap

Fig. 4. (d) Color point cloud

3 Digital Human Animation

Digital human animation is one of the most difficult and challenging topics in 3D animation. Research approaches can be divided into three categories: surface based method, volume based method and layer based method. Surface based method can be further subdivided into rigid model animation [18], deformation function representation [19], contour based deformation [20] skinning method [21] and sample based animation [22]. Volume based method mainly includes implicit surface method [23] and skeleton driven volumetric data method [24]. In layer based animation, the human body is decomposed into a skeleton layer, a muscle layer, a fat layer and a skin layer; independent model at each of the layer is constructed, and the ultimate skin deformation is realized by transferring the modeling characteristics from the most inner to the most outer layer [25]. Among the volume based methods, the implicit surface method can easily create digital human with nice shapes, while the volumetric data model is best suited to drive digital human model that is already represented by volumetric data. Since the volumetric data method generally utilizes a large number of data points, it is also the slowest among all the methods. Although layer based animation generally produces vivid and realistic animation, the model construction process is extremely complex. At present, the most commonly applied algorithms in digital human animation for skin deformation are those provided by the surface based method, and in particular, skinning method. The main drawback of this method is the need for user's manual interaction. Regardless of the approach to be adopted for digital human animation, the first task is to extract the 3D model's skeleton.

3.1 Extraction of Skelton of Digital Human

Interpolating an excessive number of mesh vertices to create an implicit surface representation of a 3D digital human is a very time consuming process; it is also unrealizable with ordinary computing machine. Although excessive mesh vertices and surface microstructures will create a large number of micro-segments in the resulting 3D skeleton which is beneficial to feature tracking and surface reconstruction, there is no intrinsic value to path planning and animation design. Therefore, before creating digital human's implicit representation, we apply the method proposed in [25] to simplify the mesh. In general, this approach can reduce the number of triangular mesh facets to 1% of the total mesh facets generated by the original mesh.

We apply the algorithm proposed by Ma [27] to extract the 3D skeleton from the simplified digital human mesh model. In the process of mesh simplification and skeleton extraction, each initial vertex's corresponding vertex in the simplified model, and each simplified model's vertex corresponding point in skeleton are recorded. Thus, to each initial vertex there is a corresponding skeleton point. This relationship is a many-to-one mapping. By taking this additional step, the process of binding the surface vertices to the corresponding skeleton points is simplified, and therefore the required computation time is reduced.

Fig. 5(a) shows the digital human's surface mesh model reconstructed from the contour data of a human body which is produced by the 3D laser scanning system introduced in Section 2. Fig. 5(b) depicts the resulting simplified model which contains about 1 % of the initial model's triangular mesh facets. Since the initial mesh model contains an excessive number of vertices and triangular mesh facets, simplification to only 1% of the original model leads to insignificant reduction in the digital human's fidelity. Fig. 5(c) and (d) are two different views of the resulting 3D skeleton extracted from the simplified model shown in Fig. 5(b).

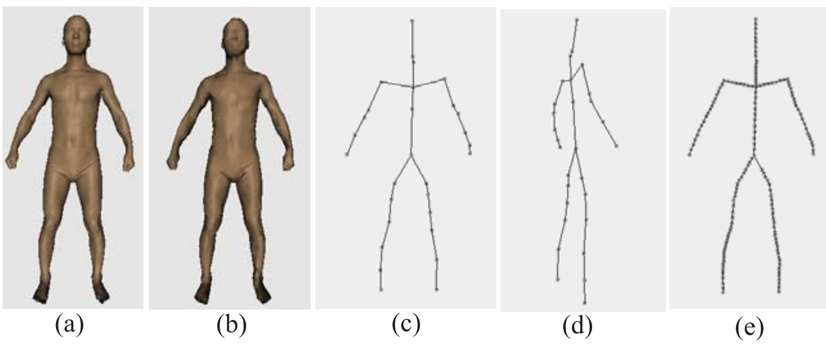


Fig. 5. Surface mesh model and the extracted 3D skeleton of a digital human

3.2 An Improved Skinning Method for Digital Human Animation

We introduce and apply an improved "skinning" method to realize digital human animation. In order to reduce the localized collapse and "candy wrapper" phenomena, the following three modifications are implemented:

(1) Extra joints on the 3D skeleton are added to minimize the distance between adjacent joints. Fig. 5(e) shows the resulting skeleton with extra equidistantly located joints added.

(2) A vertex on the surface is bound to multiple skeleton joints using the following equation:

$$T_i = \sum_{j=1}^{n} \frac{D - d_{j,i}}{(n-1) \times D} t_j \tag{2}$$

where t_j is the transformation matrix of the j[th] skeleton joint, T_i is the transformation matrix of the surface vertex i, $d_{j,i}$ is the distance between vertex i and joint j, D is total distance, and n is the number of joints that corresponds to vertex I; typically n=3. We first locate the joint j which has the minimal distance to surface vertex i; and then the surface vertex i is bound to the joint j and its adjacent joints j-1, j+1 by applying the transformation given by (2). Fig. 6 illustrates the idea where the black dot represents the surface vertex and the white circles are the skeleton joints.

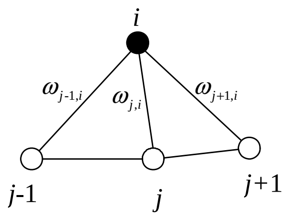


Fig. 6. Binding multiple skeleton joints to a surface vertex

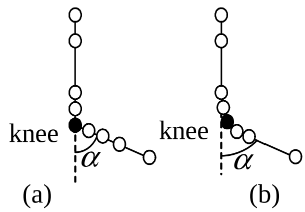
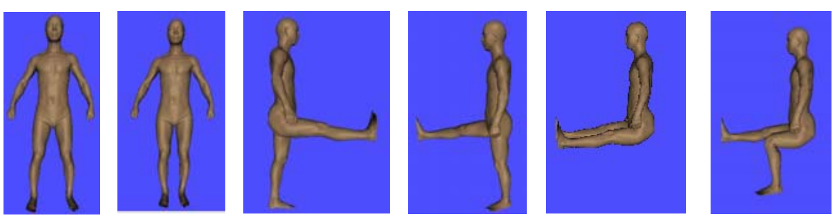


Fig. 7. An illustration of the compound joint technique

(3) Apply a compound joint technique to human body joints, such as the knee joint. Fig. 7 illustrates the kickback of the calf at an angle of α where the black dot denotes the knee joint location. Fig. 7(a) illustrates the case when all the joints below the knee joint are bent at an angle of α. In Fig. 7(b), two joints immediately above and below the knee joint are bent progressively with each of the joint bending at an angle of $\alpha/4$. This modification can effectively prevent the localized collapse and "candy wrapper" phenomena to occur. Since the distance between any two adjacent joints is very small, the compound joint technique produces minimal negative effects to overall fidelity.



(a) (b) (c) (d) (e) (f)

Fig. 8. A sequence of digital human re-posturing

Fig. 8 shows a sequence of digital human configurations where the legs are moved to present different postures. Fig. 8(a) is the digital human model in the initial stance; Fig. 8(b) indicates the posture after the feet move closer, showing a posture roughly at attention. Fig. 8(c) and (d) give two opposite views of a posture with the left leg raised to a level position. The posture showed in Fig. 8(e) result from raising both legs to level positions, and finally Fig. 8(f) records the dropping of the right calf to a down right position from the posture given in (e).

4 Applications to Anthropometry

The development on digital human modeling as reported in the previous sections has been applied to obtain key human body measurements for anthropometry applications. Due to page limitation, details are omitted herein. We shall provide these results at the paper presentation in the conference.

5 Conclusions

In this paper we described the design of a color 3D human digitalization system, and elaborated on a 3D sensor calibration method based on linearized partitioned regions, and a color sensor calibration method based on BP neural network learning. This system has been demonstrated to acquire both 3D point cloud data and simultaneously the corresponding color information for large 3D objects such as the human body. The subsequently generated surface mesh model is further simplified to improve the speed of skeleton extraction. An improved “skinning” method with three new modifications has also been developed to improve the quality of digital human animation. These new developments have also been applied to obtain key human body measurements for anthropometry applications.

Acknowledgements. The work reported herein is partially supported by the China National Science Council through Project No. 60277009, and by Tianjin City Government Targeted Technology Program Support Project No. 06GFGPX05000.

References

[1] Guoxiong, Z.: The trend in development of coordinate measuring machines [J]. Chinese Mechanical Engineering(in Chinese) 11(2), 222–227 (2000)
[2] Narkhede, S., Golshani, F.: Stereoscopic imaging: a real-time, in depth look. IEEE Potentials 23(1), 28–42 (2004)
[3] Yujin, Z.: Image Engineering [M](in Chinese). Tsinghua University Press, Beijing (2000)
[4] Jin, H., Xiaobing, Z., et al.: 3D measuring techniques based on image processing[J]. Chinese Journal of Electron Devices(in Chinese) 25(4), 364–368 (2002)
[5] Demers, M.H., Hurley, J.D., Wulpern, R.C., et al.: Three dimensional surface capture for body measurement using projected sinusoidal patterns [A]. SPIE[C] 3023, 13–25 (1997)
[6] Frankchen., Shizhen, W.: Review for the measuring of 3-D object by optical method [J]. Cloud Light Technology(in Chinese) 34(2), 36–47 (2002)
[7] Tognola, G., Parazzini, M., et al.: A fast and reliable system for 3D surface acquisition and reconstruction. Image and Vision Computing 21, 295–305 (2003)
[8] Xu, Z.Q., Sun, C.K., et al.: 3D Color Reverse Engineering, Optoelectroincs (in Chinese). Laser 12(9), 937–939 (2001)

[9] Park, J., DeSouza, G.N., Kak, A.C.: Dual-beam structured-light scanning for 3-D object modeling. In: Proc. of Third International Conference on 3-D Digital Imaging and Modeling, pp. 65–72 (2001)
[10] Yuan, Y.C., Lo, Y.H., et al.: The 3D scanner for measuring body surface area: a simplified calculation in the Chinese adult. Applied Ergonomics 34, 273–278 (2003)
[11] Young, K.D.: 3D Mannequin Database - Hong Kong Innovation and Technology Commission funded Project Final Report, The Hong Kong University of Science and Technology (2002)
[12] Xu, Z.Q., Ye, S.H., Fan, G.Z.: Color 3D reverse engineering. Journal of Materials Processing Technology 129(1-3), 495–499 (2002)
[13] Hu, H.P., Li, D.H., et al.: 3D laser color scanning system. Advanced Technology Communications(in Chinese) 11(3), 25–30 (2001)
[14] Kari, P., Shapiro Linda, G.: Surface Reconstruction and Display from Range and Color Data. Graphical Models 62(3), 165–201 (2000)
[15] Ye, Y.: Camera Calibration method and boundary detection contour tracking algorithm research, Da Lian University doctoral dissertation, (in Chinese) (2002)
[16] Faig, W.: Calibration of close-range Photogrammetry systems: Mathematical formulation. Photogrammetric Eng. Remote Sensing 41, 1479–1486 (1975)
[17] Lenz, R.K., Tsai, R.Y.: Techniques for calibration of the scale factor and image center for high accuracy 3D machine vision metrology. Pattern Analysis and Machine Intelligence, IEEE Transaction 10(5), 713–720 (1988)
[18] Zordan, B.V., Majkowsk, A., Chiu, B., Fast, M.: Dynamic response for motion capture animation. ACM Transactions on Graphics 24(3), 697–701 (2005)
[19] Magnenat-Thalmann, N., Laperriere, R., Thalmann, D.: Joint-Dependent Local Deformations for Hand Animation and Object Grasping. In: Proceedings of Graphics Interface'88, pp. 26–33 (1988)
[20] Kalra, P., Thalmann, N.M., Moccozet, L., Sannier, G., Aubel, A., Thalmann, D.: Real-time animation of realistic virtual humans. Computer Graphics and Applications 18(5), 42–56 (1998)
[21] Mohr, A., Tokheim, L., Gleicher, M.: Direct manipulation of interactive character skins. In: Proceedings of the 2003 symposium on Interactive 3D graphics, pp. 27–30 (2003)
[22] Lewis, J.P., Cordner, M., Fong, N.: Pose space deformation: a unified approach to shape interpolation and skeleton-driven deformation. In: Computer Graphics, SIGGRAPH'2000 Proceedings pp. 165–172 (2000)
[23] Leclercq, A., Akkouche, S., Galin, E.: Mixing triangle meshes and implicit surfaces in character animation. In: Animation and Simulation'01, 12th Eurographics Workshop Proceedings, Manchester, England, pp. 37–47 (2001)
[24] Gagvani, N., Silver, D.: Animating Volumetric Models. Graphical Models 63(6), 443–458 (2001)
[25] Nedel, L.P., Thalmann, D.: Anatomically modeling of deformable human bodies. The Visual Computer 16(6), 306–321 (2000)
[26] Garland, M., Heckbert, P.S.: Surface simplification using quadric error metrics. In: Proceedings of SIGGRAPH, ACM Press ACM SIGGRAPH 2000, Computer Graphics Proceedings, Annual Conference Series, ACM, vol. 1997, pp. 209–216 (2000)
[27] Ma, W.-C., Wu, F.-C., Ouhyoung, M.: Skeleton Extraction of 3D Objects with Radial Basis Functions. In: IEEE Proceedings of the Shape Modeling International 2003 (SMI'03) (2003)

Advanced Human Body and Head Shape Representation and Analysis

Afzal Godil

National Institute of Standards and Technology, 100 Bureau Dr, MS 8940,
Gaithersburg, MD 20899, USA
godil@nist.gov

Abstract. The 3D scans of human bodies contain over hundreds of thousand grid points. To be used effectively for analysis, indexing, searching, clustering and retrieval, these human bodies requires a compact shape representation. We have developed compact representations based on human body shape based on lengths mostly between joints of single large bones and in the second method silhouettes of the human body are created and then encoded as Fourier shape descriptors. We also have developed two such compact representations based on human head shape by applying Principal Component Analysis on the facial surface and in the second method the whole head is transformed to a spherical coordinate system expanded in a basis of Spherical Harmonics.

Keywords: 3D anthropometry, shape searching, PCA, Spherical Harmonics.

1 Introduction

With the wide availability of 3D scanning technologies, 3D geometry is becoming an important type of media. Large numbers of 3D models are created every day and many are stored in publicly available databases. Understanding the 3D shape and structure of these models is essential to many scientific activities. These 3D scientific databases require methods for storage, indexing, searching, clustering, retrieval and recognition of the content under study. We have developed techniques for searching a human database and have used the CAESAR anthropometric database which consists of a database of approximately 5000 human subjects. In our study we have implemented methods for similarity based retrieval from the CAESAR human database based on both human body and head shape.

Previous work on human body retrieval based on body shape was performed by [Paquet and Rioux 2003]. They performed content-based anthropometric data mining of three dimensional scanned human by representing them with compact support feature vectors. They showed a virtual environment to perform visual data mining on the clusters and to characterize the population by defining archetypes. [Paquet 2004] introduced cluster analysis as a method to explore 3D body scans together with the relational anthropometric data as contained in the CAESAR anthropometric database.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 92–100, 2007.

[Azouz 2002, 2004] analyzed human shape variability using a volumetric representation of 3D human bodies and applied a principal components analysis (PCA) to the volumetric data to extract dominant components of shape variability for a target population. Through visualization, they also showed the main modes of human shape variation. The work of [Allen 2004] demonstrated a system of synthesizing 3D human body shapes, according to user specified parameters; they used 250 CAESAR body scans for training.

Retrieval based on head shape was performed by [Ip and Wong 2002]. Their similarity measure was based on Extended Gaussian Images of the polygon normal. They also compared it to an Eigenhead approach.

The 3D scans of human bodies in the CAESAR human database contain over two hundred fifty thousand grid points. To be used effectively for indexing, searching, clustering and retrieval, this human body data requires a compact representation. We have developed two such representations based on human body shape: 1) a descriptor vector d, based on lengths mostly between single large bones. Thus, we form a 3D body description vector of fifteen distances, d, with wrist to elbow, elbow to shoulder, hip to knee etc.; and 2) three silhouettes of the human body are created by rendering the human body from the front, side and top. These silhouettes are then encoded as Fourier descriptors of features for later similarity based retrieval. These two methods are explained in more details in the Body Shape Descriptor section.

We also have developed two compact representation based on human head shape: 1) applying Principal Component Analysis (PCA) on the 3D facial surface and creating PCA based facial descriptors; and 2) in the second method the 3D triangular grid of the head is transformed to a spherical coordinate system by a least squares approach and expanded in a basis of spherical harmonics. More explanation of these two representations of human head shape follow in the Head shape Descriptor section.

2 CAESAR Database

The CAESAR (Civilian American and European Surface Anthropometry Resource) project has collected 3D Scans, seventy-three Anthropometry Landmarks, and Traditional Measurements data for each of the 5000 subjects. The objective of this study was to represent, in three-dimensions, the anthropometric variability of the civilian populations of Europe and North America and it was the first successful anthropometric survey to use 3-D scanning technology. The CAESAR project employs both 3-D scanning and traditional tools for body measurements for people ages 18-65. A typical CAESAR body is shown in Figure 1.

The seventy-three anthropometric landmarks points were extracted from the scans as shown in Figure 2. These landmark points are pre-marked by pasting small stickers on the body and are automatically extracted using landmark software. There are around 250,000 points in each surface grid on a body and points are distributed uniformly.

Fig. 1. A CAESAR body with three postures

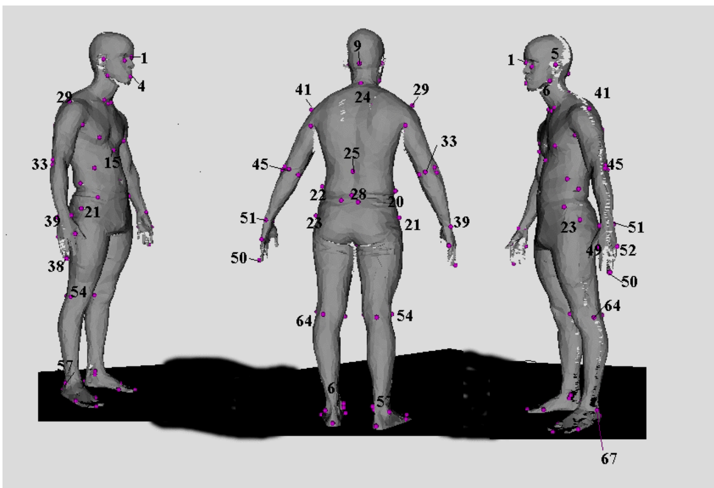


Fig. 2. A Caesar body with landmark numbers and positions

3 Body Shape Descriptor

We now describe two methods for creating descriptors based on human body shape:

3.1 Distance Based

The first method uses a descriptor vector d based on lengths mostly between single large bones. For descriptor vector purposes, we require lengths only between landmark points where their separation distance is somewhat pose-independent. The reason it is not completely pose invariant is that distance are between landmark points which are on the surface body compared to the distance between the center of the joint axis. This applies to points connected by a single large bone as shown in

Figure 3. Thus, we form a descriptor vector of fifteen distances, d, with d1 wrist to elbow, d2, elbow to shoulder, d3 hip to knee etc. For which the Euclidean distance

d = ||Pi – Pj||

is somewhat invariant across different poses. Distances such as chin-knee are avoided. The distance based descriptor is then used with the L1 and L2 norm to create a similarity matrix.

$$\text{The L1 distance: } d(P_i, P_j) = \sum_{i=1}^{k} \left| P_i - P_j \right| \tag{1}$$

$$\text{The L2 distance: } d(P_i, P_j) = \sum_{i=1}^{k} \left| \sqrt{P_i^2 - P_j^2} \right| \tag{2}$$

More details and shortcomings about this descriptor were described in the paper [Godil 2003, 2006]. To test how well the distance based descriptor performs, we studied the identification rate of a subset of 200 subjects of CESAR database where the gallery set contains the standing and the probe set contains the sitting pose of each subject. In this discussion, the gallery is the group of enrolled descriptor vector and the probe set refers to the group of unknown test descriptor vectors.

The measure of identification performance is the "rank order statistic" called the Cumulative Match Characteristic (CMC). The rank order statistics indicates the probability that the gallery subject will be among the top r matches to a probe image of the same subject. This probability depends upon both gallery size and rank. The CMC at rank 1 for the study is 40%.

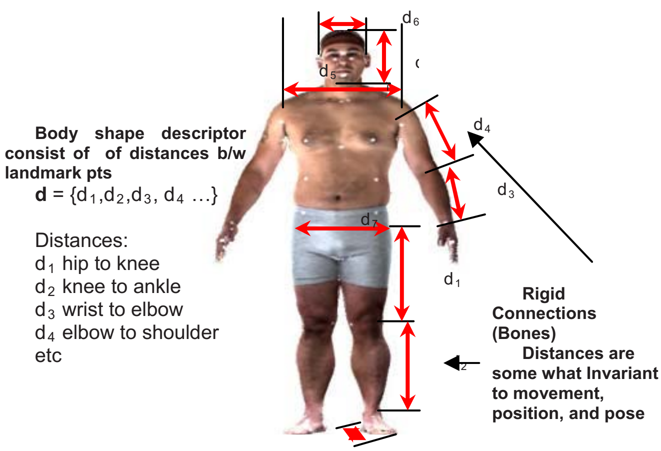


Fig. 3. A distance based body shape descriptor

3.2 Silhouette Fourier Based

The second method of body shape descriptor that we propose is based on rendering the human body from the front, side and top directions and creating three silhouettes of the human body as shown in Figure 4. The theory is that 3D models are similar if they also look similar from different viewing angles. The silhouette is then represented as R(radius) of the outer contour from the center of origin of the area of the silhouettes. These three contours are then encoded as Fourier descriptors which are used later as features for similarity based retrieval. The number of real part of Fourier modes used to describe each silhouette is sixteen (32); hence each human body is described by a vector of length ninety six (96). This method is pose dependent, so only bodies of the same pose can be compared. The Fourier based descriptor is then used with the L1 and L2 norm to create a similarity matrix.

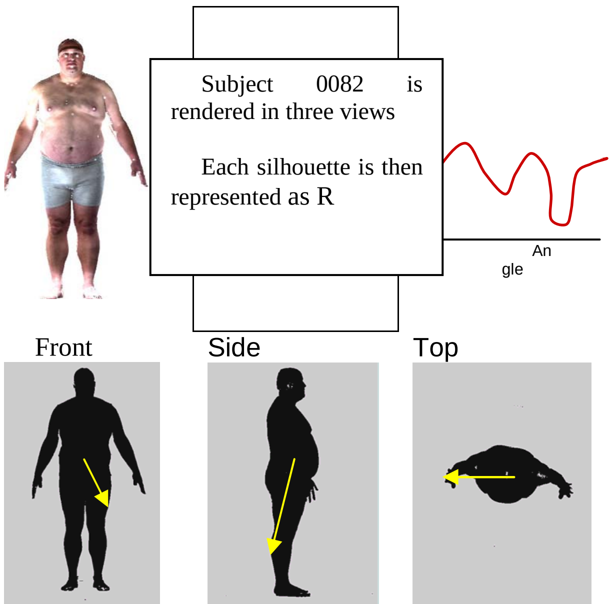


Fig. 4. Subject 00082 is rendered in three silhouette views

4 Head Shape Descriptor

We now describe two methods for creating descriptors based on human head shape:

4.1 PCA Based

The facial grid is cut from the whole CAESAR body grid using the landmark points 5 and 10 as shown in Figure 5 and listed in Table 1. Table 1 list all the numbers and names of landmark points used in our 3D face shape descriptor. The new generated facial grid for some of the subjects with two different views is shown in Figure 5. The facial grid is very coarse for some of the subjects in the seated pose.

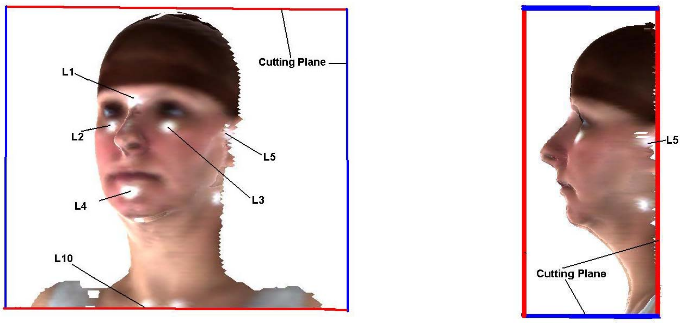


Fig. 5. Landmark points 1, 2, 3, 4, 5 and 10. Vertical and horizontal lines are the cutting plane.

Table 1. Numbers and names of landmark points used in our 3D face

1 Sellion	2 Rt Infraobitale
3 Lt Infraobitale	4 Supramenton
5 Rt.Tragion	6 Rt. Gonion
7 Lt. Tragion	8 Lt. Gonion
10 Rt. Clavicale	12 Lt.Clavicale

Next, we use four anthropometric landmark points (L1, L2, L3, L4) as shown in Figure 5, located on the facial surface, to properly position and align the face surface using an iterative method. There is some error in alignment and position because of error in measurements of the position of these landmark points. This Max error was 15 mm, obtained by taking the difference of distance between landmark points |L1-L2| and between |L3 –L4| for subjects standing compared to subjects sitting. Then we interpolate the facial surface information and color map on a regular rectangular grid whose size is proportional to the distance between the landmark points L2 and L3 (d=| L3 - L2 |) and whose grid size is 128 in both directions. We use a cubic interpolation and handle missing values with the nearest neighbor method when there are voids in the original facial grid. For some of the subjects there are large voids in the facial surface grids. Figure 6 shows the facial surface and the new rectangular grid.

We properly positioned and aligned the facial surface and then interpolated the surface information on a regular rectangular grid whose size is proportional to the distance between the landmark points. Next we perform Principal Component Analysis (PCA) on the 3D surface and similarity based descriptors are created. In this method the head descriptor is only based on the facial region. The PCA recognition method is a nearest neighbor classifier operating in the PCA subspace. The similarity measure in our study is based on L1 distance and Mahalanobis distance.

To test how well the PCA based descriptor performs, we studied the identification between 200 standing and sitting subjects. The CMC at rank 1 for the study is 85%. More details about this descriptor are described in the paper by [Godil 2004, 2006].

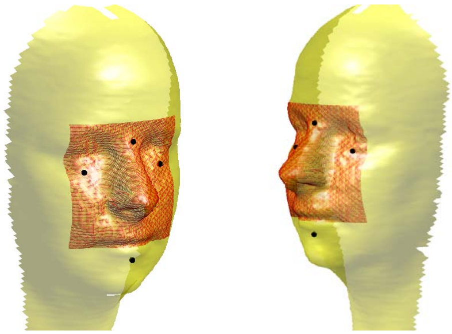


Fig. 6. Shows the new facial rectangular grid for two subjects

4.2 Spherical Harmonics Based

In the second method the 3D triangular grid of the head is transformed to a spherical coordinate system by a least square approach and expanded in a spherical harmonic basis as shown in Figure 7. The main advantage of the Spherical Harmonics Based head descriptor is that it is orientation and position independent. The spherical harmonics based descriptor is then used with the L1 and L2 norm to create similarity measure. To test how well the Spherical Harmonics Based head descriptor performs, we studied the identification of the human head between 220 standing and sitting subjects. The CMC at rank 1 for the study is 94%.

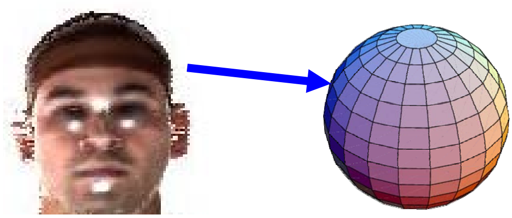


Fig. 7. 3D head grid is mapped into a sphere

5 Results

5.1 Retrieval Results

The web based interface enables us to select a particular body, or a random body or bodies, based on some criteria such as weight, age, height, etc as shown Figure 8. Subsequently, we can perform similarity based retrieval based on a single descriptor (out of the four descriptors). Using four descriptors allows users to select the best descriptor for their application, such as the use of head descriptor for helmet or eyeglasses design. The partial results from a body shape based similarity retrieval for subject number 16270 are shown Figure 9. The partial results from head shape PCA based similarity retrieval for subject number 00014 are shown in Figure 10.

The initial results show that the results and amount of time for retrieval are very reasonable.

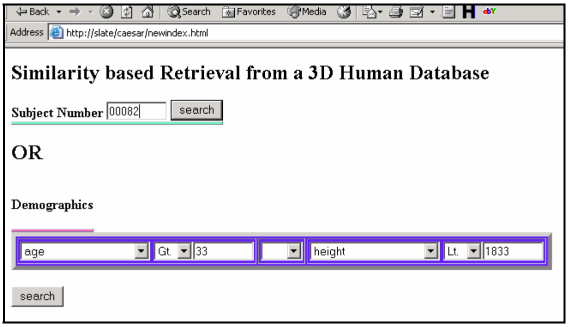


Fig. 8. The web based interface allows one to select a particular body, or a random body

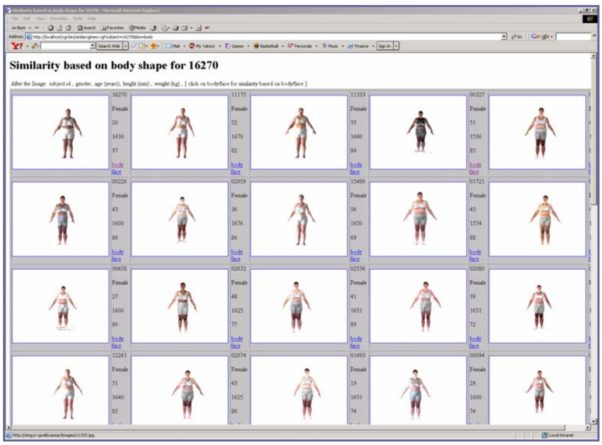


Fig. 9. Similarity based retrieval for 16270 based on body shape

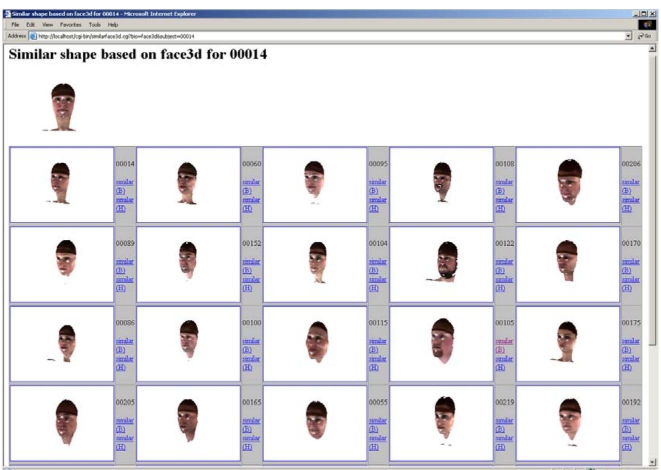


Fig. 10. Similarity based retrieval for 00014 based on PCA facial shape

6 Conclusion

We have developed four methods for searching the human database using similarity of human body and head shape. Based on some of our initial observations and from our CMC results for an identification study between 200 standing subjects and 200 sitting subjects, it can be said that the body and head descriptors represent the CAESAR bodies quite accurately.

References

1. CAESAR: Civilian American and European Surface Anthropometry Resource web site: http://www.hec.afrl.af.mil/cardlab/CAESAR/index.html
2. Paquet, E., Rioux, M.: Anthropometric Visual Data Mining: A Content-Based Approach. In: Submitted to IEA 2003 - International Ergonomics Association XVth Triennial Congress. Seoul, Korea. NRC 44977 (2003)
3. Paquet, E.: Exploring Anthropometric Data Through Cluster Analysis, Digital Human Modeling for Design and Engineering (DHM). Oakland University, Rochester, Michigan, USA. June 15-17, 2004. NRC 46564 (2004)
4. Ben Azouz, Z., Rioux, M., Lepage, R.: 3D Description of the Human Body Shape: Application of Karhunen-Loève Expansion to the CAESAR Database. In: Proceedings of the 16th International Congress Exhibition of Computer Assisted Radiology Surgery. Paris, France, June 26-29, 2002 (2002)
5. Ben Azouz, Z., Rioux, M., Shu, C., Lepage, R.: Analysis of Human Shape Variation using Volumetric Techniques. In: The 17th Annual Conference on Computer Animation and Social Agents (CASA2004). Geneva, Switzerland, July 7-9, 2004 (2004)
6. Ip, H.H.S., Wong, W.: 3D Head Model Retrieval Based on Hierarchical Facial Region Similarity. In: Proc. of 15th International Conference on Visual Interface (VI2002), Canada (2002)
7. Godil, A., Grother, P., Ressler, S.: Human Identification from Body Shape. In: proceedings of 4th IEEE International Conference on 3D Digital Imaging and Modeling, October 6-10, 2003, Banff, Canada (2003)
8. Godil, A., Ressler, S., Grother, P.: Face Recognition using 3D surface and color map information: Comparison and Combination. In: the SPIE’s symposium on Biometrics Technology for Human Identification, April 12-13, 2004, Orlando, FL (2004)
9. Godil, A., Ressler, S.: Retrieval and Clustering from a 3D Human Database based on Body and Head Shape. In: SAE Digital Human Modeling conference, 02 July to 4th August, 2006, Lyon, France (2006)
10. Allen, B., Curless, B., Popovic, Z.: Exploring the space of human body shapes: data-driven synthesis under anthropometric control. In: Proc. Digital Human Modeling for Design and Engineering Conference, Rochester, MI, June 15-17, 2004. SAE International (To appear)

Hand Grasping Motion Simulation for Astronauts Training

Qiang Guo and Yuqing Liu

China Astronaut Research and Training Center, Beijing 100094, China

Abstract. The objective of this paper is to find an appropriate method to realize the simulation of hand grasping motion for astronauts training. In this paper, on the basis of analysis and comparison of a variety of methods of hand grasping motion, we come up with an idea combined interpolation and inverse kinematics for hand motion simulation. An example demonstrates the method's effectiveness and practicability.

Keywords: Hand motion simulation, Interpolation methods, Inverse kinematics.

1 Introduction

In space mission, astronauts must finish a lot of operation. Before working in space, astronauts need to be trained on the ground. It's difficult to use physical methods on Earth to simulate the special space environment, such as weightlessness. Computer Simulation is developing very fast in recent years, many research institutes, Commercial Companies, Universities have done a lot of work on computer simulation and obtained great achievements. Therefore we consider using computer simulation to create virtual space working environment to assist astronauts to finish their trainings on the ground. In virtual environment virtual hand is used to substitute real hand to finish hand operation, which includes grasp, pull, push, and grip. If the operation can be realized in virtual environment, astronauts can practice their space work on the ground and finish their mission quickly and safely in space.

There are many ways of completing the simulation of hand grasping motion such as: Key frame, motion capture, kinematics and dynamics. Key frame is that the key picture which is designed by senior animator. The in-between frames are designed by assistant animator. With the development of computer graphics, the in-between frames are displaced by interpolation. Key frame depends on the skills of animator rather than the quality of animation system. Motion capture is using data glove or light mark to track people's motion. It converts the electrical signals into digital signals and inputs them into the model, and it can drive the model. But data glove has some technical problems by nature, sometimes the motion of virtual hand doesn't match with the real hand motion in practice. We may get unusual gestures. Furthermore, data glove is fragile to external interaction. Dynamics is that according to the strength and torque of the joints it can get the velocity and acceleration which conforms to physical constraints. But it will cost a lot to calculate the strength and torque. Kinematics comes from robotics[1]. People's body is regarded as rigid body. With the position and orientation of end-effector, the configuration of multibody can be obtained.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 101–109, 2007.

In this paper, we come up with an idea combined interpolation and inverse kinematics for hand motion simulation. During the period of real grasping, we use data glove to sampling several frames as key frames. To simulate the motion of hand grasping, we need 24 frames data of every joint in 1 second at least. The data of key frame is not enough for simulating the whole process of grasping. Therefore, we use interpolation to get the position of fingertip in-between frames, then calculate the joint angles of every joint in-between frames by Inverse kinematics solution. When we get the joint angles of all joints of the whole process, we can simulate the whole action of hand grasping.

2 Hand Structure Description

2.1 Skeleton Description of Human Hand

The proposed hand model is composed of a skeletal model. We describe the posture of a hand with the position of the bones and the joints in it. We model the hand skeleton with rigid links and joints.

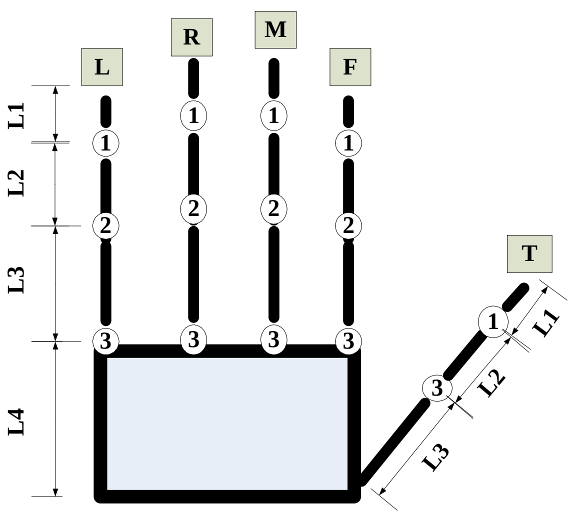


Fig. 1. Diagram for the hand skeleton

Fig.1 illustrates the hand skeleton[2]. We model the skeleton of human hand with a hierarchical structure that consists of rigid body links and joints.
T=Thumb, F=First finger, M=Middle finger,
R=Ring finger, L=Little finger.
1, 2, 3 represents the joints of distal phalange, middle phalange, proximal phalange.
L1=the length of distal phalange,
L2=the length of middle phalange,
L3=the length of proximal phalange,
L4=the length of metacarpal
The proportion of length of L1:L2:L3:L4=2:3:5:8

2.2 The Motion of Single Finger

Fig.2 shows the side view of single finger[3]. Letter A represents the position of the finger tip. Letter B represents the position of the joint of distal phalange. Letter C represents the position of the joint of middle phalange. Letter D represents the position of the joint of proximal phalange. The local reference frames of X_1BY_1、X_2BY_2、X_3BY_3 are based on the points of B, C, D. is the angle between L_1 and Y_1, is the angle between L_2 and Y_2 , is the angle between L_3 and Y_3 , is the angle between L_4 and Y.

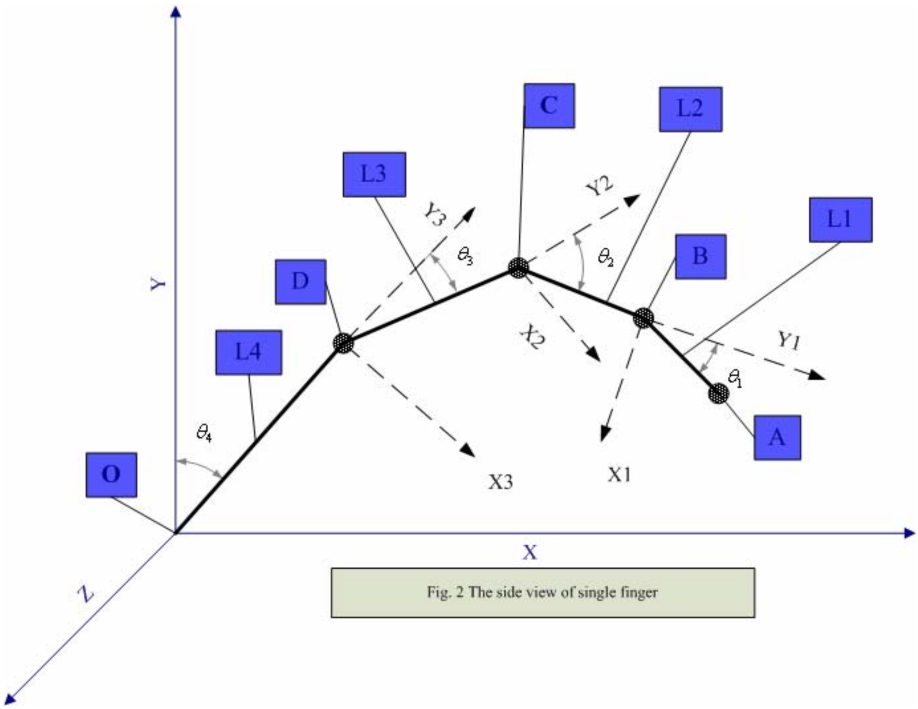


Fig. 2 The side view of single finger

2.3 Equation of Hand Motion

According to Fig.2, we get the equation of hand motion of hand tip. As follow:

$$\begin{bmatrix} X_A \\ Y_A \\ Z_A \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta_4 & \sin\theta_4 & 0 & L_4\cos\theta_4 \\ -\sin\theta_4 & \cos\theta_4 & 0 & L_4\sin\theta_4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta_3 & \sin\theta_3 & 0 & L_3\cos\theta_3 \\ -\sin\theta_3 & \cos\theta_3 & 0 & L_3\sin\theta_3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta_2 & \sin\theta_2 & 0 & L_2\cos\theta_2 \\ -\sin\theta_2 & \cos\theta_2 & 0 & L_2\sin\theta 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} L_1\cos\theta_1 \\ L_1\sin\theta_1 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{cases} X_A = L_1\sin(\theta_1+\theta_2+\theta_3+\theta_4) + L_2\sin(\theta_2+\theta_3+\theta_4) + L_3\sin(\theta_3+\theta_4) + L_4\sin(\theta_4) \\ Y_A = L_1\cos(\theta_1+\theta_2+\theta_3+\theta_4) + L_2\cos(\theta_2+\theta_3+\theta_4) + L_3\cos(\theta_3+\theta_4) + L_4\cos(\theta_4) \end{cases}$$

X_A, Y_A, Z_A is the position vector of the coordinate A. Because we just simulate the planar motion of single finger, so we set Z_A=1 in the matrix.

3 Interpolation and Inverse Kinematics

3.1 Interpolation Methods

In order to get the position of finger tip (A) in-between frame in Fig.2, we adopt interpolation methods. First, we get the value of X (A). Then, according to different interpolation methods, we get the value of Y (A). In this paper, we compare these interpolation methods such as Linear interpolation, Lagrange interpolation, Newton interpolation [4]. Linear interpolation is a basic method which it can obtain linear relationship between two sampling points to get the interpolation function. For example: the sampling point is $(x_0,y_0),(x_1,y_1)$, so the interpolation function is

$$L(x) = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0) \tag{1}$$

Lagrange interpolation polynomial is

$$L_n(x) = y_0 l_0(x) + y_1 l_1(x) + \cdots y_n l_n(x) \tag{2}$$

$$l_k(x) = \frac{(x - x_0)\cdots(x - x_{k-1})(x - x_{k+1})\cdots(x - x_n)}{(x_k - x_0)\cdots(x_k - x_{k-1})(x_k - x_{k+1})\cdots(x_k - x_n)} \tag{3}$$

Newton interpolation polynomial is

$$N_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \cdots + a_n(x - x_0)(x - x_1)\cdots(x - x_{n-1}) \tag{4}$$

Newton method divides into Difference Method[4] and Difference Quotient Method[3]. Difference Method is suitable for the condition of equidistant point $x_k = x_0 + kh$. When the sampling points are not equidistant point, we adopt Difference Quotient Method.

To compare Newton interpolation with Lagrange interpolation, Newton interpolation doesn't need to recalculate when adding a new point. According to the interpolation polynomials uniqueness theorem, Lagrange interpolation polynomial and Newton interpolation polynomial are the same interpolation polynomial. In order to make the calculation cost litter, we select Newton interpolation. In a word, we select Linear interpolation or Newton interpolation to get the value of Y.

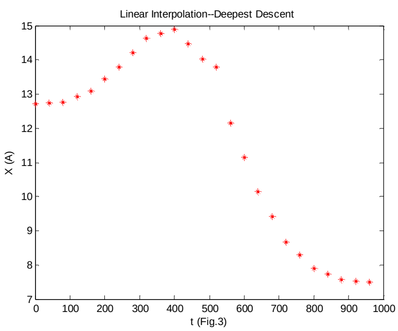
3.2 Inverse Kinematics Methods

After getting the data of finger tip (A), we can calculate the positions of B, C, D in-between frames. In order to get the data of B, C, D, we use Inverse dynamic solution. Inverse kinematics solution comes from robots. With the position of articulation link's tip, it gets other joints' data. In this paper, we compare three Inverse kinematics methods. They are Steepest descent Method, Newton Method, Quasi-Newton Method[5]. Steepest descent Method is that it starts from a point, selecting the orientation to make the object function f(x) decline quickly and converge at minimizer point fast. Newton Method is that to get the optimal solutions by iterative computations along Newton orientation $d = -\nabla^2 f(x)^{-1} \nabla f(x)$. It has many advantages: the speed of convergence is quickly; it can find the global optimal solution. But it has some shortcoming at the same time, such as if the initial point is far from minimize point, this method can not converge to the optimal point . In order to conquer this shortcoming, people proposed Quasi-Newton Method. Quasi-Newton Method is that to substitute the $\nabla^2 f(x)$'s inverse matrix with a matrix which doesn't has second derivative matrix. According to different construction matrix, there are many Quasi-Newton Methods. In this paper, we adopt DFP method.

4 Experiments and Analysis

In order to decide which methods of interpolation and inverse kinematics are suitable to this research, we conduct the following experiments. For the sake of predigesting the experiments, we just simulate planar motion of single finger. Here we assume that first finger, middle finger, ring finger, and little finger are in the same length. We take grasping motion for example. And the grasped object is not considered in this research. The period of motion is 1 second and sampling the data of key frames every 80 ms. The bound of $\theta_1, \theta_2, \theta_3, \theta_4$ is $0 \le \theta_1 \le 80^\circ$, $0 \le \theta_2 \le 95^\circ$, $0 \le \theta_3 \le 60^\circ$, $\theta_4 = 45^\circ$. To make the image smoothly, we need 24 frames of image. For the sake of getting 24 frames of image, firstly we use Linearity interpolation to get the value of X of A, then using Newton interpolation or Linear interpolation to get the value of Y of A, finally comparing the result. After getting the data of finger tip, we can calculate the data of B, C by Inverse kinematics solution. In this paper, we compare these three methods as follow: Steepest descent, Newton, Quasi-Newton.

Hardware platform: CPU Inter Core Cuo2 1.83G, Memory 1G. Software platform Windows XP, Matlab 7.1 and Visual C++ 6.0.



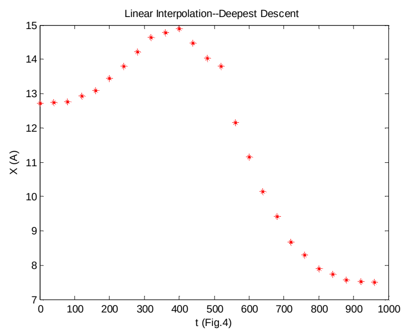
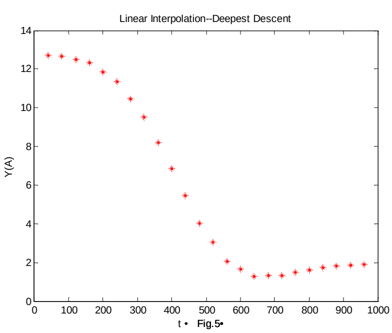
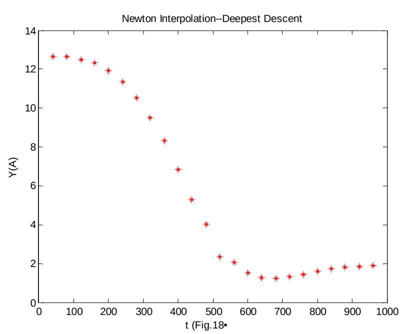
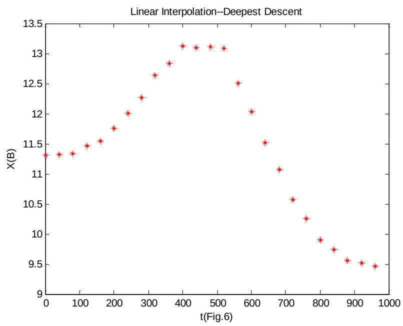
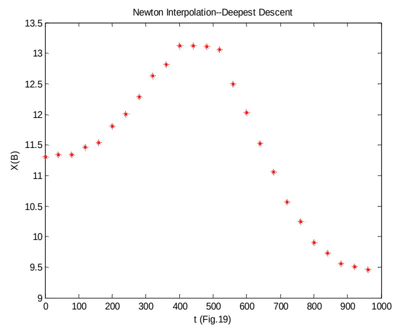


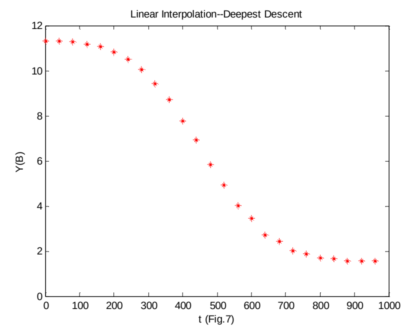
Fig.3 and Fig.4 is obtained by Linear Interpolation. They show the relationship of X between insert point(X, Y) and key frame point(X_0, Y_0), (X_1, Y_1). X=(X_0+X_1)/2.

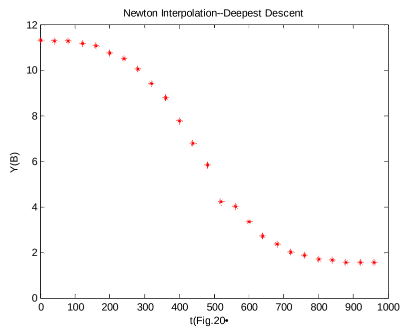


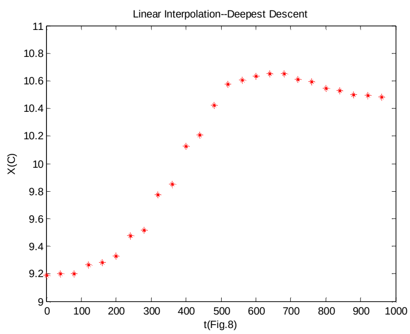


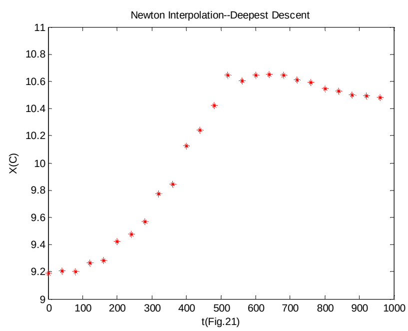

Linear Interpolation--Deepest Descent
X(B)
t(Fig.6)

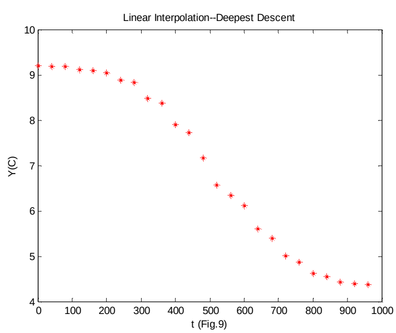

Newton Interpolation--Deepest Descent
X(B)
t (Fig.19)

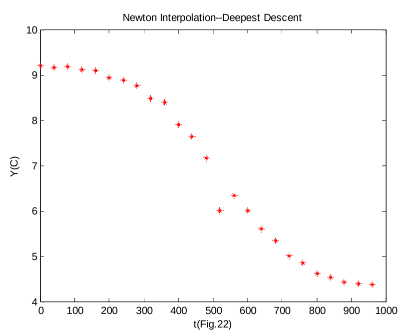

Linear Interpolation--Deepest Descent
Y(B)
t (Fig.7)


Newton Interpolation--Deepest Descent
Y(B)
t(Fig.20•

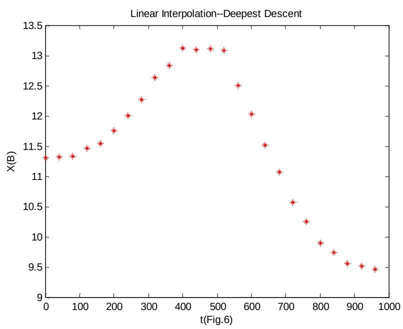

Linear Interpolation--Deepest Descent
X(C)
t(Fig.8)

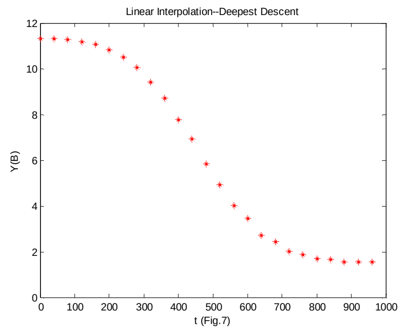

Newton Interpolation--Deepest Descent
X(C)
t(Fig.21)

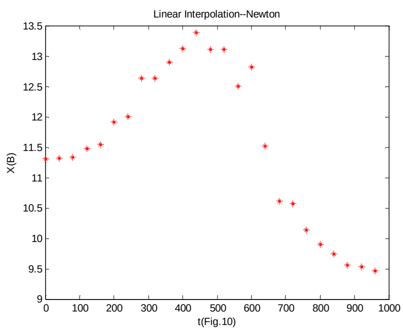

Linear Interpolation--Deepest Descent
Y(C)
t (Fig.9)

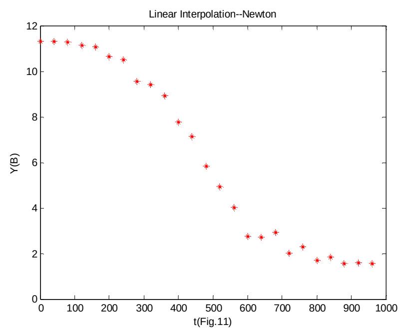

Newton Interpolation--Deepest Descent
Y(C)
t(Fig.22)

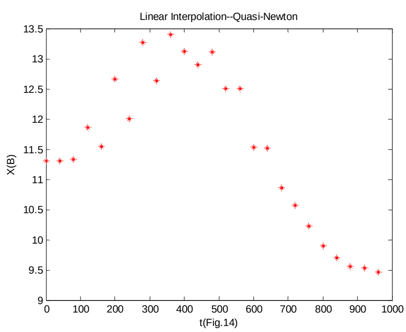
In order to get the value of Y, Fig.5 adopts Linear Interpolation and Fig.18 adopts Newton Interpolation. In Fig.6-Fig.9 the data of B and C is obtained by Deepest Descent Method which is based on the data of A in Fig.3 and Fig.5. In Fig.19-Fig.22 the data of B and C is also obtained by Deepest Descent Method which is based on the data of A in Fig.4 and Fig.18 .To compare these pictures, we find the result of Linear Interpolation and Newton Interpolation which is based on Deepest Descent Method has little difference.











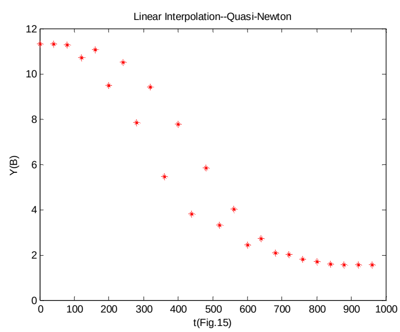
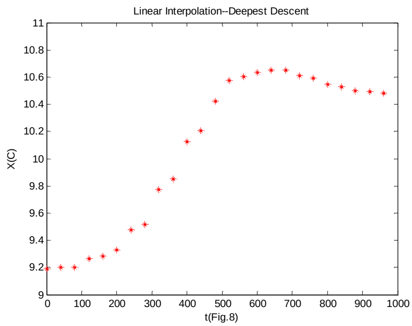
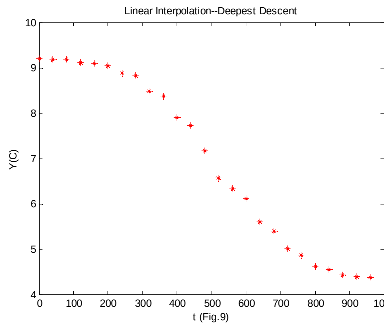
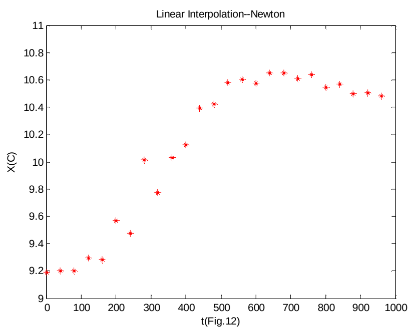
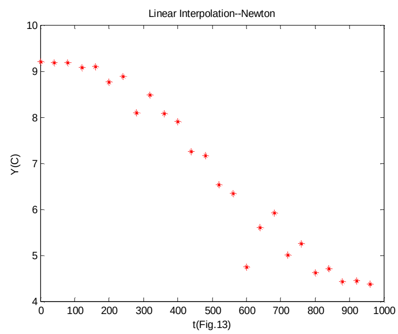


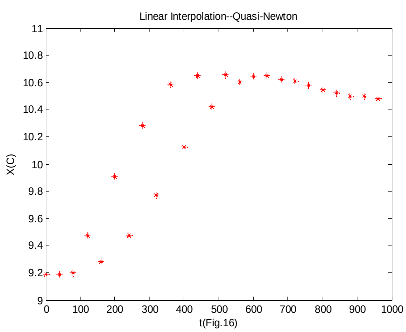
Fig.6-Fig 17 is based on the data of A which is obtained by Linear interpolation. With the image we conclude that the data of Steepest descent Method is accurate, the image is smooth. But the data of Newton Method, Quasi-Newton Method is unstable and has many fluctuations. The reason why the Newton Method doesn’t work very

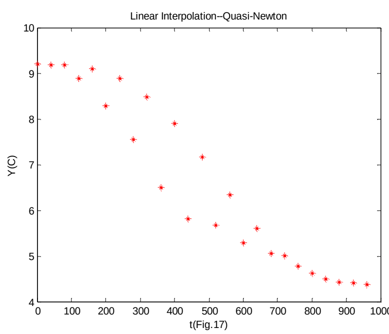






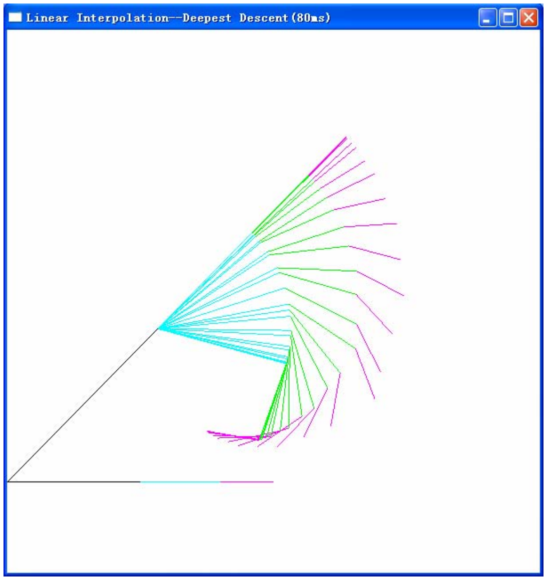






well is that Newton orientation is not descendent orientation. Furthermore, although the object function is descending, the result is not the optimal solution along Newton orientation. When the initial point is far from the target point, the result with Newton Method, Quasi-Newton Method is not good. To make the result better, let the initial point get close to the target point.

The picture is drawed by the data of Linear Interpolation—Deepest Descent. It reflects the process of hand grasping motion. From the picture, we find it has a little difference with real hand grasping motion. This means we could use this method to simulate the real action in some areas. The methodology proposed in this paper is effective and practically.



5 Conclusion

The example demonstrates that interpolation methods affect the result a little, but the Inverse kinematics methods affect the result a lot. If you select different methods, the result is different. Steepest descent Method is suitable for hand grasping motion and iterative calculation cost a little. Quasi-Newton's Method iterative calculation also cost a little, but the result is not good .Newton Method's iterative calculation cost a lot and the result is not good, it depends on many conditions such as Newton orientation and the position of initial point. The methodology proposed in this paper is effective and practically.

References

1. Cai, Z.: Robotics, Qing Hua publishing company (2000)
2. Yasumuro, Y., Chen, Q.: Three-dimensional modeling of the human hand with motion constraints. Image and Vision Computing 17, 149–156 (1999)
3. Yang, Q., Zhang, L.: Research on the rule of the process of human finger grasping. China Mechanism Engineering 2004-15-13
4. Burden, R.L., Douglas Faires, J.: Numerical Analysis, 7th edn. Higher Education Press (2005)
5. Chen, B.: Optimization Theory and Algorithm, Qing Hua publishing company (2005)

Predefined Manikins to Support Consideration of Anthropometric Diversity by Product Designers

Dan Högberg[1] and Keith Case[1,2]

[1] School of Technology and Society, University of Skövde
SE-541 28, Skövde, Sweden
[2] Mechanical and Manufacturing Engineering, Loughborough University
Loughborough, Leicestershire, LE11 3TU, UK

Abstract. The paper discusses the complexity involved in considering targeted product users' anthropometric variation in multivariate design problems, such as the design of workplaces or vehicle interiors. The authors argue for the advantages of offering designers a number of predefined digital human models to incorporate in the CAD product models. A study has been carried out in order to illustrate the use of predefined digital human models in vehicle interior design by using the Digital Human Modelling (DHM) tool RAMSIS. The paper takes a designer's view of digital human modelling and illustrates how DHM can be of great value in the design process, but also considers what implications this has on the functionality and usability of DHM tools.

Keywords: Anthropometric Diversity, Digital Human Modelling, Design Process, Product Design, Ergonomics, Human Factors.

1 Introduction

The complexity involved in considering targeted product users' anthropometric variation in multivariate design problems, such as the design of workplaces or vehicle interiors, is not new and was for example described in the 1970s by Roebuck et al. (1975), and more recently by Ziolek and Wawrow (2004) and Robinette and Hudson (2006). The use of digital human modelling (DHM) tools is expected to spread as functionality and value increases, and it is likely that many of the new users will be generalists rather than specialists in anthropometric diversity and its implications for design work. This may have major consequences when attempting to accommodate anthropometric variation amongst the end users of products. Thus, there is a need to make it 'easier to do it correctly', particularly for non-experts, and DHM tools have the potential to assist in meeting this objective. One approach is to offer designers a number of generic predefined digital human models, which might be called a *manikin family*, to incorporate in the CAD product models. Such a family would ideally represent the targeted users well enough, regardless of the design task at hand. The paper aims to illustrate how digital human modelling can be of great value to designers in the design process, but also considers what implications this has on the functionality and usability of DHM tools.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 110–119, 2007.

1.1 Use of Manikin Families in Human Simulation

Starting from the characteristics of the design problem, the DHM tool user typically determines the number of manikins required and their anthropometrical dimensions. This may for example be done by choosing manikins from a set of predefined manikins, each with different anthropometry. However, it is sometimes difficult to determine in advance which manikins are the important ones for a specific design problem, and which manikins could be left out of the analysis. This leaves it to the tool user to carry out the frequently complex selection of appropriate manikins for the design problem at hand. For an expert tool user this might be straightforward, but for a 'common' tool user this is a difficulty and a source of error, especially in multivariate design problems. One approach to this problem is to perform simulations including all members of a manikin family, regardless of whether or not some manikins are important for the specific design task. This approach would be supported by a predetermined set of manikins, e.g. as a company or project standard family of manikins that is established to correctly, or at least adequately, represent the targeted product users. This would be similar to, and indeed a complement to, having a group of real test persons within a company that would always be recruited to assess products being developed. One difference between virtual and real test persons is that the virtual test group will always be available, even concurrently at different places. A concern is that the virtual test group will only do what they are told to do, putting pressure on the tool user to set up the study properly (Ziolek and Kruithof, 2000).

2 Method

A study has been carried out in order to illustrate the use of predefined digital human models in vehicle interior design by using the DHM tool RAMSIS (Seidl, 1997). Of particular interest is how different configurations of families of human models influence the representation of targeted users and hence the accommodation. Two different manikin family configurations were used in the study. The first family is a collection of 17 manikins, specifically modelled in RAMSIS for this study, based on Bittner's A-CADRE family (Bittner, 2000), illustrated in Figure 1.

Fig. 1. One of the two manikin families in the study

Starting from a statistical treatment of anthropometric data, Bittner and his colleagues developed the CADRE manikin family, and subsequently developed it into A-CADRE (Bittner et al., 1987; Bittner, 2000). This resulted in anthropometric descriptions for 17 manikins; 16 boundary cases and one centroid case. The A-CADRE family has been validated to capably represent anthropometric diversity, and by using these manikins as user representatives in design, a high level of accommodation can be achieved (Bittner, 2000). Since not all of the 19 body variables in the A-CADRE definition are possible to enter in RAMSIS, the decision was made to just use the three key variables stature, sitting height and waist circumference. Waist circumference is not present in the A-CADRE definition, so values for weight were used instead. This assumption is believed to be adequate due to the relatively high correlation between the two dimensions (Kroemer et al., 2001).

The second family was created by using all manikins in RAMSIS Typology, i.e. a collection of 45 manikins per gender provided as standard in the DHM tool. RAMSIS Typology is based on the theory that the definition of the characterizing property of stature, proportion (ratio of the sitting height and length of the legs) and corpulence of an individual is sufficient to predict all other body dimensions for this person (Flügel et al., 1986; Bubb et al., 2006). The anthropometric database incorporated in RAMSIS was used in the study, with German as the selected nationality, using both genders and 18-70 years as age group. Table 1 show the minimum and maximum percentile value of each key variable for each family approach, for males. The values for female are quite similar, and are available in (Högberg and Case, 2005). The accommodation level that these value ranges answer to was calculated using the multidimensional analysis functionality in two separate anthropometric software packages: PeopleSize 2000 Professional (PeopleSize, 2004) and BodyBuilder (Human-Solutions, 2003).

Table 1 shows that A-CADRE has greater percentile coverage than RAMSIS Typology in stature, but less in sitting height and very similar for waist circumference. Even though different percentile ranges are covered, the two approaches result in approximately the same accommodation level, i.e. approximately 86%. Further reduction in accommodation will happen if more body dimensions limit the design problem (Roebuck et al., 1975). However, this reduction is likely to be

Table 1. Characteristic data of RAMSIS Typology and A-CADRE families

MALE		**RAMSIS Typology**	**A-CADRE**
Stature	min	3%-ile	1%-ile
	max	96.4%-ile	99%-ile
	coverage	93.4%	98%
Sitting height	min	0.9%-ile	3.4%-ile
	max	98.8%-ile	96.6%-ile
	coverage	97.9%	93.2%
Waist circumference	min	4%-ile	3.1%-ile
	max	97.6%-ile	96.9%-ile
	coverage	93.6%	93.8%
Accommodation	PeopleSize	86.8%	87.0%
	BodyBuilder	86%	86%

moderate due to the relatively high correlation of any added dimension with stature, sitting height or waist circumference (which between themselves have low correlation), i.e. the major reduction in accommodation has already been made.

Figure 2 shows skin compositions of all members in the RAMSIS Typology (black) and the A-CADRE (grey) merged into one CAD geometry to illustrate differences in the volume that the two manikin families occupy in standing postures. Their female counterparts are presented in (Högberg and Case, 2005).

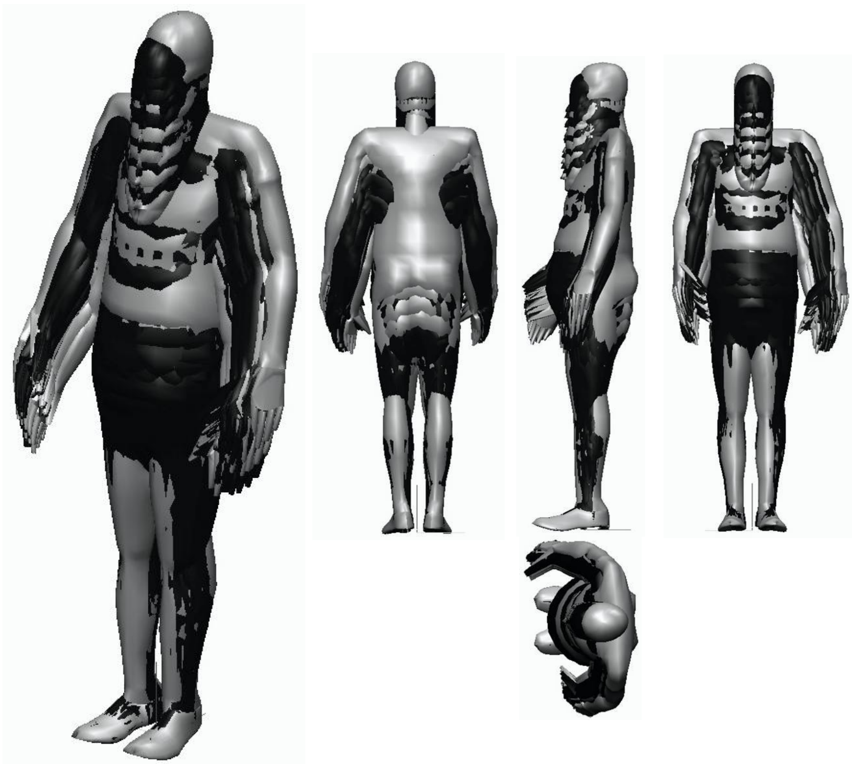


Fig. 2. Skin compositions of RAMSIS Typology (black) and A-CADRE (grey) male manikin families in standing posture

2.1 Context and Details of Study

The study is based on a fixed mid eye study for occupant packaging of the driver in a car. The approach of starting the accommodation task from a fixed eye point is common in aircraft cockpit design (Roebuck et al., 1975) and has been used in some concept cars, e.g. the Volvo Safety Concept Car (VolvoCars, 2004). This is a sensible and likely approach for future cars and requires adjustable pedals/floor, controls, steering wheel and seat.

The constraints are set to keep all manikins' mid eye points (a point right between the eyes) in a fixed position and the line of sight 5 degrees down from the horizontal line. Besides these restrictions, the manikin selects the most comfortable position according to the posture prediction functionality in RAMSIS.

3 Results

Figure 3 show the simulation results for the locations of the steering wheel grasping points, the heel and pedal points as well as the hip points (H-points). These are shown

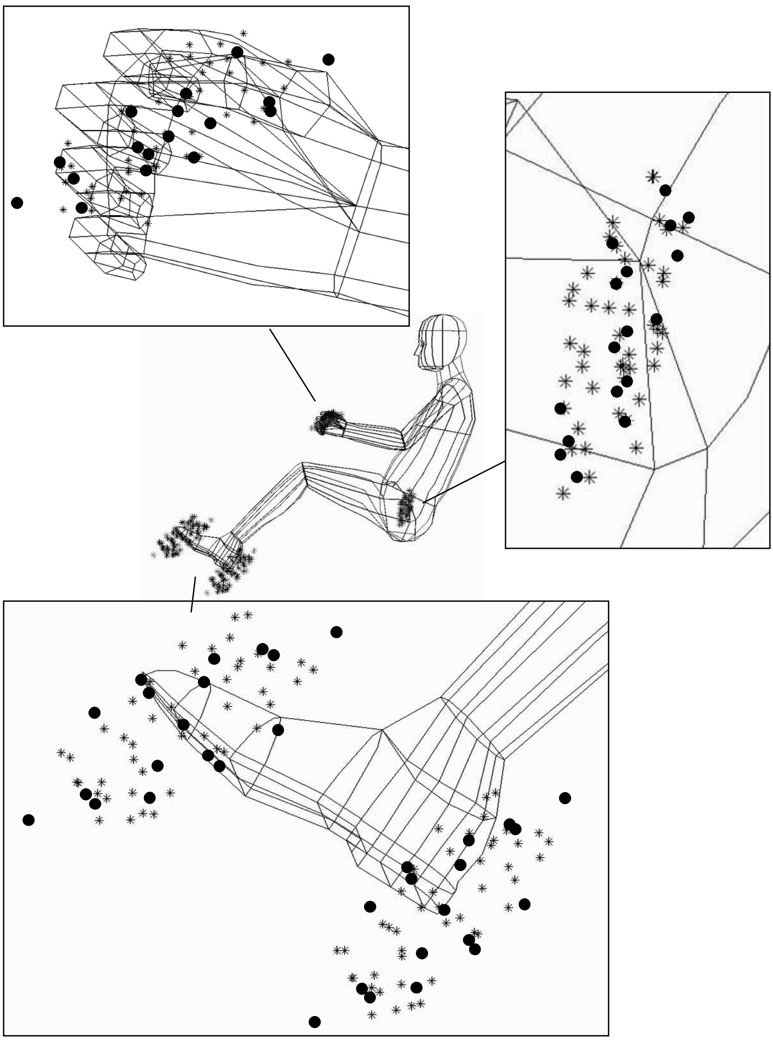


Fig. 3. Body point locations, RAMSIS Typology (Black stars) and A-CADRE (Black dots), males

for each male manikin in both families to enable visual comparison (RAMSIS Typology locations shown as black stars and A-CADRE locations shown as black dots). These points generate interior component adjustment areas required to accommodate drivers in the defined user group (as shown in Figure 4 and 5).

Figure 4 shows adjustment area dimensions generated by using the RAMSIS Typology family, where both genders are included. Figure 5 shows the corresponding results from using the A-CADRE family, both genders. The adjustment area results are compared with a similar fixed eye point study by Vogt et al. (2005) that used a variant of the RAMSIS Typology family. These values are specified in brackets in the figures.

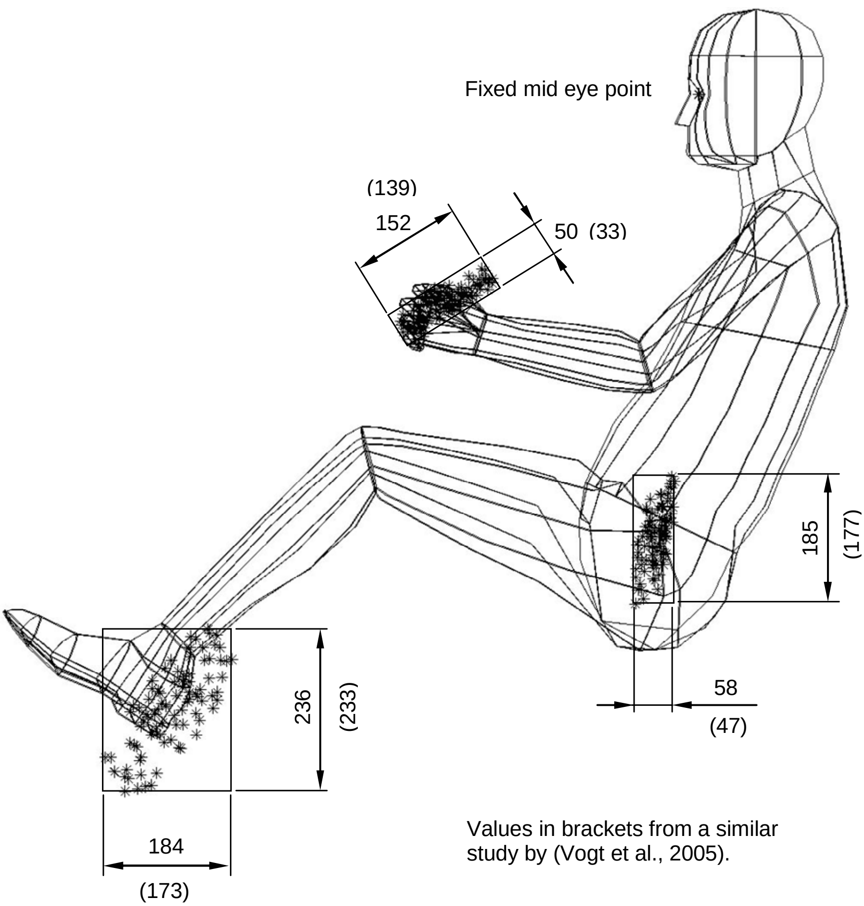


Fig. 4. Adjustment area dimensions obtained, RAMSIS Typology, males and females

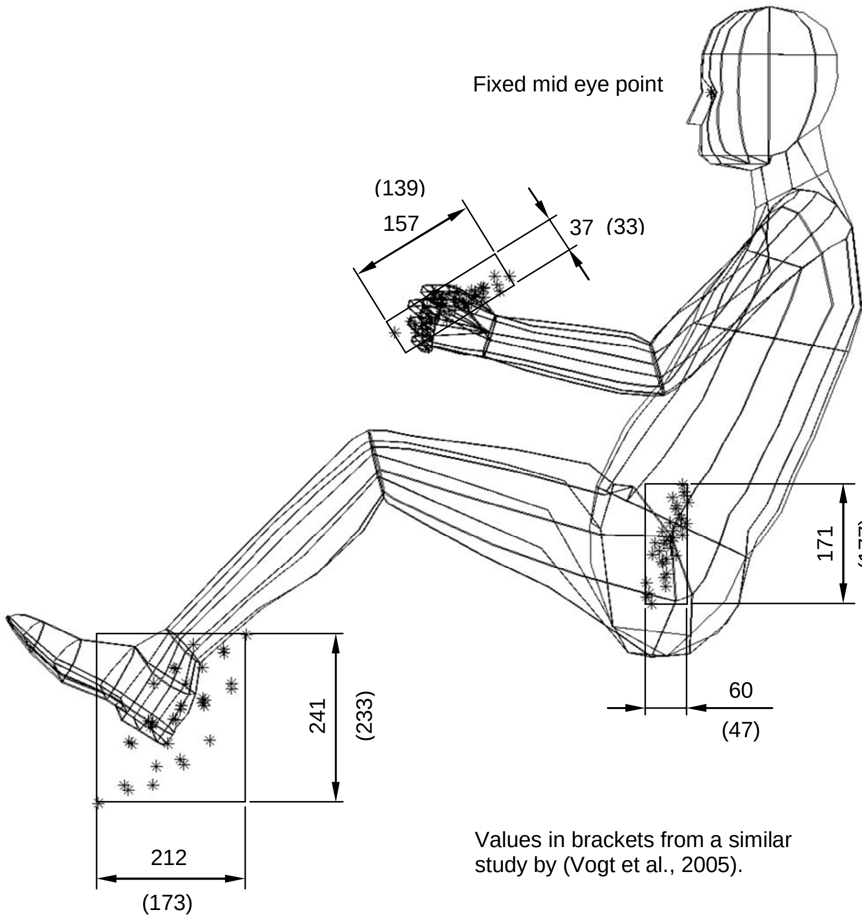


Fig. 5. Adjustment area dimensions obtained, A-CADRE, males and females

4 Discussion

The results indicate that larger adjustments areas are required compared to the study by Vogt et al (2005). This is however expected because of the larger anthropometric coverage in the RAMSIS Typology and the A-CADRE based families which were used (Table 1). For example, the tallest person in Vogt's study is 1897 mm, whereas the tallest persons in RAMSIS Typology and A-CADRE families were 1909 mm and 1947 mm respectively. The RAMSIS Typology seems to represent corpulence differently than A-CADRE. Table 1 shows similar coverage of corpulence but Figure 2 indicates that the RAMSIS Typology represents corpulence to a higher degree (hence the black abdomen), particularly for shorter persons. This may be an effect of sources of anthropometric data when creating the manikin families, and particularly of correlations between body measurements. The RAMSIS Typology is mainly based on large anthropometric surveys done in Germany by measuring civilians (Flügel et al., 1986; Human-Solutions, 2004), whereas the A-CADRE is mainly based on US Army personnel data (Bittner, 2000). It is likely that people represented in the US study are on average fitter than people in the German study, and hence that the RAMSIS Typology manikin family more accurately represents corpulence of common people. Both manikin family approaches embody human diversity in a credible way. Even though it is hard to draw major conclusions from this study, it is worth emphasising that A-CADRE gave these results by 62% fewer simulations required (34 compared to 90). A boundary cases selection approach is appropriate for the kind of design problem presented in this paper (Dainoff et al., 2003). For other types of design tasks a distributed cases method is more appropriate, e.g. for clothing design (Dainoff et al., 2003). In this sense the RAMSIS Typology approach is more general since it includes cases (manikin descriptions) within a distribution as well as on the boundary, whereas in A-CADRE all manikin description are on the boundary, except the centroid case.

From a general perspective the study is used as an illustration of DHM tools becoming everyday tools for designer. Supporting designers with a DHM tool that has appropriate functionality and usability makes the complexity involved in considering anthropometric variation in multivariate design problems less demanding. Enabling the human simulation tool user to see and operate the product user as well as the product modelled in the same virtual environment means that human-product interaction issues are more easily considered concurrently with other design issues, thereby supporting the synthesis work that is characteristic of design. The implementation of pre-defined manikin families in human simulation tool aids the tool user to consider anthropometric diversity rather straightforwardly even though not being required to know the problems in detail or the theory behind the manikin family, but rather putting effort into making sure that all manikins are accommodated by the design. Several sources report on problems and deficiencies in considering ergonomics when designing products and workstations, and how this puts pressure on the usability of the ergonomics information available to designers, e.g. (Haslegrave and Holmes, 1994; Chapanis, 1995; Burns et al., 1997; Rouse and Boff, 1998). DHM tools have the potential to act as a channel for conveying ergonomics information to designers, e.g. to support understanding and interest during the explorative, generative and evaluative activities of the design process. Careful evaluation of proposed solutions is important in order to identify and assess feasible solutions that balance all applicable requirements in a good way. Evaluation can be considered as being performed in the 'small' or 'big'. Evaluation in the 'small' is meant to convey the almost continuous evaluation of generated ideas that are performed by the designer, or the design team, as portrayed by the evaluation segment (Ev) of the 'wheel' in Figure 6. The figure is inspired by design process representations in (Pahl and Beitz, 1988) and (Cross, 2000). An evaluation in the 'big' would represent evaluations carried out more rarely, but possibly more thoroughly, and likely to also involve people other than the design engineers, e.g. specialists and managers. Typically such evaluations are done when progressing from one design process stage to the next, e.g. from conceptual design to embodiment design (Figure 6). However, although these 'big' evaluations may discover deficiencies of some sort, time and cost pressures frequently hinder the alteration of the design in any major way. This means that the required design iteration is rejected, but regularly the design process continues. Hence, there may be great process advantages of supporting designers with appropriate DHM tool functionality when performing evaluations in the 'small'. These are alleged to be particularly valuable within active conceptual design process phases, where important choices and iterations are made for finding the best overall balance of the array of product requirements. However, this puts demands on the usability of the tools, particularly for non-specialist users, who also may use the tools infrequently. In addition, there is a need for a structured process when using DHM tools, and a need for integrating the tools in the general product development processes of companies (Hanson et al., 2006).

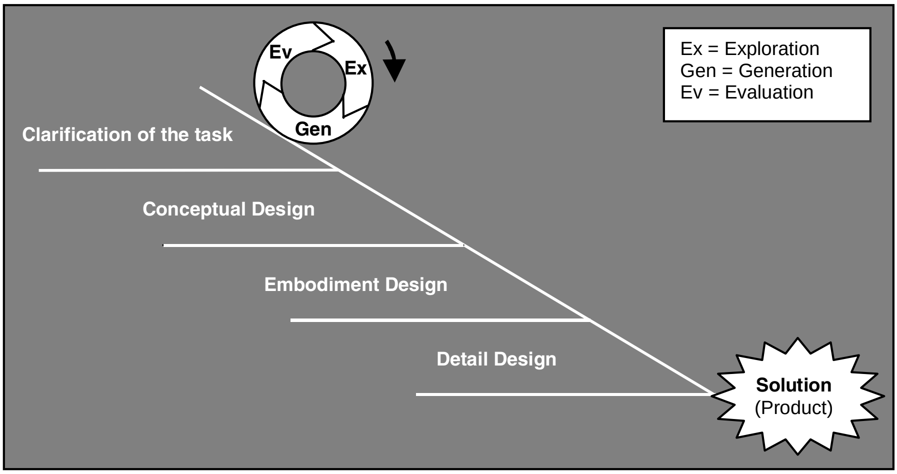


Fig. 6. Design process representation for illustrating evaluation in the 'small' and the 'big'

References

Bittner, A.C.: A-CADRE: Advanced family of manikins for workstation design. In: XIVth congress of IEA and 44th meeting of HFES, San Diego, 2000, pp. 774–777 (2000)

Bittner, A.C., Glenn, F.A., Harris, R.M., Iavecchia, H.P., Wherry, R.J.: CADRE: A family of manikins for workstation design. In: Ashfour, S.S. (ed.) Trends in ergonomics/human factors IV, pp. 733–740. Elsevier Science Publishers, Amsterdam (1987)

Bubb, H., Engstler, F., Fritzsche, F., Mergl, C., Sabbah, O., Schaefer, P., Zacher, I.: The development of RAMSIS in past and future as an example for the cooperation between industry and university. Int. J. of Human Factors Modelling and Simulation 1(1), 140–157 (2006)

Burns, C.M., Vicente, K.J., Christoffersen, K., Pawlak, W.S.: Towards viable, useful and usable human factors design guidance. Applied Ergonomics 28(5/6), 311–322 (1997)

Chapanis, A.: Ergonomics in product development: a personal view. Ergonomics 38(8), 1625–1638 (1995)

Cross, N.: Engineering design methods: strategies for product design. John Wiley & Sons, Chichester (2000)

Dainoff, M., Gordon, C., Robinette, K. M., Strauss, M.: Guidelines for using anthropometric data in product design. Human Factors and Ergonomics Society, Santa Monica (2003)

Flügel, B., Greil, H., Sommer, K.: Anthropologischer atlas: grundlagen und daten(in German) Verlag Tribüne, Berlin (1986)

Hanson, L., Blomé, M., Dukic, T., Högberg, D.: Guide and documentation system to support digital human modeling applications. International Journal of Industrial Ergonomics 36, 17–24 (2006)

Haslegrave, C.M., Holmes, K.: Integrating ergonomics and engineering in the technical design process. Applied Ergonomics 25(4), 211–220 (1994)

Human-Solutions, BodyBuilder version 1.4. Human Solutions GmbH, Kaiserslautern, Germany (2003)

Human-Solutions, Anthropometric data bases: overview. Human Solutions GmbH, Kaiserslautern, Germany (2004)

Högberg, D., Case, K.: The impact of manikin family configuration on accommodation. Nordic Ergonomics Society conference, Oslo, Norway. pp. 91–95 (2005)

Kroemer, K.H.E., Kroemer, H.B., Kroemer-Elbert, K.E.: Ergonomics: how to design for ease and efficiency. Prentice Hall, Upper Saddle River (2001)

Pahl, G., Beitz, W.: Engineering design: a systematic approach. The Design Council, London (1988)

PeopleSize: PeopleSize 2000 Professional. Open Ergonomics Ltd., Loughborough (2004)

Robinette, K.M., Hudson, J.A.: Anthropometry. In: Salvendy, G. (ed.) Handbook of human factors and ergonomics, 3rd edn., pp. 322–339. John Wiley & Sons, Hoboken (2006)

Roebuck, J.A., Kroemer, K.H.E., Thomson, W.G.: Engineering Anthropometry Methods. John Wiley & Sons, New York (1975)

Rouse, W.B., Boff, K.R.: Packaging human factors for designers. Ergonomics in Design 6(1), 11–17 (1998)

Seidl, A.: RAMSIS - A new CAD-tool for ergonomic analysis of vehicles developed for the German automotive industry. Warrendale, Society of Automotive Engineers. SAE Technical Paper 970088 (1997)

Vogt, C., Mergl, C., Bubb, H.: Interior layout design of passenger vehicles with RAMSIS. Human Factors and Ergonomics in Manufacturing 15(2), 197–212 (2005)

VolvoCars: Volvo Safety Concept Car, Volvo Car Corporation (2004) (Access: 13 February 2007) www.volvocars.com/corporation/Safety/yourfuture/VolvoSafetyConceptCar.htm

Ziolek, S.A., Kruithof, P.: Human modeling & simulation: a primer for practitioners. In: XIVth congress of IEA and 44th meeting of HFES, San Diego, 2000, pp. 825–827 (2000)

Ziolek, S.A., Wawrow, P.: Beyond percentiles: An examination of occupant anthropometry and seat design. Society of Automotive Engineers, Warrendale, SAE Technical Paper 2004-01-0375 (2004)

Comparison of Human and Machine Recognition of Everyday Human Actions

Trevor D. Jones[1], Shaun W. Lawson[1], David Benyon[2], and Alistair Armitage[2]

[1] Lincoln Social Computing (LiSC) Research Centre,
Department of Computing and Informatics,
University of Lincoln,
Brayford Pool, Lincoln, UK. LN6 7TS
{tjones,slawson}@lincoln.ac.uk
[2] School of Computing,
Napier University, 10 Colinton Road,
Edinburgh, UK, EH10 5DT
{d.benyon,a.armitage}@napier.ac.uk

Abstract. The research presented here makes a contribution to the understanding of the recognition of biological motion by comparing human recognition of a set of everyday gestures and motions with machine interpretation of the same dataset. Our reasoning is that analysis of any differences and/or correlations between the two could reveal insights into how humans themselves perceive motion and hint at the most important cues that artificial classifiers should be using to perform such a task. We captured biological motion data from human participants engaged in a number of everyday activities, such as walking, running and waving, and then built two artificial classifiers (a Finite State Machine and a multi-layer perceptron artificial neural network, ANN) which were capable of discriminating between activities. We then compared the accuracy of these classifiers with the abilities of a group of human observers to interpret the same activities when they were presented as moving light displays (MLDs). Our results suggest that machine recognition with ANNs is not only comparable to human levels of recognition but can exceed it in some instances.

Keywords: Neural network, finite state machine, moving light display, human biological motion.

1 Introduction

Gunnar Johansson [1] first illustrated how humans are skilled at visually analysing and recognising human motion from very sparse and impoverished datasets – or moving light-displays (MLDs). Since Johansson's pioneering work a great deal of literature has appeared on the subject of so-called biological motion – though an exact understanding of how humans understand MLDs has yet to be reached. In parallel with such work, the computer vision community has produced a wealth of approaches to segmentation of the spatio-temporal information held in video images as well as the classification of the feature sets determined therein (for reviews of such work see, for

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 120–129, 2007.

instance, Gavrila [2] and Essa [3]). In general, the range of approaches adopted has varied enormously depending on the application domain and its constraints. Many different pattern recognition methods have been applied to the problem including artificial neural networks (ANNs), statistical classifiers and rule based systems. Recently some studies have appeared which attempt to exploit what is known about human recognition of biological motion to inspire the development of autonomous machine recognition of the same phenomena (e.g. [4]). However, even here there is debate over whether a model based approach to the problem, which uses prior knowledge, such as a model of physical make-up of a human body, to assist classification should be favoured over classification of 'raw' spatio-temporal data.

The research presented here makes a contribution to the understanding of the recognition of biological motion by comparing human recognition of a set of everyday gestures and motions, presented as MLDs, with machine interpretation of the same dataset. Our reasoning is that careful analysis of any differences and/or correlations between the two could reveal insights into how humans themselves perceive motion and hint at the most important cues we use for this. For instance if a machine classifier can accurately recognise an action as well as a human, without any programmed knowledge of a human model or any other background or information (or context), then we could assume that we only need the information within the MLD to interpret the action. The motivation for our work is to try and understand how we might build interactive computer systems that can simulate a natural understanding of biological motion. In particular, and in the longer term, we are interested in seeking an understanding of how computer systems might be used to detect subtle changes in biological motion which indicate changes, for instance, in a human's health or emotional state.

The following section briefly describes some of the artificial classifier techniques used to classify human biological motion and problems arising from popular techniques. This is followed by an account of human biological motion in the form of MLDs. We then present our experimental work, results and conclusions.

2 Classification Techniques

Artificial neural networks are a popular means of classifier, they are well suited for appearance based representations and applications that benefit from fast response times [5]. ANNs are trained from data samples without the requirement for explicit data modelling using a weighting algorithm [6], and do not require large memory storage. The most commonly used neural network is the multiplayer perceptron [7] though other networks such as SOM (Self Organising Maps), or Kohonen map, are favoured for visual recognition systems [8].

Rule based classifiers such as finite state machines are a means of tracking transitions through sequences known as states. Each state indicates a possible path of execution reflecting input changes within the system with a transition in state occurring once a condition has been fulfilled [9]. The finite state machine (FSM) segments events into key stages that represent events within the system. It is useful for systems that have a definable flow and change according to triggered events. The diverse ways and lack of uniformity in human movement, coupled with possible signal noise, is enough to prevent a state being activated. These findings were an issue for [10] who use an FSM to track hand and head gestures, though the FSM approach does fair well with movements that are less complex and have little diversity [11].

Statistical models such as the HMM (Hidden Markov Model) have some limitations in representing non-gesture patterns [12], but are useful for modelling spatio-temporal variability where the data is temporally well aligned. Similar to the finite state machine, HMMs use state transition but with probabilistic algorithms. Another way of distinguishing behaviour from movement is using a database of possible movements for each action [13]. This type of system is reliant on predefined constraints that would require the creation of extensive profile sets for each participant. Probabilistic networks have been used with some success with recognising small movements such as those performed with the head [14] but have had improved accuracy recognising human activity when used in conjunction with a hidden markov model [15].

The problems with recognising gesture are numerous: pattern spotting [16] – locating meaningful patterns from an input data stream. Segmentation ambiguity [17] - where does the gesture begin and where does the gesture end? Spatio-temporal variability [18] – gestures vary dynamically in shape and duration, be it a large number of people or just one person making the gestures. Vision based systems suffer from their poor ability to continuously track the object of focus, mainly the hands and face. When the focus is lost, perhaps through tracking error or occlusion, the system must be recalibrated requiring the object of focus (the person) to stand in a default pose for the system to once again acquire tracking [19]. A vision based approach also suffers from location limitation – it can only track within its field of vision. It does however, allow for unencumbered freedom for those that are being tracked. A sensor based approach, on the other hand, does not suffer from occlusion, lighting conditions and other visual constraints; it can track a person's movement continuously. However, sensors are invasive and can be cumbersome.

3 Human Biological Motion and Moving Light Displays

Humans can quickly detect other humans within view and, in most cases, can determine biological aspects such as their sex, approximate age, posture, gait, weight, height and build. We recognise social significances that lead us to approximate health, strength, physical ability, and from a persons bearing can make assumptions about intent: aggressive, placid, gentle, untrustworthy, trusting etc. Human motion contains a wealth of information: actions, traits, emotions, and intentions. Our ability to extract complex visual information from the perceived world has been widely documented [20] and theorems for understanding environmental and contextual cues. Gunnar Johansson demonstrated the ability of humans to visually analyse and recognise human motion from very sparse and impoverished datasets. Information of the human form, as represented by an MLD, is considerably reduced (Fig 1). When an MLD is stationary the information presented is near meaningless, but when the MLD is in motion an observer is able to perceive the definition of a human form [1].

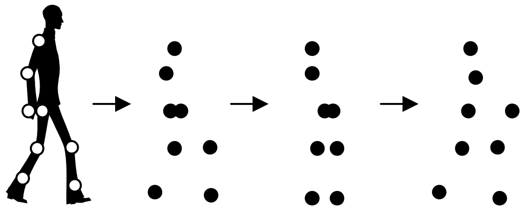


Fig. 1. An MLD representing a person walking

Experiments using MLDs have shown that humans are able to, within some degree, recognise not only the movement but also whether the form is male or female [21] and even to recognise themselves and people familiar to them [22]. Runeson and Frykholm [23] attempted to disguise the gender of the person in the point light display by asking actors to perform as though they were of the opposite sex; observers guessed the correct gender of the actor over 85% of the time. Runeson and Frykholm also showed actors throwing sandbags out to varying distances, the actor was fitted with the lights but the sandbags were not. Observers were good at judging how far the bags would have been thrown.

It is not only humans that can recognise biological motion in MLDs. In their experiment with newly hatched chicks Vallortigara et al [24], showed animated images of 'point light hens' to chicks which would be the first visual stimuli they would encounter after hatching. The chicks followed the motion of the point light hens showing that they perceived the motion. They then performed further tests to answer whether the chicks' response was innate or a learned experience. This time they used artificially induced motion patterns but the chicks were still drawn to the biological motion of not just hens, but also that of cats. Other types of models were used which used rigid structures to define body shape or models which represented a hen-like object, but again the chicks were attracted to the biological motion of the point light hen. As a control, they showed the chicks point light models of both hens and cats, for which the chicks were just as likely to approach the cats as they were the hens. They suggest that from these results chicks have evolved a predisposition to notice objects that move like vertebrates, which may maximise the probability of imprinting on the object most likely to provide food and protection after birth. They refer also to similar findings in four month old human babies and conclude that the results suggest that this preference is hard wired in the vertebrate brain.

4 Experimental Procedure

In this work our intention was to compare machine interpretation of biological motion with human interpretation of the same data. We captured biological motion data from human participants engaged in a number of everyday activities, such as walking, running and waving, and then built two artificial classifiers (a Finite State Machine and a multi-layer perceptron artificial neural network, ANN) which were capable of discriminating between activities. The intention was then to compare the accuracy of these classifiers with the abilities of a group of human observers to interpret the same activities. A sensor based approach was chosen over a visual recognition system to capture human motion, which took place in a small gymnasium. To capture walking and jogging actions the participants were required to walk and jog on a commercial running machine. Punching was performed by hitting a punchbag and throwing was performed by throwing tennis ball sized balls. Data from typical everyday full-bodied human behaviours were collected using a commercial sensor system (Polhemus Liberty) usually used for motion-capture animation work [25]. We used just five sensors – forty human participants were recruited and sensors placed on: each hand, each foot and the forehead. Participants were asked to perform the following motions: walking, jogging, shaking-hands, waving, punching, kicking, throwing a ball, looking left/right, looking up/down. During each motion, spatial (x,y,z) data was recorded for each participant. Our resultant dataset was therefore very rich in terms of the breadth of examples we recorded – however we deliberately kept the content of each recorded instance of motion very sparse and impoverished – with only 5 points recorded. Two commonly used autonomous classifiers were then trained and configured to classify the data, a Finite State Machine (FSM) and an Artificial Neural Network (ANN). The data is then converted into MLDs for classification by forty human participants.

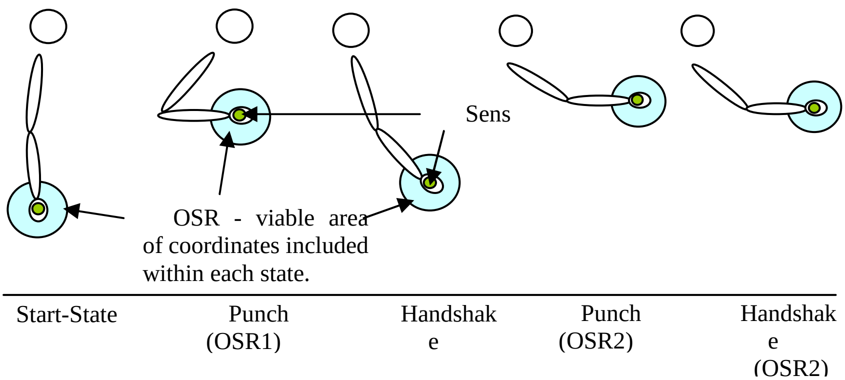


Fig. 2. FSM Optimal state ranges (OSR) for a punch and a handshake

4.1 Finite State Machine

Firstly, an FSM was configured to classify actions; this method hence combined both human and autonomous data perception (since human knowledge of the raw sensor data was incorporated into the design of the FSM). The FSM was designed to calibrate for each participant e.g. measure the length of the participants arm and setting the distance value for the states. For example, if the states were calibrated for a person with long arms the state activation area would be too far to away to be activated by a person with shorter arms. The FSM has four states: a start state, a predictive state, a hit state and a null state. The first three states comprise of an OSR (optimal state range): a spatial three dimensional sphere-shape in space which moves with the participant illustrated in Fig 2.

4.2 Artificial Neural Network

Secondly, a multi-layer-perceptron ANN was trained to classify the raw sensor data - i.e. no model of a human, or any other prior real world knowledge, was integrated into the training phase. The capture data, for all gestures, was re-sampled using linear duplication to make it temporally aligned and then converted for suitability with Trajan, a commercial ANN application (Trajan, date). The data was pre-process using PCS (Principal Component Analyses). PCA is a linear dimensionality reduction technique, which is able to identify orthogonal directions of maximum variance and project the data into a lower-dimensionality space formed of a sub-set of the highest-variance components [26]. The data was sampled using evenly distributed subsets for test, training and verification.

4.3 Human Classification

Additionally, our motion capture data was converted into MLDs by importing it into a 3D animation package and creating short video clips showing five 3D spheres which move in accordance with the captured data on a plain background. We then recruited a further forty human participants who viewed the simulated MLDs, watching each gesture in turn and stating the full-body gesture that they thought was being exhibited. The classification capabilities of the autonomous systems were then compared to those of the human participants. Motion capture data was converted into MLDs by importing the data into a 3D animation package and mapping spheres to the Cartesian coordinates, a representative sample of which is presented in Fig 3. Forty people from varying disciplines and backgrounds participated in the experiment, male and female ranging from ages eighteen to fifty eight.

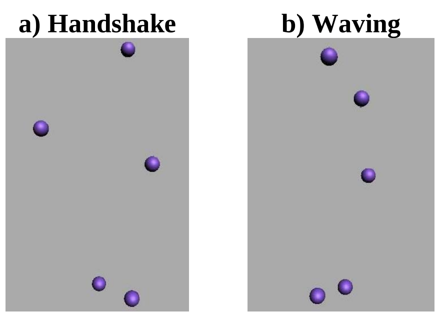


Fig. 3. a) Static image of person b) Snapshot of a person waving

The experiment comprised of three stages: Part 1: a participant viewed a static image of the spheres representing a person (upright and with their arms by their sides). They were asked to give their first impressions of what came to mind when they saw the image. Part 2: The participant was shown an animated MLD for all of the gestures, and for each, was asked to state what they thought they were seeing. Part 3: the participants are informed the spheres represent human motion and are asked to view the MLDs once again. Part 4: All of the MLDs were compiled, in a different order, into one animation medley. Random movement of the spheres was placed between actions signifying other types of movement or ambiguous action. The participants are informed of which animations represent which action and are then shown an animation medley compiled from were asked to say aloud the gesture they recognised.

5 Results

The FSM was able to recognise gestures with less variability of motion than those which can be performed with greater exertion (Fig 4). Due to the rigid OSR structure the FSM will only recognise movement that keeps within a spatial boundary relative to a fixed point. Gestures which have the potential for greater exertion e.g. punching and throwing were often performed beyond the OSR limits. Recognition for walking and jogging was considerably lower than other results mainly as a result of fluctuations in the sensor output signal registering outside the OSR. This has potential for improvement through redefining the experimental design.

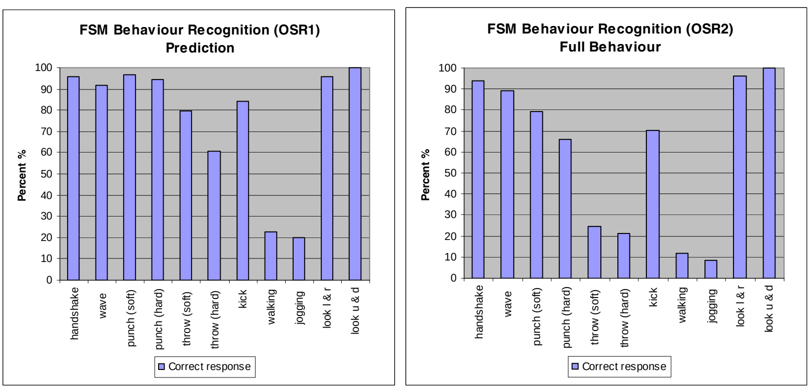


Fig. 4. Results for OSR1 and OSR2 Recognition

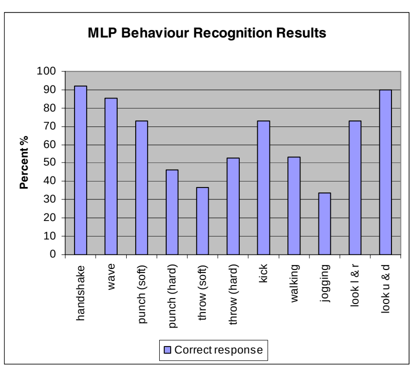
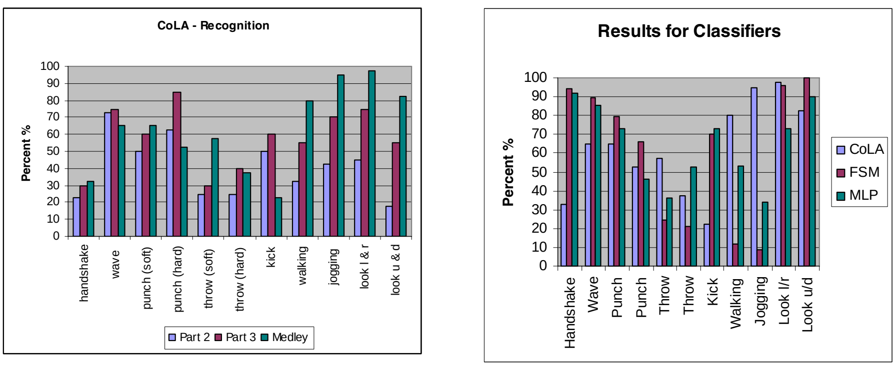


Fig. 5. Results for MLP

The MLP classification performance suffered most of all with actions such as walking, jogging, punching and throwing; for the arm actions there is more variability in these movements than the actions the MLP had less trouble classifying (Fig 5). The walking and jogging may also have posed some difficulty as there was a considerable amount of variance in the data. This variance is due to the participants' feet impacting on the running machine with each stride; the sensor effectively sustains excess motion from vibration. The difference in individual style of performing certain gestures may also be one possible reason for low recognition.

When human participants were shown the static MLD their recognition of the spheres representing a static human was low, 45% in comparison to when the animations were in motion 82%. Johansson makes this observation noting observers rarely recognise the human form in static displays [1]. The results for part three (Fig 6a) would suggest that the increase in some of the recognition results from part two is primarily due to the introduction of partial-context (knowledge the MLD represents human form), and also that of training effect. Recognition is poor for punching and throwing hard, these gestures are quite erratic when performed with pronounced physical amplitude having an adverse affect on the spatial pattern structure. The recognition for handshake was consistently poor; there are a greater number of actions performed in front of the body e.g. shaking hands, turning key in lock, manipulation of objects etc. than those performed at the sides or high up in relation to the torso e.g. waving. The viewing angle of the MLD may have contributed to low recognition of some gestures such as the kick which may need a more prominent perspective. Overall, across all actions, the results of our three classification approaches were: FSM 60.08 %, MLP 64.45 %, human 62.5 % (Fig 6b).



(a) (b)

Fig. 6. (a) Results for parts 2, 3 and animation medley (b) Results for Classifiers

6 Summary of Conclusions

Our results suggest that, at our level of abstraction, machine recognition with ANNs is not only comparable to human levels of recognition but can exceed it in some instances (for instance, in our results, the ANN and FSM showed superior classification of waving and hand shaking). However, we also found that humans were good at interpreting head and some hand gestures but particularly gestures involving the feet. It is suggested that absence of contextual cues is the main reason for the lower performance in human recognition. Additionally, our work allowed the MLP to utilise the full 3D nature of our data whilst humans were only able to view 2D projections of this data as a MLD. However overall, we believe that the finding that autonomous classifiers which make no use of prior real world information (such as a human model) can potentially perform recognition of biological motion as well as a human is a significant finding which warrants further investigation.

References

1. Johansson, G.: Visual perception of biological motion and a model for its analysis. Perception and Psychophysics 4(20), 201–211 (1973)
2. Gavrila, D.M.: The Visual Analysis of Human Movement: A Survey. Computer Vision and Image Understanding 73, 82–98 (1999)
3. Essa, I.A.: Computers Seeing People. AI Magazine 20, 69–82 (1999)
4. Laxmi, V., Carter, J.N., Damper, R.I.: Biologically-Inspired Human Motion Detection. In: Proc of ESANN - European Symposium on Artificial Neural Networks, Bruges, Belgium, 24-26, 2002, pp. 95–100 (2002)
5. Van Laerhoven, K., Aidoo, K.A., Lowette, S.: Real-time Analysis of Data from Many Sensors with Neural Networks. In: 5th IEEE International Symposium on Wearable Computers (ISWC) 2001, vol. 115 (2001)
6. Heidemann, G., Bekel, H., Bax, I., Saalbach, A.: Hand Gesture Recognition: Self-Organising Maps as a Graphical User Interface for the Partitioning of Large Training Data Sets, icpr. In: 17th International Conference on Pattern Recognition (ICPR’04), 2004, vol. 4, pp. 487–490 (2004)
7. Rumelhart, D.E., McClelland, J.: Parallel Distributed Processing. vol. 1, MIT Press, Cambridge, MA (1986)
8. Corradini, A., Gross, M.: A Hybrid Stochastic-Connectionist Architecture for Gesture Recognition. In: International Conference on Information Intelligence and Systems 1999, p. 336 (1999)
9. Hong, P., Turk, M., Huang, T.S.: Constructing Finite State Machines for Fast Gesture Recognition. In: ICPR ’00: Proceedings of the International Conference on Pattern Recognition 2000, IEEE Computer Society (2000)
10. Hong, P., Turk, M., Huang, T.S.: Gesture Modelling and Recognition Using Finite State Machines. In: IEEE, Fourth International Conference on Automatic Face and Gesture Recognition 2000, pp. 28–30 (2000)
11. El Kaliouby, R., Robinson, P.: Real Time Head Gesture Recognition in Affective Interfaces. Human-Computer Interaction Journal ’03, 950–953 (2003)
12. Lee, H., Kim, J.H.: An {HMM}-Based Threshold Model Approach for Gesture Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 21, 961–973 (1999)
13. Emering, L., Boulic, R., Thalmann, D.: Interacting with Virtual Humans through Body Actions. IEEE, Computer Graphics and Applications 18, 8–11 (1998)
14. Madabhushi, A., Aggarwal, J.K.: A Bayesian Approach to Human Activity Recognition. In: Second IEEE Workshop on Visual Surveillance 1999, p. 25 (1999)

15. Sun, X., Chen, C., Manjunath, B.S.: Probabilistic Motion Parameter Models for Human Activity Recognition. In: Proceedings of IEEE International Conference on Pattern Recognition (ICPR) Québec City, Canada,2002 (2002)
16. Rose, R.C.: Discriminant Word Spotting Techniques for Rejection Non-Vocabulary Utterances in Unconstrained Speech. In: Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing, San Francisco, 1992, vol. II, pp. 105–108 (1992)
17. Takahashi, K., Seki, S., Oka, R.: Spotting Recognition of Human Gestures from Motion Images Technical Report IE92-134, The Inst, of Electronics, Information, and Comm. Engineers, Japan, pp. 9–16 (1992)
18. Baudel, T., Beaudouin-Lafon, M.: CHARADE: Remote Control of Objects Using Free-Hand Gestures. Commun. ACM 36(7), 28–35 (1993)
19. Nickel, K., Stiefelhagen, R.: Pointing gesture recognition based on 3D-tracking of face, hands and head orientation. In: 5th international conference on Multimodal interfaces, 2003, pp. 140–146 (2003)
20. Desolneux, A., Moisan, L., More, J.: Computational gestalts and perception thresholds. Journal of Physiology - Paris 97, 311–324 (2003)
21. Kozlowski, L.T., Cutting, J.E.: Recognising the gender of walkers from point-lights mounted on ankles: Some second thoughts. Perception & Psychophysics 23, 459 (1978)
22. Kozlowski, L.T., Cutting, J.E.: Recognising the sex of a walker from a dynamic point-light display. Perception & Psychophysics 21, 575–580 (1977)
23. Runeson, S., Frykholm, G.: Kinematic specifications of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. Jounal of Experimental Psychology, General 112, 585–615 (1983)
24. Vallortigara, G., Regolin, L., Marconato, F.: Attraction to Motion. PLoS Biol. 3(7) (2005)
25. Polhemus, Liberty (2003) http://www.polhemus.com/
26. Bishop, C.: Neural Networks for Pattern Recognition. University Press, Oxford (1995)

A New Decoding Algorithm in MIMO-ZP-OFDM Systems

Rui Kang, ChaoXiong Xu, HaoChen, and HongBo Xu

Department of Electronics and Information, Huazhong Normal University,
Wuhan 430079,China

Abstract. In order to adapt for a high-rate transmission and improve the performance of operation, in this paper, a new low complex sphere decoding (SD) algorithm is proposed in the MIMO-ZP-OFDM systems. It is an approximate optimization algorithm and can be used in the space-time coding and uncoded multiple-antenna systems. ML sequence detection compared with SD algorithm the latter can reduce the complex and keep the performance of systems, especial for high-rate transmission operation and the occasion of transmit antenna beyond receive antenna. Simulation results show that the efficiency and superiority of this algorithm.

Keywords: OFDM, MIMO, sphere decoding, ZP.

1 Introduction

In recent years, “wireless and wideband” have become a hot topic of research in wireless communication. The MIMO-OFDM technique greatly improves the peak margin of noise, interference, and multipath by applying array antennas to the realization of space diversity in the transmission system. However, the optimal maximum likelihood (ML) detection incurs prohibitively high complexity and is not suitable for practical implementation, suboptimal detection algorithms are usually employed. A lot of efforts have been put into the search for algorithms achieving ML or near-ML performance with lower complexity. Recently, a new algorithm, sphere decoding, has been proposed to be applied to multiple-antenna system the sphere decoding algorithm thought was firstly proposed in 1981 as a method of find the shortest vector for lattice [1]. In 1993, E. Viterbo and J.Bountros first applied this method to communication problem in an essay on trellis decoding[2]. Sphere decoding was an algorithm that speeds up ML algorithm by limiting the region of search. It[3]-[5] can achieve an exact maximum likelihood (ML) performance with a moderate complexity by confining the search space inside a hypersphere with a given radius, centered at the received signal vector. Sphere decoding was originally designed for lattice decoding, and lately it has been shown to be applicable to space-time coded MIMO systems, resulting in lower decoding complexity. The sphere decoding algorithm was introduced in [6] assuming a single-antenna, real-valued fading channel model. Later results [7-8] generalized the algorithm to complex valued MIMO channels. A reduced complexity algorithm was proposed in [9], where the signal coordinates were sorted according to their partial metrics and explored in this order. In [10], the sphere decoding algorithm was applied to equalize frequency selective MIMO

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 130–135, 2007.

channels. All of these works considered uncoded MIMO systems and assumed quasi-static, flat fading channels. At present, sphere decoding has been widely applied to MIMO-OFDM system as well as the MIMO system[11], but mainly on the condition that the transmit antenna and the receive antenna are the same , or the receive antenna is greater than the transmit antenna[12]. Based on this, this paper proposes adopting an improved thought of sphere decoding which effectively induces the complicity of decoding and fits the condition that transmit antenna is greater than receive antenna.

In this work, the rest of this paper is organized as follows. Section II is devoted to the introduction of MIMO-ZP-OFDM system model; section III briefly explains sphere decoding, its improved arithmetic and its application in this model; section IV presents the simulation result while sections V gives the conclusion.

2 System Model

In this MIMO-ZP-OFDM system , the system adopts M_T transmit antennas and M_R receive antennas ($M_T \geq M_R$).The data stream, after constellation mapping, is divided into M_T subchannel, and is turned into OFDM through IFFT operation before being added zero-postfix (ZP) and transmit by their corresponding antenna. Zero-padded OFDM (ZP-OFDM), which prepends each OFDM symbol with zeros rather than replicating the last few samples, has been proposed. The reason that ZP is adopted here is that it is Toeplitz matrix which ensures the reversibility of matrix, and can at the same time ensure the performance of the system as the diagonal matrix of channel transmit function being zero values. On the contrary, if the CP is adopted, when the diagonal matrix of channel transmit function being zero values, thus the performance of the system will decrease. The antennas of transmit get the frequency signal through FFT after ZP is got rid of, then the channel parameter got through channel estimation makes MIMO detection in order to get the information data of the original antenna, and finally restored the original signal through the improved sphere decoding.

We assume that there are L multipath channels and each of channel time delay is l, for $l = 0,1,\cdots,L-1$, τ_l denotes l th delay time. Each OFDM block includes N subchannel and the length of the zero-postfix is greater than the delay spread to eliminate the effect of the ISI. Duration of each OFDM block, the channel characteristic is assumed invariable, and the system is synchronous.

Then the received signal can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$$
$$\mathbf{H}_{M_T,M_R}(f) = \sum_{l=0}^{L-1} h_{M_T,M_R}(l) e^{-j2\pi f \tau_l} \tag{1}$$

For transmitted symbols $\mathbf{x}=[x_1, x_2, \cdots x_{M_T}]^T$ ($[\cdot]^T$ denotes the transpose of a matrix or vector) where $x_i \in I_q^{M_T}$ denotes the transmitted signal vector whose elements are chosen from q-PAM constellation given by Iq={odd integer j|-q+1≤j≤q−1}; $\mathbf{H}$ is an $M_R \times M_T$ matrix; for received signal $\mathbf{y} = (y_1, y_2, \cdots y_{M_T})^T$,where y_j is denotes jth signal; $\mathbf{n}$ is the additive white Gaussian noise (AWGN) matrix, with zero mean and variance σ^2. In this paper, we assumed that antenna is independent. The received signal can be written as

$$\begin{bmatrix} \Re(\mathbf{y}) \\ \Im(\mathbf{y}) \end{bmatrix} = \begin{bmatrix} \Re(\mathbf{H}) & \Im(\mathbf{H}) \\ -\Im(\mathbf{H}) & \Re(\mathbf{H}) \end{bmatrix} \cdot \begin{bmatrix} \Re(\mathbf{x}) \\ \Im(\mathbf{x}) \end{bmatrix} + \begin{bmatrix} \Re(\mathbf{n}) \\ \Im(\mathbf{n}) \end{bmatrix} \tag{2}$$

where $\Re(\cdot)$ is the real part, and $\Im(\cdot)$ is the imaginary part of a complex number. $\hat{\mathbf{y}} = [\Re(\mathbf{y})\Im(\mathbf{y})]^T$ is the $2M_R \times 1$ real-valued received signal vector; $\hat{\mathbf{x}} = [\Re(\mathbf{x})\Im(\mathbf{x})]^T$ is the $2M_T \times 1$ real-valued transmit symbol vector ; and $\hat{\mathbf{n}} = [\Re(\mathbf{n})\Im(\mathbf{n})]^T$ is the $2M_R \times 1$ real-valued noise vector, also

$$\hat{\mathbf{H}} = \begin{bmatrix} \Re(\mathbf{H}) & \Im(\mathbf{H}) \\ -\Im(\mathbf{H}) & \Re(\mathbf{H}) \end{bmatrix}$$

3 Sphere Decoding and Improved Algorithm

The idea of SD is to perform minimization on the same metric, but the search is restricted to the points found within the sphere of radius centered around the received signal. That is to say SD algorithm instead of searching all possible vectors for finding **x** in the optimization problem, sphere decoders search over a hyper-sphere of radius R centered on the received signal vector to find an appropriate solution. In other words, we look for vectors **x** which satisfies $\|\mathbf{y} - \mathbf{Hx}\|^2 \le R^2$. Any time a point is actually found in the sphere, the radius of the search sphere is lowered to the distance of this point to the center. The algorithm is continued until no point is found inside the search sphere. In this paper, we take example for 4-PAM, The principle of SD algorithm is presented in Fig 1.

This paper is based on the fundamental algorithms and signal decoders in MIMO system[13] in the complex low complexly SRRC(Smart Radius Reduction Control) algorithm that propose a new method applicable to the condition that transmit antenna is greater than receive antenna.

Assuming the initialize radius is r, $\hat{\mathbf{x}}$ must satisfy the following inequality as

$$r^2 \ge \|\hat{\mathbf{y}} - \hat{\mathbf{H}}\hat{\mathbf{x}}\|^2 \qquad \hat{\mathbf{H}} = \mathbf{QR} \tag{3}$$

Here, **R** and **Q** are the $2M_T \times 2M_T$ upper triangular matrix and the $2M_R \times 2M_T$ unitary matrix, respectively, obtained by the Q-R factorization of the real-valued channel matrix $\hat{\mathbf{H}}$.

Using $\hat{\mathbf{H}} = \mathbf{QR}$, (3)can be rewritten as

$$r^2 \ge \|\mathbf{Q}^T \cdot \hat{\mathbf{y}} - \mathbf{R} \cdot \hat{\mathbf{x}}\|^2 = \|\mathbf{Q}_1^T \cdot \hat{\mathbf{y}} - \mathbf{R} \cdot \hat{\mathbf{x}}\|^2 + \|\mathbf{Q}_2^T \hat{\mathbf{y}}\|$$

And $\tilde{\mathbf{y}} = \mathbf{Q}^T \cdot \hat{\mathbf{y}}$, the matrices $\mathbf{Q}_1$and $\mathbf{Q}_2$ represent the first $2M_T$ an last $2M_R$-$2M_T$ orthogonal columns of **Q**, then we obtain

$$\|\mathbf{Q}_2^T \hat{\mathbf{y}}\| \le \|\tilde{\mathbf{y}} - \mathbf{R} \cdot \hat{\mathbf{x}}\|^2 \le r^2 \Rightarrow r^2 - \|\mathbf{Q}_2^T \hat{\mathbf{y}}\| \ge \sum_{j=1}^{2M_R} \left| \tilde{y}_j - \sum_{i=j}^{2M_T} r_{j,i} \hat{x}_i \right|^2$$

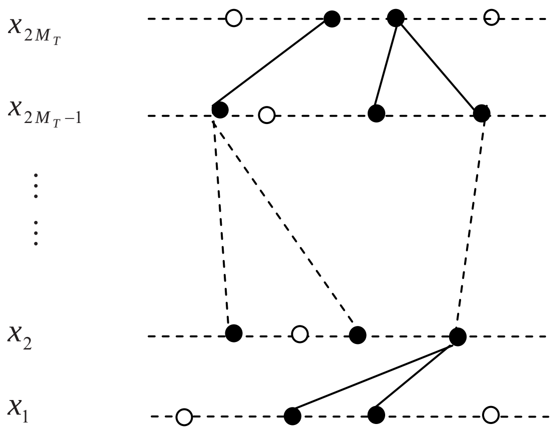


Fig. 1. Example tree formed by sphere decoder

The radius can be reduced by ZF-DFE estimated gradually and simplified calculation purposes.

The new idea of SD[13] proposed a further simplified, in this paper, this method is applied to MIMO-ZP-OFDM systems, and according to [14] made this method improvements. For simplicity, we first consider the 2-PAM constellation where the elements of x take the values from {±1}. Firstly, we partition $\tilde{y}_j - \sum_{i=j}^{2M_T} r_{j,i}\hat{x}_i$ into two subsets $\tilde{y}_j - \sum_{i=j}^{2M_T} r_{j,i}\hat{x}_i \geq 0$ and $\tilde{y}_j - \sum_{i=j}^{2M_T} r_{j,i}\hat{x}_i < 0$ according to their signs, then assumed $e_0^0 = [\tilde{y} - r\hat{x}], e_f^{(i)} = e_0^{(i-1)} + r_i \hat{x}_i \ (i \in [1, 2M_T])$ and solved above two kind of situations, separately. For *i*th detection we intend to find the smallest value $e_r^{(i)}$, making formula to meet $(\tilde{x}_i^v, r^v) = (\arg\min_{x_i^v \in I_q^{(i)}} \|e_r^{(i)}\|_{\tilde{x}_i^v = \tilde{x}_i^v})$, where $e_r^{(i)} = e_0^{(i-1)} + r_i(\hat{x}_i - \tilde{x}_i^v)$. If $r^v < r$, the current initial radius and the current estimate are updated by the new radius and the corresponding estimate, and $e_0^{(i)} = e_0^{(i-1)} + r_i(\hat{x}_i - \tilde{x}_i^v)$.

From the above, we can see that this improvement in the SD algorithm and the original SD methods of further simplifying the iterative steps, which improves the speed of operation that is more applicable to high-speed data transmission system.

4 Simulation Results

In our section, we present our simulation results for both the ML algorithms and for SD algorithms in the MIMO-ZP-OFDM systems. The simulated system has N=128 subcarriers occupying 20MHz bandwidth and all the 64 subcarriers are used for data transmission. we consider a channel of length $l = 6$ and 4-QAM modulations and deal with using 2 transmit and 1 receive antennas and using 4 transmit and 2 receive

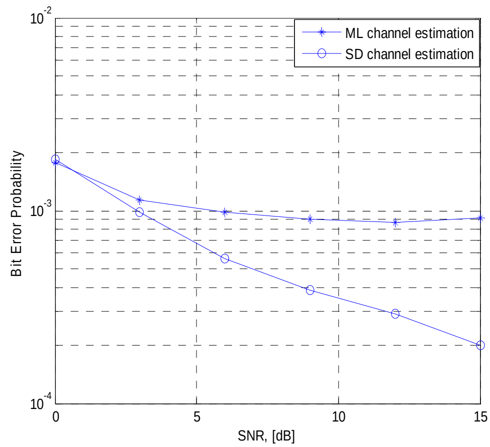


Fig. 2. Comparisons of BER performance according to the different method

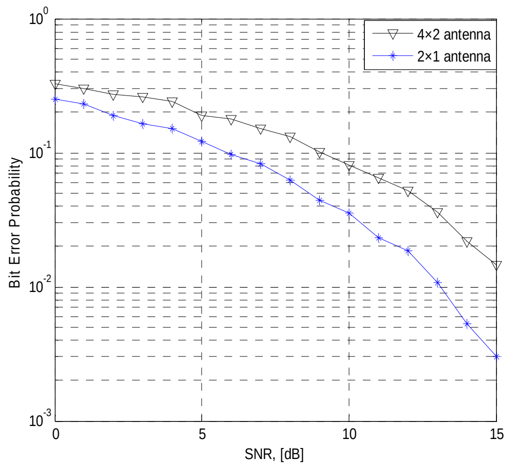


Fig. 3. Comparisons of BER performance according to the different values of antennas

antennas, respectively, the spatial domain channel **H** are assumed to be identically and independently distributed with zero mean and variance σ^2.

In Fig.2 the analytical and simulated BER results of the proposed ML algorithms and SD algorithms, respectively. According to the simulation results can be seen using this improved SD algorithm is superior to ML algorithm to the performance of systems.

In Fig. 3 shows the performance of this system with the different transmit antennas and receive antennas. Through simulation, we can see this improved algorithm applied to the transmitter antennas more than receiving antenna, but for the larger of transmit and the receiver antenna, the performance of system will be affected.

5 Conclusions

This paper proposes a new decoding method to be applied to the MIMO-ZP-OFDM system. This method is an expansion of SD algorithm in multiple-antenna system. It is applicable not only to the condition that the transmit antenna is greater than receive antenna, but also vice versa. The simulation result shows that it has the property of high speed and low complicity, thus is more applicable to the high speed data transmission system.

References

[1] POHSTM: On the computation of lattice vectors of minimal length, successive minima and reduced basis with applications. ACM SIGSAM Bull 15 (1), 37–44 (1981)
[2] VITERBOE, BOUTROSJ: A universal lattice code decoder for fading channel, [C] / /14eme colloq. France 1993, pp. 611–614 (1993)
[3] Damen, M.O., Gamal, H.E., Caire, G.: On maximum-likelihood detection and the search for the closest lattice point. IEEE Trans. Inform. Theory 49, 2389–2402 (2003)
[4] Hassibi, B., Vikalo, H.: On the sphere-decoding algorithm I. Expected complexity. IEEE Trans. Sig. Proc. 53, 2806–2818 (2005)
[5] Vikalo, H., Hassibi, B.: On the sphere-decoding algorithm II. Generalization, second-order statistics, and applications to communications. IEEE Trans. Sig. Proc. 53, 2819–2834 (2005)
[6] Viterbo, E., Boutros, J.: A universal lattice code decoder for fading channels. IEEE Transactions on Information Theory 45(5), 1639–1642 (1997)
[7] Damen, M., Chkeif, A., Belfiore, J.: Lattice code decoder for space-time codes. IEEE Communication Letters 4(5), 161–163 (2000)
[8] Damen, O., Abed-Merain, K., Belflore, J.C.: Generalised sphere decoder for asymmetrical space-time communication architecture. Electronics Letters 36(2), 166–167 (2000)
[9] Chan, A., Lee, I.: A new reduced complexity sphere decoder for multiple antenna systems. In: Proc. of IEEE ICC, 1, 460–464 (2002)
[10] Vikalo, H., Hassibi, B.: Maximum likelihood sequence detection of multiple antenna systems over dispersive channels via sphere decoding. EURASIP Journal on Applied Signal Processing 2002(5), 525–531 (2002)
[11] Damen, O., Chkeif, A.: Sphere decoding of space-time codes. In: Information Theory Proceedings, IEEE International Symposium on 25-30 June 2000, p. 362 (2000)
[12] DAI, Y.: A comparative study of QRD-M Detection and Sphere Decoding for MIMO-OFDM systems. In: PIMRC IEEE 16th International Symposium on September 11-14, 2005, vol. 1, pp. 186–190 (2005)
[13] Han, H.G., Oh, S.K.: A Computationally Efficient Sphere Decoding Algorithm with Smart Radius Reduction Control. In: VTC IEEE 63rd 2006, vol. 4, pp. 1698–1702 (2006)
[14] Yang, Z., Liu, C., He, J.: A New Approach for Fast Generalized Sphere Decoding in MIMO Systems. IEEE Signal Processing Letters 12(1), 41–44 (2005)

Computer Graphic Modeling and Simulation of Human Musculoskeletal System for Biomechanical Research*

Yoon Hyuk Kim

School of Advanced Technology, Kyung Hee University,
Yongin, Kyunggi, 449-701, Korea
yoonhkim@khu.ac.kr

Abstract. In this paper, the human musculoskeletal model based on the medical images of the Korean standard male and the computer graphics based simulation technology with biomechanical analyses are presented. The virtual human model consists of three components. The model component includes the development of anatomically accurate geometric models with physiologically relevant material data. The analysis component includes various types of mechanical. In the simulation component, task-oriented graphic simulation would be performed for virtual evaluation, test, measurement, design, and planning. Three biomechanical analyses with graphic simulation using the virtual Korean model were performed; the surgical simulation of bone fracture surgery, the biomechanical evaluation of surgical implant in knee, and the graphic visualization and simulation of tennis play. In conclusion, the developed virtual Korean model of the musculoskeletal system would have lots of potentiality for biomechanical research and development in various fields, such as medical industry, automobile, ergonomics, or nuclear industry.

Keywords: Biomechanics, Human model, Musculoskeletal, Virtual Human.

1 Introduction

The concept of the virtual human aims to describe the integrated system that can understand the human physiology by providing anatomically accurate geometry and material properties simulation with human-like environmental simulation. This concept contains from molecular scale to the organ and system scales. The necessary technologies to achieve this concept is from medical image modalities, geometric modeling, to the numerical analysis of the human motion and local deformation of the tissues. Recently, the advance of engineering technologies such as high-computing, computer graphic modeling and simulation, and numerical analysis, can provide the ability of virtual human modeling and biomechanical simulations.

There have been several commercial software which can be able to perform numerical analysis of human musculoskeletal model, such as Lifemod®(Biomechanics Research Inc.), SIMM(Motion analysis Inc.), Madymo ® (TNO Automotive Inc.). However, the above commercial software only provide a dynamic analysis module with

* This work was supported by grant No. R01-2006-000-11209-0 from the Basic Research Program of the Korea Science & Engineering Foundation.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 136–143, 2007.

simple modeling tool, so the users would be difficult to optimize the human model for diverse application such as the finite element analysis or the inverse dynamics analysis.

Several countries have already started the national level project to collect the database of the standard human and develop the virtual human model based on the collected data: Biomechanics European Lab in EU, the visible human project in USA, and the digital Korean project in Korea.

In this paper, the development of a virtual Korean musculoskeletal model and application to the biomechanical research are presented. First, computer graphic and biomechanical modeling technologies from medical images were introduced. Then, as examples, the surgical planning bone fracture surgery, the biomechanical analysis of the implant shape in revision total revision surgery, and graphic simulation of tennis play were presented.

2 Materials and Methods

To develop a graphic model of skeleton which contains medical and engineering information, high density computer tomography (CT) images are reconstructed using

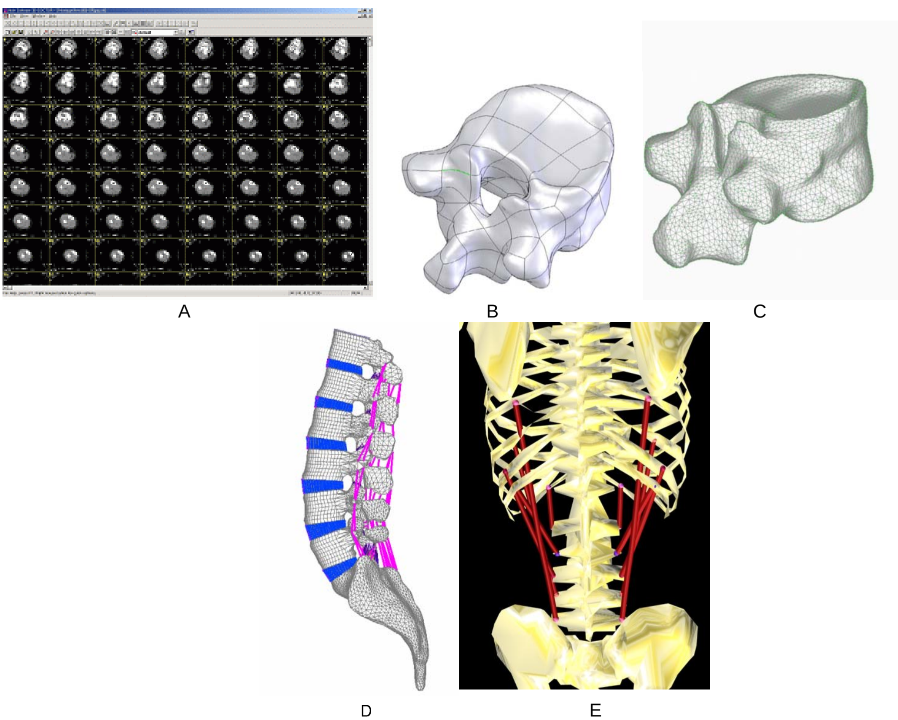


Fig. 1. A. The CT slices are reconstructed to develop the skeletal model. **B.** The solid model of the bone are developed using CT images. **C.** The finite element model developed from the solid graphic model. **D.** The finite element model of a lumber spine with intervertebral discs and spinal ligaments developed from the medical images. **E.** The trunk muscle are defined as piece-wise linear elements between the origin and insertion points of the bones.

two methods. First is a boundary edge detection method with image staking method which is extracting the boundary curve of each image by black-and-white level setting, staking the two curves to obtain the surface, and filling the surface to generate the solid model. The second is so called a voxel method which is generating the hexagonal solid element between the two surfaces. Using the second method with micro-CT images, we can obtain the micro trabecular structure of the bone. In addition, using the image density of each element, mechanical material properties of the element, such as the elastic modulus or the Poisson's ration, can be obtained by experimental relation between the bone mineral density and CT number. These estimated material properties are used for carious numerical analysis such as the finite element analysis.

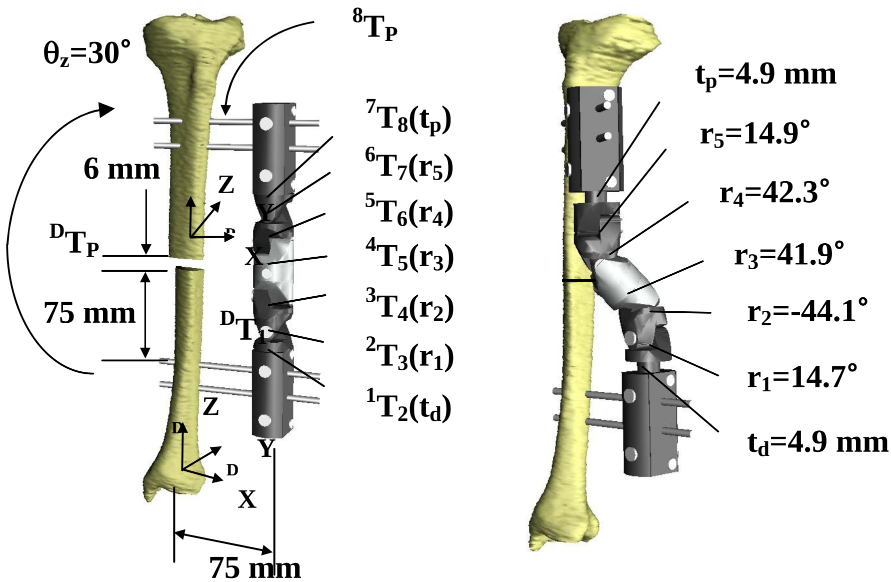
In order to develop the muscle model, there are two approaches have been used. In the first approach, the centroids of the attached and origin area of each muscle were defined as points, and then these two points were connected by piece-wise linear lines or rapping curves to represent the line of muscle force and moment arm bone. In the other approach, the full solid model of the muscle can be reconstructed using MR images. The muscle line approach has been widely used for the dynamics analysis, and the solid modeling approach has been used for the finite element analysis of the soft tissue. In addition to muscle model, ligament and cartilage model can be defined as nonlinear elastic springs and poro-visco-elastic solid materials, respectively.

3 Surgical Simulation of Bone Fracture Surgery

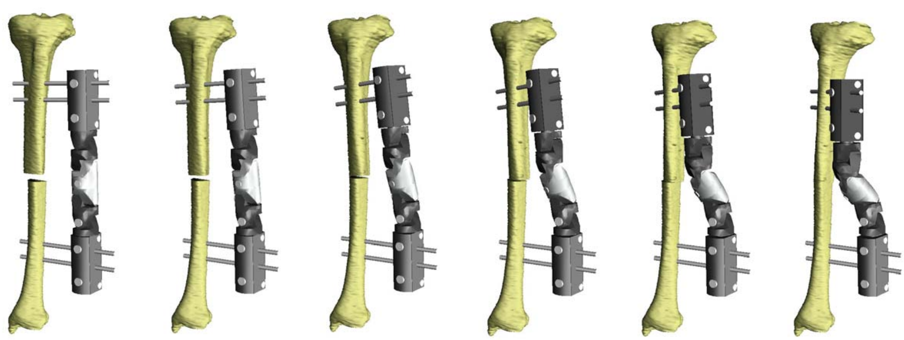
In the clinical situation of fracture treatment, re-aligning the bone segments using external fixator is often necessary to reduce the residual deformities. One of clinical difficulties in external fixation is that the exact amounts of the fixator joints to obtain the perfect alignment have to be determined. In addition, performing the execution based on the pre-operative planning has various clinical limitations depending upon the individual's skill and the configuration of the fixator. If computer simulation based pre-operative planning to reduce the given fracture deformities of the bone segments could be provided, the aligning process of the bone segments would be more accurate and manipulatable, thus obtaining the optimal clinical outcomes of the fracture treatment. The objective of this paper is to develop a computer simulation model of long bone fracture with a unilateral external fixator to perform the pre-operative planning of reducing the given fracture deformity.

A unilateral external fixator for deformity correction (Dynafix®, EBI Inc., USA) was used in this analysis. The fixator has five rotational joints with two prismatic joints. A long bone fracture model of tibia was developed by the reconstruction of CT images. Then, the analytical representation of the motion of each link of the fixator and the bone segments can be expressed as a kinematic linkage system. For the given bone deformity parameters, the necessary fixator joint variables were calculated by nonlinear least square optimization algorithm. In order to visualize the fracture reduction process, the computer graphic simulation based on the analysis results was performed from the initial position to the final reduction position of the fixator-bone system.

As an example, a rotational malalignment of 30° around the bone long axis with a 6 mm fracture gap between the bone ends was simulated. For the given malalignment, the necessary fixator joint values were $\mathbf{r_1}$=14.7°, $\mathbf{r_2}$=44.1°, $\mathbf{r_3}$=-41.9°, $\mathbf{r_4}$=-42.3°, $\mathbf{r_5}$=14.9°, $\mathbf{t_d}$=$\mathbf{t_p}$=4.9 mm. When the joint values from the inverse kinematics analysis was applied to the graphic model, the fracture reduction was completely performed (Fig. 2A).



A



B

Fig. 2. A. The tibia-fixator system for the rotational deformity of 30° and 6 mm gap and the fixator joint values to reduce the given deformity. The graphic visualization showed the calculated joint values could perfectly reduce the fracture deformity. **B.** Simultaneous adjustment of the fixator joints from the initial to the final positions. The unique adjustment path for the bone segments seemed optimal since there was no bony collision and the fracture end displacement was minimal during reduction.

The bone reduction path was also achieved for both fixators on the same bone deformity case when the pair variable values were divided into 100 increments and adjusted simultaneously (Fig. 2B). This reduction process was similar with those performed in the clinics. In addition, during this unique bone reduction path, there was no excessive soft tissue distraction or bone end collision. Therefore, by changing all joints simultaneously in small corrections, the bone seemed to take on an optimal path with no bone end collisions or excessive soft tissue disruptions.

4 Biomechanical Evaluation of Surgical Implant in Revision Artificial Knee Surgery

Stems are frequently revised in total knee arthroplasty (TKA) both to provide additional fixation and to assist in ensuring more consistent component alignment. Many clinical studies have investigated that cemented implant with longer stem length provides better stability and less micro-motion. However, previous studies have reported the incidence of end-of-stem pain on the tibial side when press-fit component with 140 mm stem extension was applied. The problem of pain related to the long stem in total knee arthroplasty is defined poorly and the biomechanical

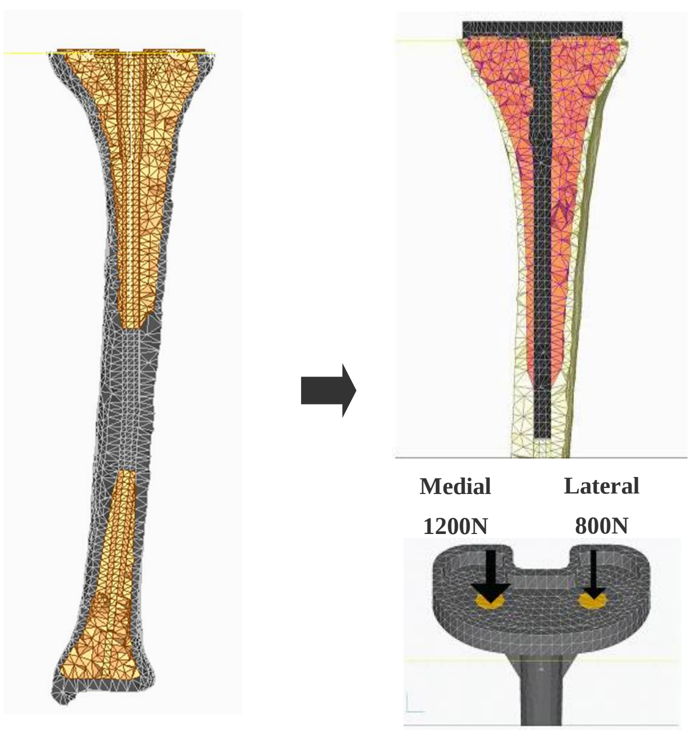


Fig. 3. Finite element model of tibia and artificial knee joint implant with loading condition at standing posture

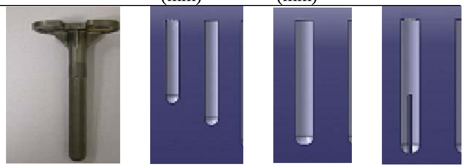
Implant	Length (mm)			Diameter (mm)		Slot	
							
Wright Medical Technologies	70	90	115	10	12	slot	solid

Fig. 4. 3D CAD models of implant for revision TKR with various stem end design parameters

rational to explain this problem is not provided. We investigated the effects of stem end design on stem pain in revision TKR.

The finite element model of tibia, including the cortical bone, the cancellous bone and canal, was developed based on CT images. The stem models with various stem lengths, diameters and frictional coefficients, and press-fit effects were considered as a parametric study. The pressure distribution on the interface of stem and the von-Mises stress around the stem end were investigated to predict the effects of stem end design on stem end pain.

The results showed that the longer stem length, the stronger press-fit, the bigger stem diameter, and the higher frictional coefficient increased both peak contact pressure and the highest von-Mises stress values.

In this study, we hypothesized that the different stem design creates different contact pressure and von-Mises stress distribution around the stem, and the locally higher pressures and stresses may be related to the stem end pain. Our clinical data supported the biomechanical findings hypothesized in this study. This study will be useful to modify the stem end design and reduce the end-of-stem pain in revision TKR.

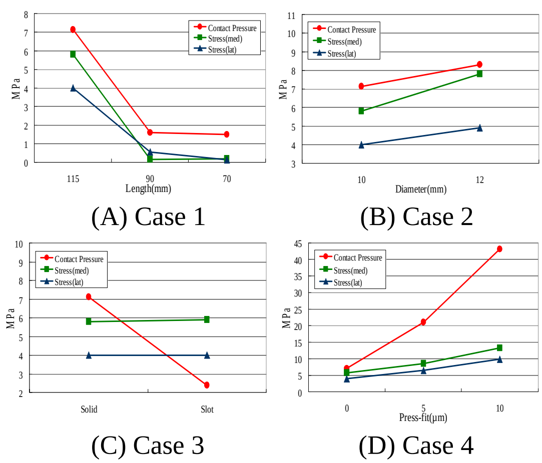


Fig. 5. Peak contact pressures and von-Mises stresses in the adjacent bone to the stem-end

5 Graphic Visualization and Simulation of Tennis Play

Fig. 6 shows the computer graphic simulation of tennis play using the developed musculoskeletal model of a Korean male. First, we captured the kinematic motion of a male subject using the motion capture system (Hawk®, Motion analysis Inc., USA). Second, the scaled the all bone segments of the musculoskeletal model fitted to the subject size. Then, the kinematic motion data was modified and inserted to the graphic model to generate the motion. Next, the necessary inverse kinematics analysis was performed to calculate the joint torques and muscle forces generating the given motion. Finally, the graphic simulation of the tennis play was generated using commercial graphic software (3D-Max®, Autodesk Inc., USA).

Fig. 6. Motion simulation of tennis play using motion captured kinematic data and virtual human model

6 Conclusion

In this paper, I present the modeling and simulation technologies of human musculoskeletal system and applications to the biomechanical researches. The potential application area of this model is not only biomechanics but also sports, surgical simulation, medical device industry, ergonomics, human safety on automotives, and so on. In addition, the model can be extended from molecular level to the system organ level. This kind of modeling and simulation can contribute research, education, and development.

For the improvement of the modeling and simulation, there are several technologies and algorithms should be improved or newly developed, such as accurate medical image modality, 3-D reconstruction, estimating the material properties form the medical images, various mechano-biology principles, multi-scale numerical analysis ant so on. In order to solve those technical problems, various types of applied mathematics principles have to be used. Finally, the present modeling and simulation techniques can be very useful to enhance the health-related quality of life.

References

1. Kim, Y.H., Inoue, N., Chao, E.: Kinematic Simulation of Fracture and Bone Deformity Correction under External Fixation. J. Biomech 25, 1047–1058 (2002)
2. Kim, Y.H., Koo, K.M., Kwon, O.S., Bae, D.K.: Effects of Stem End Design on Stem Pain. In: TKR. 06' World Congress of Biomechanics (WCB), Munchen, Germany, 2006 (2006)
3. Park, W.M., Kim, Y.H., Kim, K., Park, Y.S.: Effects of Anterior Rod Number and the Removal of Mid-Column on Anterior and Posterior Instrumentation in Thoracolumbar Burst Fracture. In: Fifteenth Annual Symposium on Computational Methods in Orthopaedic Biomechanics, San Diego, USA, 2007 (2007)
4. Kim, Y.H.: Biomechanical Modeling of Human Musculoskeletal System. In: Korean Biomedical Engineering Society Conference, Seoul, Korea, 2004 (2004)

The Strength Factor in Digital Human Modeling and Simulation: A Case for a New Framework

Kang Li and Xudong Zhang

University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
{kangli2,xudong}@uiuc.edu

Abstract. It is not well understood whether strength could ever influence performance in which only sub-maximal, moderate exertions are needed although strength may be indicative of the movement performance at or near the maximal level. In this paper, we present some evidence from two studies suggesting the strength effects on performance strategies in volitional physical tasks, and based on the evidence, a new framework was proposed for human movement modeling and simulation that incorporates the strength factor.

Keywords: Strength, Digital human modeling.

The existing digital human models are created based primarily on the gender and anthropometry information. Many physical or cognitive attributes are not included as the input variables in current software tools when rendering digital human "avatars" and their performance. Among these is the strength factor—that is, the maximum force or moment one can produce. It is well accepted that strength is posture- or movement-dependent [3]. It is also intuitive that strength may be indicative of the movement performance at or near the maximal level (e.g., the maximum acceleration producible). It is, nevertheless, not well understood whether strength could ever influence performance in which only sub-maximal, moderate exertions are needed. In this paper, we present some evidence from two studies suggesting the strength effects on performance strategies in volitional physical tasks, and based on the evidence, we propose a new framework for human movement modeling and simulation that incorporates the strength factor.

In the first study [1], we investigated the effects of individuals' dynamic (isokinetic) strength and their knowledge of strength (whether you know how strong they are) on their selection of load-pace balancing strategy. Thirty-two subjects (16 men and 16 women) first underwent dynamic strength tests of the shoulder, back, and knee joints, and then performed simulated manual materials handling tasks. In the latter tasks, they were allowed to choose any scenario to complete the handling of a batch of multiple weight plates; for example, for a batch of three plates, one may choose to handle all three in one carry, two in one carry and one in the other, or one in each of three carries. Prior to the materials handling tasks, half of the subjects were provided with knowledge feedback of their strength testing results (in percentile ranks), and the other half not. The study showed that dynamic strength and knowledge feedback both had significant effects on measures quantifying the handling strategy. Individuals with greater strength tended to adopt a strategy

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 144–146, 2007.

corresponding to heavier load per carry and fewer carries per batch. Receiving knowledge feedback evoked a tendency towards handling heavier load, and this tendency was more salient in the weaker individuals.

In the second study, we further examined whether the dynamic strength could have an effect on how people move. We asked the same 32 subjects to perform lifting tasks, measured their lifting motions, and then quantified their lifting styles (i.e., back lift vs. leg lift) using a simple index [2]. We found that dynamic strength had a significant effect on the lifting strategy: subjects with back strength greater than their total knee strength tended to use a back lift strategy, and vice versa.

The above two studies suggest that the strength factor may affect both the high-level decision making (e.g., what effort level to choose) and the low-level motor control (e.g., which body segment as the prime mover) during goal-driven (as opposed to passive) physical performance. The development of next generation digital human models ought to incorporate such factors in order to improve the realism and utility of the model-based simulation tools. We therefore propose a new, more holistic framework for digital human modeling and simulation integrating many consequential task and personal factors currently not considered as model inputs. Figure 1 illustrates the proposed framework, in which the left-side blocks and their interrelationships in fact constitute the conventional logic diagram of forward dynamic simulation of human movement production.

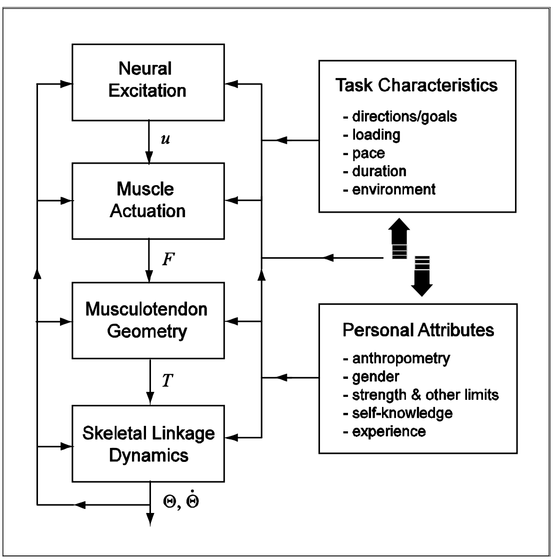


Fig. 1. A new, more holistic framework for digital human modeling and simulation of physical performance involving movements

Note this is a closed-loop system where the feedback of current state (position and velocity) may influence not only the next state but also the musculotendon geometry, muscle force production, as well as the high-level neural excitation patterns. Of note also is that the limits on muscle force production (i.e., strengths) do not play a role explicitly in the movement dynamics unless the limits become active in performance at the maximum level.

The right-side blocks represent the two groups of factors that can affect the movement form and production, and ideally should be included as input variables in digital human modeling and simulation. Presently, however, only a limited subset of these factors is included. Anthropometry and gender are the only two personal attributes known to affect how a digital human and its movements are rendered at the lower geometrical and physical property levels. A host of additional personal factors including strength and other physical capacity limits, self-knowledge of the limits, and experience can exert influence on events of all levels in movement production, as demonstrated in our studies. Task characteristics affect how a movement task is initiated and sustained at the neural and muscular levels. Currently, only explicit directions or goals are implemented as command inputs (e.g., move forward, turn left) in simulation of a movement act. More implicit goals, such as intrinsic objective of minimizing a certain cost function (e.g., energy expenditure, muscle stress) are subjects of debates and not yet embodied in any existing simulation tools. Evidence has also accumulated to suggest a multitude of task factors including the load handled and speed in a lifting task can influence the movement behavior [4]. Task duration certainly relates to muscle fatigue; however, the muscle fatigue effect on movement patterns has been difficult to model in digital human modeling.

Perhaps more important, the interaction between the task characteristics and personal attributes could play an important role in movement performance. As evidenced in the second study, given load-pace requirement of a task, one may adapt to the best multi-level strategies to complete the task, taking into account one's physical capabilities and prior knowledge.

Clearly, many of the effects outlined in the proposed framework remain poorly understood. We intend this new framework to guide our future investigations aimed to better understand and quantify these effects, one at a time, and then incorporate them into the next generation digital human models.

Acknowledgments. This work is supported by Centers for Disease Control and Prevention and National Institute of Occupational Safety and Health.

References

1. Bartlett, D., Li, K., Zhang, X.: A relation between dynamic strength and manual materials handling strategy affected by knowledge of strength. Hum Factors (to appear, 2007)
2. Burgess-Limerick, R., Abernethy, B.: Toward a quantitative definition of manual lifting postures. Hum. Factors 39, 141–148 (1997)
3. Chaffin, D.B., Andersson, G.B.J., Martin, B.J.: Occupational Biomechanics, 3rd edn. John Wiley & Sons, New York (1999)
4. Zhang, X., Xiong, J., Bishop, A.M.: The effects of load and speed on lumbar vertebral kinematics during lifting motions. Hum. Factors 45, 296–306 (2003)

The Effects of the False Vocal Fold Gaps in a Model of the Larynx on Pressures Distributions and Flows

Sheng Li, MingXi Wan, and SuPin Wang

The Key Laboratory of Biomedical Information Engineering of Ministry of Education, and Department of Biomedical Engineering, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an 710049, People's Republic of China
mxwan@mail.xjtu.edu.cn

Abstract. Human phonation does not merely depend upon the vibration of the vocal folds. The false vocal fold (FVF), as an important laryngeal constriction, has also been found by more and more research both in clinically and computer simulations that it plays an important role during phonation and contributes significantly to the aerodynamics and sound generation processes of human voice production. Among many parameters which are used to determine and describe the geometry of the false vocal folds, the false vocal fold gap (GFVF), which means the minimal distance between the two false vocal folds, is regarded as an important and dominant parameter. Therefore, this study explores the effects of the FVF gaps on the intralaryngeal pressure distributions, laryngeal resistance and flows using both three-dimensional Plexiglas model and commercially available computational fluid dynamics code.

Three glottal angles, divergent 40°, uniform 0°, and convergent -40°were used for this study to explore the effects of FVF gaps, as they represent the basic glottal shapes typically expected in phonation, the angle values also were typically expected for most phonation in modal Register. A wide variety of FVF gaps (GFVF) were parameterized with 12 different values: 0.02, 0.04, 0.06, 0.08, 0.09, 0.1, 0.11, 0.12, 0.16, 0.2, 0.4, 0.6 cm to represent important geometries often appearing within phonatory vibratory cycles. These gaps were used for each glottal angle. The specific design of the FVFs followed prior literature. The minimal glottal diameter (Dg) was constantly at 0.06 cm in this study for each FVF gaps, and the translaryngeal pressure were held constant at 8 cm H2O. A nonvibrating laryngeal airway Plexiglas model, which had linear dimensions 1.732 times of a normal male larynx, was used in this study. In order to measure pressures inside the Plexiglas model, twelve cylindrical ducts were made on the midline of the laryngeal wall of the model. The diameter of each duct was 0.07 cm human size (0.12 cm in the model), so that the connector of an Entran EPE-551 pressure transducer could fit snugly into the holes. The distance between the centers of each hole was 0.14 cm human size. FLUENT (Fluent, Inc., Lebanon, NH), a commercially available computational fluid dynamics code was also used to obtain estimates of the normal wall pressures along the laryngeal surfaces (including the surfaces of vocal folds, ventricles, and false vocal folds) as a function of the FVF gaps and the glottal angles. The code is based on the control-volume technique and was used to solve the Navier-Stokes equations for constant shapes (not for vibrating vocal folds), laminar, incompressible airflow physics occurring inside the symmetric

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 147–156, 2007.

laryngeal geometries. The flow field was assumed to be symmetric across the midline of the glottis in this study, and therefore only the half flow field was modeled.

The results suggest that (1) the intralaryngeal pressure was lowest and the flow was highest (least flow resistance) when the FVF gap was 1.5-2 times of Dg, the intralaryngeal pressures decreased and flows increased as smaller FVF gaps increased, and the intralaryngeal pressures increased and flows decreased as larger FVF gaps increased, indicating that the least pressure drop for any given flow (that is, the least flow resistance) was found to correspond to the 1.5-2 times of Dg for different glottal angle. Suggesting that the 1.5-2 times of Dg might be the optimal gap for pressure, and efficient phonation may involve laryngeal shaping of this condition. Therefore, the positioning and existing structure of the FVFs can aid in phonation by reducing energy losses and increasing airflow in the larynx when positioned appropriately; (2) both the pressure and flow were unaffected when the FVF gaps larger than 0.4 cm; (3) the divergent glottal angle gave lower pressure and greater flow than the convergent and uniform glottal angle as no FVF conditions; (4) the present of the FVF decreased the effects of the glottal angle on both the intralaryngeal pressure and flow to some extent, and the smaller the FVF gaps, the smaller this effect. Perhaps more important, (5) the present of the FVF also moving the separation points downstream, straitening the glottal jet for a longer distance, decreasing overall laryngeal resistance, and reducing the energy dissipation, suggesting that the FVF would be of importance to efficient voice production; (6) the empirical pressure distributions were supported by computational results. The results suggest that the intralaryngeal pressure distributions and the laryngeal flow resistance are highly affected by the presence of the FVFs, and the FVFs can aid in phonation when by reducing energy losses positioned appropriately. Therefore, the results might be helpful not only in maintaining healthy vocal habits, but also in exploring surgical and rehabilitative intervention of related voice problem. The results also suggest that they may be incorporated in the phonatory models (physical or computational) for better understanding of vocal mechanics.

Keywords: Speech production, false vocal folds, model, pressure distributions.

1 Introduction

It is well know that the laryngeal geometry plays an important role during speech production(Li et al., 2006a; Li et al., 2006b; Scherer et al., 2001a; Scherer et al., 2002; Scherer et al., 2001b). The false vocal folds, as an important laryngeal constriction, have been discovered by more and more research both in clinically and computer simulations that it may also contribute significantly to the aerodynamics and sound generation processes of human voice production (Berg, 1955; Kucinschi et al., 2006; Rosa et al., 2003; Zhang et al., 2002).

There are many parameters to determine and describe the geometry of the false vocal folds (Agarwal et al., 2003), however, the false vocal fold gap (GFVF), which means the minimal distance between the two false vocal folds, is regarded as an important and dominant parameter (Berg, 1955), since this parameter may lead to most of the shape change of the false vocal folds, thus affect both physiology and acoustics of voice production, especially at very narrow diameters, the intraglottal pressure distributions and flow resistance created by them will not only be a dominant feature, but also affect true vocal fold vibrations by nullifying the Bernoulli effect (Berg, 1955).

In order to adequately study the function of the false vocal folds and test its theories of phonation, the pressure distributions on the laryngeal surfaces need to be known for a wide range of FVF gaps that are expected in phonation, since these distributions define the external driving forces that interplay with the vocal folds and the false vocal folds surface configuration as well as tissue biomechanics to allow vocal folds and false vocal folds oscillation (Guo and Scherer, 1993; Titze, 1988). Furthermore, the intraglottal pressure distributions are accompanied by and depended upon the characteristics of the volume flow, and the laryngeal airflow resistance (Agarwal, 2004).

Previous studies showed in both the computational models ((Iijima et al., 1992; Pelorson et al., 1995; Rosa et al., 2003), and physical models (Zhang et al., 2002; Agarwal, 2004; Kucinschi et al., 2006) that the false vocal fold played a more positive role in both the physiology and acoustics of speech production, they also suggested that the geometry differences within the laryngeal airway might significantly changed the aerodynamic and aeroacoustic properties, and thus the whole results for the system. These studies, however, were limited in the phonatory geometries considered, and the detailed laryngeal pressure distributions for a wide range of laryngeal geometries were not specified.

The purpose of the present study is to examine the effects of a wide variety of the false vocal fold gaps with specific laryngeal geometries on the intraglottal wall pressure distributions, flows, and the translaryngeal airflow resistance for different glottal angles. If the variation of the FVF gap is to create substantial changes with the transglottal, glottal wall pressures, as well as flow resistance, both the motion of the vocal folds and the false vocal folds and the shape and size of the glottal volume velocity signal may be significantly affected, and the physical and computer models for vocal fold and false vocal fold motion would need to be extended to take such effects into account.

2 Methods

2.1 Plexiglas Model

The model used in this research was a nonvibrating laryngeal airway model made from Plexiglas. Relative to a similitude analysis (Streeter and Wylie, 1975) and for the convenience of fabrication of the model and recording experimental data, the model had linear dimensions 1.732 times that of a normal male larynx. Figures 1(a) and (b) show the planform or coronal (a) and side elevation or lateral (b) views of the model, respectively. The model was symmetric about two axes. Two laryngeal pieces were embedded into the airway duct. The laryngeal airway itself is the slit between the two laryngeal pieces. The false vocal fold gap was adjustable by a movable handle outside the laryngeal duct (Fig. 1). The indentations of the larynx at the anterior and posterior ends partially mimic the presence of the arytenoid cartilages posteriorly and the anterior commissure anteriorly (Hirano et al., 1988). There were three pairs of vocal fold pieces to provide three symmetric glottal angles, namely, 40 degree divergent, 0 degrees (uniform), and –40 degrees convergent. The minimal glottal diameter was set constantly at 0.06 cm, and the translaryngeal pressure was held constantly at 8 cm H2O.

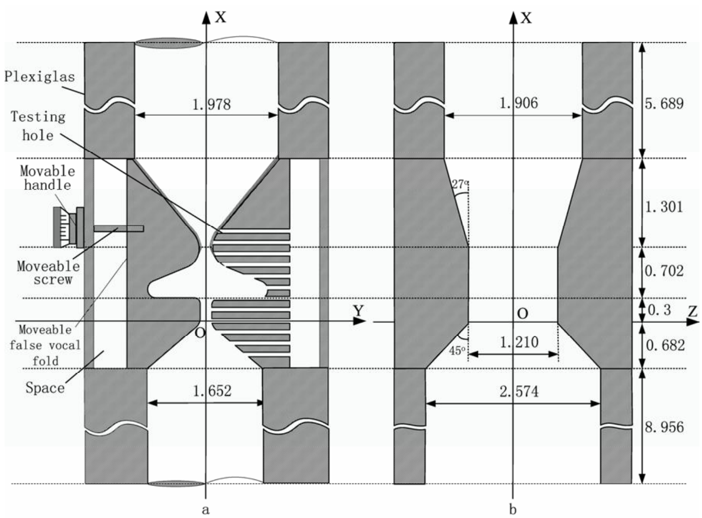


Fig. 1. The planform (a) and side elevation (b) of the Plexiglas model, the dimensions of the model were 1.732 times greater than human size values

In order to measure pressures inside the Plexiglas model, twelve cylindrical ducts were made on the midline of the one of the laryngeal wall of the model. The diameter of each duct was 0.0704 cm human size (0.1220 cm in the model) so that the cylindrical extension of a calibrated Entran EPE-551 pressure transducer (0.1040 cm outside diameter and 6 cm long) could fit snugly into the ducts. The distance between the centers of each hole was 0.1440 cm human size.

2.2 Computational Method

A commercially available computational fluid dynamics code called FLUENT, based on the control-volume technique, was used to numerically solve the Navier-Stokes equations for laminar airflow physics occurring inside the three-dimensional symmetric geometries. The input data for FLUENT included grid point coordinates, element topology, and boundary conditions. The grids used in the simulations contained both structured and unstructured meshes. In these simulations, grids containing between 1020 000 and 1030 000 nodes were used to obtain first- or second-order solutions for momentum, pressure, and velocity, with residuals less than 10-4. Since the glottal section in the model was most important for this study, the interval size of the mesh element was selected to be 0.015 cm for the laryngeal section, whereas it was 0.03 cm for the inlet and outlet tunnel. The flow field was assumed to be symmetric across the midline of the glottis in this study, and therefore only the quarter flow field was modeled.

3 Results

Figure 2 shows the empirical one-dimensional pressure distribution along the midline surface of the laryngeal wall for the FVF gaps for the uniform glottis. The corresponding predicted pressure distributions using FLUENT are also shown in Fig. 2 as the lines, which match the data points closely. The whole laryngeal airway can be divided into five regions: the subglottal region, the TVF (or glottal) region, the ventricle region, the FVF region and the superior FVF region. It can be seen from Figure 2 that there are two pressure drops, two minimum pressures (and largest velocity), and two pressure recovers (that is, pressure increase) occurred for most of the pressure distributions.

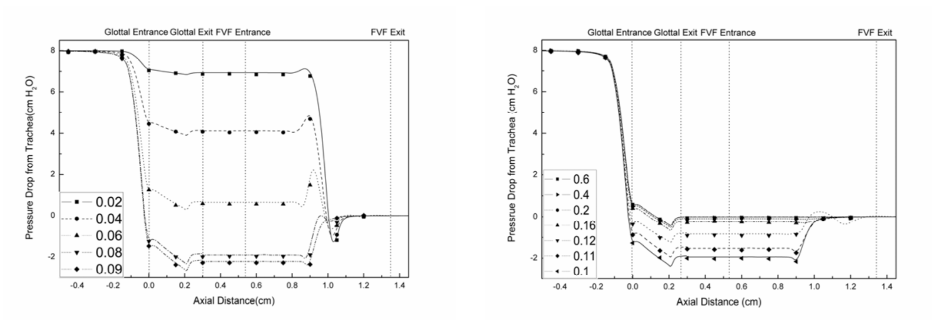


Fig. 2. Empirical one-dimensional pressure distributions on the midline surface of the laryngeal wall of the model for twelve FVF gaps for 0.06 cm minimal glottal diameter and 8 cm H2O transglottal pressure. The estimated pressure distributions using FLUENT for each conditions are also shown as the lines. (a) For five FVF gaps which are less than 1.5 times of Dg; (b) For seven FVF gaps which are larger than 1.5 times of Dg.

The pressure distributions of FVF gaps which are less than 0.09 cm (1.5 times of the minimal glottal diameter) are shown in Figure 2(a). It can be seen from the figure that the pressure distributions decreased greatly as the FVF gap became larger, with a reversal of this pattern at the minimal FVF diameter location. For example, an increase of gaps from 0.02 cm to 0.09 cm dropped the glottal exit pressures from approximately 6.93 cm H2O to -2.22 cm H2O, this pressure drop is larger than the translaryngeal pressure and nearly a sextuple of the outward driving pressure on the TVF, suggesting not only a significant pressure difference in pulling or pushing the TVF, but also strongly influencing the pressure distributions in the TVF regions, and then the whole larynx. However, for the larger gaps which are larger than 0.09 cm, the pressure distributions increased as the FVF gap became larger (Figure 2(b)). It is because of the decrease of the translaryngeal air flow resistance (Figure 3, which will be discussed below). It also can be seen from this figure that the pressure distributions of 0.4 and 0.6 cm gaps are almost superposed, suggesting that the present of the FVFs might have less effects on intralaryngeal pressure distributions for the gaps which are larger than 0.4 cm (6.7 times of Dg).

For the FVF gaps smaller than the minimal glottal diameter, the pressure dips in FVF region are larger than the one in the TVF region, but it is smaller when the FVF gaps are larger than the minimal glottal diameter. For example, the pressure difference between the pressure dips in TVF region and FVF region are 8.198 cm H2O for 0.02 cm gap, suggesting a great pressure difference acting upon the TVF and FVF. If the pressure were set at zero (atmospheric) at the outlet to the laryngeal airway, the pressure would be positive (6.93 cm H2O) for the 0.02 cm FVF gap condition, but negative (-2.22 cm H2O) for the 0.09 cm FVF gap condition at the position of their glottal exit, suggesting that the FVF gap not only has great effects on the pressure distributions in the FVF regions, but also in the TVF regions, and cause a great difference in pulling or pushing both the TVF and the FVF.

For the gaps larger than 1.5 times of Dg (Figure 2(b)), the pressure distributions changed less than smaller gaps. For example, the pressure at the location of the glottal exit increased by only 1.89 cm H2O when the gaps increased six times from 0.1 to 0.6 cm, and the larger the gaps, the less the changes. Suggesting that for larger gaps (larger than 1.5 times of Dg), the present of the FVF have less effects on the intralaryngeal pressure distributions than smaller FVF gaps. It also can be seen from the figure that for the gaps larger than 0.4 cm, the pressure distributions have no change.

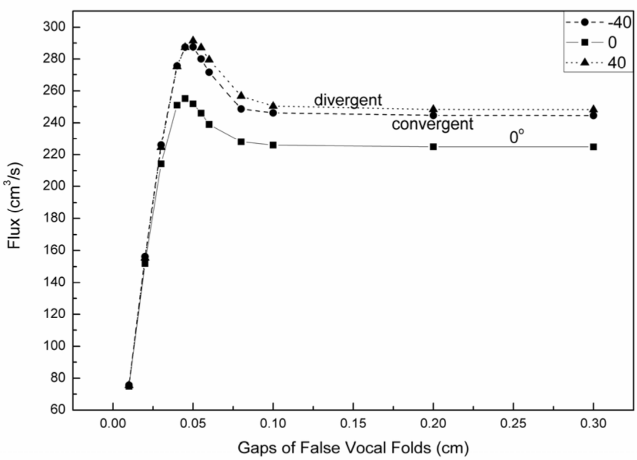


Fig. 3. Experimental flow rates for twelve false vocal fold gaps for different glottal angles for 8 cm H_2O translaryngeal pressure drop

The laryngeal airflow is another important parameter which is affected by the FVF gaps. Figure 3 shows the airflow distributions for different FVF gap for three glottal angles. It can be seen from Figure 3 that the airflow increased dramatically when the FVF gap smaller than the Dg. For example, the airflow increased by about 2.90 times as the FVF gap increased from 0.02 cm to 0.06 cm Dg separately, and had little difference among different glottal angle, this difference, however, was increased as the FVF gap increased, and the larger the FVF gap, the larger this difference.

It also can be seen from Figure 3 that for the same FVF gaps, the divergent glottal angle gave the maximize flow, followed by the convergent, then the uniform glottal angle. This result was corresponding to the laryngeal pressure distributions which were shown in Figure 2, suggesting that the divergent glottal angle was corresponding to the least laryngeal resistance. This result is also consistent with the results of Agarwal et al. (2004) and Li et al. (2006), the latter study is of no FVF cases, suggesting that the glottal angle still have important effects on laryngeal flow (and pressure distributions) under conditions of the present of the FVF.

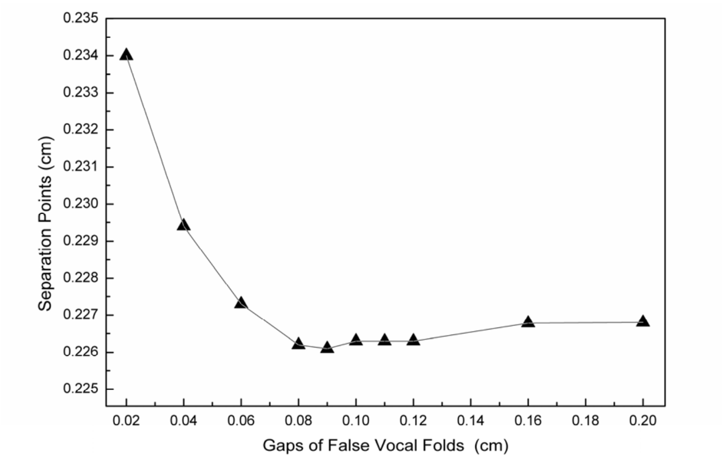


Fig. 4. Separation point locations for ten FVF gaps (predicted by FLUENT)

Figure 4 shows the separation point locations of the uniform glottis for each FVF gaps. It can be seen from the figure that (1) for the FVF gaps smaller than 0.09 cm (1.5 times of Dg), the separation point tend to move downstream toward the glottal exit with decreasing FVF gaps (decreasing flow rate), however, the separation point remained almost constant (slightly downstream) for the larger gaps. Suggesting that the present of the FVF would move the separation point downstream in the glottis (but not too much), thus would not only reduce dissipation loses above the glottis, but also might affect the dynamics of separation points.

4 Discussion

In general, when the FVF gaps are equal to or smaller than the Dg, the present of the FVF hinders the path of the glottal flow, causing more energy dissipates, and increasing the intralaryngeal pressures. It is important to note that the intralaryngeal minimal diameter are migrated from TVF region to FVF region, suggesting that the great pressure dip in the laryngeal are acting upon the FVF, this may have great difference for the fundamental frequency, vocal intensity, vocal register and vocal efficiency in phonation, due to the FVF is found to be more viscous and less stiff than the true vocal folds biomechanically.

For the FVF gaps larger than the minimal glottal diameter, the presence of the FVF showed decreased pressure drops in this study. This result is also shown by Pelorson et al. (1995) and Kuchinschi (2006). However, the results cannot be compared because their studies have not provided enough information on intralaryngeal pressure distributions. Moreover, their results have not indicated that the decreased pressure drops are only happened under conditions of the FVF gaps larger than 1.5-2 times of Dg and smaller than 0.4 cm (or smaller than about 7-10 times of Dg).

The present of the FVF would also increase the intralaryngeal airflow. It can be seen from Figure 3 that the maximal airflow is larger by about 12-17% than the no FVF conditions for different Dg and different glottal angle. This suggests that the amount of airflow produced is greater than it otherwise would have been. In other words, when the true vocal folds are constricted in phonation and allow a certain amount of airflow through, the presence of the false vocal folds might increase this amount of the airflow.

Perhaps more important, the FVF might be favorable in guiding the glottal jet to retain smooth laminar flow. It is well known in fluid dynamics that a spike or forward jet can reduce drag by contouring the free stream about the body, which is known as drafting or breaking the wind (Blevins, 1984). Suggesting the present of the FVFs might decrease overall resistance. Kucinschi et al. (2006) indicated that the glottal jet remains laminar for a longer distance, thereby again reducing the energy dissipation associated with an otherwise turbulent field. Flow visualization by (Shadle et al., 1991) and Kucinschi et al. (2006) also showed that the skewed glottal jet was straightened due to the presence of the false vocal folds.

It can be seen from previous study (Guo and Scherer, 1993; Li et al., 2006b) that the divergent glottal angle gave the least resistance (the maximize flow) and the greatest pressure drops for the same value of Dg, followed by the convergence, then the uniform one. Unfortunately, their studies had not taken into account of the effects of the FVFs. This study shows more detailed results under conditions of the present of the FVF.

In general, when the FVF are highly constricted or hyperadducted, the geometric configuration of the FVFs, thus, seems to be the most appropriate to reduce energy losses by decreasing the overall system resistance, increasing the airflow and decreasing the energy dissipation in the larynx. This effect, furthermore, would be greatest when the FVF gaps at 1.5-2 times of Dg. This does not mean that the phonation in such cases is normal, but the FVF can aid in phonation when by reducing energy losses by positioned appropriately. This might be one of the reasons why FVF adduction is observed so consistently with glottal pathology.

Lower pressures and increased airflows which are caused by the present of the FVF are also indicative of both biological as well as acoustical efficiency. As discussed earlier, the voice intensity might be increased, the voice quality can be improved and the vocal efficiency may be enhanced by modifying laryngeal geometry appropriately. Furthermore, it is possible to rehabilitate a patient of glottal pathology or improve their voice quality like ventricular dysphonica, or hyperfunctional voice in general by modifying the shape of the false vocal folds or supraglottal laryngeal musculature appropriately during phonation in addition to the proper treatment of the true vocal folds.

5 Conclusions

In this study, a wide range of FVF gaps with well specified laryngeal geometry was explored on intralaryngeal pressure and flow distributions for uniform, 40º divergent and -40ºconvergent glottal angles separately, by using a Plexiglas model and the computational fluid dynamics code FLUENT. The results suggested that the intralaryngeal pressure was lowest and the flow was highest (least flow resistance) when the FVF gap was 1.5-2 times of Dg, the intralaryngeal pressures decreased and flows increased as smaller FVF gaps increased, and the intralaryngeal pressures increased and flows decreased as larger FVF gaps increased. However, for the FVF gaps larger than 0.4 cm, both the pressure and flow were unaffected. This study also found that the divergent glottal angle gave lower pressure and greater flow than divergent and uniform glottal angle as no FVF conditions, but the present of the FVF decreased the effects of the glottal angle to some extent. Furthermore, the present of the FVF also moving the separation points downstream, straitening the glottal jet for a longer distance, decreasing overall laryngeal resistance, and reducing the energy dissipation, suggesting that the FVF would be of importance to efficient voice production.

The models used in this study were limited to one minimal glottal diameter (0.06 cm) and symmetrical laryngeal flows and pressures, which means that these results are conditional. However, this study does appear some meaningful effects of FVFs on phonation, and suggest that the results may be incorporated in the phonatory models (physical or computational) for better understanding of vocal mechanics. The results might also be helpful in exploring surgical and rehabilitative intervention of related voice problem.

References

1. Agarwal, M.: The false vocal folds and their effect on translaryngeal airflow resistance. Doctor of Pilosophy. Bowling Green State University. Bowling Green (2004)
2. Agarwal, M., Scherer, R.C., Hollien, H.: The false vocal folds: shape and size in frontal view during phonation based on laminagraphic tracings. J. Voice 17, 97–113 (2003)
3. Berg, J.v.d.: On the role of the laryngeal ventricle in voice production. Folia. Phoniatrica. 7, 57–69 (1955)
4. Blevins, R.D.: Applied fluid dynamics handbook Nostrand Reinhold Company Inc. New York (1984)
5. Guo, C.G., Scherer, R.C.: Finite element simulation of glottal flow and pressure. J Acoust Soc Am. 94, 688–700 (1993)
6. Hirano, M., Kiyokawa, K., Kuria, S.: Laryngeal muscles and glottis shaping. Raven O.Fujimura, New York (1988)

7. Iijima, H., Niki, N., Nagai, N.: Glottal impedance based on a finite element analysis of two-dimensional unsteady viscous flow in a static glottis. IEEE Transactions on Signal Processing 40, 2125–2135 (1992)
8. Kucinschi, B.R., Scherer, R.C., DeWitt, K.J., Ng, T.T.M.: Flow visualization of air moving through a model of the larynx. J. Biomechanical Engineering 128, 380–390 (2006)
9. Li, S., Scherer, R.C., Minxi, W., Wang, S., Wu, H.: The effects of inferior and superior vocal fold surface angles on vocal fold pressure distributions. J Acoust Soc Am. 119, 3003–3010 (2006)
10. Li, S., Scherer, R.C., Wan, M., Wang, S., Wu, H.: The effect of glottal angle on intraglottal pressure. J. Acoust. Soc. Am. 119, 539–548 (2006b)
11. Pelorson, X., Liljencrants, J., Kroeger, B.: On the aeroacoustics of voiced sound production. In: 15th International congress on acoustics edited by Trondheim, Norway, 1995, pp. 26–30 (1995)
12. Rosa, M.d.O., Pereira, J.C., Greelet, M., Alwan, A.: A contribution to simulating a three dimensional larynx model using the finite element method. J. Acoust. Soc. Am. 114, 2893–2905 (2003)
13. Scherer, R.C., Shinwari, D., Witt, K.J.D., Zhang, C., Kucinschi, B.R., Afjeh, A.A.: Intraglottal pressure profiles for a symmetric and oblique glottis with a divergence angle of 10 degrees. J. Acoust. Soc. Am. 109, 1616–1630 (2001a)
14. Scherer, R.C., Shinwari, D., Witt, K.J.D., Zhang, C.R., Kucinschi, B., Afjeh, A.A.: Intraglottal pressure distributions for a symmetric and oblique glottis with a uniform duct (L). J. Acoust. Soc. Am. 112, 1253–1256 (2002)
15. Scherer, R.C., Witt, K.J.D., Kucinschi, B.R.: The effect of exit radii on intraglottal pressure distributions in the convergent glottis (L). J. Acoust. Soc. Am. 110, 2267–2269 (2001b)
16. Shadle, C.H., Barney, A.M., Thomas, D.W.: An investigation into the acoustics and aerodynamics of the larynx. In: Gauffin, J., Hammarberg, B. (eds.) Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms, pp. 73–82. Singular Publishing Group, Inc, San Diego (1991)
17. Streeter, V.L., Wylie, E.B.: Fluid Mechanics. McGraw-Hill, New York (1975)
18. Titze, I.R.: The physics of small-amplitude occillation of the vocal folds. J. Acoust. Soc. Am. 83, 1536–1552 (1988)
19. Zhang, C., Zhao, W.H., Frankel, S., Mongeau, L.: Computational aeroacoustics of phonation Part II: Effects of flow parameters and ventricular folds. J. Acoust. Soc. Am. 112, 2147–2154 (2002)

A Robust Algorithm for a System Identification Approach to Digital Human Modeling: An Application to Multi-fingered Hand Movement

Kang Li, Sang-Wook Lee, and Xudong Zhang

University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
{kangli2,slee52,xudong}@uiuc.edu

Abstract. A recent study [2] proposed a forward bio-dynamic model of multi-fingered hand movement. The model employed a physics-based heuristic algorithm for system identification of the model parameters, and succeeded in replicating measured multi-fingered flexion-extension movement. However, while the model itself is general and readily applicable to other bodily movements, the heuristic algorithm required empirical adjustments to initial setups, and was therefore difficult to generalize. This paper introduces a rigorous and more robust parameter estimation algorithm to enhance the intended general modeling approach for digital human movement simulation. The algorithm is demonstrated by solving the same modeling problem posed in [2].

Keywords: hand movement, system identification, forward dynamics.

1 Introduction

Despite the successful applications of digital human modeling (DHM) and simulation in ergonomics analysis and workplace design, very little work has been done to enable designers to simulate hand and finger postures or movements. Existing collision-detection based posture prediction approaches, used in commercial software packages, are not capable of generating realistic hand representations. There is a need for computational models that reproduce naturalistic hand movements or postures. A recent study [2] proposed a forward bio-dynamic model of multi-fingered hand movement. The model employed a physics-based heuristic algorithm for system identification of the model parameters, and succeeded in replicating measured multi-fingered flexion-extension movement. However, while the model itself is general and readily applicable to other bodily movements, the heuristic algorithm required empirical adjustments to initial setups, and was therefore difficult to generalize. We hereby introduce a rigorous and more robust parameter estimation algorithm to enhance the intended general modeling approach for digital human movement simulation. We demonstrate the algorithm by solving the same modeling problem posed by Lee and Zhang [2].

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 157–160, 2007.

2 Methods

The dynamics of a 3-segment, 3-DOF finger system can be represented by the following equations of motion:

$$M(\theta)\ddot{\theta} = V(\theta,\dot{\theta}) + G(\theta) + \tau \tag{1}$$

where θ is the three-joint-angle vector, $\tau \in \Re^{3\times1}$ is the torque vector at the joints, $M(\theta) \in \Re^{3\times3}$ is a positive definite mass matrix, $V(\theta,\dot{\theta}) \in \Re^{3\times1}$ is the centrifugal (square of joint velocity) and Coriolis force, and $G(\theta) \in \Re^{3\times1}$ is the gravitational component. The model did not consider the effect of the gravitational force $G(\theta)$.

It was assumed that all the joints (DIP, PIP and MCP) of digits 2-5 were controlled by torque actuators which could be modeled as proportional-derivative (PD) feedback controllers situated at the beginning of the joint movements, i.e. the torque at joint i, τ_i, could be computed by

$$\tau_i = -K_i^p(\theta - \theta_i^f) - K_i^d(\theta_i - \theta_i^f)) \quad \text{subject to } \tau_i \le M_i^{\max} \tag{2}$$

$$K_i^p \ge 0 \tag{3}$$

$$K_i^d \ge 0 \tag{4}$$

where K_i^p is the proportional gain of the joint actuator i affecting the movement speed; K_i^d is the velocity gain of the joint actuator i; $\theta_i^f, \dot{\theta}_i^f$ are the final joint angle and final angular velocity; $M_i^{\max}$ represents the relative magnitudes of the torque generated at joint i, which serves as an auxiliary parameter to control the initiation or the temporal coordination of finger segments.

An optimization problem minimizing the discrepancy between the model-predicted and measured angle profiles can be formulated to estimate the parameters (K_i^p, K_i^d, and $M_i^{\max}$) as follows:

$$min\int_0^{t_{max}} (\theta(t) - \tilde{\theta}(t))^2\, dt \tag{5}$$

$$\text{s.t. } M(\theta(t))\ddot{\theta}(t) = V(\theta(t),\dot{\theta}(t)) + \tau(t) \tag{6}$$

$$\tau_i(t) = \min(M_i^{max}(\theta_i), -K_i^p(\theta_i(t) - \theta_i^f(t)) - K_i^d(\theta_i(t) - \theta_i^{\,f}(t))) \tag{7}$$

$$K_i^p \ge 0 \tag{8}$$

$$K_i^d \ge 0 \tag{9}$$

where t is time, $\theta(t)=\left[\theta_1(t) \quad \theta_2(t) \quad \theta_3(t)\right]^T$ is the angle vector at time t; $\tilde{\theta}(t)=\left[\tilde{\theta}_1(t) \quad \tilde{\theta}_2(t) \quad \tilde{\theta}_3(t)\right]^T$ is the observed values at time t; $\ddot{\theta}(t)=\left[\ddot{\theta}_1(t) \quad \ddot{\theta}_2(t) \quad \ddot{\theta}_3(t)\right]^T$ is the angle acceleration vector at time t; $\dot{\theta}(t)=\left[\dot{\theta}_1(t) \quad \dot{\theta}_2(t) \quad \dot{\theta}_3(t)\right]^T$ is the angle velocity vector at time t; $\tau(t)=\left[\tau_1(t) \quad \tau_2(t) \quad \tau_3(t)\right]^T$ is the torque vector at time t.

This problem is essentially a simulation-based optimization problem [4]. It aims to find reasonable values for K_i^p, K_i^d and $M_i^{\max}$ so that the system, (5) and (6), can replicate the given angular trajectories $\tilde{\theta}(t)$. However, because of non-linearity and non-smoothness of the constraints and high dimensionality of the system model, the gradient and Hessian of the system are difficult if not impossible to obtain or even to approximate, which may exclude the feasibility of using derivative-based methods in our study. Thus the most favorite method to solve this problem may be the direct search (DS) method [3]. This method is able to find optimal solution without computing derivative but evaluating function values. The method of implementing the DS follows a standard approach using the generalized pattern search (GPS) [3] algorithm for bound constrained problems. The same initial guess of the parameter in [2] was employed as the initial value of the GPS algorithm. The maximum iteration was selected to be 150. The parameter, *Complete poll*, was set to *on*, i.e. the algorithm polled all the mesh points at each iteration. The parameters, *tolerance on mesh size (TolMesh), tolerance on function (TolFun), tolerance on variable (TolX)*, and *tolerance on constraints (TolCon)*, were all set to be 1e-8.

The same experimental database employed by Lee and Zhang [1], [2] was used to test this parameter estimation approach. Twenty-eight subjects performed simulated grasping tasks. Reflective markers were placed on the dorsum of the subjects' right hands and the three-dimensional (3D) marker coordinates were recorded by a five-camera Vicon 250 system (Oxford Metrics, UK) at a sampling frequency of 120 Hz to calculate joint angles.

3 Results

The simulated grasping movement, based on the estimated parameters yielded from the GPS algorithm, successfully replicated the measured angular profiles. The

Table 1. Mean and standard deviation of RMSE values (unit: °)

Joint	Digit			
	2	3	4	5
DIP	3.39 (2.10)	4.26 (3.25)	3.92 (3.19)	3.26 (2.39)
PIP	2.63 (1.73)	2.48 (1.86)	2.91 (2.37)	2.80 (1.96)
MCP	2.19 (0.92)	2.80 (1.70)	2.59 (1.84)	2.76 (1.65)

root-mean-square error (RMSE) values for the pair-wise difference between the simulated and measured angular profiles ranged from 2.19 to 4.25° (Table 1). The grand mean of the RMSE values across 28 subjects was 3.00°. In comparison with the results from the heuristic search [2], there was marked improvement in terms of RMSE: the GPS algorithm resulted in smaller RMSE than the heuristic algorithm in 9 out of the 12 joints and smaller grand mean of the RMSE values.

4 Discussion

This was a preliminary study aimed to apply the GPS algorithm to finger movement modeling. The results indicated that the model was able to fit the experimental data closely. However, while the GPS algorithm outperformed the heuristic algorithm in estimating the model parameters, the improvement came at a cost: it took the new algorithm longer to converge on a solution.

The major advantage of the GPS algorithm is that it is not movement-dependent and could be further generalized to the highly nonlinear and non-smooth constrained problems and other difficult optimization problems; the parameter estimation approach could therefore become a very versatile tool in digital human modeling. Studies are underway to explore adapting the proposed algorithm along with the modeling framework itself for other human movements of a higher level of complexity.

Acknowledgments. This work is supported by Centers for Disease Control and Prevention and National Institute of Occupational Safety and Health.

References

1. Lee, S.-W., Zhang, X.: Development and validation of an optimization-based model for power-grip posture prediction. J BIOMECH 38, 1591–1597 (2005)
2. Lee, S.-W., Zhang, X.: Biodynamic modeling and system identification of multi-finger movements. Submitted to J BIOMECH (2006)
3. Torczon, V.: On the convergence of pattern search algorithms. SIAM J OPTIMIZ 7, 1–25 (1997)
4. Kolda, T.G., Lewis, R.M., Torczon, V.: Optimization by direct search: New perspectives on some classical and modern methods. SIAM REV 45, 385–482 (2003)

Mathematical Methods for Shape Analysis and form Comparison in 3D Anthropometry: A Literature Review

Jianwei Niu[1], Zhizhong Li[1], and Gavriel Salvendy[1,2]

[1] Department of Industrial Engineering, Tsinghua University, Beijing, 100084, China
[2] School of Industrial Engineering, Purdue University, West Lafayette, IN., 47907, USA
zzli@tsinghua.edu.cn

Abstract. Form comparison is a fundamental part of many anthropometric, biological, anthropological, archaeological and botanical researches, etc. In traditional anthropometric form comparison methods, geometry characteristics and internal structure of surface points are not adequately considered. Form comparison of 3D anthropometric data can make up the deficiency of traditional methods. In this paper, methods for analyzing 3D other than 2D objects are highlighted. We summarize the advance of form comparison techniques in the last decades. According to whether they are based upon anatomical landmarks, we partition them into two main categories, landmark-based methods and landmark-free methods. The former methods are further sub-divided into deformation methods, superimposition methods, and methods based on linear distances, while the latter methods are sub-divided into shape statistics-based methods, methods based on function analysis, view-based methods, topology-based methods, and hybrid methods. Examples for each method are presented. The discussion about their advantages and disadvantages are also introduced.

Keywords: 3D anthropometry, form comparison.

1 Introduction

There were controversy and suspicion towards quantitative form compassion at one time [1]. As stated by Adams et al. [2], the reasons to adopt quantitative methods are as follows. First of all, although the difference among forms may appear obvious, the significance of the difference cannot be detected accurately by the naked eye. Moreover, quantitative form comparison also has important meaning for the analysis in ontogenetic and phylogenetic sequence. With the emergence and development of several large three-dimensional (3D) anthropometric surveys [3, 4], the past decades have witnessed rapid progress of mathematical methods for form comparison, which leads Rohlf and Marcus [5] to proclaim a “revolution in morphometrics”. Whereas controversy also has arisen over which method is the optimal one for quantifying the morphological difference between forms and statistical analyzing properly [6]. In addition, a generalization framework for 3D modeling has not emerged yet, and it's sometimes difficult or even impossible to interconvert between those existing modeling

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 161–170, 2007.

methods [7]. This brings more challenges for form comparison. The purpose of this paper is to briefly summarize the recent development in the field of 3D form comparison, and bring forward discussion several possible future directions as well.

This paper is organized as follows. Section 1 introduces traditional form comparison methods. Section 2 presents an introduction of various landmarks-based form comparison methods. In section 3, we put our emphasis on landmarks-free form comparison methods. Discussions and speculations for future directions are given in section 4.

2 Traditional Methods

As illustrated by Adams et al. [2], in the 1960's and 1970's, quantitative description of shape, combined with statistical analyses, was adopted by biometricians to describe patterns of shape variations. Such methods, now called traditional morphometrics [8], are usually regarded as the application of multivariate statistical analyses to set of quantitative variables.

Several difficulties remained for methods in this category. For example, a set of linear distances is usually insufficient to capture the geometry of the shape, because the geometric relationships among the variables were not preserved. In the area of population grouping and human accommodation, geometric characteristics and internal structure of human surface points are not adequately considered in traditional methods, leading to design deficiency on fitting comfort [9].

3 Landmark-Based Methods

A summary of the current methods for form comparison using anatomical landmarks were put forward by Richtsmeier et al. [6]. Readers can also refer to other summaries given by Richtsmeier et al. [1], Marcus et al. [10] and Adams et al. [2]. Adapted from the classifying methods for form comparison used by Richtsmeier et al. [6], three main categories of form comparison methods are reviewed in the following section, i.e., deformation methods, superimposition methods, and methods based on linear distances. Piccus et al. [11] provided some simple examples using these methods. All of these methods rely on the identification of anatomical landmarks. The most outstanding disadvantage is that curvature and other geometric characteristic of the surfaces between the landmarks are not preserved in the analysis. Another limitation is that landmarking requires a great deal of physical effort, especially when landmarks not clearly identifiable without palpation of the skin.

3.1 Deformation Methods

Deformation methods deform a form to correspond with the reference form [1, 6]. Three steps should be followed: choose the reference object, employ elastic deformation of templates to deform the target object so that it matches the reference object exactly, and study the deformation to quantify the form difference. Deformation methods include energy matrix [11], Finite Element Scaling Analysis (FESA) [12], Free Form Deformation (FFD) [13] and Thin Plate Splines (TPS) [14], etc.

FESA is widely used in engineering and a deformation method that can be used to determine the difference between forms [15]. In FESA, it's required to subdivide the landmarks located on an object into groups to form so called “elements”. There exists a one-to-one correspondence between landmarks on the initial reference form and the target form. FESA determines the amount of "morphometric strain" required to produce the target model from the reference model [16, 12]. Mochinaru et al. [17, 18] used the FFD method, one computer graphics technology of deforming the shape of objects smoothly by moving control lattice points set around the object [19], to analyze and classify the 3D foot forms of 56 Japanese females and faces for spectacles design. The distortion of the FFD control lattice points that transform the reference form into the target form is defined as the dissimilarity between two forms.

One limitation of deformation methods is that they lack of appropriate statistical procedures and are dependent on landmarks. Richtsmeier et al. [1] pointed out that care must be taken when choose element types in partitioning form for study.

3.2 Superimposition Methods

According to some specified rules, landmark data of the reference form and the target form are arranged into one same coordinate space. Form variation is determined by the displacement of landmarks in the target form from the corresponding landmarks in the reference form. How to choose the specified rule of superimposition has no universal principia, dependent on the particular application [1]. According to Richtsmeier et al. [6], superimposition methods involves three steps: fix one form in a particular orientation and use it as the reference object; translate and rotate the other form so that it matches the reference object according to some criterion; and study the magnitude and direction of difference between forms at each landmark. Based on different criteria for matching, superimposition methods consist of Procrustean approaches [20], Bookstein’s edge matching [21], and roentgenographic cephalometry [22], etc.

One advantage to the method is that interpretable graphic of the difference between forms can be easily accessible. Unfortunately, different matching criteria can result in different superimpositions [6]. Lele and Richtsmeier [23], as well Lele and McCulloch [24] have shown in a mathematically precise fashion that no amount of data can ever tell us how to choose between different superimposition methods. Furthermore, sperimposition methods explicitly depend on landmark data and ignore the geometric information about the area between landmarks [25].

3.3 Methods Based on Linear Distances

These category methods compare linear distances that connect landmark pairs in one form with the corresponding linear distances in another one. Compare each linear distance of the form matrices as a ratio, or an absolute difference or some other metric [6]. Study the matrix of linear distance comparisons to determine the difference between the forms. Examples include Euclidean Distance Matrix Analysis (EDMA) [23] and its variations [26]. The statistics of EDMA have been formally developed [27].

Drawbacks of Linear distance-based methods have been argued by researchers from different aspects [6], e.g., the loss of geometric characteristic of the surfaces between landmarks, a great deal of physical effort in landmarking, short of providing the intelligible graphics available from other methods. It has also been argued that the information contained within the form matrix is redundant [6].

4 Landmark-Free Methods

Adapting the classifying methods of Zhang [28] and Cui & Shi [29], we divide landmark-free methods into five categories, namely shape statistics-based methods, methods based on function analysis, view-based methods, topology--based methods and hybrid methods.

4.1 Shape Statistics-Based Methods

Shape statistics-based methods extract the global characteristics of 3D objects, which are applicable to compare 3D shapes on coarse granulations. They mainly describe the feature relationship between randomly selected surface vertices. Such kind of shape descriptors include shape properties such as number of vertices falling into the subdivision grid [30], weighted points set [31], Gaussian curvatures [32], Extend Gaussian Image (EGI) [33], principal axes of the tensor of inertia [34], shape distribution [35], Parameterized Statistics [7], 3D Shape Spectrum Descriptor [36], point signatures [37], etc. Wong et al. [38] introduced EGI for 3D shape comparison of heads. Because normal vectors in EGI methods are sensitive to 3D postures, researchers made improvements on EGI, e.g., Complex Extend Gaussian Image [39, 40], HEGI (Hierarchical Extended Gaussian Image) [41], and MEGI (More Extended Gaussian Image) [42]. EGI based methods need normalization before feature extraction to be rotation invariant. Since the information of the patch's area and normal vector are involved, these methods have a poor robustness correlative with noise, grid sub-division and grid reduction.

Shape statistics-based methods are simple and fast calculative. Their intrinsic randomicity results in the poor calculation stabilizations. Also it's difficult to visualize the shape characteristics. Generally speaking, curvature-based histogram is sensitive noise, histogram made up by using amount of vertices, patch area or voxel volume is sensitive to grid division of 3D object.

4.2 Methods Based on Function Analysis

Function analysis has been adopted in 3D shape comparison, e.g., Fourier analysis, moment invariants, spherical harmonics analysis and wavelet analysis, etc.

Fourier transform produces orthogonal coefficients which allows for comparatively straightforward statistical analysis. The similarity between shapes is the normalized difference between the Fourier descriptors of the shapes. Ratnaparkhi et al. [43] used the Fourier transform to model 2-D horizontal cross-sections of the human head. Dietmar and Dejan [44], Dekubgette et al [45] implemented spherical parameterization and then regularly sampling both along with longitude and woof directions. Finally they got the spherical Fourier coefficients as the feature. This spherical representation can be implemented in multi-resolution style, and not sensitive with little disturbance of the 3D object's surface. Fourier descriptor based methods ground on the well-established FFT theory and easy to implement. One disadvantage is its un-unique representation of some surface vertices. As stated by Paquet et al [46], Vranic and Saupe [47], and Elad et al. [48], Fourier descriptor is sensitive to the change of the object's edge.

Moment invariants are usually thought as one method based on function analysis as well. Several forms of invariant moments have appeared in the literatures, e.g., Zernike moments [49, 50], and Geometrical moments [48]. Moment invariants give a compact mathematical representation for shapes. One deficiency is that high order moment invariants are too sensitive with little disturbance of the 3D object's surface. Another deficiency is it's difficult to build the relationship between high order moment invariants and shape features. Finally, as shown in the works of Paquet et al [51] and Elad et al. [48], moments are sensitive to the mass distribution of object. Spherical harmonics is often deemed as 2D FFT on unit sphere [47] and introduced for form comparison [52]. The spherical harmonic coefficients can be used to reconstruct an approximation of the underlying object at different levels.

Researchers also presented some other methods based upon function analysis, e.g., wavelet transform [34, 53], Angular Radial Transform [54], 3D Radon transform [55], and Optimal 3D Hough Transform [36].

In order to get invariability to geometry transforms, PCA, e.g., Karhunen-Loeve transform, is usually adopted in most cases. However, the principal axes may not be the ideal representatives for shape differences; normalization will make the object sensitive to local distortion; and much information will be lost due to the substitution of volume features with surface features, etc.

4.3 View-Based Methods

View-based methods are based on the common sense that if two 3D objects look similar to each other from any different projection view, then they are regarded as similar. It’s believed 2D images comprise abundant information for 3D shape. Furthermore, 2D image retrieval methods have been successfully applied in many fields. This invokes some researchers to borrow such methods to 3D shape comparison. Ohbuchi et al. [56] calculated the orthogonal depth projections of a 3D object from 42 different views, and then compare the Generic Fourier Descriptors (GFD) of each projection image one by one. Chen et al [57] presented the idea of Light Fields, using side images other than margin silhouettes. They took 100 photos in 10 light fields for each 3D object from the vertices of one dodecahedron.

View-based methods are sensitive to coordinates' rotation. The robustness is related with the number and distribution of projections.

4.4 Topology-Based Methods

Topology-based methods are usually used in CAD database retrieval, e.g., Reeb graphs [58] or augmented Reeb graphs [59], shock scaffold [60] and 3D sketch [61]. Some intrinsic shortcomings of these methods include difficult to describe the genuine topology of 3D, high calculation cost, huge storage requirements, low differentiating ability between miniature shape variance, low level stabilization [62], and incompatibility for large size database [63]. What is more, 3D object must satisfy some strict conditions, e.g., connectivity. Consequently topology-based methods are scare in real application [28].

4.5 Hybrid Methods

Hybrid feature representation will combine several shape features from different aspects. For example, 3D Hough transform methods [36] convert the surface points from the Euclidean coordinate system to polar coordinates system and qualify the polar coordinates values by using histogram. Morphing-based feature methods [64] build radials starting from the origin to intersect with the surface and calculate the statistic attributes of the intersections. In order to eliminate the influence of posture, the researchers implemented FFT on the extracted features and used the Fourier coefficients as the final shape feature. Hybrid methods must be used carefully since on one hand they may materialize the advantages of one type method, on the other hand they may also enlarge the disadvantages of another type method. One important problem is how to assign the weights of different features. Bustos et al. [65] introduced Entropy Imputity in decision tree to solving the weights' assignment problem.

5 Conclusions and Future Directions

In the last decades, much progress has been made for form comparison. More than ten years ago, there were few logical criteria for choosing the optimal method. As stated by Adams et al. [2], several basic properties, including consistency, systematic bias or veracity of estimation, and statistical power, should be taken into consideration. Now a standard protocol based on Procrustes methods with various extensions has been widely accepted for most applications [2]. The development of special extensions to the selected methods has allowed addressing particular explanatory hypotheses [1], and found applications such as developmental or evolutionary trajectories, hypothetical geometries, and reconstruction of forms when landmarks are missing, etc. It is clear that techniques for form comparison deserve further attention and investigation, and they have the potential to significantly impact researches in a multitude of various fields as well. We anticipate that new exciting improvements of the aforementioned methods will continue.

Acknowledgements. The study is supported by the National Natural Science Foundation of China (No.70571045).

References

1. Richtsmeier, J.T., Cheverud, J.M., Lele, S.R.: Advances in anthropological morphometrics. Annual Review of Anthropology 21, 283–305 (1992)
2. Adams, D.C., Rohlf, F.J., Slice, D.E.: Geometric morphometrics: ten years of progress following the 'revolution'. Italian Journal of Zoology 71, 5–16 (2004)
3. Robinette, K.M., Blackwell, S., Daanen, H., Fleming, S., Boehmer, M., Brill, T., Hoeferlin, D., Burnsides, D.: CAESAR: Final Report, Volume I: Summary. AFRL-HE-WP-TR-2002-0169. United States Air Force Research Lab., Human Effectiveness Directorate, Crew System Interface Division, Dayton Ohio (2002)
4. Bougourd, J., Treleaven, P., Allen, R.M.: The UK national sizing survey using 3d body scanning. In: Eurasia-Tex Conference, Donghua University, Shanghai, China, 2004 (2004)

5. Rohlf, F.J., Marcus, L.F.: A revolution in morphometrics. Trends in Ecology and Evolution 8, 129–132 (1993)
6. Richtsmeier, J.T., Deleon, V.B., Lele, S.R.: The Promise of Geometric Morphometrics. Yearbook of Physical Anthopology 45, 63–91 (2002)
7. Ohbuchi, R., Otagiri, T., Ibato, M., et al.: Shape-similarity search of three-dimensional models using parameterized statistics. In: Proceedings of the Pacific Graphics, Beijing, 2002, pp. 265–273 (2002)
8. Marcus, L.F.: Traditional morphometrics. In: Rohlf, F.J., Bookstein, F.L. (eds.): Proceedings of the Michigan morphometrics workshop. Special Publication. No. 2. University of Michigan Museum of Zoology, Ann Arbor, pp. 77–122 (1990)
9. Li, Z.Z.: Anthropometric topography. In: Karwowski, W. (ed.) International Encyclopedia of Ergonomics and Human Factors, 2nd edn., Taylor & Francis, London (2006)
10. Marcus, L.F., Corti, M., Loy, A., Naylor, G.J.P., Slice, D.E.: Advances in morphometrics. Plenum Press, New York (1996)
11. Piccus, M.E., Smith, G.A., Standley, B.K., Volk, T.L., Wildes, L.B.: Creation of Prototype Aircrew Protection Equipment based On Face Anthropometry. MS Thesis, Air Force Institute of Technology (1993)
12. Cheverud, J., Lewis, J., Bachrach, W., Lew, W.: The measurement of form and variation in form: an application of three-dimensional quantitative morphology by finite-element methods. American Journal of Physical Anthropology 62, 151–165 (1983)
13. Mochimaru, M., et al.: Analysis of 3D human foot forms using the FFD method and its application in grading shoe last. Ergonomics 43(9), 1301–1313 (2000)
14. Bookstein, F.L.: Principal warps: thin-plate splines and the decomposition of deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence 11(6), 567–585 (1989)
15. Richtsmeier, J.T., Morris, G., Marsh, J., Vannier, M.: The biological implications of varying element design in finite-element scaling analyses of growth. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society 1990. pp. 387–388 (1990)
16. Lewis, J., Lew, W., Zimmerman, J.: A non-homogeneous anthropometric scaling method based on finite element principles. Journal of Biomechanics 13, 815–824 (1980)
17. Mochimaru, M., Kouchi, M., Dohi, M.: Analysis of 3-D human foot forms using the free form deformation method and its application in grading shoe lasts. Ergonomics 43(9), 1301–1313 (2000)
18. Mochimaru, M., Kouchi, M.: Proper sizing of spectacle frames based on 3d digital faces. In: Proceedings of XVth Triennial Congree of the International Ergonomics Association, Seoul, Korea, August 24-29, 2003 (2003)
19. Sederberg, T.W.: Free From Deformation of Solid Geometric Models. In: Proceedings of ACM SIGGRAPH'86 in Computers & Graphics, vol. 20(4), pp. 151–160 (1986)
20. Rohlf, F.J.: Shape statistics: Procrustes superimpositions and tangent spaces. of Classification 16, 197–223 (1999)
21. Bookstein, F.L.: Foundations of morphometrics. Annual Reviews of Ecology and Systematics 13, 451–470 (1982)
22. Broadbent, B.S., Broadbent, B.J., Golden, W.: Bolton standards of dentofacial developmental growth. Mosby, St. Louis (1975)
23. Lele, S., Richtsmeier, J.T.: An invariant approach to the statistical analysis of shapes. Chapman & Hall/CRC, Boca Raton (2001)
24. Lele, S., McCulloch, C.: Invariance, identifiability and morphometrics. Journal of the American Statistical Association 971, 796–806 (2002)

25. Rohlf, F.J., Slice, D.: Extensions of the Procrustes method for the optimal suerimposition of landmarks. Systematic Zoology 39(l), 40–59 (1990)
26. Rao, C., Suryawanshi, S.: Statistical analysis of shape of objects based on landmark data. In: Proceedings of the National Academy of Sciences of the USA 1996, vol. 93, pp. 12132–12136 (1996)
27. Lele, S., Richtsmeier, J.T.: Euclidean distance matrix analysis: a coordinate free approach to comparing biological shapes using landmark data. American Journal of Physical Anthropology 98(3), 73–86 (1991)
28. Zhang, Z.Y.: The research for 3D model geometry shape similarity matching. A dissertation submitted to Zhejiang University in the partial fulfillment of the requirements for the degree of Ph.D (2005)
29. Cui, C.Y., Shi, J.Y.: Analysis of Feature Extraction in 3D Model Retrieval. Journal of Computer-aided Design & Computer Graphics 16(7) (2004)
30. Ankerst, M., Kastenmüller, G., Kriegel, H.P., et al.: 3D shape histograms for similarity search and classification in spatial databases. In: Proceedings of the 6th International Symposium on Large Spatial Databases, Hong Kong, 1999, pp. 207–226 (1999)
31. Tangelder, J.W., Veltkamp, R.C.: Polyhedral model retrieval using weighted point sets. International journal of image and graphics 4(1), 209–229 (2003)
32. Suikerbuik, C.A.M., Tangelder, J.W.H., Daanen, H.A.M., Oudenhuijzen, A.J.K.: Automatic feature detection in 3D human body scans. In: Proceedings of the conference SAE Digital Human Modelling for Design and Engineering 2004 (2004)
33. Horn, B.: Extended Gaussian Images. Proceeding of the IEEE 72(12), 1671–1686 (1984)
34. Paquet, E., Rioux, M.: The MPEG-7 standard and the content-based management of three-dimensional data: A case study. In: IEEE international conference on multimedia computing and systems. IEEE Proceedings, Florence, Itatly, 1999, pp. 375–380 (1999)
35. Osada, R., Funkhouser, T., Chazelle, B., et al.: Shape distributions. ACM Transactions on Graphics 21(4), 807–832 (2002)
36. Zaharia, T., Preteux, F.: Hough transform-based 3D mesh retrieval. In: Proceedings of SPIE Conference on Vision Geometry, San Diego, 2001, vol. 4476, pp. 175–185 (2001)
37. Chua, C.S., Jarvis, R.: Point Signatures: A New Representation for 3D Object Recognition. International Journal on Computer Vision 25, 63–85 (1997)
38. Wong, H.S., Cheung, K.K.T., Horace, H.S.I.: An evolutionary optimization approach for 3D human head model classification. In: Proceedings of the 5th ACM SIGMM international workshop on multimedia information retrieval, Berkeley, California, 2003, pp. 94–101 (2003)
39. Kang, S.B., Ikeuchi, K.: Determining 3D object pose using the complex extended Gaussian Image. In: International conference on computer vision and pattern recognition, pp. 580–585 (1991)
40. Kang, S.B., Ikeuchi, K.: The complex EGI: A new representation for 3D pose determination. IEEE Transaction on pattern analysis and machine intelligence 15(7), 702–721 (1993)
41. Xu, J., Suk, M., Ranka, S.: Hierarchical EGI: A new method for object representation. In: Proceedings of the 3rd international conference on signal processing, Beijing, China, 1996, vol. 2, pp. 926–929 (1996)
42. Matsuo, H., Iwata, A.: 3D object recognition using MEGI model from range data. In: IEEE international conference on pattern recognition. Jerusalem, Israel, 1994, pp. 843–846 (1994)
43. Ratnaparkhi, M.V., Ratnaparkhi, M.M., Robinette, K.M.: Size and shape analysis methods for design. Applied Ergonomics 23, 181–185 (1992)

44. Dietmar S., Dejan, V.: 3D model retrieval with spherical harmonics and moments. In : Proceedings of the DAGM 2001, Munichm, Germany, 2001, pp. 392–397 (2001)
45. Dekubgette, H., Hebert, M., Ikeuchi, K.: A spherical representation for the recognition of curved objects. In: Proceedings of IEEE 4th International Conference on Computer Vision, Berlin, 1993, pp. 103–112 (1993)
46. Paquet, E., Murching, A., Naveen, T., et al.: Description of shape information for 2D and 3D objects. Signal Processing: Image Communication 16(1), 103–122 (2000)
47. Vranic, D.V., Saupe, D.: 3D shape descriptor based on 3D Fourier Transform. In: Proceedings of the EURASIP Conference on Digital Signal Processing for Multimedia Communications and Services (ECMCS 2001), Budapest, Hungary, pp. 271–274 (2001)
48. Elad, M., Tal, A., Ar, S.: Content based retrieval of VRML objects: An iterative and interactive approach. In: Proceedings of the 6th Eurographics Workshop in Multimedia, Manchester, 2001, pp. 97–108 (2001)
49. Novotni, M., Klein, R.: 3D Zernike descriptors for content based shape retrieval. In: Proceedings of ACM Symposium on Solid Modeling and Applications, Seattle, Washington, USA, 2003, pp. 216–225 (2003)
50. Marcin, N., Reinhard, K.: 3D Zernike descriptors for content based shape retrieval. In: Proceedings of the eighth ACM symposium of solid modeling and applications, Seattle, Washington, USA, 2003, pp. 216–225 (2003)
51. Paquet, E., Murching, A., Naveen, T., et al.: Description of shape information for 2D and 3D objects. Signal Processing: Image Communication 16(1), 103–122 (2000)
52. Michael, K., Thomas, F., Szymon, R.: Rotation invariant spherical harmonic representation of 3D shape descriptors. In: Proceedings of the Eurographics/ACM SIGGRAPH symposium on geometry processing, 2003, pp. 156–164 (2003)
53. Gain, J., Scott, J.: Fast polygon mesh querying by example. In: International conference on computer graphics and interactive techniques (SIGGRAPH 99), California, USA, 1999, pp. 241–247 (1999)
54. Richard, J., Coeurjolly, D., Baskurt, A.: ART extension for description, indexing and retrieval of 3D objects. In: International conference on pattern recognition, 2004, pp. 23–26 (2004)
55. Daras, P., Zarpalas, D., Tzovaras, D., Strintzis, M.G.: Shape matching using the 3D Radon transform. In: International symposium on 3D data processing. Visualization and Transmission, 2004, pp. 953–960 (2004)
56. Ohbuchi, R., Nakazawa, M., Takei, T.: Retrieving 3D shapes based on their appearance. In: Proceedings of the 5th ACM SIGMM international workshop on multimedia information retrieval, Berkeley, Californica, USA 2003, vol. 63, pp. 39–45 (2003)
57. Chen, D.Y., Tian, X.P., Shen, Y.T., et al.: On visual similarity based 3D model retrieval. Computer Graphics Forum 22(3), 223–232 (2003)
58. Hilaga, M., Shinagawa, Y., Kohmura, T. et al.: Topology matching rfor fully automatic similarity estimation of 3D shapes. In: Computer Graphics Proceedings, Annual conference series, ACM SIGGRAPH, Los Angeles, CA, 2001, pp. 203–212 (2001)
59. Tung, T., Schmitt, F.: Augmented Reeb Graphs for content-based retrieval of 3D mesh model. In: International conference on shape modeling, 2004, pp. 157–166 (2004)
60. Leymarie, F.F., Kimia, B.B.: The Shock Scaffold for representing 3D shape. LNCS, vol. 2059, pp. 216–228. Springer, Heidelberg (2001)
61. Ma, W.C., Wu, F.C., Ouyong, M.: Skeleton extraction of 3D objects with radial basis functions. In: International conference on shape modeling and application, 2003, pp. 207–216 (2003)

62. Funkhouser, T., Min, P., Kazhdan, M., et al.: A search engine for 3D models. ACM Transactions on Graphics 22(1), 83–105 (2003)
63. Shokoufandeh, A., Dickinson, S., Jonsson, C., Bretzner, L., Lindeberg, T.: The representation and matching of qualitative shape at multiple scales. In: Proceedings of European Conference on Computer Vision, Copenhagen, Denmark, 2002, pp. 759–775 (2002)
64. Yu, M., Atmosukarto, I., Leow, W.K., et al.: 3D model retrieal with morphing-based geometric and topological feature maps. In: International conference on computer vision and pattern recognition, 2003, pp. 656–661 (2003)
65. Bustos, B., Keim, D.A., Saupe, D.: Using Entropy Impurity for improved 3D object similarity search. In: International Conference on Multimedia and Expo, 2004, pp. 27–30 (2004)

A Case Study of Multi-resolution Representation of Heads

Jianwei Niu[1], Zhizhong Li[1], and Gavriel Salvendy[1,2]

[1] Department of Industrial Engineering, Tsinghua University, Beijing, 100084, China
[2] School of Industrial Engineering, Purdue University, West Lafayette, IN., 47907, USA
zzli@tsinghua.edu.cn

Abstract. A wavelet analysis based multi-resolution representation method is adopted to establish mathematical description of three-dimensional anthropometric head data in this paper. This method provides flexible description of shapes at different resolution levels. Three-dimensional anthropometric data analysis can then be performed with coarse resolutions, which preserve the major shape components but ignore micro shape components. In a case study of 510 3D head scans, ,quantitative approximation errors, which reflect the approximation of the low-resolution surface to the original one, have been investigated and demonstrated with respect to various decomposition levels.

Keywords: 3D Anthropometry, wavelet analysis, multi-resolution representation.

1 Introduction

In the area of population grouping and human accommodation, many attempts have been made to measure and categorize the scope of human body variation [1, 2, 3]. However, geometric characteristics and internal structure of human surface points are not adequately considered in traditional univariate and bivariate methods [4]. This leads to design deficiency on fitting comfort. It is necessary to understand and characterize the range of human body shape variation using three-dimensional (3D) anthropometric data, since it would strongly support better product and workspace design.

In recent years, several large 3D anthropometric surveys have been conducted [5, 6, 7, 8]. The 3D human body data can have wide application in industry, such as design of shoes [9] and spectacle frames [10], apparel design [11, 3], and helmet evaluation [1, 2]. Unfortunately, there are two great problems in data analysis for 3D scanning anthropometry, i.e., the computational load and enormous storage memory. People look forward to the balance between the conflicting requirements of low data size and high accuracy [12, 13]. One way is to find a method enabling data manipulation and analysis progressively in a coarse-to-fine fashion. Multi-resolution analysis, e.g., wavelet-based analysis, provides such a useful and efficient tool for representing shape and analyzing features at multiple levels of detail and is rapidly becoming an attractive technique in 3D computer graphics and 3D geometric modelling [14, 15]. Wavelet analysis can produce a multi-resolution pyramid representation of the original data. In

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 171–178, 2007.

the wavelet domain, it is easier to manipulate shape features at different scales and analyze data on different resolutions. In this paper a multi-resolution description of human head shape based on wavelet decomposition is introduced. We restrict our attention to the case of cubic B-splines defined on a knot sequence that is uniformly spaced everywhere except at its ends, where its knots have multiplicity of 4. Such B-splines are commonly referred to as endpoint-interpolating cubic B-splines [16]. Approximation errors under different resolutions are illustrated in a case study of 510 head samples.

This paper is organized as follows. Section 2 introduces fundamentals of B-spline wavelets. Section 3 presents the detail procedure of wavelet transformation of 3D anthropometric data. A case study of multi-resolution description of 3D head data is then presented in section 4. Conclusions are given in section 5.

2 Fundamentals of B-Spline Wavelets

For any integer $j \geq 0$, let V^j be the space spanned by the endpoint-interpolating B-spline basis functions of degree 3 with 2^j uniform intervals in domain [0, 1]. In other words, for any $f(x)$ in V^j, $f(x)$ is a polynomial spline function of degree 3 with continuous second derivative and the knot vector defining the function is $(0, 0, 0, 0, 1/2^j, 2/2^j, \ldots, 1 - 1/2^j, 1, 1, 1, 1)$. Since $f(x)$ is also in V^{j+1}, $V^j < V^{j+1}$ holds. The B-spline basis functions in V_j are called scaling functions and the integer j is called resolution level. In this research, the cubic endpoint-interpolating B-spline functions defined on a closed interval are of interest in particular. These functions are discussed in detail in many texts on computer-aided design [17, 18, 19].

For any $f(x)$ and $g(x)$ in V^j, we define their inner product. For each j, we can define W_j as the space that consists of all functions in V_{j+1} that are orthogonal to all functions in V_j under the chosen inner product. The 2^j basis functions of W_j are so-called endpoint-interpolating cubic B-spline wavelets [20]. For any resolution level j and any B-spline function f_{j+1} in V_{j+1}, f_{j+1} can be represented uniquely by the sum of function f_j in Vj and function g_j in W_j [16]. Such decomposing procedure can be applied recursively to the new spline function f_j. Wavelets can be used to represent parametric curves and surfaces simply by computing the wavelet decomposition of each coordinate function, separately [16].

3 Wavelet Analysis of 3D Surface

It's well known that endpoint-interpolating cubic B-spline wavelets defined on a closed interval is one of the important wavelets used in MRA of curves and surfaces [16, 20]. If not specially mentioned, we will regard them as default wavelets in our paper.

3.1 Transformation from Arbitrary Bi-cubic to Quasi-uniform B-Spline Surfaces

In order to ensure high computation efficiency in calculating the control knots, semi-uniform B-Spline surface will be used, since its wavelets base on constant coefficient matrices, which results in no need to calculate the base each time when calculating the control knots of the surface on each decomposition level [16, 20].

Let $S(u, v)$ $(0 \le u, v \le 1)$ be arbitrary B-spline surface defined by m control points and knot vector $(u_0, u_0, u_0, u_0, ..., u_{m-4}, u_{m-3}, u_{m-3}, u_{m-3}, u_{m-3})$ in u direction, and n control points and knot vector $(v_0, v_0, v_0, v_0, ..., v_{n-4}, v_{n-3}, v_{n-3}, v_{n-3}, v_{n-3})$ in v direction. Elements of each knot vector correspond to the parametric knots in u, v direction respectively where the interpolation takes place. Due that we adopt accumulated chord length parameterization, we choose $u_0 = v_0 = 0$, $u_{n-3} = v_{n-3} = 1$. Then select two integers $j_1 > 0$ and $j_2 > 0$ which satisfy $m \le 2^{j_1} + 3$, $n \le 2^{j_2} + 3$. Calculate the data points $S(k/2^{j1}, l/2^{j2})$ $(k = 0, 1, \ldots, 2^{j1}, l = 0, 1, \ldots, 2^{j2})$ for each $(u, v) = (k/2^{j1}, l/2^{j2})$. From the $(2^{j1} + 1) \times (2^{j2} + 1)$ data points $S(k/2^{j1}, l/2^{j2})$, together with additional boundary conditions, the generation of a quasi uniform B-spline surface $S_{j1, j2}(u,v)$ with $(2^{j1} + 3) \times (2^{j2} + 3)$ control points can be implemented by interpolation in U direction and then interpolation in V direction. Finally, we obtain quasi-uniform cubic B-spline surface of the original one.

3.2 Wavelet Decomposition of Quasi-uniform Cubic B-Spline Curves

For any integer $j \ge 0$, let $B_0^{(j)}, B_1^{(j)}, \ldots, B_{2^j+2}^{(j)}$ be cubic B-spline basis functions defined by knots vector $(0, 0, 0, 0, 1/2^j, 2/2^j, \ldots, 1-1/2^j\ 1, 1, 1, 1)$, and V^j be the linear space spanned by the 2^j+3 basis functions. Let the basis functions of V^{j-1} and V^j denoted by $B_k^{(j-1)}(k=0,1,\ldots,2^{j-1}+2)$ and $B_l^{(j)}(l=0,1,\ldots,2^j+2)$ respectively. From the recursive formula of DeBoor [21], each basis function of V^{j-1} could be expressed by linear combination of basis functions of V^j. According to linear algebra theorem [22], there exists one and only one orthogonal complement of V^{j-1} in V^j under the presupposition of standard inner product, denoted by W^{j-1}. Let the basis functions vectors of V^j, W^j denoted by Φ_j, Ψ_j respectively.

For any quasi-uniform cubic B-spline curve $\gamma(t) = (x(t), y(t), z(t))$ $(0 \le t \le 1)$ with $2^j + 3$ control points, where $x(t)$, $y(t)$, and $z(t)$ are in V^j, and j is a non-negative integer,

$$\gamma_j(t) = \Phi_j C_j \ . \tag{1}$$

where C_j is the ith column vector composed of the control points. Assuming Φ_j given, the curve $\gamma_j(t)$ corresponds to the control points sequence C_j uniquely.

If wavelet decomposition is implemented in each coordinate of $\gamma_j(t)$, the decomposition process of $\gamma_j(t)$ can be regarded as solving the linear system of [20]

$$C_j = P_j C_{j-1} + Q_j D_{j-1} = \left[P_j \middle| Q_j \right] \left[\frac{C_{j-1}}{D_{j-1}} \right] . \tag{2}$$

where $[P_j | Q_j]$ denotes a block matrix with left part $[P_j]$ and right part $[Q_j]$, and $\left[\frac{C_{j-1}}{D_{j-1}} \right]$ denotes a block matrix with upper part $[C_{j-1}]$ and nether part $[D_{j-1}]$. In this way, C_j can be decomposed into a low-resolution part C_{j-1} and a detail part D_{j-1}, which is equivalent to decomposing $\gamma_j(t)$ into a low-resolution part $\Phi_{j-1}C_{j-1}$ and a detail part $\Psi_{j-1}D_{j-1}$.

3.3 Wavelet Decomposition of Quasi-uniform Bi-cubic B-Spline Surfaces

Let $S_{j1,j2}(u,v)$ $(0 \le u, v \le 1)$ be a bi-cubic quasi-uniform B-spline surface defined by m scaling functions in V_{j1}, n scaling functions in V_{j2}, and $m\times n$ control points $P_{i,k}$ $(0\le i< m$, $0\le k< n)$, where

$$m = 2^{j_1} + 3, n = 2^{j_2} + 3 . \tag{3}$$

Then

$$S_{j_1,j_2}(u,v) = \Phi_{j_1}(u)P\left[\Phi_{j_2}(v)\right]^T . \tag{4}$$

where Φ_{j1} and Φ_{j2} are row vectors consisting of scaling functions in V_{j1} and V_{j2}, respectively. P is an $m \times n$ matrix generated from control points $P_{i,k}$ $(0\le i< m$, $0\le k < n)$. Using the algorithm introduced by Quak & Weyrich [23], we obtain

$$S_{j_1,j_2}(u,v) = S_{j_1-1,j_2-1}(u,v) + T_{j_1-1,j_2-1}(u,v) . \tag{5}$$

where

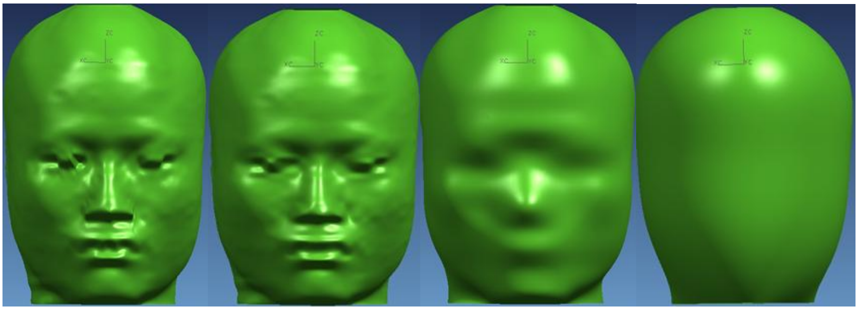
$$S_{j_1-1,j_2-1}(u,v) = \Phi_{j_1-1}(u)(A_{j_1}PA_{j_2}^T)\Phi_{j_2-1}^T(v) . \tag{6}$$

is a bi-cubic quasi-uniform B-spline surface with $(2^{j1-1}+3)\times(2^{j2-1}+3)$ control points and regarded as the low resolution part of the original surface $S_{j1,j2}(u,v)$.

4 Preliminary Analysis of 3D Head Data

510 head samples have been analyzed. Approximation errors under different resolutions are illustrated. In Fig. 1, surfaces (a), (b), (c) and (d) correspond to the original surface and the decomposed surfaces at the first, third and fifth levels, defined by 12468, 4489, 361 and 49 control points, respectively. It can be seen that the decomposed surface can describe the main shape of the original form.

Distances between the original surface and the low-resolution surface can help to evaluate the approximation error of the decomposition. Let V be the set of (u_k, v_l) $(k=0,1,\ldots,m\text{-}3;\ l=0,1,\ldots,n\text{-}3)$, $P(u,v)$ be the data point at the knot of (u_k, v_l) on original surface $S(u,v)$ and let S' be the new surface, then we introduce the idea of



(a) (b) (c) (d)

Fig. 1. Multi-resolution modeling of an arbitrary bi-cubic surface

Averaged Distance (*AD*) between the original surface and the new surface. *AD* can be defined as

$$AD = \underset{(u,v)\in V}{average}\left|P(u,v) - S^{'}\right| \tag{7}$$

Averaged Relative Error (*AvgRE*) can be defined as

$$AvgRE = \frac{AD}{\underset{(x,y,z)\in(u,v)}{\max}(\max(x)-\min(x),\max(y)-\min(y),\max(z)-\min(z))} \tag{8}$$

Table 1. Decomposition results of one sample

surface	decomposition level	*MD* (mm)	*MaxRE* (%)	*AD* (mm)	*AvgRE* (%)	control points number
$S_{6,6}(u,v)$	1	1.813	0.788	0.109	0.047	$(2^6+3)*(2^6+3)=$ 4489
$S_{5,5}(u,v)$	2	2.754	1.198	0.169	0.073	$(2^5+3)*(2^5+3)=$ 1225
$S_{4,4}(u,v)$	3	5.696	2.477	0.447	0.194	$(2^4+3)*(2^4+3)=361$
$S_{3,3}(u,v)$	4	13.899	6.043	1.078	0.469	$(2^3+3)*(2^3+3)=121$
$S_{2,2}(u,v)$	5	14.791	6.431	3.367	1.464	$(2^2+3)*(2^2+3)=49$

Note: *AD-Average Distance; AvgRE*-average relative error

Table 2. Descriptives of distance and relative error at various decomposition levels (*N*=510)

Descriptives of averaged distance (mm)	Decomposition level (control points number)				
	1 (4489)	2 (1225)	3 (361)	4 (121)	5 (49)
Min	0.100	0.134	0.308	0.782	2.792
Max	0.125	0.235	0.622	1.376	3.732
M	0.108	0.170	0.431	1.047	3.212
SD	0.004	0.015	0.044	0.090	0.172
Descriptives of averaged relative error (%)					
Min	0.040	0.057	0.134	0.348	1.210
Max	0.056	0.098	0.254	0.625	1.617
M	0.047	0.074	0.186	0.453	1.389
SD	0.002	0.007	0.019	0.041	0.079

AD and *AvgRE* both together reflect the overall approximation of the new surface to the original surface. In this research, AD_{j1j2} denotes *AD* between the original surface and the low-resolution surface, whose control points' number is $(2^{j1}+3)\times(2^{j2}+3)$.

Table 1 shows the approximation errors at various decomposition levels for one sample from our database. The calculation environment is Intel Pentium 4 CPU (2.40GHz), 512MB RAM, and Unigraphics, V18.0. It can be seen that with the recursive decomposition, the number of control points decreases while the approximate errors increase.

The descriptives of $AD_{j1,j2}$ and $AvgRE_{j1,j2}$ at various decomposition levels are shown in Table 2. The table shows that the decomposed surfaces can preserve the main shape of the original surface. However, there is a significant increase of the descriptives at level 5 (7×7 poles), which may suggest level 4 as the appropriate decomposition level of head for future data analysis, depending on specific applications and still needs more investigation.

5 Conclusions

In this paper, we proposed a multi-resolution pyramid description for 3D anthropometric data, which enables data analysis of 3D anthropometric data progressively in a coarse-to-fine fashion in the wavelet domain. The fundamental properties of wavelets guarantee that the lower dimensionality can preserve the major information and describe the main shape. The multi-resolution description of 3D anthropometric data makes use of not only the coordinates of data points, but also the geometric characteristics and internal structure among the points. The proposed method in this paper is demonstrated through 510 head data with the aid of a Unigraphics user function. One quantitative approximation error, i.e. MRE, which reflects the approximation of the low-resolution surface to the original one, has been investigated and demonstrated with respect to decomposition levels. Though we only examine head data, which can be used in helmet design, it's also expected to be a helpful universal method for data analysis of other human body surfaces, sizing of shape-fitting wearing items, clinical practice or ergonomically work place design in the future, etc.

Acknowledgements. The study is supported by the National Natural Science Foundation of China (No.70571045).

References

1. Whitestone, J.J.: Design and evaluation of helmet systems using 3D data. In: Proceeding of the Human Factors and Ergonomics Society 37th Annual Meeting. Seattle, USA, 1993, pp. 64–68 (1993)
2. Meunier, P., Tack, D., Ricci, A.: Helmet accommodation analysis using 3D laser scanning. Applied Ergonomics 31(4), 361–369 (2000)
3. Wang, M.J., Lin, Y.C., Wu, W.Y.: The development of automated tailoring system using 3D scanning data. In: IEA2003, August 24-29, 2003, Seoul, Korea (2003)
4. Li., Z.Z.: Anthropometric topography. In: Karwowski, W. (ed.) International Encyclopedia of Ergonomics and Human Factors, 2nd edn., Taylor & Francis, London (2006)
5. Richtsmeier, J.T., Cheverud, J.M., Lele, S.: Advances in anthropological morphometrics. Annual Review of Anthropology 21, 283–305 (1992)
6. Robinette, K.M., Vannier, M.W., Rioux, M., Jones, P.R.M.: 3-D surface anthropometry: review of technologies, agard-ar-329, Dayton Ohio (1997)
7. Robinette, K.M., Blackwell, S., Daanen, H., Fleming, S., Boehmer, M., Brill, T., Hoeferlin, D., Burnsides, D.: CAESAR, Final Report, Volume I: Summary. AFRL-HE-WP-TR-2002-0169. United States Air Force Research Lab, Human Effectiveness Directorate, Crew System Interface Division, Dayton Ohio (2002)
8. Bougourd, J., Treleaven, P., Allen, R.M.: The UK national sizing survey using 3d body scanning. In: Eurasia-Tex Conference, Donghua University, Shanghai, China, 2004 (2004)
9. Mochimaru, M., Kouchi, M., Dohi, M.: Analysis of 3-D human foot forms using the free form deformation method and its application in grading shoe lasts. Ergonomics 43(9), 1301–1313 (2000)
10. Kouchi, M., Mochimaru, M.: Analysis of 3D human face forms and spectacle frames based on average forms. In: Proceedings of the Digital Human Modeling Conference. Munich, 2002, pp. 69–89 (2002)
11. Paquette, S.: 3D scanning in apparel design and human engineering. Computer Graphics and Application 5, 11–15 (1996)
12. Heckbert, P.S.: Multiresolution surface modeling. In: Proceedings of ACM SIGGRAPH 97 (1997)
13. Garland, M.: Multi-resolution modeling: survey and future opportunities. In: Eurographics'99. In: State of The Art Reports of the Annual Conference of the European Association for Computer Graphics, pp. 111–131 (1999)
14. Kobbelt, L., Campagna, S., Vorsatz, J., Seidel, H.P.: Interactive multiresolution modeling on arbitrary meshes. In: Proceedings of ACM SIGGRAPH 98, pp. 105–114 (1998)

15. Hahmann, S., Elber, G.: Constrained multiresolution geometric modeling. In: Dodson, N.A., Sabin, M.A., Floater, M.I. (eds.) Advances in Multiresolution for Geometric Modeling, Springer, London (2004)
16. Finkelstein, A., Salesin, D.H.: Multi-resolution curves. In: Proceedings of the 21st Annual Conference on Computer Graphics and interactive Techniques SIGGRAPH '94, ACM Press, New York (1994)
17. Farin, G.: Curves and Surfaces for Computer Aided Geometric Design, 3rd edn. Academic Press, New York (1992)
18. Farin, G.: Curves and surfaces for CAGD: a practical guide. Kaufmann, San Francisco (2002)
19. Hoschek, J., Lasser, D.: Fundamentals of computer aided geometric design, 3rd edn. A. K. Peters, Wellesley (1992)
20. Stollnitz, E.J., DeRose, T.D., Salesin, D.H.: Wavelet for computer graphics: a primer. IEEE Computer Graphics and Applications 15(3), 76–84 (1995) and 15(4), 75-85 (1995)
21. DeBoor, C.: A Practical Guide of Splines, Applied Mathenatical Sciences Series, 27 (1978)
22. Chui, C.K.: An Introduction to Wavelets. Academic Press, Boston (1992)
23. Quak, E., Weyrich, N.: Decomposition and reconstruction algorithms for spline wavelets on a bounded interval. Applied and Computional Harmonic Analysis 1, 217–231 (1994)

The Application of Kane Equation in the Impact Prediction of Human Motion

Mu Qiao, Chunxin Yang, and Xiu-gan Yuan

Institute of Human-Machine & Environment,
Beijing University of Aeronautics and Astronautics,
100083 Beijing, P.R. China
Mu Qiao sonic7656@hotmail.com

Abstract. This paper presents a computational procedure for deriving and solving the governing dynamical equations of multi-body human systems subjected to impact. The procedure is developed on the basis of the assumption that the duration of the impact is very short, during which the configuration of the system remain the same, although the velocities of the system have undergone significant changes. These assumptions lead to a set of linear algebraic equations for the velocity increments. These equations may then be solved to obtain initial conditions for the analysis of the subsequent motion of the multi-body human system. A calculating task of ejecting seat problem is presented to illustrate and validate the procedure.

Keywords: multi-body, Kane equation, impulse, and ejection.

1 Introduction

The prediction of human motion via classic multi-body system dynamics has received considerable attention from dynamicists in the last three decades. References [l-9] and their bibliographies summarize a vast number of research efforts and literatures. The application of biomechanical whole body models, which are normally composed of rigid segments and linked by simple kinematic connections, ranges from jumping, landing to automobile crash. [12-17]. Admittedly, along with the development and mature of the commercial computational multi-body dynamics software, the pure analytical and derivative process has been overshadowed, even the core algorithm of these software, which is based on the equation of analytical mechanics such as Lagrange equation of first kind and Kane equation, has been obscured.

The purpose of this paper is to clearly present the core of this algorithm in human motion mechanics and its application in the task of motion prediction. We take the phenomena of impaction into consideration and formulate a procedure to make efficient computational analysis of it. With the use of vector differentiating algorithm Kane equation [10] is ideally suited for the analysis of multi-body system dynamics [9]. They lead to equations which are readily converted into algorithms for numerical analyses.

By integrating the dynamical equations obtaining from the impact time interval we achieve the balance of generalized impulses and generalized momentum. These equations in turn lead to a set of linear algebraic equations for velocity increments.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 179–188, 2007.

The solution of these equations consequently provides the basis for predicting the movement of the system during the post-impact period. The fundamental assumption is that during impact 'impulsive' forces occurs over a short time interval. It is assumed that the system configuration remains essentially unchanged during this short impact interval, while that system velocities change dramatically.

This paper falls into four parts with the first part providing a useful summary of multi-body dynamics analysis to the sequel. The next part illustrates the examples and concluding remarks.

2 Preliminary Considerations

2.1 Multi-body Systems

A multi-body system is a collection of bodies with a given connection configuration. The system contains a chain, tree or loop topology and may be subjected to both geometric and kinematic constraints. Forces can be applied to the bodies of the system both externally (such as through gravity and contact forces) and internally between the bodies (such as through spring, damping, and actuator forces).

2.2 Kinematics

Suppose a multi-body system has N bodies with n degrees of freedom (DOFs). Let the movement of such a system be described by n generalized coordinates x_j (j = 1,…, n). (These coordinates will, in general, include both translation and rotation variables.) Name n variables $\dot{x}_j$ (j = 1,…, n), called 'generalized speeds' [10], being introduced such as that $\dot{x}_j$ are the derivatives of the generalized coordinates, respectively. It is often convenient to have $\dot{x}_j$, associated with the rotational DOFs, be the components of the relative angular velocities of the bodies. In Ref. [9] it shows that the angular velocity $\boldsymbol{\omega}_k$ and the center of mass (COM) velocity $\mathbf{v}_k$ of a typical body B_k (k = 1,…, N) of the system may be expressed as:

$$\boldsymbol{\omega}_k = \sum_{i=1}^{3}\left(\sum_{j=1}^{n}\omega_{i,j,k}\dot{x}_j\right)\mathbf{e}_i^r \quad (1)$$

and

$$\mathbf{v}_k = \sum_{i=1}^{3}\left(\sum_{j=1}^{n}v_{i,j,k}\dot{x}_j\right)\mathbf{e}_i^r \quad (2)$$

Here, the $\mathbf{e}_i^r$ (i = 1, 2, 3) are mutually perpendicular unit vectors fixed in an inertia frame R. The coefficients $\omega_{i,j,k}$ and $v_{i,j,k}$ (i = 1, 2, 3, j = 1, …, n, k = 1,…, N) are called 'partial angular velocity components' and 'partial velocity components' [9], forming arrays which depend merely upon the geometrical parameters of the system.

By differentiating in eqn (1) and (2), the angular acceleration and the COM acceleration of B_k in R are expressed as

$$\boldsymbol{\alpha}_k = \sum_{i=1}^{3}\left(\sum_{j=1}^{n}\omega_{i,j,k}\ddot{x}_j + \sum_{j=1}^{n}\dot{\omega}_{i,j,k}\dot{x}_j\right)\mathbf{e}_i^r \quad (3)$$

and

$$\mathbf{a}_k = \sum_{i=1}^{3}\left(\sum_{j=1}^{34}\dot{v}_{i,j,k}\dot{x}_j + \sum_{j=1}^{34}v_{i,j,k}\ddot{x}_j\right)\mathbf{e}_i^r \quad (4)$$

We could observe that $\omega_{i,j,k}$ and $v_{i,j,k}$ are coefficients of the generalized speeds $\dot{x}_j$. That is

$$\sum_{i=1}^{3}\omega_{i,j,k}\mathbf{e}_i^r = \frac{\partial \boldsymbol{\omega}_k}{\partial \dot{x}_j} = \frac{\partial \boldsymbol{\alpha}_k}{\partial \ddot{x}_j} \stackrel{D}{=} \boldsymbol{\omega}_k{}^j \quad (5)$$

and

$$\sum_{i=1}^{3}v_{i,j,k}\mathbf{e}_i^r = \frac{\partial \mathbf{v}_k}{\partial \dot{x}_j} = \frac{\partial \mathbf{a}_k}{\partial \ddot{x}_j} \stackrel{D}{=} \mathbf{v}_k{}^j \quad (6)$$

where $\boldsymbol{\omega}_k^j$ is the ‘partial angular velocity’ of B_k and $\mathbf{v}_k^j$ is its ‘partial velocity’ of COM in R.

2.3 Kinetics

Let the multi-body system be subjected to a force field which may be represented on a typical body B_k by a single force $\mathbf{F}_k$ passing through COM G_k together with a couple $\mathbf{M}_k$. Then the contribution of $\mathbf{F}_k$ and $\mathbf{M}_k$ from all the bodies to the generalized active force F_j (j = 1,…, n) associated with generalized speed $\dot{x}_j$ of the entire system is

$$F_j = \sum_{k=1}^{N}\left(\frac{\partial \mathbf{v}^k}{\partial \dot{x}_j}\cdot\mathbf{F}_k + \frac{\partial \boldsymbol{\omega}^k}{\partial \dot{x}_j}\cdot\mathbf{M}_k\right) = \sum_{k=1}^{N}\sum_{i=1}^{3}\left(v_{i,j,k}F_k{}^i + \omega_{i,j,k}M_k{}^i\right) \quad (7)$$

F_k^i and M_k^i are the $\mathbf{e}_i^r$ components of $\mathbf{F}_k$ and $\mathbf{M}_k$.

Similarly, let the inertia forces on B_k be represented by a single force $\mathbf{F}_k^*$ passing through G_k together with a couple with torque $\mathbf{T}_k^*$. Then $\mathbf{F}_k^*$ and $\mathbf{T}_k^*$ may be expressed as:

$$\mathbf{F}_k^* = -\, m_k\mathbf{a}^k . \quad (8)$$

$$\mathbf{T}_k^* = -\, \mathbf{I}^k \cdot \mathbf{a}^k - \boldsymbol{\omega}^k \times \mathbf{I}^k \cdot \boldsymbol{\omega}^k . \quad (9)$$

where m_k is the mass of B_k and $\mathbf{I}^k$ is the central inertia dyadic.

The generalized inertia force F_j^* associated with generalized speed $\dot{x}_j$ of the entire system is given as

$$F_j^* = \sum_{k=1}^{N}\left(\mathbf{F}_k^*\cdot\frac{\partial \mathbf{v}^k}{\partial \dot{x}_j} + \mathbf{T}_k^*\cdot\frac{\partial \boldsymbol{\omega}^k}{\partial \dot{x}_j}\right) = \sum_{k=1}^{N}\left[\left(\sum_{i=1}^{3}v_{i,j,k}\,{}^*F_k{}^i\right) + \left(\sum_{i=1}^{3}\omega_{i,j,k}\,{}^*T_k{}^i\right)\right] \quad (10)$$

where F_k^i and T_k^i are the $\mathbf{e}_i^r$ components of $\mathbf{F}_k^*$ and $\mathbf{T}_k^*$.

Finally, from Kane equations [8], the dynamical equations of motion may be written as:

$$F_j^* + F_j = 0\ . \tag{11}$$

3 Impact Analysis

An 'impulsive force' is defined as a force which is large but acts only over a short time interval. The term 'large' means that the impulsive force has a significantly greater magnitude than other forces exerted on the system during the impact time interval. Suppose a multi-body system has N bodies and n DOFs which can be characterized by n generalized coordinates x_j (j = 1,…, n). The definition of n generalized speeds $\dot{x}_j$ is the same as the above, which represents the relative angular velocity components and relative displacement derivatives among adjoining bodies and between a reference body of the system and the inertia frame. Let the impulsive forces on a typical body B_k be equivalent to a force $\mathbf{F}_k$ passing through COM G_k, together with a couple with torque $\mathbf{M}_k$. Then the generalized impulse I_j acting on the system and associated with the generalized speed $\dot{x}_j$ is defined as:

$$I_j = \sum_{k=1}^{N}\left(\frac{\partial \mathbf{v}^k}{\partial \dot{x}_j}\cdot\int_{t_1}^{t_2}\mathbf{F}_k\,dt + \frac{\partial \boldsymbol{\omega}^k}{\partial \dot{x}_j}\cdot\int_{t_1}^{t_2}\mathbf{M}_k\,dt\right)\ . \tag{12}$$

where $[t_1, t_2]$ is the impact time interval.

The system generalized momentum $P_j(t)$ for the generalized speed $\dot{x}_j$ is defined as:

$$P_j(t) = \sum_{k=1}^{N}\left[m_k\frac{\partial \mathbf{v}^k}{\partial \dot{x}_j}\cdot\mathbf{v}_k(t) + \frac{\partial \boldsymbol{\omega}^k}{\partial \dot{x}_j}\cdot\mathbf{I}^k\cdot\boldsymbol{\omega}_k(t)\right]\ . \tag{13}$$

By integrating Kane equation [eqn (11)] over the impact time interval we have:

$$\int_{t_1}^{t_2}F_j^*dt + \int_{t_1}^{t_2}F_j dt = 0\ . \tag{14}$$

By using eqn (7), the second term in eqn (14) is written as

$$\int_{t_1}^{t_2}F_j dt = \sum_{k=1}^{N}\left(\mathbf{v}_k{}^j\cdot\int_{t_1}^{t_2}\mathbf{F}_k dt + \boldsymbol{\omega}_k{}^j\cdot\int_{t_1}^{t_2}\mathbf{M}_k dt\right) = I_j\ . \tag{15}$$

By using eqn (8), (9) and (11), the inertia force term of eqn (14) may be written as:

$$\int_{t_1}^{t_2}F_j^*dt = \int_{t_1}^{t_2}\left[\sum_{k=1}^{N}\mathbf{v}_k{}^j\cdot\left(-m_k\mathbf{a}^k\right) + \boldsymbol{\omega}_k{}^j\cdot\left(-\mathbf{I}^k\cdot\boldsymbol{\alpha}^k - \boldsymbol{\omega}^k\times\mathbf{I}^k\cdot\boldsymbol{\omega}^k\right)\right]dt\ . \tag{16}$$

By using the previous assumption of essentially unchanged configurations during the impact, the first term may be expressed as:

$$\int_{t_1}^{t_2}\sum_{k=1}^{N}\mathbf{v}_k{}^j\cdot\left(-m_k\mathbf{a}^k\right)dt = -\sum_{k=1}^{N}m_k\mathbf{v}_k{}^j\cdot\int_{t_1}^{t_2}\mathbf{a}^k dt = -\sum_{k=1}^{N}m_k\mathbf{v}_k{}^j\cdot\left[\mathbf{v}_k(t_2) - \mathbf{v}_k(t_1)\right]\ . \tag{17}$$

where $\mathbf{v}_k(t_1)$ and $\mathbf{v}_k(t_2)$ are the velocities of G_k at the beginning and the end of the impact.

Similarly, the second term on the right side of eqn (16) can be denoted as:

$$\int_{t_1}^{t_2}\sum_{k=1}^{N}\boldsymbol{\omega}_k{}^j\cdot\left(-\mathbf{I}^k\cdot\boldsymbol{\alpha}^k\right)dt=-\sum_{k=1}^{N}\boldsymbol{\omega}_k{}^j\cdot\mathbf{I}^k\cdot\int_{t_1}^{t_2}\boldsymbol{\alpha}^k dt=-\sum_{k=1}^{N}\boldsymbol{\omega}_k{}^j\cdot\mathbf{I}^k\cdot\left[\boldsymbol{\omega}_k\left(t_2\right)-\boldsymbol{\omega}_k\left(t_1\right)\right]. \quad (18)$$

where $\boldsymbol{\omega}_k(t_1)$ and $\boldsymbol{\omega}_k(t_2)$ are angular velocities of B_k at the beginning and the end of the impact.

Finally, since the duration of impact is very short the third term on the right side of eqn (16) can be definitely neglected according to its order of magnitude which is less than the first two,

$$\begin{aligned}&\int_{t_1}^{t_2}F_j^*dt\\&=-\left[\sum_{k=1}^{N}m_k\mathbf{v}_k{}^j\cdot\left(\mathbf{v}_k\left(t_2\right)\right)+\sum_{k=1}^{N}\boldsymbol{\omega}_k{}^j\cdot\mathbf{I}^k\cdot\left(\boldsymbol{\omega}_k\left(t_2\right)\right)\right]+\left[\sum_{k=1}^{N}m_k\mathbf{v}_k{}^j\cdot\left(\mathbf{v}_k\left(t_1\right)\right)+\sum_{k=1}^{N}\boldsymbol{\omega}_k{}^j\cdot\mathbf{I}^k\cdot\left(\boldsymbol{\omega}_k\left(t_1\right)\right)\right]\\&=-P_j\left(t_2\right)+P_j\left(t_1\right).\end{aligned} \quad (19)$$

where $P_j(t_1)$ and $P_j(t_2)$ are generalized momentum of the system with respect to the generalized speeds $\dot{x}_j$ at beginning and end of the impact.

By substituting from eqn (16) and (19) into eqn (14) we obtain

$$P_j(t_2)-P_j(t_1)=I_j\,. \quad (20)$$

That is, during the impact time interval, the changes in generalized momentum are equal to the generalized impulses for each generalized speed.

Substituting eqns (1), (2), (5) and (6) into eqn (13) the generalized momentum P_i may be expressed as:

$$P_i\left(t\right)=\sum_{k=1}^{N}\left[m_k\sum_{l=1}^{3}\left(v_{l,i,k}\sum_{j=1}^{N}v_{l,j,k}\dot{x}_j\right)\right]+\sum_{k=1}^{N}\left\langle\sum_{m=1}^{3}\omega_{m,i,k}\left\{\sum_{l=1}^{3}\left[{}^rI_{m,l}^k\left(\sum_{j=1}^{N}\omega_{m,j,k}\dot{x}_j\right)\right]\right\}\right\rangle. \quad (21)$$

where the ${}^rI_{m,l}{}^k$ are the $\mathbf{e}_m{}^r$ and $\mathbf{e}_l{}^r$ components of central inertia dyadic $\mathbf{I}^k$

Substituting eqn (21) into eqn (20) the governing dynamical equation may be written as:

$$\sum_{j=1}^{n}A_{i,j}\dot{x}_j\left(t_2\right)-\sum_{j=1}^{n}A_{i,j}\dot{x}_j\left(t_1\right)=\sum_{j=1}^{n}A_{i,j}\Delta\dot{x}_j=I_i \quad (i,j=1,\ldots,n)\,. \quad (22)$$

where the $A_{i,j}$ are generalized masses given by

$$A_{i,j}=\sum_{k=1}^{N}\left[m_k\sum_{l=1}^{3}v_{l,i,k}v_{l,j,k}+\sum_{m=1}^{3}\omega_{m,i,k}\left(\sum_{l=1}^{3}{}^rI_{m,l}^k\omega_{l,j,k}\right)\right] \quad (i,j=1,\ldots,n)\,. \quad (23)$$

where $\dot{x}_j(t_1)$ and $\dot{x}_j(t_2)$ are generalized speeds $\dot{x}_j$ at time t_1 and t_2, respectively, and where $\Delta\dot{x}_j$ is the change in the generalized speed during the impact time interval.

Equations (22) form a set of linear algebraic equations, which in turn form the basis for the analysis of the effects of impulsive forces upon multi-body systems. By solving these equations for the $\Delta \dot{x}_j$ we obtain

$$\Delta \underline{\dot{x}} = \underline{A}^{-1} \underline{I} \quad . \tag{24}$$

$\Delta \dot{x}_j$ can be used to determine COM velocity changes and body angular velocity changes. These quantities in turn produce initial conditions for the subsequent motions of the system. A hallmark and benefit of eqn (24) is that there are exactly the same numbers of equations as there are DOFs of the system. Hence, the unknown generalized speeds are uniquely determined. The core algorithm of governing equation for most commercial software is no more than converting the formulas mentioned above into calculating program.

4 Calculating Task

Specifically, provide that a 15 (N) rigid segments Hanavan human body model (Fig. 1) represent the torso, head, hand, foot and upper and lower arms and legs. It is a three-dimensional model with ball-and-socket joints except for the elbows, knees, waist and ankle which were modeled as hinge joints. The ball-and-socket joint is described by three Cardano angles α, β, γ relative to the principal axis ${}^k\mathbf{e}_1$, ${}^k\mathbf{e}_2$, ${}^k\mathbf{e}_3$ of the body segment, and the hinge joint by one rotational variable. The whole DOF of this human model is 34 (n) and x_1, x_2, x_3 and x_4, x_5, x_6 are the translation and rotation of B_1 (upper torso) relative to the inertia reference frame R, respectively.

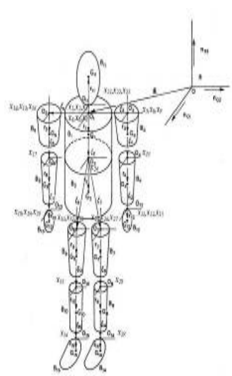


Fig. 1. Hanavan human model

The model is subject to impulsive forces, in this example we employ the equations developed in the preceding section and examine the model's dynamic response on impact forces. A basic configure -- sitting position, was used. An impulse of 742 N·s was applied in several directions acting through a point located in the midway on a line connecting the hip joints. This configuration could be used to study ejection seat problems in a jet fighter or problems encountered in automobile crashes [18], as depicted in Fig. 2. When the hip is struck the body segment will begin to rotate relative to each other. The rate of rotation is determined by eqn (23). Specifically eqn (24) we use needs not be integrated but which can be solved algebraically, i.e. based upon specification of the impulsive force the resulting generalized speed increments in the bodies of the model are obtained. (This solution for joint velocities was determined from link rotations using a self-developed FORTRAN 90 multi-body dynamics simulation computer program.) The results presented in Table 1 shows the increment change in $\dot{x}_j$, say $\Delta\dot{x}_j$.

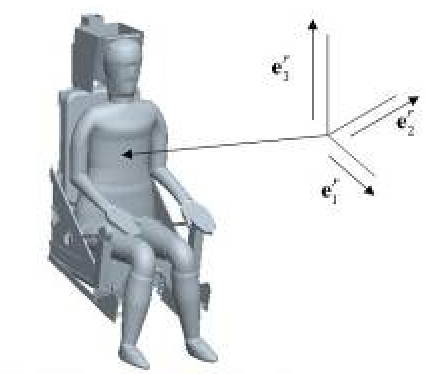


Fig. 2. Human-Seat Model

4.1 Discussion

As eqn (24) is a set of linear equations, it is useful to obtain the three sets of the general velocity increment when the unit impulsive forces is acted respectively along with each three unit vector, say $\mathbf{e}_1^r$, $\mathbf{e}_2^r$ and $\mathbf{e}_3^r$ in reference R. In more general case when the impulse is of large magnitude and has different acting direction, the increment of generalized speed is obtained by summing the three sets which are multiplied with its own coefficient which is exactly the $\mathbf{e}_i^r$ (i = 1, 2, 3) component of the impulse.

Special note should be paid to the velocities of head for the five different cases tabulated. The most sever case results in the head sustaining an angular velocity change of 130 rad/s when in the impulse is in the $\mathbf{e}_1^r$ direction (defined in Fig. 2). When the skull is subject to associated centrifugal inertia forces due to angular velocities of such magnitude, the resulting tensile and shear stresses in the brain are of magnitudes that exceed the ultimate strength of dura matter and of the arteries. Hence it is necessary to restrict the head of the pilot to the backrest of the eject seat.

Table 1. Response to impulse forces, a These values are in m/s; the remainder are in rad/s

		$\mathbf{I}=742\mathbf{e}_3^r$N·s	$\mathbf{I}=742\mathbf{e}_2^r$N·s	$\mathbf{I}=742\mathbf{e}_1^r$N·s	$\mathbf{I}=524\mathbf{e}_1^r$ $+524\mathbf{e}_3^r$N·s	$\mathbf{I}=428\mathbf{e}_1^r$ $+428\mathbf{e}_2^r$ $+428\mathbf{e}_3^r$ N·s
j	x_j	$\Delta\dot{x}_j$	$\Delta\dot{x}_j$	$\Delta\dot{x}_j$	$\Delta\dot{x}_j$	$\Delta\dot{x}_j$
1	-	-0.000340a	0.0000457a	18.9a	13.4a	10.9a
2	-	0a	17.2a	0a	0a	9.93a
3	-	13.0a	0a	0.00157a	9.20a	7.51a
4	-	0	-7.52	0	0	-4.34
5	-	0.0101	0	25.0	17.7	14.4
6	-	0	0.00441	0	0	0.00255
7	-	0	-38.6	0	0	-22.3
8	-	0.00662	-0.00212	60.3	42.6	34.8
9	-	0	-84.1	0	0	-48.6
10	-90	112.	0	-0.00456	79.2	64.7
11	-	0	184.	0	0	106.
12	-	62.5	0.00901	0.00250	44.2	36.1
13	-	0	0.0311	0	0	0.0180
14	-	0	-38.6	0	0	-22.3
15	-	0.00662	0.00658	60.3	42.6	34.8
16	-	0	-84.1	0	0	-48.6
17	-90	112.	-0.0121	-0.00456	79.2	64.7
18	-	0	184.	0	0	106.
19	-	62.5	0.0206	0.00250	44.2	36.1
20	-	0	0.0311	0	0	0.0180
21	-	0	95.5	0	0	55.1
22	-	0.0167	0	-130	-91.9	-75.0
23	-	0	-0.00307	0	0	-0.00177
24	-	0	-0.00240	0	0	-0.00139
25	-	0.0317	-0.000134	-2.64	-1.84	-1.51
26	-90	33.0	-0.50	-25.0	5.66	4.33
27	-	0	-0.000138	0	0	-0.0000800
28	90	-33.0	8.48	29.8	-2.26	3.05
29	-	0.0806	18.5	-78.8	-55.7	-34.8
30	-	-0.0123	-0.000134	-2.37	-1.68	-1.38
31	-90	33.0	1.50	-25.0	5.66	5.48
32	-	0	-0.000130	0	0	-0.0000750
33	90	-33.0	-8.48	29.8	-2.26	-6.74
34	-	-0.0806	18.5	-78.8	-55.7	-34.9

5 Conclusion

We have presented a general procedure for the research on the prediction of human body motion in impact. This procedure and the resulting governing dynamical equations are also applicable in a wide variety of situations, i.e. given any force field applied to the human body, or given any prescribed motion of the human body, or given any combination of these the governing dynamical equations will lead to a determination of the unknown forces and motion. What is more complicated computer simulation focuses on the following core formula,

$$\underline{A}\ddot{\underline{x}} = \underline{f} + \underline{F} \quad . \tag{25}$$

Column matrix f represents the third term of eqn (16), which has not been neglected in more accurate analysis. $\underline{F}$ represents the item (10).

$$f_l = -\sum_{k=1}^{N}\left[\sum_{i=1}^{3} v_{i,j,k}\left(\sum_{j=1}^{n}\dot{v}_{i,j,k}\dot{x}_j\right)\right]m_k - \sum_{k=1}^{N}\left\langle\sum_{m=1}^{3}\omega_{m,j,k}\left\{\sum_{i=1}^{3}{}^{r}I_{m,i}^{k}\left(\sum_{j=1}^{n}\dot{\omega}_{i,j,k}\dot{x}_j\right)\right\}\right\rangle$$
$$-\sum_{k=1}^{N}\left\langle\sum_{m=1}^{3}\omega_{m,j,k}\left\{\sum_{l=1}^{3}\left[\left(\sum_{i=1}^{3}\varepsilon_{l,m,i}\sum_{p=1}^{n}\omega_{i,p,k}\dot{x}_p\right)\left(\sum_{i=1}^{3}{}^{r}I_{l,i}^{k}\sum_{j=1}^{n}\omega_{i,j,k}\dot{x}_j\right)\right]\right\}\right\rangle \tag{26}$$
$$(l = 1,\ldots,n) \quad .$$

In forward kinetic issues what we concerns is the integration of eqn (25), while in reverse kinetic issues we care more about the differating of $\underline{x}$ in order to get $\dot{\underline{x}}$ and $\ddot{\underline{x}}$. This comprehensive nature of the analysis is owing to a large extent to using Lagrange's form of d' Alembert's principle to develop the governing equations, say Kane equation. It possesses the feature of Lagrange equation in the non-working constraint forces are automatically eliminated and avoids the drawbacks of Lagrange equation, which is cumbersome and unwieldy to differentiate the energy function. Indeed, this principle essentially reduces the derivation of the governing equations to the evaluation of kinematical coefficients. Therefore all of the numerical work beginning with the derivation of these coefficients and ending with the numerical solution of the governing equations can be easily performed on computer. The whole process is semi-analytic and semi-numerical.

5.1 Future Work

Traditional human body models are normally composed of rigid segments linked by simple kinematic connections, which are maintained for a long time [21]. However, the segments of the human body are not absolutely rigid and such an assumption can lead to some errors in both inverse and forward dynamics analyses. These types of activity are often associated with injuries or discomfort [22]. It may be meaningful in respect that in future work the algorithm can account for some soft tissue motion and the kinetic effects it has on the body. Hence, additional biodynamic experimental data are necessary to improve the accuracy of model.

We can also employ the core algorithm developed here to meet this challenge as follows: Divide the mass of each rigid body segment into rigid part and soft tissue part and calculate their inertia of rotation respectively. Then calculate the load on the joints caused by each part. The final result is the sum of these two parts. As the existence of the soft tissues it takes more time for the system to get stable in the same motion, hence it is reliable to validate the algorithm by comparing its result with its total rigid body counterpart during the same movement by finding a decrease in joints load.

References

1. Wittenburg, J.: Dynamics of Systems of Rigid Bodies. Stuttgart, Germany (1977)
2. Huston, R.L., Passerello, C.E., Harlow, M.W.: Dynamics of multi-rigid-body systems. Journal of applied Mechanics 45, 889–894 (1978)
3. Kane, T.R., Levinson, D.A.: Formulation of equations for motion for complex spacecraft. Journal of Guidance Control 3, 99–112 (1980)
4. Kane, T.R., Levinson, D.A.: Multi-body dynamics. Journal of applied Mechanics 50, 1071–1078 (1983)
5. Nikravesh, P.E.: Computer-Aided Analysis of Mechanical Systems. Prentice Hall, Englewood Cliffs, NJ (1988)
6. Roberson, R.E., Schwertassek, R.: Dynamics of Multi-body Systems. Springer, Berlin (1988)
7. Shabana, A.A.: Dynamics of Multi-body Systems. Wiley, New York (1989)
8. Schielen, W.: Multi-body Systems Handbook. Springer, Berlin (1990)
9. Huston, R.L.: Multi-body Dynamics. Butterworth-Heinemann, Stoneham, MA (1990)
10. Kane, T.R., Levinson, D.A.: Dynamics: Theory and Applications. McGraw Hill, New York (1985)
11. Amirouche, F.M.L., Shareef, N.H.: Gain in computational efficiency by vectorization in the dynamic simulation of multi-body systems. Computers & Structures 41(2), 293–302 (1991)
12. Huston, R.L., Passerello, C.E.: On the dynamics of a human body model. Journal of Biomechanics 4, 369–378 (1971)
13. Adrian, M.J., Copper, J.M.: Biomechanics of Human Movement, Benchmark Press, Inc. (1989)
14. Glitsch, U., Farkas, R.: Applications of a multi-body simulation model in human movement studies. Journal of Biomechanics 27(6), 778 (1994)
15. Bobbert, M.F., van Soest, A.J.: Effects of muscle strengthening on vertical jump height: a simulation study. Medicine and Science in Sports and Medicine 26, 1012–1020 (1994)
16. Gerritsen, K.G.M., Nachbauer, W., van den Bogert, A.J.: Computer simulation of landing movement in downhill skiing: anterior cruciate ligament injuries. Journal of Biomechanics 29, 845–854 (1996)
17. Huang, S.-C.: Analysis of human body dynamics in simulated rear-end Impacts. Human Movement Science 17, 821–838 (1998)
18. Hardin, E.C., Su, A., van den Bogert, A.J.: Foot and ankle forces during an automobile collision: the influence of muscles. Journal of Biomechanics 37, 637–644 (2004)
19. Kane, T.R.: Impulsive motions. Journal of applied Mechanics 29, 715–718 (1961)
20. Keller, J.B.: Impact with friction. Journal of applied Mechanics 53, 1–4 (1986)
21. Nigg, B.M., Bobbert, M.F.: On the potential of various approaches in load analysis to reduce the frequency of sports injuries. Journal of Biomechanics 23(Suppl. 1), 3–12 (1990)
22. Pain, M.T.G., Challis, J.H.: The influence of soft tissue movement on ground reaction forces joint torques and joint reaction forces in drop landings. Journal of Biomechanics 39, 119–124 (2006)

A Fast Motion Estimation Algorithm for H.264

Jianbin Song, Bo Li, and Qinglei Meng

Digital Media Laboratory, School of Computer Science and Engineering,
Beihang University, Beijing, 100083, China
buaasjb@yahoo.com.cn, boli@buaa.edu.cn

Abstract. Compared with the previous video standard, the performance of H.264 has improved significantly because of multi-block-size motion estimation. However, encoder using above technology is very time-consuming. Making full use of the spatiotemporal correlation of video and similar motion between the different-size block, according to the central-biased characteristic of motion vector, a fast motion estimation algorithm for H.264 is proposed in this paper. By effective prediction of search window center point, adaptive selection of search patterns and termination of the search process, the proposed algorithm speedups 123-186 times than full search algorithm and 4.6-5.9 times than the fast motion estimation algorithm recommended in H.264 while maintains similar rate distortion performance.

1 Introduction

Compared With the video standard MPEG-4 [1], the H.264 [2] can save more than 40% of the rate with similar distortion [3]. However, the encoder based on the H.264 is more time-consuming and the motion estimation account for roughly 60% to 80% encoding time especially. The implementation of real-time encoder based on the H.264 is facing extraordinary challenge [4]. Thus, it is more valuable to research the fast algorithm of motion estimation for H.264.

The full search motion estimation (FS) has good distortion rate but low efficiency and it is used for the analysis of the fast algorithm' performance. The fast motion estimations (TSS [5],NTSS [6],FSS [7],DS [8] and HS [9]) just study the search patterns and the speed of the above fast algorithm is not so fast and is need to be improved. Based on DS, the time of fast algorithm MVFAST [10] and PMVFAST [11] has decreased, however, neither of the algorithms is suitable for the multi-block-size motion estimation of H.264. An Unsymmetrical-cross multi-resolution motion search algorithm [4] is adopted by H.264 and implemented in reference software JM9.0. However, the performance of the algorithm is not fast enough because of its conservative search strategy.

After analyzing the spatiotemporal correlation of the video and mode correlation of the different size of block, a novel fast motion algorithm based on prediction of search start point, selection of search patterns and termination of search procedure is proposed in this paper.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 189–196, 2007.

2 Multi-Block-Size Motion Estimation in H.264

In H.264, multi-block-size motion estimation plays an important role in improving coding performance. There are seven different kinds of block and each block is labeled from one to seven, as show in Fig.1. Coding a macro block in JM9.0, the encoder performs motion estimation with different block size in turn within a certain scope of the search window and Lagrange' function is used to obtain the optimum of distortion rate. From above procedure, it can be see that the time consumption of multi-block-size motion estimation in H.264 is increased by several times than that of previous video standard .The multi-block-size motion estimation in H.264 does not use the similarity of motion vector of different mode (mode correlation). This paper uses the motion vector of upper-mode (for instance, mode 1 is the upper mode of the mode 2 and 3) to predict the search window center of current motion estimation.

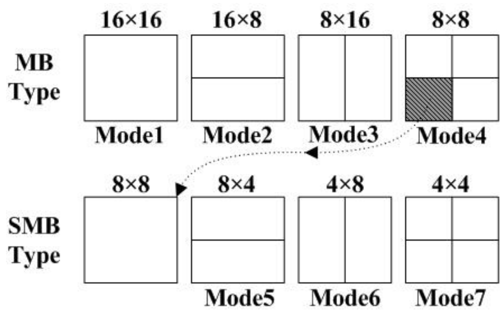


Fig. 1. Partitions of a macro block in H.264

3 The Proposed Fast Motion Estimation Algorithm

3.1 Prediction of Search Window Center

Statistics show that 66% of the coding block has the same motion vector predicted from spatial neighboring block and 64% has the same motion vector predicted from temporal neighboring block. There exists spatiotemporal correlation between or within video frames. Statistics also show that the ratio that the upper motion vectors are identical to the below motion vectors is as high as 97%. The motion trend of different kind of blocks within the same macro block is always similar which is called mode correlation in this paper and is used in the proposed algorithm to predict the search window center.

In this paper, spatial neighboring block such as block 1\2\3 and temporal neighboring block such as block 4 in Fig.2 (a) are used to predict the search window center. And the fast algorithm starts a new search within a scope from the prediction window center. When inter mode is not mode 1, the motion vector of upper mode and the motion vector of spatiotemporal neighbor block are used to predict the window center. There are several methods such as the median method, the weighted average method and the error cost comparison method which are used to determine the search

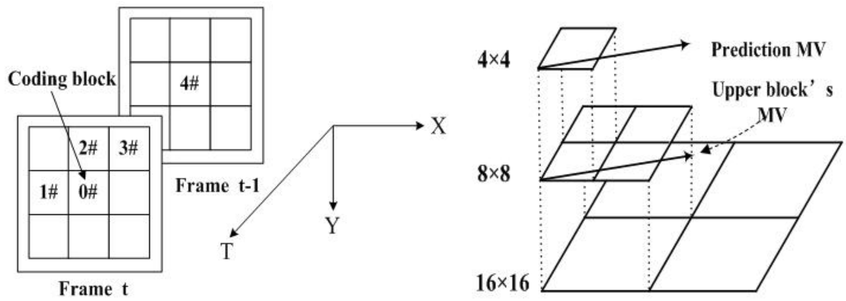


Fig. 2. Neighboring blocks and Upper blocks

start point. In this paper, error cost comparison method is selected because of the following reasons: first, in most cases the effect of prediction is the most important, accurate forecasts can provide a good starting point for the follow-up search and can reach search destination point quickly; Second, because of the motion correlation of the video, the point corresponding to prediction motion vector may become the optimal match point likely. Statistics show that large proportion (64%-97%) of the optimal destination point is the point corresponding to the prediction motion vectors. Thus, it is necessary to search for the points corresponding to the prediction MV.

3.2 Adaptive Selection of Search Patterns

The motion vectors gotten by the motion estimation have center-biased characters [7], namely most of MV' s module is small and the point corresponding to the MV always is near the search window center point. After analyzing characteristic of the MV, we get a conclusion that the X-coordinate is always larger than Y-coordinate which accords with the motion characteristic of the real world video. New search patterns include large diamond search (LDS) and small diamond search (SDS) are used in this paper. The horizontal search step of the LDS is 3 points while the vertical search step is 2 points which is different from the large diamond search in literature [9]. Different search step of large diamond is fit for the motion direction of real world video likely. The LDS is appropriate for the fast search of large-motion video while the SDS (1 points search step) is used to search the block which has small MV.

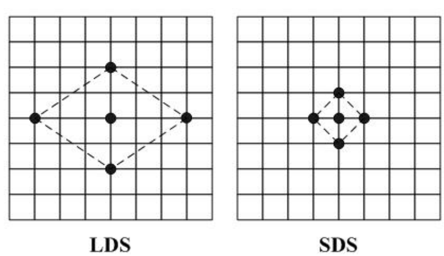


Fig. 3. Large and small diamond search patterns

In this paper, the search patterns are selected adaptively based on the characteristic of video motion. If current inter mode is not equal to 1 or $V_1 = V_2 = V_3 = V_4$, which show that the block is stillness or has small-range motion. The algorithm need to do dense search around the window center and the search mode is set as SMS. If current inter mode is 1 or V_1, V_2, V_3, V_4 are not all equal which show that the block has medium or large scale motion, the algorithm must search the optimal match point with lager step, so the initial search mode is set as LDS. If the search starts with LDS and when the error cost is lower than the threshold, it can be concluded that the search result is good enough, so the algorithm set the search mode as SMS and start a refined search around the latest optimal point. Experiments show that the above method of the search mode setting is not only simple but also effective.

3.3 Termination of Search Process

The motion estimation algorithm in literature [9] completes one step SDS, the search process stops immediately in despite of the search results, thus it may not find the optimal point. In this paper, after SDS search, if the current optimal point is the center of the small diamond, the search process stops, otherwise the algorithm continues to do the SDS search. Many fast algorithms such as MVFAST and PMVFAST use the thresholds to terminate the search process as early as possible to speed up the search process. However, MVFAST uses the fixed threshold and PMVFAST uses the neighboring block's error cost as current coding block's termination threshold. Fixed threshold is not always appropriate for all the blocks and the error cost of two spatial neighboring blocks do not have the obvious correlation. Statistics illustrate that there exists strong correlation of error cost between the temporal neighboring block which is reflected in Fig.4 (the X-coordinate is mode and the Y-coordinate J is error cost).

Based on the analysis above, thresholds used to terminate the search process in our paper are decided as follow: using the error cost of the same position in the reference

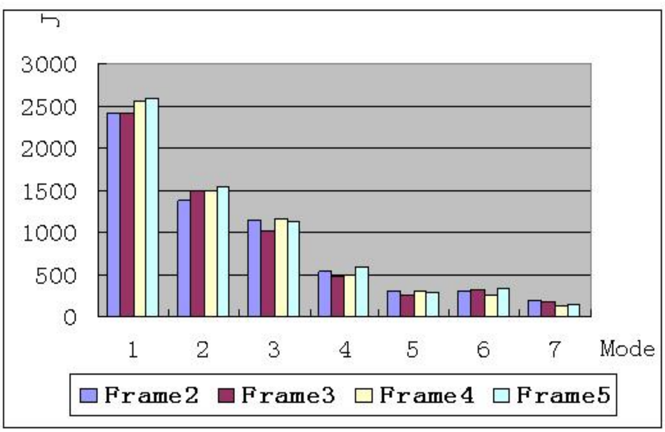


Fig. 4. Error cost column diagram of the 164 macro-block within frame 2-5

multiplies a coefficient as the threshold. Two level search termination rules is used and there are two thresholds T_1 and T_2 used in the proposed algorithm, $T_1 = \alpha \times PrevJ$, $T_2 = \beta \times PrevJ$, PrevJ is the error cost of the same macro-block or sub-block in the reference frame. (1) If $J \le T_1$,the search results is good enough and the search process is terminate immediately; (2) If $T_1 < J \le T_2$, it needs refined search, so the search patterns changed from LDS to SDS. After experiments, we assigned T_1=500 and T_2=750 if the 4# neighboring block (Fig.2 (a)) is not inter block. The value regions of α and β are : $0.9 < \alpha < 1.4$, $1.2 < \beta < 2.5$, and $\alpha < \beta$. The larger the value of α and β, the faster the proposed algorithm and the worse the performance of the motion estimation. After a lager number of experiments, we concluded that when $\alpha = 1.05$, $\beta = 1.5$, the algorithm can get the good tradeoff between the speed and performance. The above method deciding the thresholds is so adaptive that can terminate the search process with good performance.

3.4 Description of the Proposed Algorithm

The proposed algorithm include several steps such as prediction of search window center, decision of the termination, the LDS search, the SMS search, the selection of the search patterns. V_0, V_1, V_2, V_3, V_4 and V_u is the motion vector of the 1#, 2#, 3#, 4# and upper block (Fig.2) and the error cost corresponding to the above motion vector is J_0, J_1, J_2, J_3, J_4 and J_u.The variable Modetype represents the inter mode and has the value between 1 and 7. Searchtype represents the type of search pattern, value 0 represents LDS search and value 1 represents SDS search and value 2 represents just doing one step SDS search and the motion estimation is terminated immediately. T_1 and T_2 are thresholds used to stop the motion estimation, MaxJ is maximum value in the algorithm and PrevJ is error cost after current search step.

step1. Getting the prediction MVs and the corresponding error cost : if 1# (2#,3#,4#) macro block exists and its coding type is inter mode, get V_1 (V_2, V_3, V_4) and calculate J_1 (J_2, J_3, J_4), otherwise let J_1 (J_2, J_3, J_4) =MaxJ. If Modetype>1, getting V_u and calculate J_u, otherwise J_u =MaxJ.

To calculate J_0 for the V_0 = (0,0) .

step2. Prediction of search window center : MinJ=Min {J_0, J_1, J_2, J_3, J_4, J_u}, the point corresponding to MinJ is search window center.

step3. Decision of the termination threshold : If Modetype of 4# block is intra mode, set T_1=500, T_2=750. If Modetype of 4# block is inter mode, $T_1 = \alpha \times PrevJ$, $T_2 = \beta \times PrevJ$,when Modetype=1, PrevJ is set as J_u, when Modetype>1, PrevJ is set as $J_u / 2$.

step4. Initialization of search pattern : If ModeType>1 or $V_1 = V_2 = V_3 = V_4$, set SearchType = 1, then go to step6 ; otherwise set SearchType =0, then go to step5.

step5. The LDS search : Doing LDS search, record current minimum error cost MinJ. If the current optimal point is the center of the large diamond, go to step6 ; Otherwise go to step7.

step6. The SMS search : Doing SDS search, record current minimum error cost MinJ. If the current optimal point is the center of the small diamond, go to step8 ; Otherwise go to step7.

step7. Selection of search pattern : If MinJ≤T_1, go to step8 ; otherwise if SearchType = 0 and T_1 < MinJ≤T_2, set SearchType as1, go to Step6 ; otherwise SearchType = 0 and T_2 < MinJ, go to step5 ; otherwise if SearchType =1 and T_1 < MinJ≤T_2, let SearchType=2, go to step6 ; otherwise if SearchType =1 and T_2 < MinJ, go to step6.

step8. Termination of search : To record the motion vector, inter mode and error cost, the integer pixel motion estimation is completed.

4 Experiment and Analysis

In order to test the performance of the proposed algorithm, experiments have been done on JM9.0 which is reference software of H.264 and FS and JMFAST (Recommended by H.264 and implemented in JM 9.0) is selected as compare algorithms. Four standard video sequences “Coastguard” (176×144), “Foreman” (176×144), “Table” (176×144), “Mother” (176×144) are used and the selected video sequences have different-extent motion and different texture, so they could reflect the real performance of every algorithm. In table 1, column1, 2, 3 is the name of the sequence, quant parameter and name of algorithm, column 4, 6, record the rates, the PSNR, column5, 7, 8 records the increase percentage of the rate, the increase value of PSNR and the multiple of speedup.

From table 1, it can be seen that the speed of the proposed algorithm is 123-186 times of the full search and 4.6-5.9 times of the JMFAST recommended in H.264 while maintains similar distortion performance. The excellent performance gotten have several reasons that are effective prediction of search window center point, adaptive selection of search patterns and termination of the search., namely the proposed algorithm make full use of the spatiotemporal correlation of video, similar motion between the different-size block and the central-biased characteristic of motion vector.

At present, the research of inter mode decision is also an effective way to speed up the inter frame coding, literature [12] have research the inter mode decision and get some research fruits. Based on the proposed fast algorithm, using inter mode decision policy, the inter frame coding would speed up extraordinary and the above research will be done by author at next step.

Table 1. Pce of the proposed algorithm an the other algorithm

Seq.	*QP*	*Algorithm*	*Rate (Kbps)*	*R↑ (%)*	*PSNR*	*△PSNR*	*Speedup*
Coastguard	40	FS	26.45	0	26.76	0	1
		JMFAST	26.36	-0.3	26.75	-0.01	28
		Proposed	26.42	-0.1	26.76	0	164
Foreman	40	FS	71.10	0	26.75	0	1
		JMFAST	73.18	2.9	26.75	0	23
		Proposed	75.05	5.6	26.80	0.05	123
Table	36	FS	65.06	0	30.71	0	1
		JMFAST	65.10	0.1	30.71	0	41
		MSFAST	64.81	-0.4	30.69	-0.02	189
Mother	32	FS	87.58	0	33.20	0	1
		JMFAST	88.56	1.1	33.21	0.01	38
		MSFAST	89.12	1.8	33.20	0	180

Acknowledgements. Supported by the NSFC (60505007), the Equipment Advanced Research Project, the Program for New Century Excellent Talents in University and ERIPKU. The research was made in the State Key Lab of Software Development Environment.

References

1. Information Technology—Generic Coding of Audio-Visual Objects Part 2: Visual, ISO/IEC 14496-2 (MPEG-4 Video) (1999)
2. Joint Video Team (JVT) of ISO/ IEC MPEG and ITU-T VCEG, Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H. 264/ISO/IEC 14496-10 AVC), JVT-G050 (2003)
3. Wiegand, T., Sullivan, G.J., Bjøntegaard, G., Luthra, A.: Overview of the H.264 Video Coding Standard. IEEE Trans. Circuits Syst. Video Technol. 13(7) (2003)
4. Yang, P., He, Y., Yang, S.Q.: An Unsymmetrical-Cross Multi-Resolution Motion Search Algorithm For Mpeg4-Avcm.264 Coding. In: IEEE International Conference on Multimedia and Expo (ICME), pp. 531–534 (2004)
5. Koga, T., Iinuma, K., Hirano, A., Iijima, Y., Ishiguro, T.: Motion compensated interframe coding for video conferencing. In: Proc. Nat. Telecommun. Conf., New Orleans, LA, November 29–December 3, 1981, pp. G5.3.1–5.3.5 (1981)
6. Li, R., Zeng, B., Liou, M.L.: A new three-step search algorithm for block motion estimation. IEEE Trans. Circuits Syst. Video Technol. 4, 438–442 (1994)
7. Po, L.M., Ma, W.C.: A novel four-step search algorithm for fast lock motion estimation. IEEE Trans. Circuits Syst. Video Technol. 6, 313–317 (1996)
8. Tham, J.Y., Ranganath, S., Ranganath, M., Kassim, A.A.: A novel unrestricted center-biased diamond search algorithm for block motion estimation. IEEE Trans. Circuits Syst. Video Technol. 8, 369–377 (1998)
9. Zhu, C., Lin, X., Chau, L.-P.: Hexagon-Based Search Pattern for Fast Block Motion Estimation. IEEE Trans. Circuits Syst. Video Technol 12, 349–355 (2002)

10. Ma, K.K., Hosur, P.I.: Performance Report of Motion Vector Field Adaptive Search technique (MVFAST) [S]. In: ISO/IEC JTC1/SC29/WG11 MPEG99/m5851, Noordwijkerhout, NL (March (2000)
11. Tourapis, A.M., Au, O.C., Liou, M.L.: Fast Block-Matching Motion Estimation using Predictive Motion Vector Field Adaptive Search Technique (PMVFAST)[S]. In: ISO/IEC JTC1/SC29/WG11 MPEG2000/m5866, Noordwijkerhout, Netherlands (March 2000)
12. Tourapis, A.M., Wu, F., Li, S.: Direct mode coding for bipredictive slices in the H.264 standard. IEEE Trans. Circuits Syst. Video Technol. 15, 119–126 (2005)

Validating Optical Motion Capture Assessments of the Dynamic Aspects of Work

Jackie Sutherland[1] and Vincent Duffy[1,2]

Purdue University
[1] School of Industrial Engineering, 315 N. Grant Street
[2] Regenstrief Center for Healthcare Engineering and Agricultural and Biological Engineering
West Lafayette, IN 47907 USA
{jasuther,duffy}@purdue.edu

Abstract The objective of this study is to validate two three-dimensional motion systems used capture human movement; the Lumbar Motion Monitor (LMM) and optical motion capture. Marras et al. captured the accuracy and reliability of optical motion capture and the LMM in a 1992 validation study, and found several benefits of using the LMM for ergonomic evaluations. However, since 1992, several advances have been made in the field of digital human modeling and optical motion capture, and it is believed that a modern validation of the two systems could serve others in academic and industry alike when choosing a methodology toward capturing ergonomic data. The purpose of this research is to validate the methods of collecting dynamic data and predicting injury risk in humans during lifting procedures.

Keywords: optical motion capture, lumbar motion monitor, dynamic lifting, digital human modeling, ergonomics, validation.

1 Introduction

Low-back disorders (LBDs) continue to serve as detrimental components of health and safety in the workplace [1]. In 2005, the occupational incidence rate of nonfatal injuries remained around 6 total recordable cases per 100 full-time workers in the manufacturing sector. Manufacturing alone consisted of 20.2% of all nonfatal workplace industries in 2005, with health care and social assistance following with 15.7% of all industry injuries. Nonfatal injuries to the back in industry reached 270,890 in 2005 [2].

2 Early Lifting Guidelines

Numerous efforts have been made to modify work practices and the methods of which industrial tasks are performed. As early as 1930, laws limited the weights that women and children could handle. In 1962 the International Labour Organization published a listing of suggested limits for "occasional weight lifting" for both women and men. These weight restrictions were based on previous injury statistics of manual materials

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 197–204, 2007.

handling (MMH) tasks that had led to increased injuries in the spinal, knee, shoulder, elbow, and hip injuries [3].

2.1 Risk Factors for Lifting Injury

In 1974, Herrin, et al. identified seven risk factors towards lifting injury. These factors included:

Weight of material lifted.
Position of the center of gravity of the load relative to the worker.
Frequency, duration, and pace of lifting.
Stability of the load.
Coupling of the load.
Workplace geometry, including direction of movement and distance.
Environmental factors such as temperature, humidity, vibration, and stability [4].

Herrin's work was influential for its time, and paved the way for research in the field of industrial safety and dynamic lifting. However, a more explicit method was required to identify specific lifting procedures that were detrimental to the worker.

2.1 NIOSH Work Practices Guide and Lifting Equation

Later, more quantified research led to the 1981 publication of the Work Practices Guide for Manual Lifting (WPG) by the National Institute for Occupational Safety and Health (NIOSH). The NIOSH WPG was generated by a team of scientists, engineers, and physicians, based on several investigations of the human physiological system. In addition, the team relied upon research from countless others in the areas of neurology, physical therapy, biomechanics, musculoskeletal surgery, industrial hygiene, and beyond. In the 1981 WPG, NIOSH stressed the susceptibility of the lower-back to overstress and injury from flawed lifting procedures. NIOSH utilized previous studies to draw knowledge about compressive and shear forces on the spine, disc moments, and other psychological measurement of the musculoskeletal system to develop the NIOSH lifting guidelines [3].

Since its development in 1981, NIOSH has revised its original lifting guidelines to include asymmetrical lifting tasks and lifts with varying coupling levels. This resulted in the Revised NIOSH Lifting Equation and its principal component, the Recommended Weight Limit (RWL). The RWL, as defined by NIOSH, is the maximum load weight that a healthy worker should be able to utilize over an eight-hour work day without increasing their risk of lift-related lower back pain (LBP). NIOSH defined the RWL as the multiples of the load constant, horizontal multiplier, vertical multiplier, distance multiplier, asymmetric multiplier, frequency multiplier, and coupling multiplier. Each multiplier has specific definitions and methods of measurement as defined by NIOSH. In addition, application of the lifting equation is limited to certain conditions, specifically, two-handed lifting tasks which do not require more energy expenditure than can be handled repetitively from day to day. NIOSH also assumes that lifting occurs in workplaces that provide adequate lifting space, stable objects, adequate temperature and humidity, and a sufficient coefficient of friction between the worker's shoes and the floor [1].

3 Marras' Lumbar Motion Monitor

Despite NIOSH's previous efforts, the sources for low-back injuries in the workplace are not always easy to identify. Other analytical methods have therefore been developed to assess the risk of lumbar injury in the workplace. One such tool is the Lumbar Motion Monitor (LMM), a triaxial electrogoniometer worn on the back and used to measure an individual's three-dimensional angular position in space. Developed by William Marras and colleagues at the Ohio State University Biodynamics Laboratory, the LMM can be considered for predictions of back injury based on the type of movement performed. This exoskeleton of the spine has the capability to process lumbar position, velocity, and acceleration in the frontal, sagittal, and transverse planes by differentiating the value of the subject's position in space [5]. Figures of the LMM can be found on the Ohio State Biodynamics Laboratory website [6].

By utilizing the lumbar motion monitor, Marras was able to quantify LBD injury rates by creating a LBD risk model. Marras' model approximates the probability that a particular lifting task could be classified as a 'high risk' task. In addition, Marras validated his model in 2000 by measuring the trunk movement of 142 employees and judging this against their LBD incident rates. Marras found a statistically significant correlation between the changes in LBD incident rates and LBD risk values predicted by the Marras LBD risk model [7].

3.1 Incorporation of Electromyography

Other physical systems have been integrated into lumbar injury prevention studies as well. Marras, Granata, Davis, Allread, and Jorgensen incorporated an electromyograph (EMG) into their analysis of low back disorder risk using the LMM. Electromyographic activity was measured through bipolar electrodes spaced 3 cm apart at ten major trunk muscle sites. By including an EMG into their analysis, Marras was able to calculate internal moments and forces for specific lumbar muscles, and determine the specific exertion of lumbar muscles when a downward force was applied [8].

4 Optical Motion Capture

Although NIOSH and others have developed helpful tools used to evaluate possible risk found in the workplace, digital human modeling methods propose innovative and accurate approaches for examining the musculoskeletal system and possible risks to the spine during lifting and movement. This research aims to test the validity and accuracy of two biodynamic systems during pre-defined lifting and movement techniques, the lumbar motion monitor, and optical motion capture. This study utilizes the LMM and updated, three-dimensional optical motion capture technology in efforts to capture a comparison between the two systems.

4.1 Envision Center

Purdue University's Envision Center for Data Perception (Envision Center) provides the opportunity for students, faculty, and industry professionals to utilize cutting-edge facilities in the field of human computer interaction [9]. Partially funded by a National Science Foundation grant, Envision Center features an extensive list of both hardware and software technologies [10]. Current hardware resources include the FLEX™ Virtual Reality Theater, a 12' x 7' tiled display wall, an Access Grid which allows participants to collaborate with others from distant geographical areas quickly and easily, and several haptic devices including the CyberGlove®. Recent projects include incorporating tactile feedback into terrain modeling, development of virtual learning environments for hearing impaired children, and creating bio-terror communication training modules for public relations students [11]. In 2006 Purdue's Envision Center received a national award from Campus Technology magazine as a winner in its annual Innovators Competition [10].

Motion capture technology is commonly thought of as a tool found in the entertainment industry to capture and incorporate human movement into films. However, motion capture is proving to have a valuable impact in commercial industry in the areas of industrial safety and ergonomics. By capturing human movement during lifting techniques involving various loads and twisting motions, industries can determine the safe limit that can be performed in a typical work setting. In addition, industries can utilize motion capture in the design and development of future lifting procedures and work tasks.

4.2 Motion Capture Configuration

Motion capture technology uses multiple STT Motion Captor cameras to capture human position and movement in space to create a 3D simulation. A typical motion

Fig. 1. One of six infrared cameras used in Envision Center's STT Motion Captor optical motion capture system (left) and a participant demonstrating a basic LED configuration (right)

capture setup includes the placement of nineteen markers, or LEDs placed on pre-specified areas of the body, and uses a circular four- or six-camera configuration to capture all degrees of the subject's movement. One of the cameras used during optical motion capture is shown in the left image in Figure 1. The LED configuration includes placing markers on the foot, ankle, knee, hip, wrist, elbow, shoulder, waist, neck, and head, as shown in the right image in Figure 1 [12]. Markers are essentially LEDs whose brightness point is over a specified value. Therefore, it is recommended that bright clothes, jewelry, and reflective material be avoided during the capture process [13].

5 Previous LMM Validation

This research was inspired by previous work by Marras concerning the accuracy and repeatability of a LMM. In his 1992 publication, subjects moved in one of three angular ranges of motion while wearing the LMM and motion capture equipment. To capture human movement through motion capture, several LEDs were placed on the upper and lower plate of the LMM. The values recorded by both the LMM and LEDs were then compared to the actual position in space. Marras used the deviation between actual position in space and specified position for both the LMM and motion capture system to indicate their accuracy. The data taken by Marras verified that deviations between the actual and predicted range of motions in the optical motion capture system were twice that of the LMM. This research demonstrated that the LMM can assess the risk of low back disorders more accurately and effectively than optical motion capture, and serves as a reliable and cost-effective means to monitor lumbar motion [5]. In fact, the Ohio State Biodynamics Laboratory states that the LMM boasts three times the predictive power of that given by the original NIOSH lifting guide [6].

5.1 Other Model Validation

Research has also been performed to validate the models used to measure and prevent injuries in the workplace. In 2001, Robson, Shannon, Goldenhar, and Hale described validity and reliability as it applies to injury statistics and analytical equipment measures. They believe that experimental control has a significant impact on the results of a biomechanical experiment using analytical equipment, particularly calibration and maintenance of equipment, the equipment's interaction with the environment, and the proper use of the equipment by the operator. It has been shown that validity and reliability of collecting data with analytical equipment can be improved by minimizing variation in equipment operation, and, if the cost of measuring allows, by taking multiple measurements [14]. It is therefore believed that this experiment will cover multiple repetitions of each trial in order to assess the equipment's accuracy.

5.2 Assessment of the LMM and Optical Motion Capture

Marras' development of the job risk classification model has sparked additional research in the area of classifying lifting kinematics. In 2005, a study by Tara

Cappelli examined two- and three-plane lifts through the use of motion capture in relation to the Marras *et al. 1993* Model (Marras model). Cappelli found that motion capture technology and the LMM were capable of collecting similar types of data. One major aspect of Cappelli's research was the calculation of a participant's angular velocity and acceleration by differentiating the position data points with respect to time [15].

However, since Marras' 1992 validation of the LMM and optical motion capture system, there have been several advances in the area of optical motion capture, including the reliability, accuracy, and ease of capturing detailed three dimensional motion data. Although Marras' work is thorough and comprehensive, it compares the LMM to the version of optical motion capture that was available in 1992. In addition, Marras' capture of motion using optical motion equipment consisted of LEDs placed on the LMM; one on the top, dynamic portion, and another on the bottom, static portion. It is believed that a more recent and detailed validation of the LMM and optical motion capture would be useful in the context of current research on the measurement of lumbar movements and prediction of lower back disorders.

5.3 Experimental Method

With a modern assessment of the validity and accuracy of the LMM and optical motion capture, a comparison can be made between the two systems. It is not currently known whether one system is clearly superior to the other; however, it is believed that both systems offer effective measurements in various dimensions of human movement. Additional insights into the use of optical motion capture for industrial ergonomic assessments can be found in [16]. One goal of this research is to determine the varying degrees of accuracy, reliability, and practicality in determining lumbar movements. For example, while one system may exhibit lower error rates when determining subject position in space, another system may demonstrate superiority for velocity and acceleration measurements. This research hopes to capture an impartial analysis of the accuracy and validity of the LMM compared to optical motion as well as their practicality in industrial settings.

In order to capture the accuracy of the LMM and motion capture systems, the reported position value from the LMM and motion capture can be compared to a frame of reference that restricts the subject through a clamp. By restricting the subject to pre-specified ranges of motion, it is possible to measure the 'error' of the position value recording by either system. In addition to measuring the accuracy through error rates, the reliability of each system will be determined by performing several repetitions for each specified range of motion [5].

Testing will be somewhat similar to that of Marras et al., but may include EMG data in order to standardize lifting and lumbar movements. Independent variables will include the angular range of motion (ROM) in degrees specified by the experiment, as well as vertical position in space (also pre-specified). In addition,. Task candidates will be considered such that different levels of velocity and acceleration can be monitored and reported in relation to validity. Dependent variables include the ROM and area in space reported by the LMM and motion capture equipment. For each trial, an error will be reported for the difference between the subject's actual position in space specified by the clamp, and the reported position in space given from the LMM and motion capture systems.

5.4 Results

Preliminary results from this study will be presented at the HCI International Conference in Beijing, China in 2007.

5.5 Future Work

A considerable range of future research could be performed in the study of lifting procedures and the tools used to evaluate them. One area of research is the study of lifts in non-standard postures, such as one-handed lifts and those performed more infrequently. In addition, it would be beneficial to utilize the LMM and optical motion capture in an industrialized setting to test the system's compatibility with their environment. The LMM has already been tested and applied in industrial settings, and by further incorporating the LMM or motion capture system into a factory environment, the practicality of either system will be greater recognized. Possibilities also exist for further study of lifting procedures in health-care settings such as the task of lifting and transporting patients. Use of the LMM and motion capture systems promise rewarding benefits in these and many other commercial settings, particularly where standard repetitive strain related models may not be sufficient.

References

1. NIOSH 1994, Applications Manual for the Revised NIOSH Lifting Equation, NIOSH Technical Report No. 94-110 (Cincinnati: US Department of Health and Human Services, National Institute for Occupational Safety and Health) (1994)
2. US Department of Labor 2006, Workplace Injuries and Illnesses in 2005, Bureau of Labor Statistics, Washington, D.C. 20212 (2006) Internet address: http://www.bls.gov/iif/home.htm
3. NIOSH 1981, Work Practices Guide for Manual Lifting, NIOSH Technical Report No. 81-122 (Cincinnati: US Department of Health and Human Services, National Institute for Occupational Safety and Health) (1981)
4. Herrin, G.D., Chaffin, D.B., Mach, R.S.: Criteria for Research on the Hazards of Manual Materials Handling. NIOSH Contract Report, CDC 99-74-118 (1974)
5. Marras, W.S., Fathallah, F.A., Miller, R.J., Davis, S.W., Mirka, G.A.: Accuracy of a three-dimensional lumbar motion monitor for recording dynamic trunk motion characteristics. International Journal of Industrial Ergonomics 9, 75–87 (1992)
6. Biodynamics Laboratory, The Ohio State University (cited 2007 February 7) Available from: http://biodynamics.osu.edu/research.html
7. Marras, W.S., Allread, W.G., Burr, D.L., Fathallah, F.A.: Prospective validation of a low-back disorder risk model and assessment of ergonomic interventions associated with manual materials handling tasks. Ergonomics 43, 1866–1886 (2000)
8. Marras, W.S., Granata, K.P., Davis, K.G., Allread, W.G., Jorgensen, M.J.: Effects of box features on spine loading during warehouse order selecting. Ergonomics, 42(7), 980–996
9. Envision Center for Data Perception at Purdue (cited 2007 February 10) Available from: http://www.envision.purdue.edu/

10. Campus Technology Innovators: Technology Area: Virtual and Immersive Learning. Campus Technology (August 2006)
11. Envision Center for Data Perception at Purdue. Envision Center for Data Perception Summary of Projects, Spring 2006 (2006) Available from the Envision Center and online from: http://www.envision.purdue.edu/projects_details.html
12. Standard-19 Markers (Ref.: 7) Envision Center for Data Perception at Purdue (cited 2007 February 10) Available from: https://www.envision.purdue.edu/wiki/index.php/MoCap%20Images%20Files%20and%20Documents
13. Motion RT Captor User's Guide. STT Systems Engineering (cited 2007 February 7) Available from: www.simtechniques.com
14. Robson, L.S., Shannon, H.S., Goldenhar, L.M., Hale, A.R.: Measuring Outcomes. In: Guide to evaluating the effectiveness of strategies for preventing work injuries: how to show whether a safety intervention really works. Department of Health and Human Services, Center for Disease Control and Prevention, National Institute for Occupational Safety and Health, pp. 51–66 (2001)
15. Cappelli, T., Duffy, V.G.: Motion capture for job risk classifications incorporating dynamic aspects of work. In: Proceedings of the Society of Automotive Engineers, Conference on Digital Human Modeling for Design and Engineering, SAE-DHM-2006-01-2172 Lyon, France (2006)
16. Du, C.J., Duffy, V.G.: Development of a Methodology for Assessing Industrial Workstations Using Optical Motion Capture Integrated with Digital Human Models. Occupational Ergonomics, manuscript in press (2004)

Modeling Human Bipedal Navigation in a Dynamic Three Dimensional Virtual Environment

Mark. D. Thomas[1,2], Daniel W. Carruth[1,2], Bryan Robbins[1,3], John A. McGinley[1], and Alex Morais[1,3]

[1] Center for Advanced Vehicular Systems
[2] Department of Psychology
[3] Department of Computer Science and Engineering
Mississippi State, Mississippi 39759
{mthomas,dwc2,bryanr,jam,amorais}@cavs.msstate.edu

Abstract. The current research sought to construct a computational model of human navigation for virtual three dimensional environments. The model was implemented within the ACT-R cognitive architecture [1]. The navigation model incorporates visual search, encoding object features and spatial relationships, motion, obstacle avoidance, and incidental visual memory.

Keywords: walking navigation, ACT-R, digital human model, incidental visual memory, visual search.

1 Introduction

One can likely stroll through a shopping mall and observe people walking while talking on cell phones, listening to music, or even chewing gum. Since people have the ability to do a wide variety of tasks while walking, people seem to be able to navigate without much thought. The apparent ease with which humans navigate begs one question for digital human modelers. *Why bother?* After all, end effectors can be used to motivate digital human models in virtual environments and many biomechanical and ergonomic constructs can be researched. However, a model of human navigation would be beneficial for other applications. For example, an architect may want to model an office building complete with digital human model (DHM) workers to test the emergency exits so changes could be made to the building design or escape route signs before construction begins. While such an application may seem unrealistic, researchers have claimed knowledge doubles every five to ten years in some scientific fields [2]. If digital human modeling advances at one-half the speculated rate, this and other diverse DHM applications are probably in the future.

1.1 Kinesthesis and Proprioception

Researchers have speculated navigation consists of motion and wayfinding with wayfinding defined as the cognitive aspects of navigation that do not involve motor actions [3], [4]. Proprioception is the sum of the kinesthetic and vestibular systems

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 205–214, 2007.

and it gives one the internal sensations of directional movement. Kinesthesis refers to the knowledge one has about ones body regarding the relationships between body parts due to sensations from muscles, tendons, and joints. The vestibular system relies on organs in the inner ear and provides feedback about one's orientation and movement. To navigate and return to roughly the same point of origin a blindfolded person could rely on proprioceptive cues by completing four series of paces and 90° turns.

In order to navigate from and return to a point of origin, one must have some strategy or internal representation of the point of origin. Sadeghiam, Kantardzic, Lozitskiy, and Sheta proposed several cognitive elements of wayfinding that parallel the psychological notion of spatial knowledge [4], [5], [6]. Researchers have also speculated about the existence of internal maps [7], [8], [9], [10], [11], [12]. While people may rely on internal representations, they may use external cues as well. For example, external cues can be used for navigation by following a sidewalk around the block to arrive at the initial point of origin, or a shopper may exit a store and depend on cues, such as a sign for a specific parking row, to locate his or her car. Rather than intentionally learning a landmark to be later recalled (e.g., a sign), the car could be located by recognizing a particular landmark, such as a streetlight with a bent pole, that was not intentionally learned when entering the store. The latter is an example of incidental visual memory.

1.2 Incidental Visual Memory

Incidental visual memory refers to memory for visual information that is not actively learned. In navigation, landmarks may be intentionally or incidentally learned in order to retrace a previously traveled route. Castelhano and Henderson investigated whether the intention to remember visual details impacted visual memory [13]. They had participants view scenes in an intentional learning condition and in an incidental learning condition (visual search). Participants remembered details of the scenes regardless of how the scenes were viewed initially. They proposed that memory for objects depended on whether the object was looked at, how often it was looked at, and the length of time it was viewed, not what instructions were given to the participant.

Using a conjunction visual search task, Williams, Henderson, and Zacks had participants count predefined target objects in a display of real-world distractor objects [14]. Williams et al. found that target objects were generally viewed more often and remembered better than distractor objects. Target memory rate was 85% while objects related to the target were remembered at approximately 60% and objects not related to the target were remembered slightly above chance. They also replicated Castelhano and Henderson's finding of a relationship between visual memory and viewing behavior for distractor objects [13].

The Williams et al. and the Castelhano and Henderson search tasks required participants to remain stationary and look at stimuli presented on computer monitors with the goal of searching a confined space [13], [14]. Other researchers have demonstrated incidental memory effects in real-world environments. Change blindness is a phenomenon in which people fail to detect changes in scenes if those changes occur during visual disruptions. Simons, Chabris, Schnur, and Levin had a confederate ask passersby for directions [15]. While receiving directions, a group of confederates passed between the direction giver and the person receiving the directions. While the direction received was occluded by the crowd, a member of the crowd exchanged a basketball with the direction receiver. Simons et al. reported that some participants were able to report visual details of the basketball although they did not report the exchange. Simons et al. found incidental memory for details of the exchanged object to be about 54%.

1.3 Visual Navigation

In order to navigate one must also have some knowledge of one's relationship to the objects in the environment and this knowledge has to be updated when moving through the environment [16]. Gibson speculated that the constructs of a *focus of expansion* and *optic flow* (or retinal flow) were the basis of human navigation [17], [18]. Focus of expansion refers to the region in the visual field at which an object appears to increase in size as one moves toward it. As one gets closer, the object covers more of the visual field. Using the focus of expansion to navigate can be accomplished by overlapping the focus of expansion on a desired goal and walking toward it [16], [17], [18], [19], [20]. (See Warren for a thorough review of focus of expansion and optic flow [19]).

Optic flow refers to the displacement of the environment on the retina as a person moves. Navigation can also be accomplished with optic flow by using a flow-equalization strategy in which the radial displacement of images on the retina is consistently maintained at proportional rates relative the visual field [16], [21]. In other words, distance from an object can be kept constant by maintaining a consistent movement of the object in the periphery of the retina. An example is walking parallel to a wall without looking directly at the wall.

Research has demonstrated that optic flow can be used for navigation and computer simulations have shown that optic flow is capable of providing the impetus for navigation [16], [22], [23], [24], [25]. Kim and Turvey proposed the *"units" of an optic array are projections of facets, faces, and surfaces of the environmental layout to the point of observation* [23]. Mathematically, optic flow is the relative velocity of points across the visual field as a person moves in the environment. Optic flow should operate whether one looks in the direction one is moving, or not, because optic flow depends on external environmental cues that change with the direction of movement. Using oscillating dots presented in linear and rotating fashions, Regan and Beverley demonstrated differences between processing linear and rotating stimuli, which suggested unique processing mechanisms may exist for the curl of velocity (i.e., vorticity) and the divergence of velocity (i.e., dilation), which in turn implies special processing mechanisms may exist for optic flow [26]. Regan and Beverley further speculated the mechanism for processing optic flow may include an ability outside one's awareness that parallels vector calculus to determine and track the extent of optic flow changes when moving.

Other research has brought the importance of optic flow into question. Cutting, Reading, and Wang examined whether people veered in the direction they looked [27]. Cutting et al. had participants walk straight while looking to the side in an illuminated condition and in a darkened condition. They found that people veered in the direction looked. Cutting et al.'s results appear to contradict optic flow theory.

Because participants never had the opportunity to look at a destination goal, an argument could be made the veering behavior resulted from participants using an egocentric-direction strategy for navigation [16], [28], [29]. During egocentric-direction strategies, a person is believed to visually mark, or tag, a goal and walk toward it. Cutting et al. found no difference between the extent of veering and the lighting conditions. If visual-based navigation depended on optic flow, one would have anticipated differences because almost no visual input was available in the darkness condition. Conversely, if navigation depended on checking headings against a focus of expansion, one would have expected the veering behavior to be the same. However, navigation in low illumination environments and environments with few external cues has been demonstrated and sighted individuals may rely on more than vision to navigate [29], [30], [31]. Participants in Cutting et al.'s research may have used navigation strategies that depended very little, or not at all, on vision (e.g. kinesthesis and proprioception) [27].

2 Incidental Visual Memory and Navigation

The following data were previously reported by Thomas et al. [32]. Space constraints do not allow a complete recapitulation. This brief summary is included to provide some details on the methodology and results of the human data that the current modeling effort is attempting to match. Please consult the original study for a better understanding of the methodology.

2.1 Participants

The participants were 12 males with a mean age of 19.08 years (R = 18 - 24) who received monetary compensation for their participation and were in self-reported "average" or better health status. No participants reported a history of head or brain anomalies (traumatic injuries or illnesses). Participant visual acuity and color vision was screened. Immediately prior to their participation, the participants took part in a motion capture study that required wearing a portable motion capture suit and the eye tracker.

2.2 Materials, Apparatus, and Procedure

Materials. The navigation task included three rooms: the laboratory, a workshop, and a break area. See Figure 1 for an illustration of the layout. The laboratory served as the starting and ending point of the navigation task. The workshop floor dimensions were 230 feet × 49.5 feet with a ceiling height of 35 feet in one area and a dropped ceiling with a height of 9.5 feet in an area 7.5 feet from the wall with various laboratory entry doors spanning the entire length of the workshop. A 30-feet section spanning the breadth of the workshop at the area farthest from the lab also had the dropped ceiling. The break area consisted of an open floor plan with a glass block curved wall.

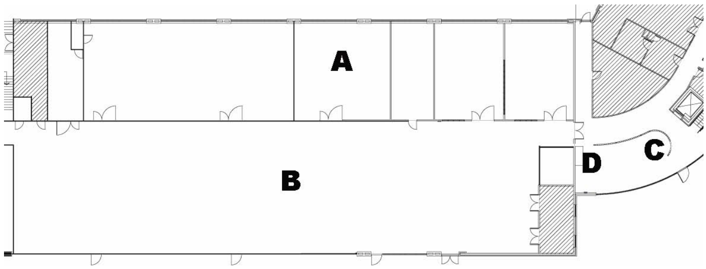


Fig. 1. Floor Layout: A. Human Factors and Ergonomics Lab, B. Workshop, C. Curved glass wall, D. Soft drink vending machine

The memory test stimuli were composed by taking advantage of natural occlusions in the workshop. The photographs were constructed so that none of the manipulated objects overlapped in the photographs. When possible, the photographs were taken from 3 different angles of incidence (left, right, and head-on). The images were reduced and cropped to 300 x 300 pixels. Three versions were photographed of each viewing angle when possible: (a) the original, unchanged scene, (b) with an object added to the scene, (c) and with an object removed from the scene. A fourth version of each scene was created by digitally changing the color of an object(s) in the scene. Thus, each scene had as many as 3 views and each view had as many as 3 foils which is a possible 9 foils for each scene. However, in order to prevent overlap of the manipulated scene regions and due to naturally occurring obstructions, composing 3 views of each scene was not always feasible, nor was adding or removing objects always feasible.

Apparatus. Right eye point-of-gaze was recorded with an Applied Science Laboratories Mobile-Eye tetherless infrared-video-based eye tracker. The eye tracker camera output at 60 Hz and the video capture was 29.97 frames per second. Therefore, the actual sampling rate is approximately 30 Hz. Memory test stimuli were displayed at a resolution of 800 × 600 pixels × 24-bit color on a flat screen LCD computer monitor viewed at approximately 75 cm (35.43° × 26.57° of visual angle). Test scenes were 300 × 300 pixels and subtended approximately 8.89° × 8.89° of visual angle at 75 cm viewing distance.

Procedure. Participants were escorted to the laboratory without passing through the portion of the workshop they would later be asked to navigate. Once in the laboratory, informed consent was obtained. Demographic data, anthropometric data, and a brief medical history were collected. Visual acuity and color vision were screened. Participants were fitted with a portable motion capture suit and the eye tracker. We interacted with participants for approximately 1 hour during another experiment before the navigation task began. Participants wore the eye tracker for the other tasks as well as the navigation task. Participants were not told about the memory test until immediately before test administration and they were not told we were investigating navigation until debriefing. We gave each participant change to buy a soft drink from a vending machine and told the participant to buy a drink of his choice. We pointed in the general direction of the vending machine and instructed the participant to, "Just go that way until you find the drink machine." If participants sought clarification, the researcher repeated the hand gesture and verbal instructions. No other clarification was offered.

Navigation consisted of walking out of the lab through a door and turning left, then walking 110 feet through a closed pair of double doors. Although the break area was not visible through the doors, it was visible through a small opening in the wall immediately after exiting the workshop via the double doors. A barricade was placed to prevent participants from using the immediate path. Thus, participants had to walk an additional 30 feet around the curving glass wall, turn 90° and walk another 25 feet to reach the vending machine. After interacting with the vending machine, participants had to return around the glass block wall, pass through the double doors, walk partially through the workshop, and pass through the lab door.

Upon returning to the lab, participants were immediately given an unannounced two-alternative forced-choice recognition test. Participants were tested on 18 scenes from the workshop. Test objects were presented to the right and left of a fixation point. The foil item was the same scene with either an item added to the scene, removed from the scene, or with a color replaced for an object in the scene. No scene was tested more than once for any participant. Task instructions were presented on the initial memory test screen and responses were input by participants with a standard button box. Participants were instructed to guess if they were uncertain. No feedback was given.

2.3 Analyses and Results

Video coding. All videos were independently coded by a senior undergraduate student and the primary author of the current paper. From a possible 31,944 coded frames, 26,325 were agreed upon. Interobserver agreement was 82.4% and the disagreements were resolved by conference. The reported eye movement and viewing behavior results are for 9 participants only. Two were excluded based on frame drop rates (≥ 90%) and the third file was corrupted. Percentages of the lengths of time spent looking at a focus of expansion were calculated by adding all the instances of looking at what appeared to be a focus of expansion for the entire navigation sequence minus the time spent interacting with the vending machine.

Scene viewing behavior was erratic. Participants often appeared to make ballistic saccades from one object to another within a scene. If this was the case, apparent fixations consisted of a single frame, but without at least 2 frames, an apparent fixation may have been a sample taken in mid-saccade. All calculations are based on a minimum of 2 consecutive frames and are believed to underestimate actual viewing behavior. Of the 9 videos retained for analyses, the average rate of frame drop was 22.17% and, other than eye blinks, it appears most of those were due to participants looking beyond the angle capabilities of the camera. Despite possible frame rate concerns, some objects were viewed as many as 53 frames (nearly 2 s).

Incidental Visual Memory. The average time spent navigating through the environment that was tested for memory was 60.5 s. (R = 44.47 - 78.88 s). The average time spent viewing the tested scenes was 6.9% (calculated by using a minimum of 2 consecutive frames). Memory test results indicated the incidental memory for the scenes was 58.82% [$t(11) = 2.39$, $p = .036$, $SE = .037$] with a standard deviation of 12.79%. Performance for 1 participant was more than 2 standard deviations below the mean and performance for another was more than 2 standard deviations above the mean. Both were included in the reported memory results. Performance excluding the two possible outliers ranged from 47.06% to 64.71%.

Navigation Behavior. The average total time spent navigating (total time from start to finish excluding the machine interaction time) was 118.4 seconds (*R* = 86.2 - 182.7 s). The 182.7 s time is inflated because the gentleman walked past the break area and into other areas of the building. Although 2 participants did not view any of the same scenes during the return trip through the workshop, an average of 2.55 objects (4.7%) viewed during the initial trip through the workshop were viewed again during the return trip through the workshop. Six participants viewed an initial object immediately outside the lab, but only 2 viewed the object again when returning to the lab.

Vertical architectural elements were viewed more often than any other aspect of the environment with 11.24% of the total navigation time spent viewing walls, columns, and other vertical spatial boundaries. Participants appear to have defined an average of 6 areas of reference that may have been foci of expansion. Including the initial reference gaze, participants averaged 21.1 looks at prospective foci of expansion for an average of 9.9% of the time spent while walking. In some instances the possible foci of expansion included walls and these instances were counted as both possible foci of expansion and vertical spatial boundary looks. The subjective ranges of times per looks at a possible focus of expansion ranged from a short 0.07 s glance to a long 5.2 s stare. The average reference look at a possible focus of expansion was 0.52 s.

3 Navigation Model

The navigation model is a computational model implemented within the ACT-R cognitive architecture [1]. The ACT-R cognitive architecture is a production system symbolic architecture that incorporates models of memory, learning, perception, and action. Models of human performance constructed within ACT-R specify if-then rules that coordinate the goal, memory, perception, and action modules. The ACT-R architecture and incorporated modules are formal implementations of psychological theories.

3.1 The Tasks

Navigation. The navigation task can be divided into direction inference, foci of expansion target identifications and reference looks, obstacle avoidance, and incidental visual memory. When using a focus of expansion, the actor selects a heading target to serve as a focus of expansion and moves toward that target. The actor periodically checks the focus of expansion to verify the heading and once the navigation target is reached, another navigation target is selected.

Obstacle avoidance. When an obstacle blocks the participant's desired path, the space at the extents of the obstacle can be tested to determine if a viable path exists to attempt to bypass the obstacle. In the navigation task, the only open space is at the left-most extent of the glass wall. The participant navigates to this space with a goal of finding an opening in the wall. In this case, the participant does find an opening. In a more complex, maze-like environment, other spatial based memory strategies would be needed.

Incidental Visual Memory. As the actor navigates through the environment and searches for the goal target (e.g., vending machine or lab), the actor is not only looking at and encoding obviously relevant objects but also other interesting objects in the environment. The shifting of attention to “irrelevant” objects is based on a combination of bottom-up and top-down factors. The visual features of some objects are particularly salient and will attract attention. Also, the visual features driving the participant’s search for the goal (i.e. color: red, category: box) will influence what objects are considered worthy of focused attention. The current model simulates these factors via feature-based search of the visual array.

3.2 The Sub-Tasks

Visual Search. Search is modeled as a feature-based conjunction search of the visual array [33]. As suggested by Williams, et al., the actor has some concept of a goal target (e.g., a soft drink vending machine) [14]. The actor searches the visual array for the features that are associated with their mental representation of the goal target. In this task, the actor has not seen the actual goal target and is assumed to be using some exemplar representation.

Encoding. The encoding of visual objects is accomplished by integrating features into chunks. The subsequent chunks are stored in declarative memory. Once visual features are encoded as objects, the model can compare them to representations of the goal target or evaluate them as potential targets for navigation.

Spatial Relationships. In navigation, the spatial relationship between the self and the environment is particularly important [16]. Spatial information is encoded from the visual array as part of the visual object chunk. This spatial information is egocentric in nature. The visual object chunk encodes the observer's bearing to and distance from the encoded object. An allocentric representation of the environment is built by constructing object-to-object relationships from the egocentric spatial relationships.

4 Conclusion

The current research seeks to construct computational models of human navigation in virtual three dimensional environments. The current model, within the ACT-R cognitive architecture [1], incorporates visual search, encoding of object features and spatial relationships, motion, obstacle avoidance, and incidental visual memory.

The navigation model is a computational model implemented within the ACT-R cognitive architecture which is a production system symbolic architecture that incorporates models of memory, learning, perception, and action [1]. The visual search implementation is a feature-based conjunction search of the visual array. The encoding of visual objects is accomplished by integrating features into chunks that are stored in declarative memory. Spatial information is encoded from the visual array as part of the visual object chunk. The spatial information is egocentric and the visual object chunk encodes the observer's bearing to and distance from the encoded object. The allocentric representation of the environment is constructed from the egocentric object-to-object spatial relationships. Navigation tasks used multiple foci of expansion target identifications and reference looks, obstacle avoidance, and incidental visual memory. The extents of the object were used to select a new navigation targets to bypass obstacles. Information about interesting objects in the environment is encoded.

References

1. Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., Qin, Y.: An Integrated Theory of the Mind. Psychological Review 111, 143–166 (2004)
2. Mitchell, M., Jolley, J.: Research Design Explained, 4th edn. Harcourt College Publishers, Fort Worth, Texas (2001)
3. Darken, R., Sibert, J.: Navigating Large Virtual Spaces. International Journal of Human-Computer Interaction 8, 48–72 (1996)
4. Sadeghiam, P., Kantardzic, M., Lozitskiy, O., Sheta, W.: The Frequent Wayfinding-Sequence (FWS) Methodology: Finding Preferred Routes in Complex Environments. International Journal of Human-Computer Studies 64, 356–374 (2006)
5. Thorndyke, P.W., Hayes-Roth, B.: Differences in Spatial Knowledge Acquired From Maps and Navigation. Cognitive Psychology 14, 560–589 (1982)
6. Thorndyke, P.W., Stasz, C.: Individual Differences in Procedures for Knowledge Acquisition From Maps. Cognitive Psychology 12, 137–175 (1980)
7. Hirtle, S.C., Jonides, J.: Evidence of Hierarchies in Cognitive Maps. Memory & Cognition 13, 208–217 (1985)
8. Huttenlocher, J., Hedges, L.V., Duncan, S.: Categories and Particulars: Prototype Effects in Estimating Spatial Location. Psychological Review 98, 352–376 (1991)
9. McNamara, T.P.: Mental Representations of Spatial Relations. Cognitive Psychology 18, 87–121 (1986)
10. Stevens, A., Coupe, P.: Distortions in Judged Spatial Relations. Cognitive Psychology 10, 422–437 (1978)
11. Rosenbaum, R.S., Ziegler, M., Winocur, G., Grady, C.L., Moscovitch, M.: I have often walked down this street before: fMRI Studies on the Hippocampus and Other Structures During Mental Navigation of an Old Environment. Hippocampus, 14, 826–835 (2004)
12. Taylor, H.A., Tversky, B.: Descriptions and Depictions of Environments. Memory & Cognition 20, 483–496 (1992)
13. Castelhano, M.S., Henderson, J.M.: Incidental Visual Memory for Objects in Scenes. Visual Cognition 12, 1017–1040 (2005)
14. Williams, C.C., Henderson, J.M., Zacks, R.T.: Incidental Visual Memory for Targets and Distractors in Visual Search. Perception & Psychophysics 67, 816–827 (2005)
15. Simons, D.J., Chabris, C.F., Schnur, T., Levin, D.T.: Evidence for Preserved Representations in Change Blindness. Consciousness and Cognition 11, 78–97 (2002)
16. Turano, K.A., Lei Hao, D.Y., Hicks, J.C.: Optic-flow and egocentric-direction strategies in walking: Central vs peripheral visual field. Vision Research 45, 3117–3132 (2005)
17. Gibson, J.J.: Perception of the Visual World. Houghton-Mifflin, Boston, MA (1950)

18. Gibson, J.J.: The Ecological Approach to Visual Perception. Houghton-Mifflin, Boston, MA (1979)
19. Warren Jr, W.H.: Visually Controlled Locomotion: 40 Years Later. Ecological Psychology 10, 177–219 (1998)
20. Warren, W.H., Kaye, B.K., Zosh, W.D., Duchon, A.P., Sahuc, S.: Optic Flow is Used to Control Human Walking. Nature Neuroscience 4, 213–216 (2001)
21. Duchon, A.P., Warren, W.H.: A Visual Equalization Strategy for Locomotor Control: Of Honey Bees, Robots, and Humans. Psychological Science 13, 272–278 (2002)
22. Harris, M.G., Carre, G.: Is Optic Flow Used to Guide Walking While Wearing a Displacement Prism? Perception 30, 811–818 (2001)
23. Kim, N.-G., Turvey, M.T.: Eye Movements and a Rule for Perceiving Direction of Heading. Ecological Psychology 11, 233–248 (1999)
24. Wilkie, R.M., Wann, J.P.: Driving as Night Falls: The Contribution of Retinal Flow and Visual Direction to the Control of Steering. Current Biology 12, 2014–2017 (2002)
25. Wilkie, R.M., Wann, J.P.: Eye-Movements Aid the Control of Locomotion. Journal of Vision 3, 677–684 (2003)
26. Regan, D., Beverley, K.I.: Visual Responses to Vorticity and the Neural Analysis of Optic Flow. Journal of the Optical Society of America 2, 280–283 (1985)
27. Cutting, J.E., Reading, W.O., Wang, R.F.: Walking, Looking to the Side, and Taking Curved Paths. Perception & Psychophysics 64, 415–425 (2002)
28. Harris, J.M., Bonas, W.: Optic Flow and Scene Structure do not Always Contribute to the Control of Human Walking. Vision Research 42, 1619–1626 (2002)
29. Rushton, S.K., Harris, J.M., Llyod, M.R., Wann, J.P.: Guidance of Locomotion on Foot Uses Perceived Target Location Rather Than Optic Flow. Current Biology 8, 1191–1194 (1998)
30. Hollands, M.A., Patala, A.E., Vickers, J.N.: Look where you're going!: Gaze Behavior Associated with Maintaining and Changing the Direction of Locomotion. Experimental Brain Research 143, 221–230 (2002)
31. Land, M.F., Tatler, B.W.: Steering With the Head: The Visual Strategy of a Racing Driver. Current Biology 11, 1214–1220 (2001)
32. Thomas, M.D., Carruth, D.W., McGinley, J.A., Follett, F.: Task Irrelevant Scene Perception and Memory During Human Bipedal Navigation in a Genuine Environment. In: Proceedings of the 25th Army Science Conference, Orlando, FL (2006)
33. Anderson, J.R., Matessa, M., Lebiere, C.: The Visual Interface. In: Anderson, J.R., Lebiere, C. (eds.) The Atomic Components of Thought, pp. 143–166. Erlbaum, Mahwah, NJ (1998)

A Data-Based Modeling Approach of Reach Capacity and Discomfort for Digital Human Models

Xuguang Wang, Elodie Chateauroux, and Nicolas Chevalot

LBMC, French National Research Institute for Transports and Safety (INRETS), 25 avenue François Mitterrand, Case 24, 69675 Bron, France
{xuguang.wang,elodie.chateauroux,nicolas.chevalot}@inrets.fr

Abstract. In this paper, we have proposed a unified data based approach which aims to predict both reach envelops and reach discomfort for a digital human model. Reach envelops can be obtained by applying the existing reach posture data to a new subject according to simulation scenario. Four reach surfaces are proposed according to the radial distance from the shoulder. The discomfort of a target on each surface needs to be defined at first. Then, the discomfort of an intermediate distance between two reach distances is interpolated. The proposed approach is illustrated by the data of a previous study. In this study, 38 young and elderly subjects were instructed to reach 94 targets for each from a seated position, covering a large reachable space.

Keywords: Reach, discomfort, digital human modeling, reach posture.

1 Introduction

For ergonomic design of workplace, most of the studies on arm reach were either limited to the determination of reach envelopes [1-4] or to motion analysis for understanding motion control strategies or/and their simulation [5-8]. Recently, a few investigators studied the discomfort (or difficulty) of arm reach ([9], [10]). The proposed discomfort predictive models were either based on target position or on reach posture (joint angles). Both models cannot take into account maximal reach limits directly, as target location and joint angles are continuous variables and no criteria were defined for testing if a target is out of reach. The predictive models have to be able to differentiate reachable targets from those of out-of-reach. Discomfort evaluation makes sense only for reachable targets. Reach envelopes were studied in the past. But only statistical models for predicting population envelopes were developed. Though they are useful especially for car interior design, they cannot be used for predicting an individual reach capacity in case of DHM (Digital Human Models) applications. In this paper, a unified data based approach for digital human models will be presented for predicting both individual reach envelopes and discomfort of a reachable target.

In order to illustrate the proposed approach, the data from a previous study on seated reach discomfort will be used [10].

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 215–223, 2007.

2 Data and Preliminary Analysis

2.1 Data Collecting

Thirty eight subjects participated in the experiment and were paid for it. They were divided in two age groups: 18 young subjects (9 women and 9 men) aged between 20 and 33 years and 20 elderly people (10 women and 10 men) aged between 64 and 76 years. In order to reduce the effects of environment (seat type, cloth, etc…) on reach capacity, the subjects were seated on a flat stool and were asked to push a toggle switch. The seat height was adjusted so that the knees were flexed around 90 degrees. Subjects were strapped on the seat to restrict pelvis motions relative to the seat. They were instructed to reach the toggle switch naturally with the right hand from a same starting position for all targets and then to go back to the starting position. The experimental device is shown in Figure 1. The subjects were asked to keep both arms along the body and the torso upright at the beginning.

Target locations covered a wide workspace within the reaching limits of the subjects:

- 5 plane orientations: -45° (left), 0°, 45°, 90° and 135°
- 5 heights
- 4 distances

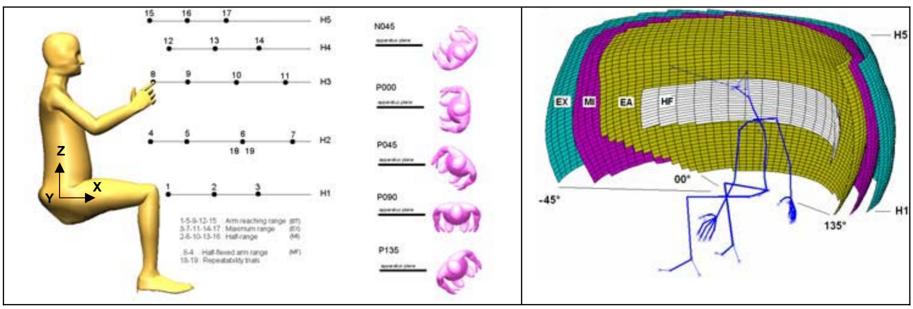


Fig. 1. Target location and illustration of the four reach surfaces of a subject

The five plane orientations were the sagittal plane through the right shoulder (P000), 45° on the left (N045), the planes (P045, P090, P135) 45°, 90° and 135° on the right with respect to P000. Target heights and distances were defined with respect to one’s own anthropometry and reach capacity. The five heights were seat height H1, shoulder height H3, mid height between H1 and H3 (H2), height corresponding to an arm elevation of 135° (H5), mid height between H3 and H5 (H4). The four radiale distances were defined with respect to the index fingertip maximum reach distances without (Extended Arm EA) and with (Extreme distance EX) torso participation for each height. MI was the mid distance between EA and EX. HF (Half flexed arm) was defined by subtracting the hand length from EA. Reach distances EA and EX for each height in each orientation plane were determined prior to the experiment for every subject. Due to the interference with the left knee, the target N045-H1-EA was removed. In total, eighty four different positions (17x5-1) were defined. In order to test the repeatability of discomfort rating and movement, the target H2-MI was repeated three times for every plane orientation (6-18-19 in Figure 1). In total, each subject was asked to reach ninety four (94) targets. The trial order was randomly chosen in an orientation plane which was also randomly fixed. Movements were captured using the optoelectronic system VICON. After each reach movement, the subjects were asked to rate the associated discomfort using a slightly modified category partition scale CP-50 [10]. The ratings were ranged from 0 to 50.

To ensure a good quality of motion reconstruction, the markers trajectories provided by VICON were smoothed and missed markers were recovered as much as possible thanks to redundant number of markers. Then, for each subject, a digital model was manually defined by superimposing the model on the subject's photos in a referential posture. The positions of markers in local body segment coordinate systems were identified on the digital model from a referential posture. Finally, joint angles were calculated by inverse kinematics by minimizing the distance between markers positions of the model and measured ones. Refer to the work by Wang et al. [11] and Ausejo et al. [12] for more details. Figure 2 shows an example of reconstructed reach posture with a digital human model.

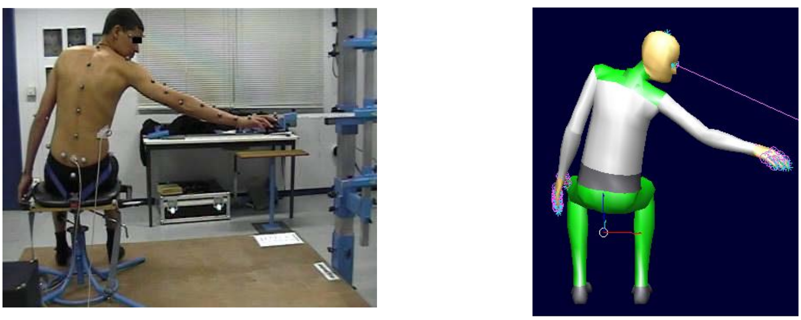


Fig. 2. Real and reconstructed reach postures

2.1 Preliminary Analysis

A preliminary analysis was carried out to examine the effects of age, gender and target location on perceived discomfort (Table 1). Target position was defined in a seat spherical coordinate system centered at the mean shoulder center of the starting postures of all trials. Therefore a target is defined by its longitude (α), latitude (β) and its radial distance (ρ_n) from the shoulder center normalized by the upper limb length (see also Figure 4). From Table 1, one can see that there was a strong effect of age. The gender had negligible effect. As expected, target location had a strong effect. The target latitude and longitude had a strong interaction with age. However, target distance had the same effect for both younger and older groups. This is probably due to the fact that the target distance was defined with respect to one's reach capacity.

In what follows, we will divide the data only according to age group without distinction of the gender.

Table 1. ANOVA table showing effects of age, gender and target location on perceived discomfort of reach posture

Source	Sumof squares	Dof	Mean sq.	Ratio F	Proba.
Age (A)	459.234	1	459.234	6.88	0.0087
Gendre (G)	14.4652	1	14.4652	0.22	*0.6416*
Distance (ρ_n)	34464.9	1	34464.9	516.23	0.0000
Longitude (α)	7081.79	1	7081.79	106.07	0.0000
Latitude (β)	39.6942	1	39.6942	0.59	*0.4407*
A x ρ_n	2.96247	1	2.96247	0.04	*0.8332*
A x α	383.843	1	383.843	5.75	0.0165
A x β	2656.3	1	2656.3	39.79	0.0000
ρ_n x α	926.417	1	926.417	13.88	0.0002
ρ_n x β	943.301	1	943.301	14.13	0.0002
α x β	101.05	1	101.05	1.51	*0.2186*
Residu	224324.	3360	66.763		
Total	391186.	3371			

3 A Unified Reach and Reach Discomfort Modeling Approach

As stated in Introduction, our aim is to propose a unified approach which predicts both reach capacity and reach discomfort for a digital human model. The proposed approach consists of three steps. As it is a data-based approach, the first step is to collect data for both reach postures and discomfort ratings. In order to make discomfort ratings comparable between subjects, target location should be defined with respect to one's anthropometry and reach capacity. Four reach surfaces were defined according to reach distance in our experiment: nearly half flexed distance without moving the torso (HF), extended arm length (EA) without moving the torso, maximum reach distance with torso participation (MX), mid distance (MI) between the distances EA and MX. Once the data are collected, the next step is to structure the data of postures for each of these four reach distances in terms of subject's anthropometric characteristics (sitting height, upper limb length, age, gender, joint mobility …). The database is then used for generating these four reach surfaces for an individual digital human according to simulation scenario. The same method was used for constituting an in-vehicle reach motion database [13]. The third step is to fit the discomfort ratings in terms of target position (longitude and latitude) on each of these four reach envelops using a surface regression fitting method. The discomfort of an intermediate distance between two reach distances will be interpolated (Figure 4). The reason why we fit discomfort ratings on reach envelope instead of target position is that a target distance is judged with respect to one's reach capacity. A far target for a short person is certainly not as far as for a tall person. The short and tall persons will certainly not rate the target in the same way. We believe that the discomfort for a target defined with respect to one's reach surfaces like EA and MX is rated in the same way between different people. This makes it possible to compare the ratings of the subjects of different reach capacity and to reduce the number of variables in the discomfort predictive model.

Figure 3 shows how the proposed approach can be used for simulating a subject's reach envelop as well as his(her) reach discomfort for a target. The basic idea is to re-use the structured database for extracting the reach postures of the most similar subject who participated in the real experiment. As only the posture data are re-used, the four reach envelops are generated by applying these posture data using direct kinematics to the anthropometric dimensions of the virtual subject. Therefore, the anthropometric dimensions are indirectly taken into account for predicting reach capacity and discomfort.

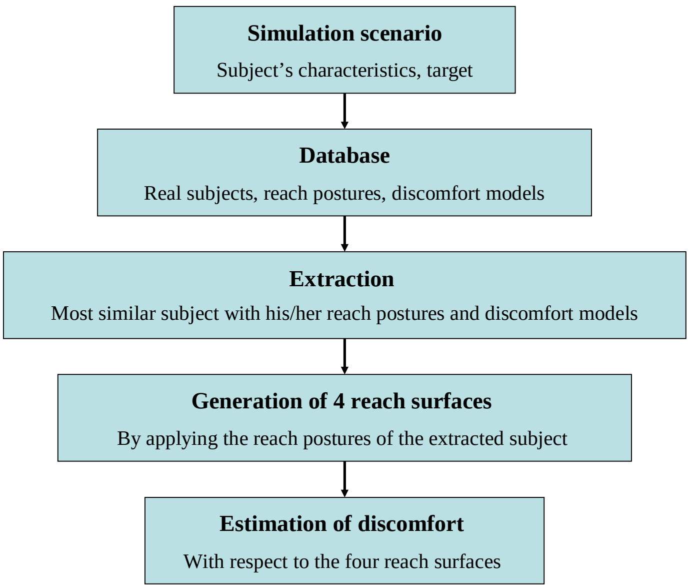


Fig. 3. Application the proposed approach to simulate a subject's reach envelopes and reach discomfort

4 Discomfort Models

In the proposed approach, we need to know how discomfort varies in terms of target position on each of these four reach envelops at first. Then, reach discomfort of a target is obtained by interpolation as function of its distance with respect to the intercepted surfaces (Figure 4). The discomfort models for each reach surface can be obtained using a surface regression fitting method. The aim of model fitting here is to express the discomfort rating as a function of target longitude α and latitude β for a reach surface. The position of zero longitude and zero latitude is defined when the target is located in the frontal plane at the shoulder height (Figure 4). Here the method of surface fitting with orthogonal polynomials was applied to our data. The same method was also used to model the upper arm axial rotation limits [14]. As in [14], a transformation on β is necessary so that the model should not be sensitive to the variation of longitude when the target position is near to the poles β =±π/2:

$$x(\alpha,\beta) = \alpha\cos\beta, \quad y(\alpha,\beta) = \beta \tag{1}$$

making the variable x independent of the longitude angle α at the poles. We assume that the discomfort of a target on a reach surface can be described by the following staistical regression model

$$D(\alpha,\beta) = \sum_{j=0}^{k} a_j P_j(x,y) + \varepsilon \tag{2}$$

where P_j is a j-th ordered polynomial basis with two independent variables x and y, k is the highest order of the polynomial basis (the degree of regression) and ε is a normal random error.

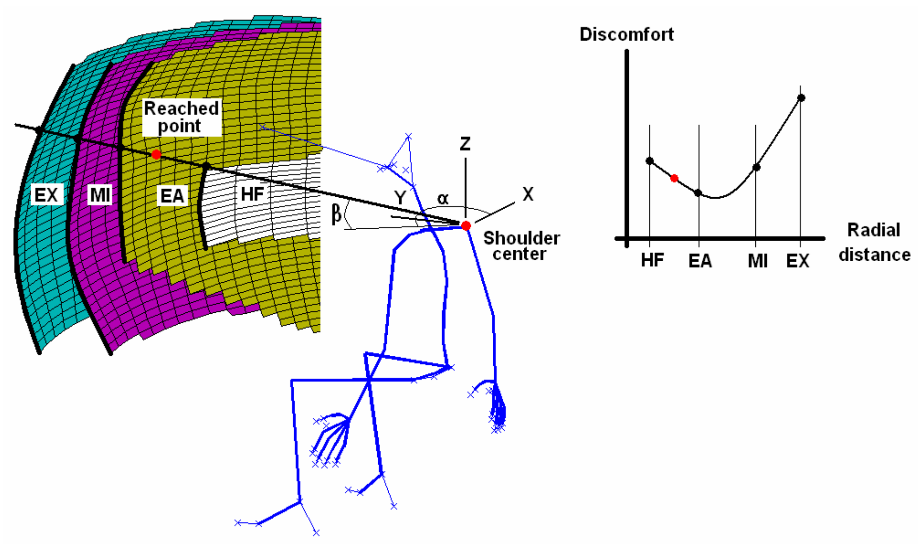


Fig. 4. Estimation of the reach discomfort of a target by interpolation as a function of its distance to the intercepted surfaces

Instead of using a direct polynomial basis, which may suffer from computational complications and may have oscillatory behavior particularly near to the boundary of the data area, an orthogonal homogeneous polynomial basis can be used,

$$P_j(x,y) = \sum_{m=0}^{j} b_{jm} x^{j-m}(\alpha,\beta) y^m(\alpha,\beta) \tag{3}$$

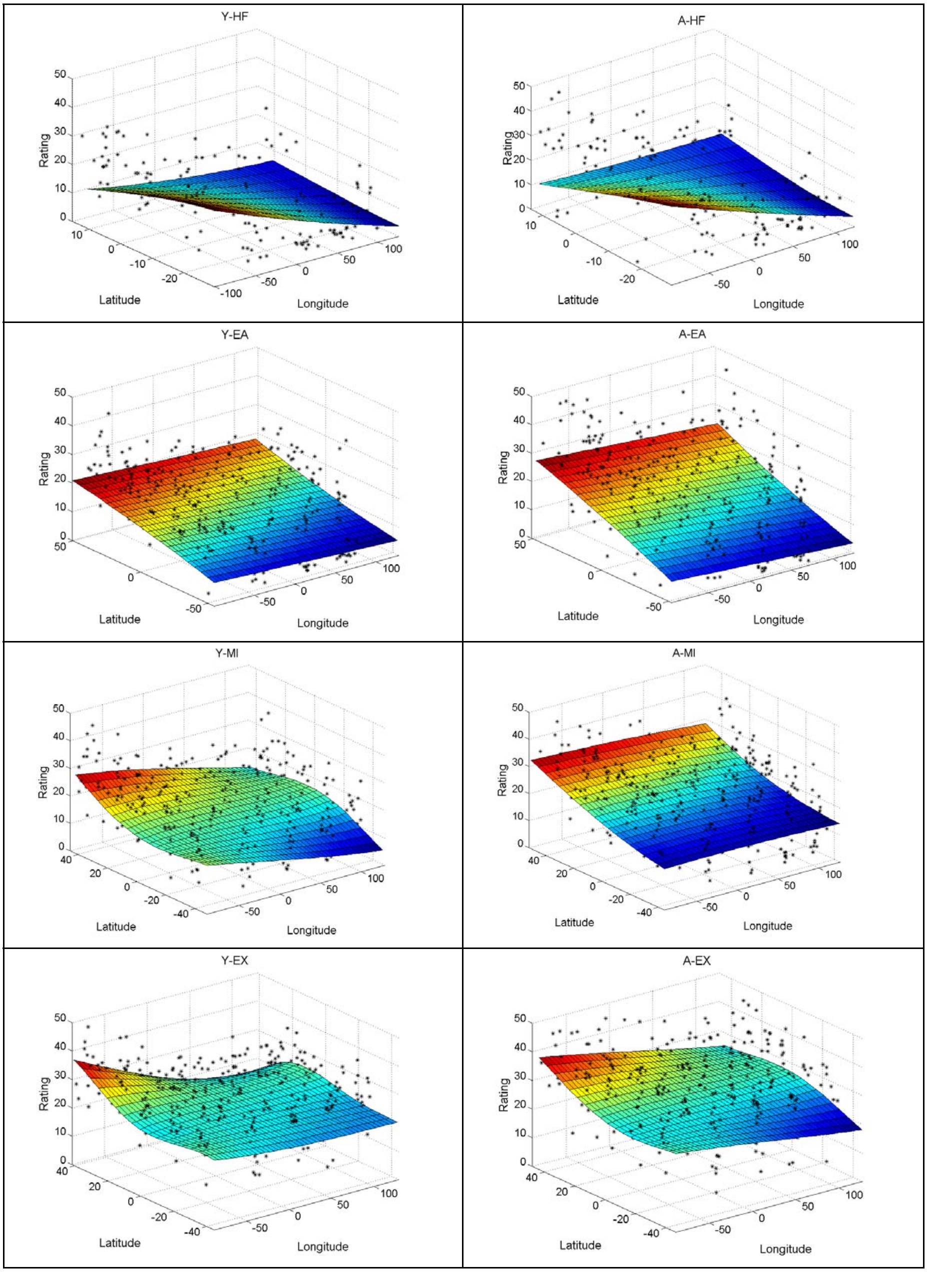


Fig. 5. Modeling of discomfort rating of a target located in 4 reach surfaces (HF, EA, MI and EX). The models for the young subjects (on the left column) and the aged ones (on the right column) are compared.

with b_{j0}=1 for each j and $P_0(x,y)=0$. The following conditions of orthogonality allow the calculation, in a step-by-step way, of the j unknown coefficients b_{j1}, b_{j2}, …, b_{jj} for every P_j ($j \geq 1$) based upon the n pairs of recorded data (α_i, β_i), i=1,2,...,n,

$$\sum_{j=1}^{n} P_j(x_i, y_i) P_m(x_i, y_i) = 0 \text{, m = 0, 1, 2, …, j-1} \tag{4}$$

The coefficients a_j and k in (2) can be determined by classical multiple regression procedure. Refer to [14] for more details.

Figure 5 compares the discomfort models for the 4 reach surfaces. From the preliminary analysis, a strong effect of age group was observed. The data were therefore separated according age groups. In Figure 5, the models for the young and elderly subjects are also compared. One can see that subjects preferred the targets which were located forwardly below the shoulder height, as expected. Notice that there is a slight difference between young and elderly groups for the surfaces MI and EX. The elderly subjects were more sensitive to target latitude.

5 Concluding Remarks

In this paper, we have proposed a unified data based approach which aims to predict both reach envelops and reach discomfort for a digital human model. Like all data based approaches, the prediction depends on the wealth of data. Besides, it depends on how a referential subject is selected from database. In addition to classical anthropometric dimensions, other criteria are required for describing one's characteristics. Clearly, the reach envelops and discomfort of reaching a target depend on one's reach capacity. This should be included in the selection criteria. However, it is difficult to define a global index that can be used for the definition of reach capacity. One solution may be to take all subjects contained in the same age group and to simulate the reach discomfort from these data, thus the inter-individual variability being taken into account. In workplace design, the question often posed is how many percent of target population can reach a control under a certain limit of discomfort rating. This requires the simulation of a population, called 'population simulation'. This concept has been proposed and implemented in our ergonomic simulation software RPx ([13], [15]). The basic idea behind is to generate a representative sample of digital human subjects at first and then to evaluate the overall discomfort based on individual responses using the proposed approach.

References

1. Chaffee, J.W.: Methods for determinating driver reach capability. SAE report 6901205, New York (1969)
2. Kennedy, F.W.: Reached capability of men and women: A three dimensional analysis. Aerospace medical research laboratory, Wright-Patterson Air force base, Yellow springs, Ohio (July 1978)
3. Boydstun, L.E., Kessel, D.S.: Maximum reach models for the seated pilot. Center for ergonomics, college of engineering, University of Michigan (March 1980)

4. Sengupta, A.K., Das, B.: Maximum reach envelope for the seated and standing male and female for industrial workstation design. Ergonomics 43(9), 1390–1404 (2000)
5. Chaffin, D.B., Faraway, J.J., Zhang, X., Woolley, C.: Stature age and gender effects on reach motion postures. Human factors 42(3), 408–420 (2000)
6. Wang, X., Verriest, J.P.: A geometric algorithm to predict the arm reach posture for computer-aided ergonomic evaluation. The Journal of Visualization and Computer Animation 9, 33–47 (1998)
7. Zhang, X., Chaffin, D.B.: A three-dimensional dynamic posture prediction model for in-vehicle seated reaching movements: development and validation. Ergonomics 43(9), 1314–1330 (2000)
8. Jung, E.S., Choe, J.: Human reach posture prediction based on psychophysical discomfort. Internaltional Journal of Industrial Ergonomics 18, 173–179 (1996)
9. Reed, M.P., Parkinson, M.B., Chaffin, D.B.: A new approach to modeling driver reach, 2003 SAE World Congress, Detroit, Michigan, March 3-6, 2003. SAE Paper N° 2003-01-0587 (2003)
10. Chevalot, N., Wang, X.: Experimental investigation of the discomfort of arm reaching movements in a seated position. SAE Transactions 2004 Journal of Aerospace 1, 270–276 (2004)
11. Wang, X., Chevalot, N., Monnier, G., Ausejo, S., Suescun, Á., Celigüeta, J.: Validation of a model-based motion reconstruction method developed in the Realman project. In: SAE International conference and exposition of Digital Human Modeling for Design and Engineering, June 14-16, 2005, SAE paper 2005-01-2743 (2005)
12. Ausejo, S., Suescun, Á., Celigüeta, J., Wang, X.: Robust Human Motion Reconstruction in the Presence of Missing Markers and the Absence of Markers for Some Body Segments. In: SAE International conference and exposition of Digital Human Modeling for Design and Engineering, Lyon, France, July 4-6, 2006, SAE paper 2006-01-2321 (2006)
13. Wang, X., Chevalot, N., Monnier, G., Trasbot, J.: From motion capture to motion simulation: an in-vehicle reach motion database for car design. In: SAE International conference and exposition of Digital Human Modeling for Design and Engineering, Lyon, France, July 4-6, 2006, SAE Paper 2006-01-2362 (2006)
14. Wang, X., Maurin, M., Mazet, F., De Castro Maia, N., Voinot, K., Verriest et M Fayat, J.P.: Three-dimensional modeling of the motion range of axial rotation of the upper arm. Journal of biomechanics 31(10), 899–908 (1998)
15. Monnier, G., Renard, F., Chameroy, A., Wang, X., Trasbot, J.: Motion simulation approach integrated into a design engineering process. In: SAE International conference and exposition of Digital Human Modeling for Design and Engineering, Lyon, France, July 4-6, 2006, SAE Paper 2006-01-2359 (2006)

Motion Retrieval Based on Temporal-Spatial Features by Decision Tree

Jian Xiang[1] and HongLi Zhu[2]

[1] ZheJiang University of Science and Technology, School of Information and Electronic Engineering, 310023, HangZhou, China
[2] ZheJiang University City College, College of Information & Electrical Engineering, 310015, HangZhou, China
freexiang@gmail.com

Abstract. In order to retrieve similar motion data from Mocap database, each human joint's motion clip is regarded as a bag, while each of its segments is regarded as an instance. 3D temporal-spatial features are extracted and data driven decision trees are automatically constructed to reflect the influence of each point during the comparison of motion similarity. At last we use the method of multiple instance retrieval to complete motion retrieval. Experiment results show that our approaches are effective for motion data retrieval.

Keywords: Motion Capture, 3D Temporal-Spatial, Data Driven, Decision Tree, Retrieval.

1 Introduction

Now more and more motion capture systems are used to acquire realistic human motion data. Due to the success of the Mocap systems, realistic and highly detailed motion clips are commercially available and widely used for producing animations of human-like characters in a variety of applications, such as simulations, video games and animation files[1]. Therefore an efficient motion data retrieval technique is needed to support motion data processing, such as motion morph, edition and synthesis, etc. At present, most motion data are stored in Mocap database with different length of motion clips, which is convenient for manipulating in animation authoring systems and retrieval based on keyword or content.

Until now, several motion features have been proposed: Y.Chui et al.[3][4] proposed local spherical coordinates relative to the root orientation as the segments posture of each skeletal segment; Liu et al.[5] constructed a motion index tree based on a hierarchical motion description for motion retrieval; Lee et al.[6] described a two-layer structure for representing human Mocap data. The lower layer is a Markov process that retains the details of the original Mocap data, while the higher layer is a statistical model that provides support for the user interfaces by clustering the data to capture similarities among character states; Asaa et al[7] proposed a method of Action Synopsis, which used original motion capture data, such as joints positions, angles, speed, angular velocity, etc. , to make some affinity matrices. Then a low dimension

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 224–233, 2007.

motion curve is produced from these affinity matrices by an RMDS method. The above-mentioned methods, which extracted features directly from original data, need some dimension reduction algorithms to avoid “Curse of dimensionality”. Mueller et al.[8] introduced 31 Boolean features expressing geometric relations between certain body points of a pose. These geometric features seem to be a kind of local 3D semantic features. But they are extracted by the relationship of some neighbor joints, so geometric features is a kind of 2D features. The geometric features are too complex, which is 31 dimensions.

The temporal-spatial feature is defined in this paper first, which describes 3D space relationship of each joint. Unlike traditional 2D motion features, the 3D temporal spatial feature of each joint can represent a part of the whole motion independently. after extraction of temporal-spatial features, this paper builds 16 index lists for 16 human joints and similarity measure is implemented in each index list by DTW algorithm[5] separately. In order to learn the contribution of each joint during similarity measure, a data-driven decision tree is automatically constructed to reflect each joint’s weight. Thus, we can measure similarity of joints with largest weight between motion example Q and motion A in the motion database. If these joints are dissimilar, motion A is skipped; if and only if similar, similarities of other joints between Q and A are calculated. It is obviously that this method can save computing time dramatically.

So it is possible to use machine learning method to study the relationship of each joint of human motion.

2 Temporal-Spatial Features

In this paper, a simplified human skeleton model is defined, which contains 16 joints that are constructed in the form of tree. Joint *root* is root of the tree and those paths from root to all endmost joints in human skeletal model from sub-trees of root.

World coordinate of each joint can be represented as follow:

$$\vec{p}_i^{(j)} = T_i^{(root)} R_i^{(root)} ... T_0^{(grandparent)} R_i^{(grandparent)} (t) T_0^{(parent)} R_i^{(parent)} \vec{p}_0^{(j)} \qquad (1)$$

where $\vec{p}_i^{(j)}$ is world coordinate of joint j at time i, $T_i^{(root)}, R_i^{(root)}$ are the position matrix and orientation matrix of root at time i, $T_0^{(k)}$ is position matrix of joint N_k (N_k is an arbitrary joint in the paths from Root to endmost joint in human skeleton tree) in local coordinate system of its parent at start time; $R_i^{(k)}$ is orientation matrix of joint N_k at time i, made by r_i^k , $\vec{p}_0^{(j)}$ is position matrix of joint N_j in local coordinate system of its parent at start time.

2.1 Temporal-Spatial Feature Extraction

According to Equation (1), we can calculate world coordinate of each joint and get 48 dimensional data.

Given a motion M consisting of n sampling frames, each motion can be represented as follow:

$$F_i = (p_{i1}, p_{i2}, ..., p_{ij}, ..., p_{i16})$$
$$M_s = (F_1, F_2, ..., F_i, ..., F_n) \quad (2)$$
$$p_{ij} = (x, y, z)$$

where n is the number of frames of motion data, p_{ij} is world coordinate of joint j at i^{th} frame.

Now space transformations of each joint are calculated. Firstly, we define a space transformation set of upper body S_{up}, and a space transformation set of lower body S_{down} as following:

$S_{ui} \in S_{up}$, i=1,2…m; $S_{dj} \in S_{down}$, j=1,2…m; where m is the number of spaces in space transformation set, S_{up} and S_{down} have the same number of spaces. If we take Root as benchmark, then space transformations of joints above Root belong to S_{up}, and others belong to S_{down}, if a joint on upper body enters into space S_{ui}, its space transformation is S_{ui}.

Four space partition rules are defined as follow:

$$front(N_i, N_j) = \begin{cases} 1, N_i \text{ in front of } N_j \\ 0, N_i \text{ behind of } N_j \end{cases} \quad left(N_i, N_j) = \begin{cases} 1, N_i \text{ left to } N_j \\ 0, N_i \text{ right to } N_j \end{cases}$$

$$high(N_i, N_j) = \begin{cases} 1, \ N_i \text{ above } N_j \\ 0, \ N_i \text{ below } N_j \end{cases}$$

$$far(N_i, N_j) = \begin{cases} 1, \ N_i \text{ distance from } N_j \ > \lambda \\ 0, \ N_i \text{ distance from } N_j \ < \lambda \end{cases}$$

where rules of front, left and high depend on space relationship of up/down and left/right between joint N_i and N_j, rule of far depends on range of motion. As usual, in rules of front and left, N_j is Root, but in rules of high and far, N_j on upper and lower body are different. N_i, N_j are both at the same sampling frame.

Now we define motion space transformations:

$$B=(S_1, S_2, ..., S_{16})', S_i = (s_{i1}, s_{i2}, ..., s_{in}) \quad (3)$$

where S_i is space transformation vector of joint i, n is the number of frames, s_{ip} is space transformation of joint i at p th frame. Suppose S_a is space transformation vector of joint a on lower body, S_a =(s_{a1} , s_{a2} … s_{aj} … s_{an}):

Table 1. space rule table to calculate s_{aj} , N_{aj} is joint a at j th frame, N_{rj} is joint root at j th frame, N_{kj} is joint knee at j th frame

S_{aj}	$front(N_{aj},N_{rj})$	$left(N_{aj},N_{rj})$	$high(N_{aj},N_{kj})$	$far(N_{aj},N_{kj})$
$S_{aj}=S_{d1}$	1	1	1	1
$S_{aj}=S_{d2}$	0	1	1	1
...	...	...	...	...
$S_{aj}=S_{dm}$	0	0	0	0

In table 1, some rules can be concluded:

If $s_{aj} = s_{d1}$ $\Leftrightarrow$ rule:

$$front(N_{aj}, N_{rj}) \wedge left(N_{aj}, N_{rj}) \wedge$$

$$high(N_{aj}, N_{kj}) \wedge far(N_{aj}, N_{kj})$$

……

If $s_{aj} = s_{dm}$ $\Leftrightarrow$ rule:

$$\neg front(N_{aj}, N_{rj}) \wedge \neg left(N_{aj}, N_{rj}) \wedge$$

$$\neg high(N_{aj}, N_{kj}) \wedge \neg far(N_{aj}, N_{kj})$$

The rules cited above are calculated by 48 dimension data from Equation (2). Because these rules are all calculated at same frame, time and space complexity are not high. Moreover, space transformations of each joint are independent.

For example, we extract local space transformations of motion run’s left foot and right foot as following:

$$S_{leftfoot} =(S_{dk}, S_{dj}, S_{dk}, S_{dj}, \ldots);$$

$$S_{rightfoot} =(S_{di}, S_{dl}, S_{di}, S_{dl}, \ldots);$$

Up to now, motion’s space transformations are extracted, which is a kind of the reflection of motion spatial characteristic. But first of all, a complete motion is a group of time series data. Without time property, temporal-spatial features cannot represent motion clearly.

So the time property of motion is calculated as a part of temporal-spatial features.

The first time property is space transformation speed. Because of independence of each joint's space transformations, space transformation speed is independent either. The algorithm can be summarized as follow:

Procedure SpaceSpeed()

Input: local space transformation vector of kth joint $s_k = (s_{k1}, s_{k2}, ..., s_{kn})$, n is the number of frames.

Output: $SP_k = (SP_{k1}, …, SP_{ki}, …)$, SP_{ki} is space transformation S_{ki}'s speed of kth joint.

(1) Initialization: num_j=0,i=1,j=0,L=S_{ki}

(2) if $s_{ki} \neq s_{k(i+1)}$, {$spacespeed_{kl}$ = num_j ,l=$S_{k(i+1)}$,j=j+1}
else num_j=num_j+1;

(3) i=i+1,if meet end of frames goto (4) else goto (2)

(4) return SP_k

This spacespeed is actually the speed of a joint moving from a space to another. The weighted sum of every joints' spacespeeds consists of the whole motion's spacespeed.

During similarity measure, because of irregularity and contingency of human motion, there are odd space transformations that cannot be matched. Therefore spacenoise is defined to measure some odd space transformations.

Procedure SpaceNoise()

Input: local space transformation vector of kth joint $s_k = (s_{k1}, s_{k2}, ..., s_{kn})$, n is the number of frames

Output: $SpaceNoise_k$

(1) Initialization: num_j=0,i=1,j=0,l=1

(2) if $s_{ki} \neq s_{k(i+1)}$

Noise= num_j, j=j+1,

if $\frac{Noise}{n} < \varepsilon$ add S_{ki} to $SpaceNoise_k$
else
num_j=num_j+1;

(3) i=i+1,if meet the end of frames goto (4) else goto (2)

(4) return $SpaceNoise_k$

As space transformations, spacespeeds and spacenoises of 16 joints are gotten, complete temporal-spatial features are formed through the merger of them.

2.2 Keyspace Construction and Indexing Mulas

Since space is continuous, a joint is always in same space at some consecutive frames, which is spatial invariant. Therefore frames with the same temporal-spatial features can be merged into a keyspace. The keyspace is similar to the keyframe, but the keyspace is the local concept for each joint and keyframe is the global one for the whole motion. The following algorithm is proposed to calculate keyspace:

Procedure keyspace()

Input: local temporal-spatial features of k th joint $st_k = (st_{k1}, st_{k2}, ..., st_{kn})$. n is the number of frames

Output: keyspace $_k$ =(..., $st_{ki}, st_{k(i+1)}, st_{k(i+2)}$ ...)

($st_{ki} \neq st_{k(i+1)}, st_{k(i+1)} \neq st_{k(i+2)}$)

(1) Initialization: i=0, $s_{ki} = \phi$

(2) i=i+1,if meet the end of frames,goto (4) else goto (3)

(3) if $st_{ki} \neq st_{k(i-1)}$ {add st_{ki} to keyspace $_k$,goto (2)} else goto (2)

(4) return keyspace $_k$,S $_p$

Now the space transformation is simplified to the keyspace of each joint and similarity measure can implement in keyspaces directly that can reduce time and space complexity.

We build a motion retrieval database with 16 index lists by temporal-spatial features of each joint. For a query, the similarity measure is for each joint and so it is the local measure. In this study, it can apply local DTW algorithm [7].

Temporal-spatial features can be used to extract semantic features of common motions.

For running, man's left foot is moving from S $_{dk}$ to S $_{dj}$, then moving from S $_{dj}$ to S $_{dk}$, and so on. And the movement of right foot is between S $_{di}$ and S $_{dl}$. Furthermore, moving speeds of two feet are almost the same, so are moving ranges. It is widely known that a man's movement of two hands is similarity to his feet when he is running or walking. So we can summarize the rules of temporal-spatial features for some important joints from running. These rules are semantic features. As mentioned above, most kinds of motions have their own semantic features.

By semantic features, some common motions are recognized and retrieved automatically, an inverse retrieval system is made by using keyword (e.g. walk, jump) instead of query example to implement motion retrieval and retrieval time is reduced.

Now our system can extract semantic features of some simple motions (such as walk, run, wave, jump) from temporal-spatial features.

3 Decision Tree Induction

The earliest system of decision tree is CLS designed by Hunt[9].This method with the help of the intuition is short of mathematics foundation. Now the methods based on probabilistic theory are universally adopted. One of the well-known methods is a decision tree induction algorithm: ID3, proposed by Quilan in 1986[10]. After ID3, Quilan also gave some improved algorithm, such as C4.5 [11], C5.

The principle of ID3 is that the choice of attribute should maximize the information gain and minimize the Entropy while the average test times from root to leaf node is minimal in probabilistic.

Given some classes, the number of classes is c, if attribute A=v in all examples, its Entropy is defined to be:

$$\text{Entropy (A=v)}=\sum_{i=1}^{c} - p_i \log_2 p_i$$

The information gain of attribute A is defined as the difference between the original Entropy and the new Entropy after classifying train examples by attribute A.

$$\text{Gain(T,A)}=\mathrm{E(T)} - \sum_{j=1}^{V} \frac{|T_j|}{|T|} E(T_j)$$

where T is train set of examples, T_j is a train set of examples when A=j, $T_j \subset T$

A decision tree is constructed automatically by training six hundred human motions in database. (see Fig.1).

Now can conclude that two feet’s temporal-spatial features have the greatest effect on human motion similarity, and two hands are the second. And the influences of other joints cannot be compared with. Given a query example A and a motion B in the database, during the similarity measure, two feet’s temporal-spatial features should be calculated firstly; if they are similar, two hands are calculated secondly; if dissimilar, Motion B is skipped and a new motion is taken from database for motion retrieval. Other joints’ temporal-spatial features are calculated only if temporal-spatial features of two feet and two hands are similar. Consequently, Retrieval based on decision tree learning avoids a great deal of meaningless DTW calculating and becomes more efficient.

4 Experiment Results

We implement our retrieval algorithm in matlab. It is more than 1300 real human motion captured sequences with different types in motion database for test. Different types of query examples and key words of some common motions can be used. We also summarize some semantic features by temporal-spatial features to recognize some types of motions.

Now a methodology for assessing prediction quality after the fact is introduced [12]:

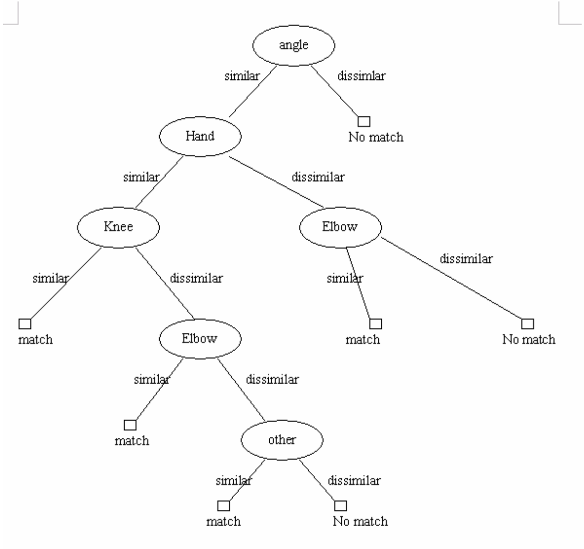


Fig. 1. Decision tree during motion retrieval

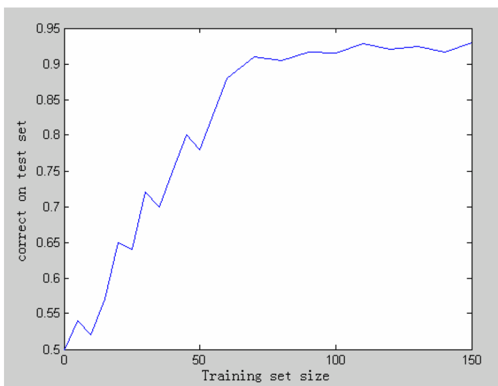


Fig. 2. The learning curve of decision tree built by motion capture data

The learning curve for the decision tree with motion capture data is shown in Fig.2.

Notice that as the train set exceeds 70, the prediction quality remains stable. So the algorithm of data driven decision tree is reliable.

Table 2 shows that the retrieval time of our method by decision tree is so much less than time of conventional method (such as the method of motion retrieval based on bone feature and keyframe) that is measuring all human joints with different weight that the retrieval performance has been improved significantly. To compare motion retrieval efficiency of the proposed method with that of the method based on bone

Table 2. performance comparison of some retrieval systems. KF is the method of motion retrieval based on bone feature and keyframe, SF&DT is the method of motion retrieval based on temporal-spatial features and data driven decision tree.

method	Precision	Recall	Retrievaltime (500clips database)
SF&DT	92.6%	96.2%	8.1292s
KF	79.1%	83.5%	35.6185s

angles feature and keyframe extraction, table 1 also shows the recall and precision of these two methods in the same motion database. It is obvious that the retrieval accuracy of the proposed method is better than the other one.

5 Conclusion

In this paper, temporal-spatial features are proposed which describe 3D space relationship of each joint and each joint has its own independent temporal-spatial features. If consecutive frames yield the same temporal-spatial features vectors, we refer to them as keyspaces. The index lists of 16 joints are then built by their keyspaces. The data-driven decision tree is constructed automatically and the joint in higher level in this tree means greater influence on global motion match. During motion retrieval, more important joints have higher priority on similarity measure. At last motion retrieval is speed up significantly. Semantic features of some common motions can be extracted from motion temporal-spatial features.

References

1. Wolf, W.: Key frame selection by motion analysis[J] acoustics. Speech, and Signal Processing 2, 1228–1231 (1996)
2. Lim, I.S., Thalmann, D.: Key-Posture Extraction Out of Human Motion Data by Curve Simplification[C]. In: Proc. EMBC2001, 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2001, pp. 1167–1169 (2001)
3. Chui, Y., Chao, S.P., Wu, M.Y., Yang, S.N., Lin, H.C.: Content-based Retrieval for Human Motion Data[J]. Journal of Visual Communication and Image Representation 16(3), 446–466 (2004)
4. Chao, S.-P., Chiu, C.-Y.: Motion retrieval and synthesis based on posture features indexing. In: The 5th International Conference on Computational Intelligence and Multimedia Applications, Xi'an china (September 27-30, 2003)
5. Liu, F., Zhuang, Y., Wu, F., Pan, Y.: 3d Motion Retrieval with Motion Index Tree[J]. Computer Vision and Image Understanding 92(2-3), 265–284 (2003)
6. Lee, J., Chai, J., Reitsma, P.S.A., Hodgins, J.K., Pollard, N.S.: Interactive Control of Avatars Animated with Human Motion Data[A]. In: proceedings: SIGGRAPH 2002[C], San Antonio, Texas, pp. 491–500 (2002)
7. Assa, J., Caspi, Y., Cohen-Or, D.: Action synopsis: pose selection and illustration. ACM Trans. Graph. 24(3), 667–676 (2005)

8. Mueller, M., Roeder, T., Clausen, M.: Efficient Content-Based Retrieval of Motion Capture Data[A]. In: Proceedings of ACM SIGGRAPH 2005[C], vol. 24(3), pp. 677–685 (2005)
9. Hunt, E.B., Marin, J., Stone, P.T.: Experiments in induction. Academic Press, New York (1966)
10. Quilan, J.R.: Generating Production rules from decision tree. In: Proceeding of IJCAI-87 (1987)
11. Quilan, J.R.: 4.5: Programs for machine Learning. Morgan Kaufmann, San Mateo, CA (1993)
12. Russell, S., Norvig, P.: Arfificial Intelligence: A Modern Approach, 2nd edn. Prentice Hall, Englewood Cliffs, New Jersey (2002)

Motion Retrieval Based on an Efficient Index Method for Large-Scale Mocap Database

Jian Xiang[1] and Hongli Zhu[2]

[1] ZheJiang University of Science and Technology,
School of Information and Electronic Engineering, 310023, HangZhou, China
[2] ZheJiang University City College, College of Information & Electrical Engineering,
310015, HangZhou, China
freexiang@gmail.com

Abstract. In this paper, a novel approach is presented for motion retrieval based on double-reference inde(DRI) in Motion Capture(Mocap) database. Due to high dimensionality of motion's feature, the Isomap nonlinear dimensionality reduction is used. For handling new motion data, Isomap is generalized based on the estimation of underlying eigenfunctions. Then DRI is build based on selecting a small set of representative motion clips in the database. So we can get candidate set by abandoning most unrelated motion clips to reduce the number of costly similarity measure significantly. Experimental results show that our methods are effective for motion data retrieval in large database.

Keywords: Motion retrieval, reference index, Isomap, Mocap database.

1 Instruction and Related Work

Now more and more motion capture systems are used to acquire realistic human motion data. Most of motion data are stored in Mocap database with different length of motion clips, which is convenient for manipulating in animation authoring systems.

Since original features of motion clips lie in high-dimensional space and on a high-dimensional manifold which is highly contorted, we use a non-linear dimensionality reduction technique(Isomap)[1] to map them into low-dimensional manifold.

However, geo-distance of Isomap is only defined on training sets and raw Isomap cannot map new samples to the embedding space because it requires a whole set of points in the database to calculate geo-distance. And new queries outside the given database can't be handled by raw Isomap. Since new queries outside the given database can't be handled by raw Isomap, an extend Isomap in [2] is proposed to treat increasing data. A RBF(Radial Basis Function) neural network [3] is trained to approximate the optimal mapping function from input space to the embedding space.

In this paper, we use a generalization of Isomap [2], this generalization with a unified framework in which these algorithms are seen as learning eigenfunction of a kernel can obtain embeddings for new samples.

When motion clip data are embedded to low dimensional spaces, a number of index structures have been developed to reduce the cost of such searches. F. Keogh[4]

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 234–242, 2007.

shows a new indexing of human motion capture data technique with global scaling. One reference indexing method, Venkateswaran[5] gives MVD-index for selecting references as well as a new strategy for assigning references to database sequences. But it need to traverse whole database to get candidate set and is infeasible. Reference-based indexing method can select a small fraction of clips referred to as the set of reference clips. The distances between reference and database clips are pre-computed. Given a query, clips that are too close to or too far away from a reference clip are removed from the candidate set based upon those distances by LB(lower bound) and UB(upper bound) with the help of the triangle inequality. Although Computing LB and UB for each pair data is very simple, searching all data for large-scale databases is time-consuming and infeasible. So we extend reference index to double-reference index(DRI). DRI builds a bi-directional mapping between reference and database clips. Mapping from references to database can be used to make a new candidate motion set for search process based on reference index. Then our system can complete motion retrieval with DRI based on new reduced candidate motion set far less than the original database.

Finally we test our method on a large collection of motion capture clips with a large variety of actions and compare the performance with that of other methods. The results are very good and the DRI is more appropriate for large-scale database than conventional methods.

2 Motion Representation and Isomap Dimensionality Reduction

In this paper, a simplified human skeleton model is defined, which contains 16 joints that are constructed in the form of tree. Joint *root* is root of the tree and those paths from root to all endmost joints in human skeletal model from sub-tree of root

World coordinate of each joint can be represented as follow:

$$M = \{F(1), F(2),..., F(t),..., F(n)\}\text{; } F(t) = \{p(t), q_1(t),..., q_m(t)\} \quad (1)$$

where $F(t)$ is the *t-th* frame in motion clip M, $p(t)$ is the rotation of the root joint and $q_i(t)$ is the rotation of joint i at frame t. m is the number of joints used in human skeleton.

Hence we generate additional features of the skeletal motion by 16 joints of original motion data. In our implementation they include: (1) joint positions, (2) joint angles, (3) joint velocities.

First we use Isomap to do dimensionality reduction. Isomap constructs neighborhood graph and computer shortest path distances (geodesic distances) for each pair of points in graph. Then classical MDS is used with geodesic distances.

To obtain an embedding for a new data point, we need to generalize Isomap for new samples, that is learning the principal eigenfunction of a kernel[2] and the functions are from a function space whose scalar product is defined with respect to a density model.

Let D={ $x_1, x_2, ..., x_n$ } be a data set sampled from an unknown distribution with continuous density p and let $\hat{p}$ be the corresponding empirical distribution. Consider a Hilbert space H_p of functions with following inner product:

$$\langle f, g \rangle_p = \int f(x) g(x) p(x) dx$$

where $p(x)$ is a weighting function.

So the kernel K can be associated with a linear operator K_p in H_p:

$$(K_p f)(x) = \int K(x, y) f(y) p(y) dy$$

Now an "empirical" Hilbert space H_p is defined using the empirical distribution $\hat{p}$ instead of p. Let $\tilde{K}(a,b)$ be a kernel function that gives rise to a symmetric matrix $\tilde{M}$ with entries $\tilde{M}_{ij} = \tilde{K}(x_i, x_j)$ upon D. Let (v_k, λ_i) be an (eigenvector, eigenvalue) pair that solves $\tilde{K}_{\hat{p}} f_k = \lambda_k^{'} f_k$ with $\hat{p}$ the empirical distribution over D. Let $e_k(x) = y_k(x)\sqrt{\lambda_k}$ or $y_k(x)$ denote the embedding associated with a new point x. Then:

$$y_k(x_i) = y_{ik}, e_k(x) = \sqrt{\lambda_k} y_k(x), e_k(x_i) = e_{ik} \tag{2}$$

Here the definition of $D_G(a,b)$ is geo-distance of Isomap, which only uses the training points in the intermediate points on the path from a to b. We obtain a normalized kernel as follow:

$$\begin{aligned}\tilde{K}(a,b) = -\frac{1}{2}(D_G(a,b) - E_x\left[D_G{}^2(x,b)\right] \\ - E_{x'}\left[D_G{}^2(a,x^{'})\right] + E_{x,x'}\left[D_G{}^2(x,x')\right])\end{aligned}$$

Then the follow formula is applied by Eq(2) for the extension of Isomap to a new point x . It can yield the projection of x on the principal components of the corresponding low-dimensional data points.

$$e_k(x) = \frac{1}{2\sqrt{\lambda_k}} \sum_i v_{ik} (E_{x'}\left[\tilde{D}^2(x', x_i)\right] - \tilde{D}^2(x_i, x)) \tag{3}$$

If we have access to a huge amount of data to estimate the eigenfuntions of K_p in H_p, the optimal out-of-sample embedding would be obtained at a new sample. The consequence of the Eq(3) is that Isomap which only provided an embedding for the training examples can be extended to provide an embedding for new samples. So our system can handle new queries outside the given motion database by the generalization of Isomap.

3 Double-Reference Index

For motion retrieval, DTW distance is widely used in similarity measure of sequence database. In this paper, DTW distance is the measure for calculating distances.

Let $R = \{r_1, r_2, ..., r_m\}$ denote a reference motion clip set, S is a motion database, q is a motion query example with range ε, where $r_i \in S$ and m is the number of set. Distances set $\{s_i r_j \mid (\forall s_i \in S) \wedge (\forall r_j \in R)\}$ between reference and database clips are pre-calculated to build reference index. For each clip s_i, a lower bound(LB) and an upper bound(UB) for $s_i q$ are given by:

$$LB = \max(\vee_{r_j \in R} \mid qr_j - r_j s_i \mid) ; UB = \min(\vee_{r_j \in R} \mid qr_j + r_j s_i) ;$$

The range need to be searched within selected set is [LB, UB], so for each query pair (q, s_i):

$$\begin{cases} 1)\ \varepsilon < LB \Rightarrow s_i \in \text{abandoned set} \\ 2)\ \varepsilon > UB \Rightarrow s_i \in \text{selected set} \\ 3)\ LB \le \varepsilon \le UB \Rightarrow s_i \in \text{candidate set} \end{cases}$$

So the one-way reference index[5] can partition the database into three sets by query examples. The motion matching by DTW during motion retrieval should only be done in the candidate set, which is much less than original database. Computing time of similarity measure is dramatically saved.

Figure 1 illustrate reference index in a hypothetical two dimensional space. Here, point $r1$ and $r2$ are reference clips. $\varepsilon < LB = \mid r1s - r1q \mid < sq$, so the clip s can't be similar with query q and be abandoned.

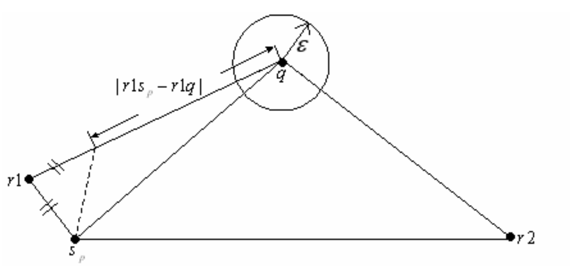


Fig. 1. Reference index example by $r1$ and $r2$

3.1 Choosing Reference Clips

Here we describe how to choose some appropriate reference clips inside the database, which can represent all parts of the database, to build the reference index.

The follow algorithm first selects a sample database S' by Nearest Neighbor rules, which can approximately represent the distribution of the database. Then mean and variance of $Dist$ are calculated. So the database S can be sorted in descending order of their variances. The clip with maximum variance is selected as the next reference to add to reference set R and the clip are close to or far away from the new reference are removed from S. At each iteration, a new reference that is neither close to nor far away from the existing reference clip is selected until the number of references reach to m.

Procedure choose-reference

Input: motion database S, N is the number of motion clips in S, m is the number of references, ω is a range to measure the closeness of two clips.

Output: references set $R = \{r_1, ..., r_m\}$

Initialize R;

For all $s_i \in S$

Use Nearest Neighbor rules to partition the database into some clusters.

use centroid of every clusters to make a sample set of clips, $S' \in S$

calculate distances $Dist_i = \{s_i s_j \mid \forall s_j \in S'\}$ and mean μ_i, variance σ_i of $Dist_i$

if the number of references in R < m

add the motion clip s_i of the biggest variance in the database to R.

Remove s_i from S, then remove all clips in ω-neighborhood and all clips in ω-farthermost from s_i too.

If the number of references in R<m goto 3)

return R

3.2 Mapping Between References and Database

In this section, we discuss how to build relationship between references and database, this relationship means efficient index by reference clips for motion database.

The follow algorithm above returns a mapping from each database clip to local reference sets. For each database clip s, the algorithm repeats mapping references until m reference clips are mapped. The procedure selects a reference clip by which s is abandoned for maximum number of queries.

```
Procedure: map_DBtoR

Input: motion database S, the number of S is N;
Reference set R, the number of R is M;

     Sample queries Q, the number of Q is q; m is the
number of references per clip.

Output: M_n is mapping from database clip to
references.

Initialize Be_i (benefit from each reference) and M_i

For all s in S

       If the number of references in M : |M| < m
            Count[]=0;

            Test all r∈R,  q∈Q , If reference r can
abandon s for the query q

                 Count[r]=Count[r]+1;

            max=argmax_x(Count(x)); Be[max]+=Count(max);

           Remove max label reference from R and add
to M;

           Remove queries from Q for which s is
abandoned with reference max.

       Re-insert all deleted entries from R and Q.

3) For all r∈R

   If Be[r]≤|Q|

       Remove r from R

4) Return the reference sets M_i,∀i∈S
```

Then we extend one-way reference index to double reference index(DRI). In the follow algorithm, mapping from references to database is built. So we can use this mapping to prune most unrelated motion clips from the database before motion retrieval. This procedure means much to process for large-scale database.

```
Procedure map_RtoDB
Input: motion database S, the number of S is N;
Reference set R, the number of R is M;
      n is the number of clips per reference; ω is a
range to measure the closeness of two clips.
Output: NN_n, NF_n are two mappings from references to
database clips.
1) for all r ∈ R
     Select sample clips set S' (as section *);
     Compute mean μ_i for distance between reference r_i
and S'
     For all s ∈ S
         If   r_i s_j > μ_i + ω
              Add s_j to NN_i;
         If   r_i s_j < μ_i − ω
              Add s_j to NF_i;
2) return NN_n, NF_n
```

Now DRI can be used to in motion retrieval system. For a new motion query example q and range ε, we first calculate the distances $Dist_i$ from q to the reference clips. Compare $Dist$ with mapping from references to database clips: NN_n, NF_n and first candidate set S_{new} is built for next step. For each $s_i \in S_{new}$, the lower and upper bounds LB and UB are computed for (q, s_i). If $UB \le \varepsilon$, s_i is added to Result set. If $LB \le \varepsilon \le UB$, the last candidate set S_{last} include s_i and if otherwise, s_i is abandoned. At last, we do the similarity measure based on DTW between q and S_{last} to get final result.

4 Experimental Results and Analysis

We implement out algorithm in matlab. It is more than 5000 motion clips with 30 common types in the database for test. Most of the typical human motions are performed by actors, such as walking, running, kicking, punching, jumping, washing floor, etc.

Fig. 2 shows the comparison of our method with existing reference-like indexing methods using same numbers of references. So motion retrieval based on DRI runs much faster than other methods.

To compare motion retrieval efficiency of the proposed method with the Liu's[6] method and Keogh's[4] index method. Table 1 shows that the performances of these methods in same database. It is obvious that the time of our method is so much less than time and higher precision and recall than other methods.

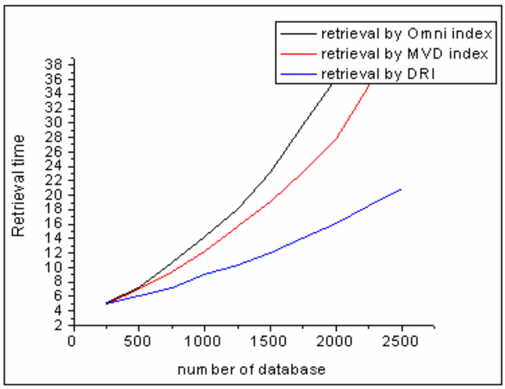


Fig. 2. Comparison of performance for Omni, MVD and DRI on varying databases

Table 1. Performance comparison

method	Precision	Recall	Retrieval time
DRI	92.6%	96.2%	8.1292s
IT	79.1%	83.5%	35.6185s
GSI	82.3%	87.2%	17.2169s

5 Conclusion

In this paper, a double-reference index method is proposed for motion retrieval system. Before retrieval, some motion features are extracted from motion data and generalization of Isomap is used to reduce dimension of these features and embed original motion space and new motion data into low-dimensional space. Then DRI selects a number of reference clips to represent the whole database and builds a bi-directional mapping between references and database for choosing candidate sets before similarity measure in retrieval. At last, the whole motion database is efficiently and accurately indexed. The motion retrieval system is also sped up significantly.

References

1. Tenenbaum, J., de. Silva, V., Langford, J.: A global geometric framework for nonlinear dimensionality reduction. Science 290(5500), 2319–2323 (2000)
2. Bengio, Y., Paiement, J.F., Vincent, P.: Outofsample extensions for LLE, ISOMAP, MDS, Eigenmaps, and Spectral Clustering. In: Proceedings of Neural Information Processing Systems (2003)
3. Hei, X.F., Ma, W.Y., Zhang, H.J.: Learning an image manifold for retrieval. In: Proceedings of 12th ACM International Conference on Multimedia, pp. 17–23 (2004)
4. Keoph, E., Palpanas, T., Zordan, V.B.: Indexing large Human-motion database. In: Proceedings of the 30th. VLDB Conference, pp. 780–791 (2004)
5. Venkateswaran, J., Lachwani, D., kahveci, T., Jermaine, C.: Reference-based indexing of Sequence Databases. In: Proceeding of 32th. VLDB Conference. pp. 139–148 (2006)
6. Liu, F., Zhuang, Y.T., Wu, F., Pan, Y.H.: 3d Motion Retrieval with Motion Index Tree. Computer Vision and Image Understanding 92(2-3), 265–284 (2003)

Modeling of Layered Fuzzy Facial Expression Generation

Yu-Li Xue, Xia Mao, Zheng Li, and Wei-He Diao

School of Electronic and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China
YuLi_Xue@ee.buaa.edu.cn, moukyou@buaa.edu.cn

Abstract. This paper proposed a layered fuzzy model of facial expression generation, in which the layers of physiological model at low level, emotional model at middle level and social rules at high level determine the fuzzy facial expression generation. In the layered fuzzy facial expression generation system, facial expressions of 26 emotions can be fuzzily generated, as well as social expressions and instinctive expressions. Experiment results indicated that the model of layered fuzzy facial expression generation works well in displaying lifelike facial expressions.

Keywords: Facial expression generation, fuzzy theory, layered fuzzy model.

1 Introduction

Facial expression takes an important role in human's daily life, and it has been widely researched in psychology, sociology, cognitive science, biology, neuroscience and so on. In the "Emotion View", kinds of facial expressions of emotions were studied by Darwin [1] and facial expressions of 6 basic emotions were pointed by Ekman and Friesen [2] to be universally recognized. On the contrary, the "Behavioral Ecology View" treats facial displays as expressions as social signals of intent [3]. Recently, facial expressions have been considered as emotional activators and regulators [3].

Along with the rapid development of computer science, human-computer interaction and related fields, the generation of facial expressions in computer has been actively researched. In the Improvisational Animation system [4], facial expressions such as angry, daydreaming, disgusted, distrustful, fiendish, haughty etc. were realized by relate lower level layer facial movements to higher level moods and intentions. In the fuzzy rule-based system [5], an animated agent's representations of 7 single expressions and blending expressions of 6 basic emotions of were mapped onto muscle contraction values. In the CharToon system [6], facial expressions were obtained by interpolation between the 7 known expressions (neutral, sadness, happiness, anger, fear, disgust, and surprise) positioned on the emotion disc. In the visual-speech animation system [7], a variety of emotional expressions can be obtained by blending 6 well-characterized emotional key-expressions (including neutral, happy, sad, surprised, afraid, and angry expressions). In the intelligent agent architecture [8], emotional intelligence was introduced into an animated character to express felt emotions and expressed emotions.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 243–252, 2007.

However there are some limits in the researches above. The first limit is that facial expressions are mostly constrained in the 6 basic emotions: happiness, sadness, surprise, anger, disgust and fear. Even blends of the 6 basic emotional expressions do not often appear in our daily life and can't be related to distinct emotional categories. As human can display facial expressions of many emotions, why computers do not? The second limit is that facial expressions are merely related to emotions. However, emotion is not the only source of facial expressions. For example, one winks to give a hint to someone or because he is too tired. Lifelike facial expression generation should be more complicated. The third limit is the monotone of facial expressions. Most works correlate one model of facial expression to one emotion, just with different intensities, but human don't. For example, human display kinds of facial expressions to express happiness, such as smile with mouth open or closed, symmetrically or asymmetrically, even with head wobbled.

To solve the above limits in facial expression generation, this paper proposed a model of layered fuzzy facial expression generation, in which the layers about social rules, emotions and physiological conditions were considered and fuzzy theory was used to generate mutative and rich facial expressions.

This paper is organized as follows. In the next section, a novel model of layered fuzzy facial expression generation was proposed. Then, the layered fuzzy facial expression generation (LFFEG) system was realized. Before the conclusion, the evaluation of the LFFEG system was given.

2 Model of Layered Fuzzy Facial Expression Generation

2.1 Fuzzy Facial Expression

Fuzzy is the ubiquitous character of human mind and objects. Human displays facial expressions fuzzily in real life. As human face is capable of as many as 250,000 expressions, it is impossible to relate an expression to an exact emotion or intention. Also, as human's emotions are rich and mutative, a single facial expression can not express an emotion exactly. For example, more than 60 kinds of facial expressions about anger have been found [9].

2.2 What Influence Facial Expression

There are many factors that influence mapping emotional state to facial expression, such as the intensity and type of emotion, how emotional state elicited, social display rules and whether the person was encouraged to express or suppress emotion [9]. Also, the speed, amount and duration of facial expression are different from individuals [3].

However, emotions are not the only source of facial expressions [10]. The sources of facial expressions include mental states (felt emotions, conviction and cogitation), verbal communication (illustrators, listener responses and regulators), non-verbal communication (unfelt emotions, emblems and social winks), physiological activities (manipulators, pain and tiredness).

2.3 Various Mechanisms for Generation

Human's emotion is generated not only by cognitive illation but also by non-cognitive factors. Sloman proposed a structure assuming that human's brain has three layered structure: reactive layer, deliberative layer and self-monitoring layer [11]. Reactive layer was used to generate fast and first emotion. Deliberative layer was related to second emotion generated by cognition. The self-monitoring layer was the layer where self concept worked weightily.

As human's facial expression is innate and social, influenced by emotional, physiological and social factors, facial expression generation based on various mechanisms is reasonable.

2.4 Layered Fuzzy Facial Expression Generation

Based on the fuzzy character of facial expression, kinds of factors that influence facial expression and the various-mechanisms-based generation, the model of layered fuzzy facial expression generation was proposed. As seen in fig.1, the layers of physiological model at low level, emotional model at middle level and social rules at high level determine the fuzzy facial expression generation.

Emotion is considered the primary factor for facial expression generation in human computer interaction. Once emotion is determined by the motional model, emotional facial expression can be fuzzily generated. The physiological model simulates the physiological states such as pain or tiredness, which can elicit instinctive facial expression. Social rules are related to the linguistic and nonlinguistic communications, and social facial expressions are generated, working as the filter to the emotional and instinctive facial expressions.

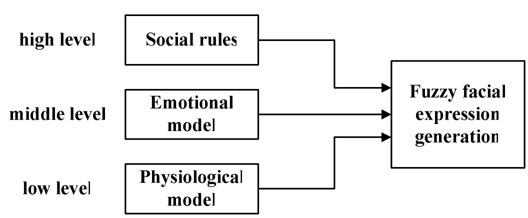


Fig. 1. Model of layered fuzzy facial expression generation

If emotion is always expressed in a changeless way, it is easy to be recognized by human or computer. However, most emotions are not mapped to a fixed form of facial expression in real life. Fuzzy facial expression generation maps one emotion to different facial expressions, making facial expressions more smart and expressive.

2.5 Layers of the LFFEG Model

The layer of physiological model includes the physiological variables which influence human's emotion and expression. Picard recognized that hunger or pain can influence the activation of emotional states [9]. For example, hunger can increase irritability, and pain can spur anger. Also, changes in brain blood temperature along with other physiological changes may lead to the feeling of excitation or depressed [3]. In the LFFEG model, the physiological variables influence the emotional expressions or lead to instinctive expressions such as grimace of pain.

The layer of emotional model includes multiple emotions based on the OCC model [12]. As the 22 emotions in OCC model are well accepted in the research of affective computing, it is reasonable to research the facial expressions to display the emotions. In the LFFEG model, multiple facial expressions can be fuzzily generated according to the emotions. For example, kinds of smile facial expressions can be elicited by the emotions such as joy, appreciation, gloating, satisfaction and happy-for, fuzzily controlled by the factors such as the intensity of the emotion, the type of the emotion and the expression personality.

The layer of social rules model includes what Ekman called social display rules and some predefined social expressions. When, where and how to express facial expressions is restricted by the social rules. A felt emotion may be masked by a fake emotion due to some display rules. In the LFFEG model, the social expressions generated by social rules can exceed the affect of the emotional model. For example, whatever a waiter felt, he should be polite to the client.

In the LFFEG model, social expression, emotional expression and instinctive expression are generated from the layers of social rules, emotional model and physiological model. The relation of them is shown in fig.2, with some cross expressions with each other. Social expressions are the facial expressions such as smile, wink and plea regardless of the emotion behind. Emotional expressions are the facial expressions elicited by kinds of emotions such as happiness, sadness and so on. Instinctive expressions are the facial expressions elicited by physiological activities, including the quick emotional expressions such as surprise, horror and other expressions such as frown, blink and gape.

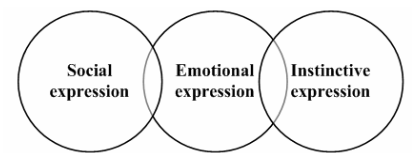


Fig. 2. The relation of facial expressions in different levels

3 The Layered Fuzzy Facial Expression Generation System

3.1 System Overview

The overview of the layered fuzzy facial expression generation system is shown in fig.3. The layers of social rules, emotional model and physiological model have respective priority, denoting the weight of the layer at a time. The module of fuzzy facial expression generation processes the output of the three layers, giving the final facial expression. Social expression is determined by the parse module from the layer of social rules. Emotional expression is determined by the fuzzy control function block from the layer of emotion model. Instinctive expression is determined by the

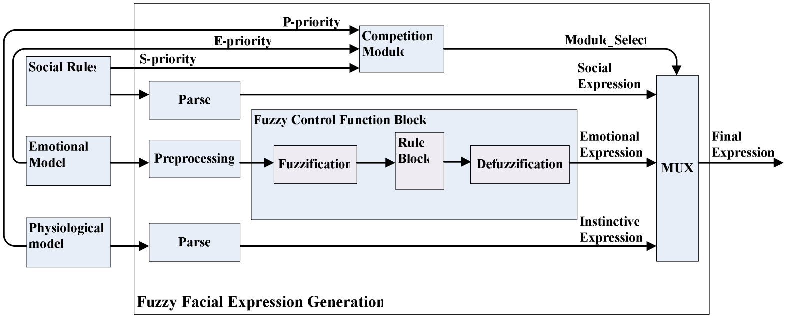


Fig. 3. Overview of the Layered Fuzzy Facial Expression Generation System

parse module from the layer of physiological model. According to the priorities of the three layers, final expression is determined from the social expression, emotional expression and instinctive expression.

The inputs of the facial expression generation are defined as followed:

1. Time: t;

2. Emotional Parameters: $E(t)=\{E_P(t), E_S(t), E_M(t), E_P(t)\}$ is the interface of the emotional model, where E_P: Priority, E_S: Specific emotions, E_M: Mood, E_P: Expression personality;

3. Social Parameters: $S(t)=\{S_P(t), S_E(t), S_R(t)\}$ is the interface of the social rules, where S_P: Priority, S_E: Social expressions, S_R:Social rules;

4. Physiological parameters: $P(t)=\{P_P(t), P_V(t)\}$ is the interface of the physiological model, where P_P: Priority, P_V: Physiological variables.

So the layered fuzzy facial expression generation function is:

$F(t)=F(E(t), S(t), P(t))$, where $F(t)$ is the fuzzy function.

3.2 The Lingual Realization of the LFFEG

The layered fuzzy facial expression generation is realized based on extensible markup language (XML), which provides an easy way to control the agent’s behavior. Previous efforts in the XML-based languages are Human Markup Language (HML), Multimodal Presentation Markup Language (MPML), Synchronized Multichannel Integration Language (SMIL), etc. In this paper, the Layered Fuzzy Facial Expression Generation Language (LFFEGL) is developed to realize the LFFEG model.

In the LFFEGL script as seen in fig.4, the tags of “social”, “emotion” and “physiological” relate to the layers of social rules, emotional model and physiological model respectively. The attribute “priority” gives the weight of the layers. Possible parameters of the LFFEGL are shown in fig.5, 5 social expressions are provided in the layer of social rules, 26 emotions are provided in the layer of emotional model and 12 physiological variables are provided in the layer of physiological model.

```
<seq>
  <Social priority="0.7" positive_threshold="0.4" negative_threshold="0.3">
        <Smile intensity="0.5"/>
  </Social>
  <Emotion priority="0.5" mood="positive" positive_weight="1">
        <joy intensity="0.5"/>
  </Emotion>
  <Physiological priority="0.9">
        <Pain intensity="1"/>
  </Physiological>
</seq>
```

Fig. 4. Sample LFFEGL Script

```
<!-- specification of possible social expressions -->

<!ENTITY % Social
"(normal   smile   plea   agreement   disagreement) ">

<!-- specification of possible emotions -->

<!ENTITY % Emotion
"(happyfor   pity   resentment   gloating   joy  distress   hope   fear   satisfaction
  fearsconfirmed   relief   disappointed   pride   selfreproach   appreciation
  reproach   gratitude   anger   gratification   remorse   liking   disliking   surprise
  disgust   sadness   sorryfor) ">

<!-- specification of possible physiological variables -->

<!ENTITY % Physiological
"(adrenaline   bloodpressure   bloodsugar   dopamine   endorphine   energy   heartrate
  respirationrate   temperature   vascularvolume   pain   tiredness) ">
```

Fig. 5. Specification of the Parameters in LFFEGL

3.3 Facial Expressions Influenced by Physiological Factors

In the layer of physiological model, the physiological variables are chosen based on literature [13], in which emotions can arise as a result of simulated physiological changes. The physiological variables are adrenaline, blood pressure, blood sugar, dopamine, endorphine, energy, heart rate, respiration rate, temperature, vascular volume, pain and tiredness, which may influence the emotional expressions or lead to instinctive expressions. For example, high levels of endorphines can increase the expressiveness of positive emotions or decrease the expressiveness of negative emotions, or trigger a state of happiness.

3.4 Fuzzy Emotion-Expression Mapping

In the layer of emotional model, the arousal-valence emotional space [14] is introduced to position the 26 emotions. Valence denotes if the emotion is positive or negative, and arousal denotes the intensity of the emotion.

The mapping of emotion to expression is many-to-many. Firstly, one emotion can be mapped to many facial expressions. For example, the fuzzy emotional facial expression of joy can be expressed as formula 1. Secondly, a predefined facial expression can express many emotions. For example, the emotions of joy and happy-for can be expressed as the facial expression of “SmileOpen” with different intensities.

$$E_{\text{happyfor}} = \begin{cases} a, SmileClosed \\ b, SmileOpen \\ \cdots \end{cases} \quad a, b, \cdots \in (0,1] \tag{1}$$

Where a, b are respectively the probabilities that joy is mapped to the facial expression of “SmileClosed” or “SmileOpen”.

Intensity control of predefined facial expression was realized in literature [15], however, the intensity of facial expression is not mapped directly to the intensity of the emotion. As seen in fig.6(a), the largest intensity of facial expression of “SmileClosed” may be not enough to express the largest intensity of happy-for. As seen in Fig. 6(b), moderate intensity of facial expression of “Disgust” may be enough to express the largest intensity of disliking.

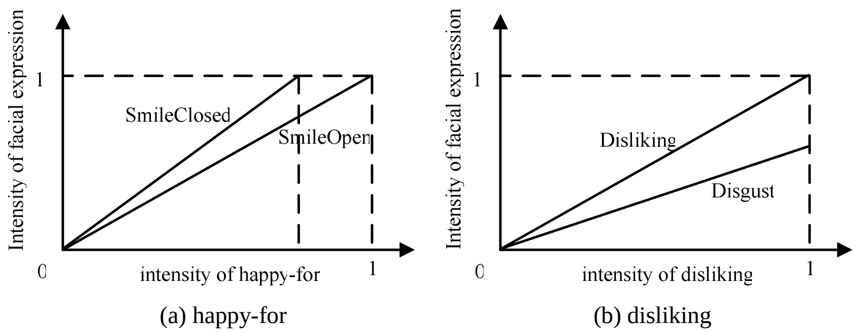


Fig. 5. Mapping of intensity of emotions to intensity of facial expressions

3.5 Expression Personality

Facial expressiveness, like other forms of sentic modulation, is influenced by a person’s innate physiology, which is related to temperament [9]. Affected by the personality, different person with the same intensity of emotion may express facial expressions with different intensities. For example, inhibited ones have lower overall facial expressiveness than uninhibited ones.

To realize the personality in facial expression, literature [16] introduced the module of Expression Personality in the signal processing structure of the robot, which is defined as a matrix of expression personality, weighting 7 emotions. Here, as there are more emotions in this system, the intensity of the emotional expression is gained by weighting the positive emotions or negative emotions with the attribute of “positive_weight” or “negative_weight”, even the weight of specific emotion can be given, thus the expression personality is realized.

3.6 Facial Expressions Influenced by Social Rules

Buck argued that social factors can facilitate or inhibit facial expression depending upon the nature of emotion being expressed and the expressor's personal relationship with the other [17].

In the LFFEG system, if the layer of social rules has higher priority than layer of emotional model, the attributes "positive_threshold" and "negative_threshold" will restrict the largest intensities of the positive facial expressions and negative facial expressions respectively, acting as inhibiting facial expression. The element such as "smile" or "agreement" specifies social expression to take the place of emotional expression. The specified social expression sometimes operates as facilitating facial expression.

4 Evaluation of the LFFEG System

To evaluate the LFFEG system, lively facial expression animations of the character should be presented to the subjects. Xface [15] was utilized to generate keyframes of facial expressions to display kinds of emotions. Some example keyframe facial expressions of disgust, surprise and disliking are shown in fig.7. In facial expression animation of some emotions such as pride and agreement, head movement was involved to strengthen the effect. Fuzzy facial expression animation can be generated through LFFEGL script.

(a) disgust1 (b) disgust2 (c) surprise1 (d) surprise2 (e) disliking1 (f) disliking2

Fig. 6. Some example keyframe facial expressions

In the first experiment, 10 subjects were asked to judge 21 keyframes of facial expressions, giving what emotions each facial expression likely to present and the corresponding score of expressiveness from 1 score (low) to 5 score (high). According to the results, each emotion was related to some facial expressions and fuzzy parameters were determined in the LFFEG system.

In the following experiment, 30 scripts were written in LFFEGL to generate fuzzy facial expression animations with middle intensity. 20 subjects were asked to run each script 5 times to evaluate the effect of the fuzzily generated facial expression animations, giving the score of satisfaction from 1 score (low) to 5 score (high). The results are showed in table 1, denoting that most facial expressions acted well. For those facial expressions with low scores, better keyframes and appropriate movement of head should be taken to strengthen the expressiveness.

Table 1. Score of satisfaction of fuzzily generated facial expressions

F.E.	S.	F.E.	S.	F.E.	S.	F.E.	S.	F.E.	S.
anger	3.9	disappointed	2.8	gratitude	2.9	sadness	3	agreement	4.8
fear	3.5	appreciation	3.3	happyfor	4	sorryfor	2.8	resentment	4.1
joy	4	fearsconfirmed	3.4	distress	3.3	relief	3	satisfaction	3.7
pity	3.2	disagreement	3.7	gloating	3.5	remorse	3.1	disgust	4.8
pride	4.2	gratification	3.7	disliking	3.2	liking	3.5	tiredness	4
hope	2.8	selfreproach	2.5	reproach	3.4	pain	4.6	surprise	4.3

Note: F.E.=facial expression, S.=score

Comparing with early works, more emotions are considered in our work, together with social and physiological factors. Facial expressions fuzzily generated by the layers in different levels are richer and reasonable, tally well with the character of human. The lingual realization of the LFFEG model provides an easy way to control the agent's facial expression. However, blend of emotional facial expressions is not the point in our work, as they seem amphibolous and often make people puzzled.

5 Conclusion

In this paper, a model of layered fuzzy facial expression generation is proposed. In the LFFEG model, the affects of the social rules, emotional model and physiological model are realized, and the generated facial expressions are lifelike, fuzzy and with personality. The lingual realization of the LFFEG makes it easy to realize and extensible. There are two primary novelties in our work: layered and fuzzy. Firstly, the factors that affect facial expression generation are considered in different layers, not only emotion but also social and physiological factors. Secondly, fuzzy facial expressions are realized to display multiple emotions, making the expression of the agent smart and rich. So the facial expression generation of the agent is more like human, making the agent intelligent in displaying facial expressions in human computer interaction.

Acknowledgments. This work is supported by the National Nature Science Foundation of China (No.60572044) and High Technology Research and Development Program of China (863 Program, No.2006AA01Z135). We also appreciate ITC-irst to provide the open source Xface toolkit for 3D facial animation.

References

1. Darwin.: The Expression of the Emotions in Man and Animals. John Murray, London (1872)
2. Ekman, P., Friesen, W.: Constants across Cultures in the Face and Emotion. J. Personality and Social Psychology 17(2), 124–129 (1971)
3. Lisetti, C.L., Schiano, D.J.: Automatic Facial Expression Interpretation: Where Human-Computer Interaction, Artificial Intelligence and Cognitive Science Intersect. Pragmatics and Cognition (Special Issue on Facial Information Processing: A Multidisciplinary Perspective), 8(1), 185–235 (2000)

4. Perlin, K.: Layered Compositing of Facial Expression. SIGGRAPH’97 Technical Sketch, New York University Media Research Lab (1997)
5. Bui, T.D., Heylen, D., et al.: Generation of Facial Expressions from Emotion using a Fuzzy Rule based System. In: 14th Australian Joint Conf. on Artificial Intelligence, pp. 83–94 (2001)
6. Hendrix, J., Ruttkay, Zs., et al.: A Facial Repertoire for Avatars. In: Proceedings of the Workshop Interacting Agents, Enschede, The Netherlands, pp. 27–46 (2000)
7. N.N.: An Example-based Approach to Text-driven Speech Animation with Emotional Expressions. EUROGRAPHICS 2003, 22(3) (2003)
8. Ochs, M., Niewiadomski, R., et al.: Intelligent Expressions of Emotions. In: 1st Int. Conf. on Affective Computing and Intelligent Interaction, pp. 707–714 (2005)
9. Picard, R.W.: Affective Computing. MIT Press, Cambridge, MA (1997)
10. Fasel, B., Luettin, J.: Automatic Facial Expression Analysis: A Survey. Pattern Recognition 36, 259–275 (2003)
11. Sloman, A., Croucher, M.: Why Robots will have Emotions. In: Proceedings of the Seventh Int. Conf. on AI, pp. 197–202 (1981)
12. Ortony, A., Clore, G.L., Collins, A.: The Cognitive Structure of Emotions. Cambridge University Press, Cambridge, UK (1988)
13. Canamero, D.: Modeling Motivations and Emotions as a Basis for Intelligent Behavior. In: Proceedings of the first international conference on Autonomous agents, pp. 148–155 (1997)
14. Lang, P.J.: The Emotion Probe. American Psychologist 50(5), 372–385 (1995)
15. Not, E., Balci, K., et al.: Synthetic Characters as Multichannel Interfaces. In: 7th Int. Conf. on Multimodal Interfaces, pp. 200–207 (2005)
16. Miwa, H., Takanishi, A., Takanobu, H.: Experimental Study on Robot Personality for Humanoid Head Robot. In: IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, vol. 2, pp. 1183–1188 (2001)
17. Buck, R.: Social Factors in Facial Display and Communication: A Reply to Chovil and Others. J. Nonverb. Behav. 15, 155–161 (1991)

An Inverse Dynamical Model for Slip Gait

Jiankun Yang[1], Dewen Jin[1,*], Linhong Ji[1], Jichuan Zhang[1], Rencheng Wang[1], Xin Fang[2], and Dawei Zhou[2]

[1] Division of Intelligent and Biomechanical System, State Key Laboratory of Tribology, Tsinghua University, Beijing (100084), P.R. China
[2] China Center of Orthopedic Technology (CHICOT), Beijing (101601), P.R. China

Abstract. A inverse dynamical model for slip gait developed by using Kane's method of dynamics is described in this paper. The modeling was based on two separate sub-systems: anti-slip and anti-fall. Both sub-systems were modeled as an open kinematic chain and formulated using the same equations, so that the whole slip and fall process could be simulated using the same equations by switching input data. In the simulation, only kinematic data need to be input to obtain the joint moments and horizontal ground reaction force in the whole slip and fall process. The kinematic input data were acquired from one health male subject who was asked to perform the normal and slip gait. The anthropometric data for each body segment in the skeletal model was calculated using the body height, weight and the national standards on inertia parameters regression equations. The kinematic and kinetic results from the simulation are discussed in the paper which are well consistent with the conclusions in previous studies .

Keywords: slip and fall, modeling, human gait; kinetics.

1 Introduction

Slip-related fall is an accident people often encounter during walking in daily life. Investigations have shown that it has become a major cause of serious injury or even death especially for elders. Most studies on slip-related fall events can be classified as either biomechanical or epidemiological. Biomechanical analyses are commonly used in the studies of the potential slip risk via ground reaction forces and the proactive and reactive reaction strategies.

Almost all previous studies in slip and fall are based on in vivo experiments, although the mathematical modeling has been widely used as a method to investigate the human gait [1]-[4]] There are always some limitations in the in vivo experiments of slip gait. The most noticeable one is that there are no "real" falls occurring, because the protect equipments are employed and the anticipants must be ensured being safe. Therefore, a mathematical model for slip gait is urgently required. But up to now there has not been any mathematical model for slip gait reported yet. The complexity of the slip and fall procedure increases the difficulties in modeling the slip gait. Thus, the aim of this paper is to present a preliminary mathematical model of slip gait which was based on Kane's method for dynamical systems.

* Corresponding author.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 253–258, 2007.

2 Methods

Skeletal modeling. The human body model consists of 7 segments representing the stance foot, stance shank, stance thigh, swing shank, swing thigh, pelvis and HAT (head and torso). The number of segments was determined on the consideration of compromise between unnecessary complexity and adequate movement simulating. These segments are separated by anatomic points, which are the bottom of foot (BOF), the lateral center of malleoli (LCM), the lateral condyle of the femurs (LCF), the greater trochanter of the femurs (GTF), the xiphoid process (XIP) and the vertex (VER). The both arms and the swing foot is neglected in this model. As a preliminary attempt, all segments are assumed to be rigid and their motions are modeled in the sagittal plane only. And all segments are assumed to be connected by hinge joints. Furthermore, it has been demonstrated in the previous papers of our lab [5] that the slip and fall procedure can be divided into two steps: anti-slip and anti-fall. So that, the whole slip and fall procedure was modeled by two separate sub-systems. Moreover, in order to reduce the complexity of the model, the posterior foot was released completely during the anti-slip sub-system; while the anterior foot was released completely during the anti-fall sub-system. And the relative motion between the support foot and the ground is assumed as translation only in the sagittal plane. In this way, the ground was treated as a segment with zero velocity and infinite mass, which makes it possible to calculate the ground reaction forces (GRF) in the same way as the joint forces.

Experimental acquisition of gait data. To obtain the necessary kinematical data for the dynamic model, a normal healthy subject (male, 24 yr, 69.5 kg, 1.74m) underwent gait analysis during both level walking and slip gait. Some anatomical dimensions of this subject were measured to define the segments model. The subject did not have a prior history of lower extremity infirmity of musculoskeletal diseases that may affect the ability to perform the experiment at the time of testing.

Depending on the total body height and weight, the mass, the position of the center of mass in the local frame and the moment of inertia are determined with the regression equations of the national standards on the inertia parameters of adult human body [6]. And in order to adapt the model in this study, some corrections have been performed. The anthropometric data for each segment are listed in **Table 1**.

The subject wore shorts and walked on bare feet. He was required to walk at a self-selected comfortable pace along a 5 m plastic walkway, and to perform walking trials on dry and oily conditions respectively. In oily condition, the motor oil (40#) was evenly applied across 2 m long in the middle of the walkway. Thus, the from-dry-to-slippery condition under which falls usually occurred could be simulated in laboratory. The subject had a priori knowledge of the surface condition, whereas he was asked to walk as naturally as possible.

In order to record the movement of body segments, reflective markers were placed on the subject's sides: over the 5th metatarsal bone heads, the lateral center of malleoli, the lateral condyle of the femurs, the greater trochanter of the femurs, the anterior superior iliac spine, and the acromion of shoulders. Qualisys Motion Capture System (Qualisys Medical AB, Sweden) was employed to record the three-dimensional movements of the whole body at 200 Hz.

Table 1. The anthropometric data for each segment of the skeletal model. Mass center positions are from the proximal endpoint and the moments of inertia are in the sagittal plane.

Segment	Proximal endpoint	Distal End point	Segment length (*mm*)	Mass (*kg*)	Mass center position (*mm*)	Moment of inertia ($kg \cdot cm^2$)
Foot	BOF	LCM	61.5	0.9200	38.9600	ignored
Shank	LCM	LCF	371.7	3.0575	192.245	256.849
Thigh	LCF	GTF	514.1	9.7750	264.835	1572.415
Pelvis	GTF	XIP	461.0	18.2990	277.590	3733.031
HAT	XIO	VER	485.5	18.8875	201.105	4727.862

Governing Equations. The definitions of coordinate system and generalized coordinates in both anti-slip and anti-fall subsystems is shown in Fag.1. By applying Kane method, the governing equations are as

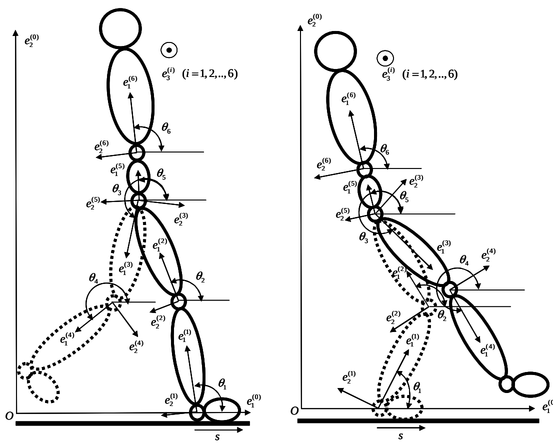


Fig. 1. The definitions of coordinate system and generalized coordinates in both anti-slip and anti-fall subsystems: (a) anti-slip; (

$$F^{(j)} + F^{*(j)} = 0 \quad (j = 1, 2, .., 7) \tag{1}$$

$$F^{*(j)} = \sum_{i=1}^{n} (\vec{R}_i^* \cdot \vec{v}_i^{(j)} + \vec{L}_i^* \cdot \vec{\omega}_i^{(j)}) \qquad (j = 1, 2, .., 7; \;\; n = 7) \tag{2}$$

$$F^{(j)} = \sum_{i=1}^{n} (\vec{R}_i \cdot \vec{v}_i^{(j)} + \vec{L}_i \cdot \vec{\omega}_i^{(j)}) \qquad (j = 1, 2, .., 7; \;\; n = 7) \tag{3}$$

Where $F^{*(j)}$ is the generalized inertia force , $F^{(j)}$ is the generalized applied force, $\vec{R}_i^*$ is the inertia force applied on the ith body segment, $\vec{L}_i^*$ is the inertia moment applied on the ith body segment, $\vec{R}_i$ is the active force applied on the ith body segment, $\vec{L}_i$ is the active torque applied on the ith body segment, ω and v are the partial angular velocities and partial velocities.

3 Results

The kinematic data in both subsystems during a slip gait were calculated and imported into the inverse dynamic model, respectively. The joint moments were obtained and normalized by the whole body weight of the subject. As a representative, the ankle joint moments and hip joint moments are shown in Fig.2 and Fig.3 respectively. Because the slip happened to the anterior right side, the anti-slip moments of right ankle are obviously increased until the posterior left leg supported the body. The posterior ankle moments are significantly large during the anti-fall period. The moments of hip joint on both sides indicate that the right hip joint moments were increased significantly during both anti-slip and anti-fall stage, while the posterior hip joint pays an important role in anti-fall stage.

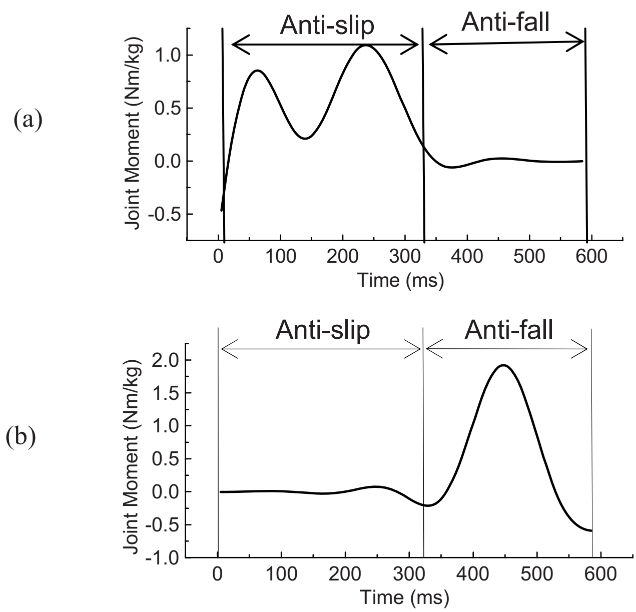
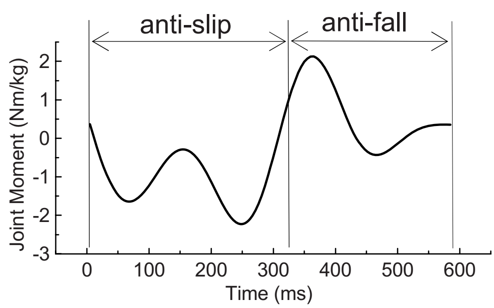


Fig. 2. The predicted ankle joint moments (a) right ankle joint; (b) left ankle joint

(a)



(b)

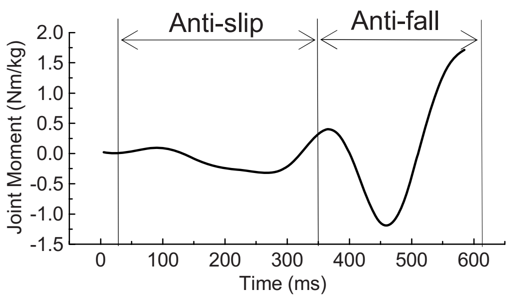


Fig. 3. The predicted hip joint moments (a) right hip joint; (b) left hip joint

4 Discussion and Conclusion

The objective of this paper was to define a mathematical model for slip gait. Since the procedure of a slip-related fall event is always complex, there is not any reported mathematical model of whole body for slip gait yet. Based on the theory of that the human reaction in slip-related fall events can be decomposed in anti-slip and anti-fall sub-procedures, the mathematical model was developed in this paper using two single support models. The governing equations of these two single support models were formulated in the same format. The joint moments and horizontal GRF in both procedures could be calculated only by switching the kinematic inputs.

The governing equations were obtained by applying Kane's dynamic method to avoid using the GRF as model inputs. In conventional applications of inverse dynamics in human gait analysis, the GRF are usually measured using force plates, which are planted in the walkway, and are inputs to the calculations [7]. But in the slip gait experiment, it is impossible to make the subject's feet to step on the force plates affirmatively in the sudden reaction. Even though the feet step on the force plates luckily, they usually slide out of the area of the force plates as well. Thus, it is very difficult to record the GRF during slip gait. Under such a circumstance, the application of inverse dynamics in this study employs the measured motions of all the major body segments as the only input data. And based on these kinematic data only, the joint moments and GRFs are predicted using by using this simulation.

The previous studies have reported that the ankle joint is very sensitive to the slip gait. The hip joint plays the most important role to maintain balance in healthy persons' unexpected slips [8]. The results from the presented procedure are consistent with these conclusions. Moreover, during the anti-fall procedure, the significantly increased joint moments of the posterior leg were consistent with the surface electromyography analysis on muscle reaction in slip-related falls [5].

It could be concluded that the inverse dynamic model established in this paper is an effective way for dynamic analysis of the human body in slip gait.

Acknowledgement

This work is supported by the Nature Science Fund (NO. 30170242) and National Hi-Tech Research and Development (863) Program (NO. 2001AA320601), of China.

References

1. Komistek, R.D., Stiehl, J.B., Dennis, D.A., Paxson, R.D., Soutas-Little, R.W.: Mathematical model of the lower extremity joint reaction forces using Kane's method of dynamics. Journal of Biomechanics 31, 185–189 (1998)
2. Koopman, B., Grootenboer, H.J., de Jongh, H.J.: An inverse dynamics model for the analysis, reconstruction and prediction of bipedal walking. Journal of Biomechanics 28(11), 1369–1376 (1995)
3. Marigold, D.S., Patla, A.E.: Strategies for dynamic stability during locomotion on a slippery surface: effects of prior experience and knowledge. Journal of Neurophysiology 88, 339–353 (2002)
4. Ren, L., Jones, R.K., Howard, D.: Dynamic analysis of load carriage biomechanics during level walking. Journal of Biomechanics 38, 853–863 (2005)
5. Yang, J.K., Jin, D.W., Wang, R.C., Zhang, J.C., Ji, L.H., Fang, X., Zhou, D.W.: The investigation on sEMG of lower extremity when a slip occurs in level walking. In: The 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Shanghai, China, September 1-4, 2005 (2005)
6. Standardization Administration of China (SAC), National standards: Inertial parameters of adult human body (GB/T 17245-2004) (2004)
7. Winter, D.A.: The biomechanics and motor control of human gait. University of Waterloo Press, Waterloo, Canada (1987b)
8. Ferber, R., Osternig, L.R., Woollacott, M.H.: Reactive balance adjustments to unexpected perturbations during human walking. Gait and Posture 16, 238–248 (2002)

Redundant Muscular Force Analysis of Human Lower Limbs During Rising from a Squat

Yiyong Yang[1], Rencheng Wang[2], Ming Zhang[3], Dewen Jin[2], and Fangfang Wu[2]

[1] School of Engineering and Technology, China University of Geosciences (Beijing), Beijing 100083, China
[2] Division of Intelligent and Biomechanical System, State Key Laboratory of Tribology, Tsinghua University, Beijing, 100084, China
[3] Jockey Club Rehabilitation Engineering Centre, The Hong Kong Polytechnic University, Hong Kong, China
yangyy@cugb.edu.cn

Abstract. Muscular coordination analysis of lower limbs during rising from a squat is one of the important categories in rehabilitation engineering and gymnastic science. This paper describes an efficient biomechanical model of the human lower limb with the aim of simulating the real human rising from a squat with lifting. To understand how intermuscular control coordinates limb muscle excitations , the optimal control technique is used to solve the muscle forces sharing problem. The validity of the model is assessed comparing the calculated muscle excitations with the registered surface electromyogramm (EMG) of the muscles. The results show that synergistic muscles are build up by the neural control signals using a minimum fatigue criterion during human squat lifting, with distinct phases that include the acceleration during the initial movement and the posture at the final specified position. Synergistic muscular groups can be used to simplify motor control, and are essential to reduce the number of controlled parameters and amount of information needing to be analyzed in the performance of any motor act.

Keywords: Redundant Muscular force, Neural control analysis, Human squat lifting.

1 Introduction

The determination of muscle forces during human movement is not only essential for a complex analysis of internal loads acting on bones and joints, it also contributes to a deeper understanding of the underlying neural control. Due to the difficulty of measuring muscle forces directly within the living system by means of invasive techniques and due to the circumstance of the mechanically redundant arrangement of the actuators, static optimization [1] or dynamic optimization[2]have been used to estimate muscle forces during the movements.

According to the optimization theory, a cost function is hypothesized, and muscle forces are calculated according to the mechanical, energetic, and physiological

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 259–267, 2007.

properties of the neuro-musculoskeletal system. Static optimization has been the most common method used to estimate muscle forces during human movement. The main disadvantage of static optimization is that the results are heavily influenced by the accuracy of the available experimental data, particularly the measured limb motions.

In this paper, we aim to develop a dynamical model which can be used to specify how individual muscles coordinate movement as humans rising from a squat with lifting. Dynamical models of the musculo-skeletal system combined with neural control models for the activation of muscles have been developed to simulate the integrative properties of the neuro-musculo-skeletal system in the movement of rising from a squat with lifting. Through a quantitative comparison of the model and experiment, we show that unconstrained, point-to-point movements such as rising from a squat with lifting are well described by a quantity which a minimum fatigue criterion is able to reproduce the major features of the movement**.**

2 Methods

2.1 Human Experiments

Five young and healthy male subjects voluntarily participated in this study. Their average age, height and body mass were 28±2 years, 1.74±0.04 m and 62±5 kg, respectively. Ethical approval from the Tsinghua University was granted for this study and all procedures were in accordance with ethical guidelines.

Each subject was asked to perform six rising from a squat with lifting tasks. The task is, from a prespecified squatting position(90 degree in knee joint), and a barbell is on his shoulder (the mass of the barbell is equaled to 50 percent of the body weight(shown in Fig 1.), Kinematic, force-plate, and EMG data were recorded simultaneously.

Fig. 1. The tasks of the rising from a squat with a barbell

The EMG signals were recorded using a Noraxon™ (ARIZONA U.S.A) Ttelemyo 2400R EMG Data Acquisition system sampling at 1500Hz. Pairs of EMG-preamplifier surface electrodes (Meditrace™, center-to-center distance 4 cm, circumference 12 mm) were attached to the left lower extremity of each subject to record activity in seven muscle groups: soleus (SOL), gastrocnemius (GAS),tibialis anterior (TA), vasti (VAS), rectus femoris (RF), hamstrings (HAMS), and gluteus maximus (GMAX). To record the body-segmental displacements, retroreflective markers (2.54 cm and 5.08 cm in diameter) were positioned over four bony prominences: lateral malleolus, lateral epicondyle, greater trochanter, and glenohumeral joint. Together, these landmaraks defined the three body segments of interest: tibia femur, and HAT segment (head, arms, and trunk). With a 3-D, video-based, kinematic, data-acquisition system (Motion Analysis Inc., Santa Rosa), absolute displacements were recorded in three dimensions at 60 Hz. These data were then reflected onto the sagittal plane to obtain the two-dimensional coordinates of each landmark relative to an inertial frame fixed on the force platform.

2.2 Musculoskeletal Model

For the task of rising from a squatting position, we modeled the human lower limb as a three-segment, articulated, planar linkage, with adjacent links joined together by frictionless revolutes. The skeleton was actuated by ten musculotendinous units with each muscle modeled as a three-element, lumped-parameter entity in series with tendon. Fig. 2 is the schematic representation of the model used to simulate the rising from a squat[3].

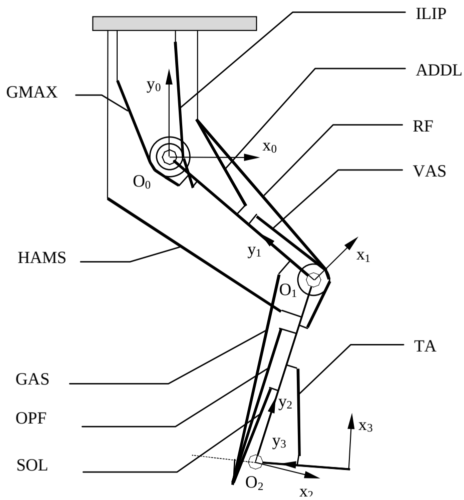


Fig. 2. Schematic representation of the model

Note that the foot was assumed to be rigidly attached to the ground. Symbols appearing in the diagram are tibialis anterior (TA), soleus (SOL), gastrocnemius (GAS), other plantarflexors (OPF), vasti (VAS), hamstrings (HAMS), rectus femoris (RF), and gluteus maximus (GMAX) and ILIP, ADDL.

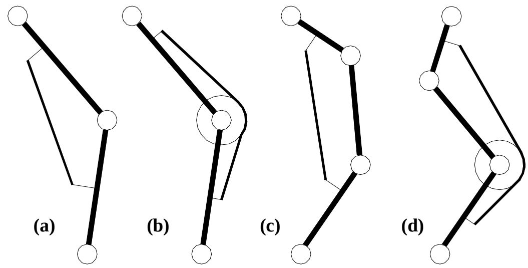


Fig. 3. The styles of the ten musculotendinous units (a) GMAX, TA, OPF, SOL, ILIP, ADDL; (b)VAS; (c)HAMS, GAS; (d) RF

The ten musculotendinous units can be classified as 4 styles*: (a), (b), (c) and (d).* The styles of the ten musculotendinous units are shown in Fig 3.

The dynamical equations of motion for the musculotendonskeletal system are:

(a) skeletal dynamics

$$A(q)\ddot{q} = B(q)\dot{q}^2 + G(q) + M(q)P^T + T(q,\dot{q}) \quad (1)$$

where $A(q)$ is 3×3 system mass matrix ; $B(q)\dot{q}^2$ is a 3×1 vector describing both Coriolis and centrifugal effects ; $G(q)$ is a 3×1 vector containing only gravitational terms ; $M(q)$ is a 3×10 the moment arm of musculotendon force ; P^T is an 10×1 vector of musculotendon actuator forces ; $T(q,\dot{q})$ is the passive movements at each joint ; q $\dot{q}$ $\ddot{q}$ are 3×1 vectors of lower limb segmental displacements, velocities, and accelerations.

(b) musculotetendon dynamics,

$$\dot{P}^T = f(q, \dot{q}, P_i^T, a_i(t)); i = 1,10 \quad (2)$$

where $a_i(t)$ is the level of activation in the *i*th muscle.

(c)neural control and muscle excitation-contraction dynamics

$$\dot{a}_i(t) = (1/\tau_r)(1-a_i(t))u_i(t) + (1/\tau_f)(1-a_i(t))(1-u_i(t)) \quad (3)$$

where $u(t)$is the control vector (input excitation) given to the muscles in the model(Zajac et al., 1989); τ_r and τ_f are the rise and decay times for muscle activation, respectively.

(d) parameterized optimal control theory and solution

The optimal control problem can be stated as follows: minimize the performance criterion subject to the dynamical equations of the model (Eqs.(1)-(3)), and the constraints imposed at the beginning and end of the simulated SQUAT-RISING movement.

In the research we have focused on evaluating 6 performance criteria (OB1-OB6) for the movement:

$$\text{OB1:} \quad J_1 = \min \int_0^{tf} dt \tag{4}$$

$$\text{OB2:} \quad J_2 = \min \int_0^{tf} \sum_{i=1}^{10} (P_i^T / P_i^{\max}) dt \tag{5}$$

$$\text{OB3:} \quad J_3 = \min \int_0^{tf} \sum_{i=1}^{10} (P_i^T / P_i^{\max})^2 dt \tag{6}$$

$$\text{OB4:} \quad J_4 = \min \int_0^{tf} \sum_{i=1}^{10} (\dot{P}^T / P^{\max})^2 dt \tag{7}$$

$$\text{OB5:} \quad J_5 = \min \int_0^{tf} \sum_{i=1}^{10} (P_i^T / P_i^{\max})^2 dt \tag{8}$$

$$\text{OB6:} \quad J_6 = \min \int_0^{tf} \sum_{i=1}^{10} (P_i^T / P_i^{\max})^3 dt \tag{9}$$

Finally, we hypothesized that the motor patterns that typify rising from a squat movement are the result of a minimization in OB6. The objective function proposed by us is based on the relationship between endurance time (T, time during which a muscle can produce a given force) and muscle stress (P/PCSA) (Y.Y. Yang, 2004) . The dynamical equations of motion for the musculotendonskeletal system and the performance criteria can be defined as a fatigue-change model.

The optimization problem was solved using the routine *constr()* of the optimization toolbox of MATLAB (The MathWorks Inc., Natick, MA, USA) on a PC computer.

3 Results

3.1 Muscle Excitation Patterns

The predicted activations from the model's sub-optimal minimum performance criterion solution were compared to the experimental activations (i.e., processed EMG). Rectified and normalized EMG activity recorded from subjects and the optimal neural excitation signals obtained by minimizing “a fatigue-change model”. The results show that the model's activation for the uniarticular knee and hip (GMAX) extensors, and biarticular leg muscles matched well with the EMG. In general, the model predicted muscle excitation patterns similar to the processed EMGs.

3.2 Musculotendon Forces Synergy

Fig. 5 shows time-varying musculotendon forces predicted by the model of the ten actuators during the rising from a squat movement.

Predicted individual musculotendon forces for the ten actuators were analyzed to determine their contribution to the rising from a squatmovement. The primary extensors (TA, RF, GMAX, HAMS) force magnitudes all increased and maintained at all time (about 85% movement cycle). They had a somewhat similar force pattern. After their peaks, the force magnitudes don't decrease to allow the limb to decelerate, the flexors group (ILIP, ADDL, VAS) then reactivated to clamp the final position.

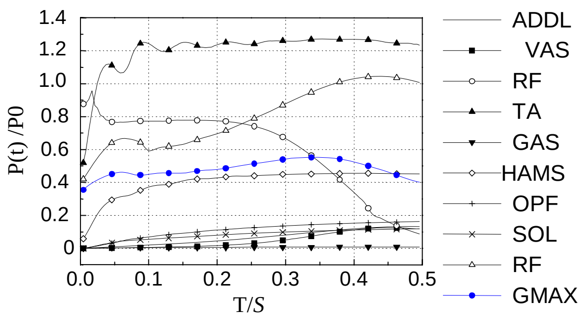


Fig. 4. Time-varying musculotendon forces predicted by the model

The biarticular leg muscles (GAS, RF, HAMS) redistributed the segmental force, producing much force and were excited first . Because of the redistribution of force, by the biarticular muscles, propulsion could be prolonged (i.e., so extension of the joints could be coordinated with full leg extension occurring near lift off).An increase in muscle strength was shown to be most effective in producing higher rising, with faster contracting muscle fibers being the next most effective. We also found that the uniarticular knee (VAS) and hip (GMAX) extensors produce high work output during leg extension[11].

In generally, all the muscles are excited and necessarily, there are not excrescent or waste actuators, the different muscles were controlled and combined into units of synergistic muscular group necessary to rise, that is group1-the primary extensors (TA, RF, GMAX, HAMS: group1), the primary coordinated muscles(GAS, SOL,VAS: group2) and the primary stabilization muscles(OPF, ILIP, ADDL: group3).

3.3 Neural Control

Fig.5, Fig.6 and Fig.6 show the corresponding controls histories which were obtained from the ten muscles.

As shown in Fig.5, it is noted that group1 is almost fully active during the rising from a squat . However, in Fig.6, group 2 is mainly active at the middle and toward

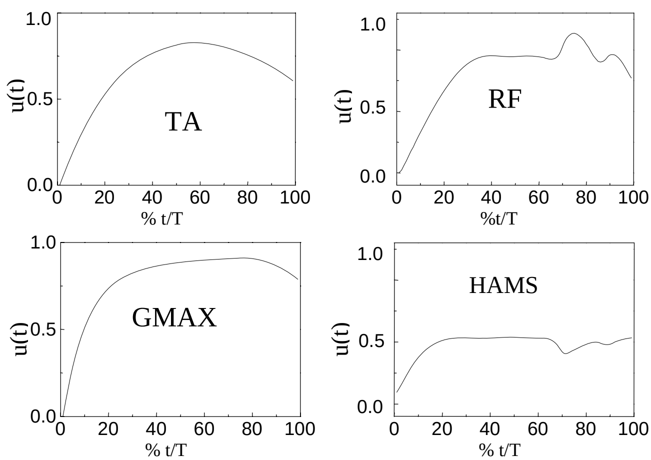


Fig. 5. The control histories of the primary extensors(group1)

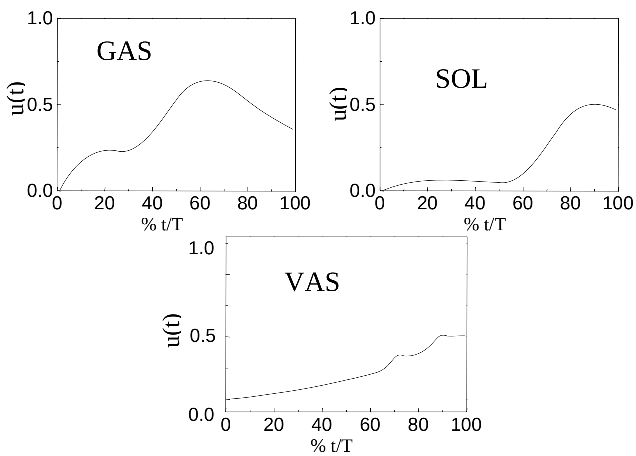


Fig. 6. The control histories of the primary coordinated muscles(group2)

the end of the interval. In contrast, the control histories predicted by group 3 shows that muscles are essentially inactive for the first 15-20 percent of the interval (Fig.7), it is active for the last 50 percent. It is noted that there is a substantial difference between the predictions of group1, group2 and group 3.

The different muscles were controlled and combined into three units of synergistic muscular groups necessary to rise, we named that the group1 is the primary extensors (TA, RF, GMAX, HAMS), it's task is mainly to rise, the group2 is the primary coordinated muscles(GAS, SOL,VAS) and the group3 is the primary stabilization muscles(OPF, ILIP, ADDL).

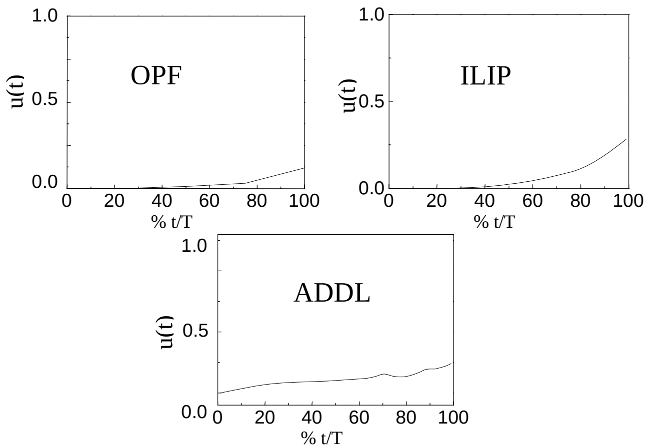


Fig. 7. The control histories of the primary stabilization muscles(group3)

4 Discussion

In this paper, we proposed a performance criterion for the movement of the rising from a squat based upon an important and previously overlooked dynamical property of muscles: fatigue-change model. Our motivation for including this quantity in a mathematical description of the goal of a motor task is founded upon the belief that the rising from a squat movement are effected by rapid fluctuations in muscle force rather than gradual increases by in force over time.

Several investigators have used optimal control theory to determine the sensitivity of muscle force predictions to changes in the performance criterion [4]. They computed the movement patterns which minimized various measures of physical cost including movement time, distance, peak velocity, energy, peak acceleration, and the rate of change of acceleration (jerk). By comparing the response of a model with the movements, and argued with skilled movements are well described by a minimum-jerk criterion or energy criterion. In a similar study, Pandy et al [5]showed, using a single-mass, single-force model, that optimal control solutions are generally sensitive to changes in the performance index. Because smooth movement patterns could be generated by minimizing the time derivative of acceleration integrated over time, they argued that non-ballistic movements are well described by minimization of jerk[6]7][8].

Our fatigue-change model criterion is consistent with the biomechanical requirements of rising from a squat position. We believe that it obeys the principle that most muscles, and especially all uniarticular extensor muscle, ought not to remain silent until just prior to rising. With the ballbell exerting a reaction force on the body, muscle, particularly the uniarticular knee and hip extensors VAS and GMAX, have to contract and exert force to counteract the effects of gravity. Our fatigue-change model criterion is able to reproduce the major features of rising from a squat position.

We argued that that synergistic muscles are build up by the neural control signals using a minimum fatigue criterion during human squat lifting, with distinct phases that include the acceleration during the initial movement and the posture at the final specified position.

Acknowledgements. The authors wish to acknowledge the support of this work by the NSF of China (50475017) and NSF of CUGB(200407).

References

1. Crowninshield, R.D, Brand, R.A.: A physiologically based criterion of muscle force prediction in locomotion. J. Biomech. 14, 793–801 (1981)
2. Pandy, M.G.: Computer Modeling and Simulation of Human Move-ment. Ann. Rev. Biomed. Eng. 3, 245–273 (2001)
3. Yang, Y.Y.: Motion synergy and Control of Human Neuromusculoskeletal System. Ph.D. thesis, Tsinghua University, Beijing, China, pp. 66–88 (2004)
4. Anderson, F.C, Pandy, M.G.: Static and Dynamic Optimization Solutions for Gait Are Practically Equivalent. J. Biomech. 34, 153–161 (2001)
5. Pandy, M.G, Garner, B.A, Anderson, F.C.: Optimal Control of Non-ballistic Muscular Movements: A constraint-Based Performance Criterion for Rising From a Chair. J. Biomech. Eng. 117, 15–26 (1995)
6. Pedersen, D.R, Brand, R.A, Davy, D.T.: Pelvic Muscle and Acetabular Contact Forces During Gait. J. Biomech. 30, 959–965 (1997)
7. Rab, G.T.: Muscle. Human Walking. In: Rose, J., Gamble, J.G. (eds.): Williams and Wilkins, Baltimore, pp. 103–121 (1994)
8. Taga, G.: A Model of the Neuromusculoskeletal System for Human Locomotion: Emergence of Basic Gait. Biol. Cybern. 73, 95–111 (1995)

Optimal Control and Synergic Pattern Analysis of Upper Limb Reaching-Grasping Movements

Yiyong Yang[1], Rencheng Wang[2], Ming Zhang[3], Dewen Jin[2], and Fangfang Wu[2]

[1] School of Engineering and Technology, China University of Geosciences (Beijing), Beijing 100083, China
[2] Division of Intelligent and Biomechanical System, State Key Laboratory of Tribology, Tsinghua University, Beijing, 100084, China
[3] Jockey Club Rehabilitation Engineering Centre, The Hong Kong Polytechnic University, Hong Kong, China
yangyy@cugb.edu.cn

Abstract. A three-dimension, neuromusculoskeletal model of the human upper limb, consisting of 30 muscle–tendon systems, was combined with dynamic optimization theory to simulate reaching-grasping movements. The model was verified using experimental kinematics, muscle forces, and electromyographic(EMG) data from volunteer subjects performing reaching-grasping movements. Despite joint redundancy, the topological invariance was observed in the trajectories of different task performance, and the linear relationships between joints covariation were exhibited. Quantitative comparisons of the model predictions and muscle activations obtained from experiment show that the minimum torque-change criterion is a valid measure of reaching-grasping performance.

Keywords: Synergic Pattern, Optimal Control, Upper limb, Reaching to grasp movements.

1 Introduction

Prehension is a popular task for studying human sensorimotor control[1]. It involves reaching and grasping an object for a purpose such as manipulating it, transporting it or just feeling it. On the kinetic level, prehension entails applying forces during interaction with an object. On the kinematic level, prehension involves the orienting and posturing of the hand and fingers, with the appropriate transportation of the limb to the correct location in space. How the reaching and grasping to be well-coordinated? Experimental studies have been conducted to explore the processes and neural mechanisms that allow coordination of the prehensile act. With the exception of a few studies using the model prediction[2], most have examined prehension only using motion analysis system[3], and focusing on single-joint movements[4] within the last 15 years.

Quantitative analyses of muscle force synergies have not been reached, and the neural control mechanisms or muscle synergies (structural units) among the whole-arm involving in reaching-grasping remain unclear.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 268–275, 2007.

In this paper, we will present a three-dimension, neuromusculoskeletal model of the human upper limb, and combined with dynamic optimization theory to simulate normal self-paced reaching-grasping movements. The model was partially verified using experimental kinematic, muscle forces, and electromyographic(EMG) data from volunteer subjects performing reaching-grasping movements. This verification lends credibility to the time-varying muscle force predictions and the synergic recruitment of muscles that contribute to both reaching and grasping manipulation.

Our specific aims are: (a) to present a physiologically supported, time-dependent, performance criterion for normal reaching-grasping; (b) to evaluate the model simulation results through comparisons with experimental data. (c) to decide whether a patterns of movement coordination and units of synergistic muscular group are involved into healthy people prehension.

2 Methods

2.1 Human Experiments

Five healthy male adults participated in this study. The average age, height, and mass of the subjects was 26 ± 3 years, 177 ± 3 cm, and 70.1 ± 7.8kg, respectively. Ethical approval from the Tsinghua University was granted for this study and all procedures were in accordance with ethical guidelines.

Subjects were instructed ‘to grasp the glass and drink from the glass’ using the right hand. No instructions about movement speed were given. As a warm-up, each subject had five practice trials prior to data collection in each condition,. The glass (height 11.8 cm and radius 5.8 cm) filled with water, and put on two heights (h_1 =710mm and $h_{2=}$=1100mm). Each subject was asked to perform four reaching-grasping tasks with different indices of difficulty (2 heights × 2 type of strap).

In the Task1 and Task4, Subjects were seated in a rigid-supported chair with their right shoulder strapped to the chair back with a wide belt. The trunk motion occurred in the condition had not the contribution to hand movement. Task1 was named as a “natural task”, and Task 4 was named as a “difficult task”.

The starting position and reach distances were standardized. At the beginning of each trial, the right arm of the subject was relaxed and rested on the chair, the hand being at the hip height. Seat height was adjusted to 100% of lower leg length as measured from the lateral knee joint line to the floor with the subject standing barefoot. Reach distances were standardized to 100% of arm length. Arm length was measured separately for each subject, and the length was defined as the distance from the shoulder marker to the most distal point of the hand, i.e. tip of the flexed middle finger around the glass for the grasp task.

Retroreflective marks (2.54 cm and 5.08 cm in diameter) were placed on the right arm and hand to measure the three-dimensional position of each segment. The joint coordinate system, as defined by Schimidt *et al.*[5], was used in this study. Subjects were video-taped from the right side as they reached for the object with their right hand. The video camera had a fixed sampling rate of 25 Hz and a shutter speed of 1/1000 s.

Surface electromyography (EMG) electrodes (Meditrace™, 11 mm, self-adhesive, silver-silver chloride) were placed over five muscles on the right arm: biceps brachii (BIC), triceps brachii (TRI), anterior deltoid (DEL), flexor digitorum superficials(FDS), and extensor pollicis longus(EPL). The electrodes were placed in a bipolar arrangement over the muscles in the recommended locations.

The right shoulder of subject was in the plane of the glass which was parallel to the subject's sagittal plane. We are going to refer to this plane as ``movement plane" in the following statement since endpoint movements as well as wrist and elbow rotations occurred mostly in this plane.

2.2 Mathematical Model of the Upper Limb

We developed a three-dimensional neuromusculoskeletal model of the human upper limb, the model allows elbow flexion/extension, forearm pronation/supination, wrist flexion/extension and radial/ulnar deviation.

The skeleton of the upper arm was articulated by joints and moved by the actions of muscles actuators. Each actuator was modeled as a 3-element, Hill-type muscle in series with an elastic tendon . Altogether, 30 muscle-tendon systems, having effect on the motion of the forearm, palm and fingers, are included. Parameters determining the nominal properties of each actuator were based on data reported by Delp. Following Chao et al.[6], additional tendons are introduced to model the extensor systems of the fingers. Values of the musculotendon parameters assumed for each actuator in the model are given in Ph.D. theses 0f Yang Yi-yong [7].

The corresponding dynamical equations for the upper limb model can be written as follows:

(a) skeletal dynamics

$$A(q)\ddot{q} = B(q)\dot{q}^2 + G(q) + M(q)P^T + T(q,\dot{q}) \tag{1}$$

where $A(q)$ is 3×7 system mass matrix; $B(q)\dot{q}^2$ is a 3×1 vector describing both Coriolis and centrifugal effects; $G(q)$ is a 3×1 vector containing only gravitational terms; $M(q)$ is a 3×30 the moment arm of musculotendon force; P^T is an 30×1 vector of musculotendon actuator forces; $T(q,\dot{q})$ is the passive movements at each joint; q $\dot{q}$ $\ddot{q}$ are 7×1 vectors of upper limb segmental displacements, velocities, and accelerations.

(b) musculotetendon dynamics,

$$\dot{P}^T = f(q,\dot{q},P^T,a_i(t)); i = 1,30 \tag{2}$$

where $a_i(t)$ is the level of activation in the ith muscle.

(c) muscle excitation-contraction dynamics

$$\dot{a}_i(t) = (1/\tau_r)(1-a_i(t))u_i(t) + (1/\tau_f)(1-a_i(t))(1-u_i(t)) \quad (3)$$

where $u_i(t)$is the input excitation given to the ith muscle in the model; τ_r and τ_f are the rise and decay times for muscle activation, respectively.

(d) parameterized optimal control theory and solution

The optimal control problem can be stated as follows: minimize the performance criterion subject to the dynamical equations of the model (Eqs.(1)-(3)), and the constraints imposed at the beginning and end of the simulated reaching-grasping movement cycle.

Examined some performance criteria (e. g. energy, torque, movement time etc.), we hypothesized that the motor patterns that typify normal grasping movement are the result of a minimization in torque-change model.

The objective function of the criterion is expressed as

$$C^T = \frac{1}{2}\int_0^{tf} \sum_{i=1}^{n} \left(\frac{dz_i}{dt}\right)^2 dt \quad (4)$$

Here, t_f indicates the motion duration. N is the number of joints, and Z_i shows the commanded torque generated in joint i.

Presuming that the objective function must be related to the dynamics, we proposed the following measure of a performance index: sum of square of the rate of change of torque integrated over the entire movement. Here, let us call this model "minimum torque-change model".

The model described above was implemented with commercially available software, Automatic Dynamic Analysis of Mechanical Systems (ADAMS^R) analyzed and solves the motions and forces of three-dimensional mechanical systems. The user enters the body parts, joints, and imposed forces or motions and the software produces and solves over time the equations of motion of the mechanical system using Lagrangian equations and predictor-corrector methods of numerical integration.

3 Results

3.1 Kinematics

Table1 shows the human movement time (HMT) affected by both height and strap condition. As *Index of difficulty* increased, there was a significant increase in HMT [Flinear(1, 5)=41.6, P=0.001]. The difficult grasp task took longer time to complete than the normal task at each trial [F(1, 5)=61.2, P=0.001]. The movement time measured for Task 1 of the subject and predicted by model subject to minimum torque-change criterion were 0.990 and 0.960 seconds, respectively, a 5% difference in the final times, which is very close to the optimal speed.

Fig. 1 shows the spatial trajectories of the hand in performing different tasks by the same subject in normalized time (for ease of comparison, each experimental and predicted results was normalized its respective final time to compare overall patterns). In the reaching phase (about 70% movement cycle), the hand was consistently moved in a relatively straight path to the object (even though the pattern of segmental motion and hence the amount of weight shift varied between conditions). In the grasping phase (about 30% movement cycle), it can be seen that the trajectories of the hand

Table 1. Means and standarded diviation of HMT for each task

Task	HMT (ms)	S.D.	Index of difficulty
Task 1	960	56	0.1
Task 2	1020	61	0.3
Task 3	1180	78	0.6
Task 4	1238	89	1.0
Model	990	0	0.0

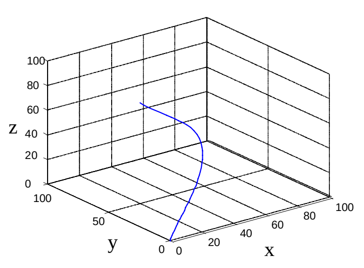


Fig. 1. Three dimension trajectories of prehension

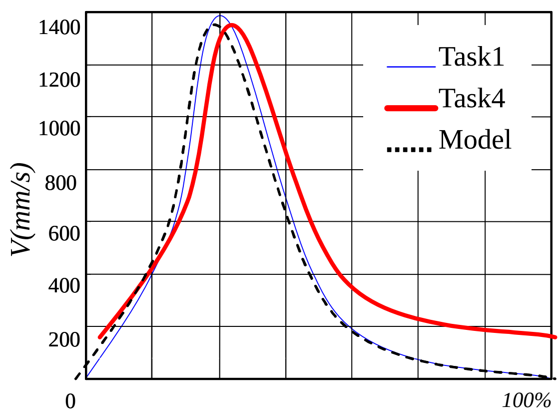


Fig. 2. The Velocity of the hand from a representative trial and model prediction

were not in straight lines (an arc).The figures also indicate that the movements in spite of different tasks have topological invariance in spatial trajectories.

Experimental (Task1, Task 4) and predicted velocity trajectories for a reaching and grasping movement are shown in figure 2. All curves were normalized with respect to its respective final time. The velocity profiles also indicate a good fit between the model's and the experiment's. The peak velocities for the Task1 are larger for the Task4. The peak occurs in task4 slightly later in the model (8%). Both velocity profiles of the subject and model showed a pattern indicating a bell-shaped profile. Also apparent in Fig. 2 is the final velocity of the hand (i.e. the velocity at contact

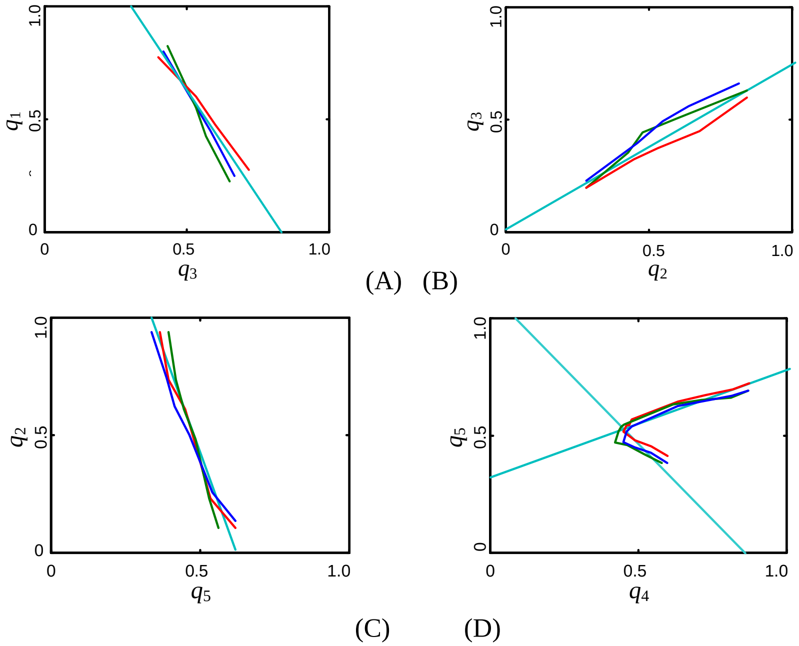


Fig. 3. The Angular covariations of prehension in normal condition (task1). q1=δ:Forearm rotation , q2=δ: Upper arm rotation ,q3=α: Upper arm azimuth, q4=β: forearm elevation, q5=ψ: Elbow flexion.

with the object). In Task4, final velocity was not zero, indicating that subjects used the object to stop the action. In contrast, in the Task1, which involved more accuracy, the hand slowed markedly as it approached the glass in order to prevent spillage.

This comparison indicates that the minimum torque-change kinemaitcs solution of the model for a hand reaching-grasping movement corresponds well with how the subject performed.

3.2 Angular Covariations

Fig.3 shows angular covariations for the normal condition (Task1). Here, q1=δ:Forearm rotation , q2=δ: Upper arm rotation , q3=α: Upper arm azimuth , q4=β: forearm elevation , q5=ψ: Elbow flexion. They (in Fig.6)are averaged values for 3 subjects.(shows single trials performed by subject 3 during the normal task). Regression lines were computed on averaged curves. The coefficients of regression were highly significant for all the single and averaged curves (p< 0.0001). This observation indicated the existence of a linear relationship between joint covarivation.

3.3 Muscle Excitation Patterns

The predicted muscle activations from the model's sub-optimal minimum performance criterion solution were compared to the experimental activations.

In general, the model predicted muscle excitation patterns similar to the processed EMGs, especially for the acceleration portion of the movement.

Results of the predicted musculotendon forces and activations indicate that it would be correct to assume that so called "synergistic muscles" at the upper limb produce similar time-varying forces and activations. Therefore, to lump these muscles together to reduce the system would be possible.

4 Discussion

Because the number of the freedom of the upper limb (DOF=7 when finger joints are neglected) exceeds those necessary to completely specify the location and orientation of an object in space (DOF=6). The mathematical relationship associating the coordinates of the object to grasp and the final posture of the arm (inverse mapping) is a priori indeterminate. Optimal control theory is potentially the most powerful method for determining redundancy muscle forces during movement.

The power of an optimal control approach derives from the scope of the modeling not only does optimal control theory allowing muscle dynamics to be included in the formulation of the problem, but also delivers a purely predictive result independent of experiment. Once an optimal control solution is found, there is a wealth of information available to compare with experimental data. Optimal control theory requires that a model of the system dynamics be formulated and that a performance criterion is specified as well.

In this paper, the model computational solutions were compared to the grasping experimental kinemaitcs and EMG data and matched well with each other. This comparison indicates that minimum torque-change kinematics solution of the model for upper limb reaching-grasping corresponds well with how the subject performed. This criterion has been proposed in a computational model of brain's activity during a reaching movement.

Several researchers measured the hand trajectories of skilled movements and found common invariant features. The integrated components for developing computational musculoskeletal models have been established for the human limbs through long time efforts[8]. The models of musculoskeletal systems have used optimal control strategies to solve the indeterminate problem using a variety of performance indices[7]. These attempts have yielded a better understanding the coordination of muscle forces.

During our trials, the joint covariations patterns were very stable, the individual variations o f the upper limb angles were systematically coupled with respect to the time (joint synergies). During the movement, joint angle variations are not controlled independently, but in a synergic way (temporal coupling). The movement trajectory observed, either in the task or joint space, results directly from this temporal coupling.

These results suggest that natural movements are mostly carried out in joint space by postural transitions.

5 Conclusion

The results show that the topological invariance was observed in the trajectories of different task performance, and the linear relationships between joints covariation were exhibited. Moreover, the different muscles were controlled and combined into units of synergistic muscular group necessary to reach and grasp the goal.

Acknowledgements. The authors wish to acknowledge the support of this work by the Key Program of NSFC (50435040) and NSF of CUGB(200407).

References

1. Bernstein, N.A.: The co-ordination and regulation of movements, pp. 15–28. Pergamon Press, Oxford, London (1967)
2. Lemay, M.A, Patrick, E.C.: A dynamic model for simulating movements of the elbow, forearm, wrist. J. Biomechanics 29, 1319–1330 (1996)
3. Gottlieb, G.L., Song, Q., Hong, D.A., et al.: Coordinating movement at two joints: A principle of linear covariance. Journal of Neurophysiology 75, 1760–1730 (1996)
4. Weinberg, A.M., Pietsch, I.T., Helm, M.B.: A new kinematic model of pro-and supination of the human forearm. Journal of Biomechanics 33, 487–491 (2000)
5. Schmidt, R., Disselhorst-Klug, C., Silny, J., et al.: A marker-based measurement procedure for unconstrained wrist and elbow motions. Journal of Biomechanicals 32, 615–621 (1999)
6. Chao, E.Y.S., An, K.N., Cooney III., et al.: Biomechanics of the Hand. A Basic Research Study, pp. 60–66. World Scientific, Singapore (1989)
7. Yi-yong, Y.: Motion synergy and Control of Human Neuromusculoskeletal System. Ph.D. thesis, pp. 57–78, Tsinghua University Press, BeiJing (2004)
8. Crowninshield, R.D., Brand, R.A.: A physiologically based criterion of muscle force prediction in locomotion. Journal of Biomechanics 14, 793–801 (1981)

Low Cost 3D Shape Acquisition System Using Strip Shifting Pattern

Li Yao[1], Lizhuang Ma[2], and Di Wu[1]

[1] Digital Media Lab, Shanghai Jiao Tong University, China
{yaolily77,malz,guanliao}@sjtu.edu.cn

Abstract. We present a simple and low cost 3D shape acquisition method that can measure the objects with interreflection and high light. The capture system employs a projector to project line shifting shadow on the object, and a digital camera to record videos (image sequence) of the object and the distorted shadow. A novel spacetime edge finding method is introduced to get the shadow edge accurately and overcome the interreflection and high light. The 3D data of the object are got by using the spacetime information. A post-processing 3D filter is proposed to filter the bad data. The 3D filter makes full use of the neighborhoods' geometric constrains and the view point constrain.

Keywords: 3D scanner, shape recovery, surface reconstruction.

1 Introduction

3D shape acquisition has been an important technology in the field of computer vision in recent years. Fig. 1 shows the taxonomy of 3D Range Acquisition according to [7]. Traditional techniques such as stereo rigs and laser scanners got competition from more flexible and scalable alternatives. Nevertheless, several challenges remain. One critical problem is that these instruments are usually large scale and expensive. Secondly, 3D capturing techniques tend to fail on certain types of surfaces. Most methods assume the surfaces are Lambertian surface. For the surface with specular reflections, all too often still, one has to resort to adapting the scene to the system, e.g. by powdering the objects to be scanned. Otherwise, the specular reflection will affect the measurement accuracy, and the 3D information of the highlight area will be lost.

Inspired by laser slit scanner, and the structured light method, we combine their advantages(laser scanner's principle is simple but it can give high accuracy, and structured light need low cost device and can operate easily), we propose a simple method of 3D information acquisition based on shiftable project pattern that is a black strip with white background. The strip can scan the whole object and then the object 3d shape is estimated by the strip distortion on the object. Our method avoids the highlight area by casting shadow on the surface. Only a projector and a digital camera are used to acquire 3D information without any manual intervention.

2 Principle

The general principle consists of projecting a shadow onto the scene with a projector and using the images of the deformed shadow to estimate the three dimensional shape

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 276–285, 2007.

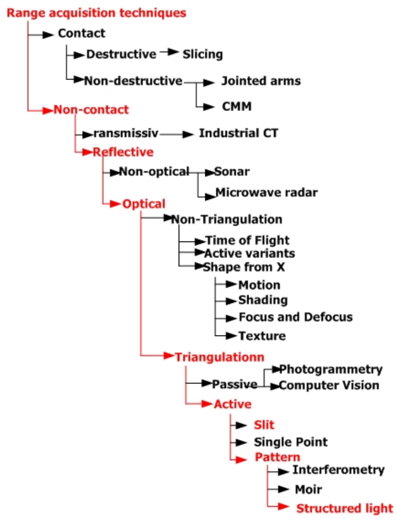


Fig. 1. A taxonomy of 3D Range Acquisition

of the scene. The objective is to extract scene depth at every pixel in the image. Figure 2 gives a geometrical description of the method that we propose to achieve that goal.

The camera acquires a sequence of images (or video) $I(x, y, t)$ as the strip moves in the screen so that the project shadow scans the entire scene. This constitutes the input data to the 3D reconstruction system. The variables x and y are the pixel coordinates (also referred to as spatial coordinates), and t the time (or frame number). The three dimensional shape of the scene is reconstructed using the spatial and temporal properties of the shadow boundary throughout the input sequence.

We denote the projector optical center with a point O_p, and denote by S_d the desk plane. O_p and S_d are known from calibration. The goal is to estimate the 3D location of the point P in space corresponding to every pixel P' in the image. Call t the time at which a given pixel P' sees the shadow boundary (later referred to as the shadow time). Denote by $S_s(t)$ the corresponding shadow plane at that time t. Let us assume that two portions of the shadow projected on the desk plane are visible on two given rows of the image (top row L_{top} and bottom row L_{bot}). After extracting the shadow boundary along those rows L_{top} and L_{bot}, we find two points on the shadow plane P_{top} and P_{bot} by intersecting S_d with the optical rays (O_c,P'_{top}) and (O_c,P'_{bot}) respectively. The shadow plane S_s is then inferred from the three points in space O_p, P_{top} and P_{bot}. Finally the point P corresponding to P' is retrieved by intersecting S_s with the optical ray (O_c, P).

So we need to understand first how to calibrate the system (extracting desk plane and the projector optical center locations S_d and O_p), and second how to localize spatially and temporally the shadow boundary on the images $I(x, y, t)$. These two stages will be described in the following sections .The final stage consists of calculating depth from geometrical triangulation.

3 Camera and Projector Calibration

The problem of camera calibration is to compute the camera intrinsic parameters (focal length, optical center, radial distortion factor), as well as extrinsic parameters (location and orientation) based on a number of points whose object coordinates in the (X_w, Y_w, Z_w) coordinate system are known and whose image coordinates (X, Y) are measured. The camera calibration process is important for a measurement system, because any error in measuring and calculating the camera location and orientation will cause a propagation error in the global coordinates, which will prevent a high overall accuracy of the final measurement.

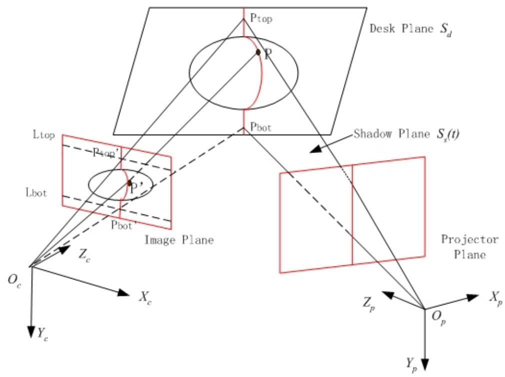


Fig. 2. Geometrical principle of the method

A standard method is to acquire an image a known 3D object (a checker board pattern, a box with known geometry) and look for the set of parameters that best match the computed projection of structure with observed projection on the image. The reference object is also called calibration rig. Since the camera parameters are inferred from image measurements, this approach is called visual calibration. This technique was originally presented by Tsai [1] [2] [3].

Fig. 7 shows the calibration images using a planar rig (checker board pattern). The principle of the calibration system is shown in fig. 2.

The transformation from (X_w, Y_w, Z_w) to (X_f, Y_f) is depicted in four steps:

Step 1: Rigid body transformation from the object world coordinate system (x_w, y_w, z_w) to the camera 3D coordinate system $(X_c, Y_c,, Z_c)$.

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = R \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + T \tag{1}$$

where $R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$ is rotation matrix and $T = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}$ is translation matrix.

Step 2: Transformation from 3D camera coordinate (X_c, Y_c, Z_c) to ideal image coordinate (X_u, Y_u) using perspective projection with pinhole camera geometry:

$$X_u = f\frac{x}{z} \qquad Y_u = f\frac{y}{z} \tag{2}$$

where f is the effective focal length.

Step 3: Radial lens distortion is

$$X_d + D_x = X_u \qquad Y_d + D_y = Y_u \tag{3}$$

Where (X_d, Y_d) is distorted or true image coordinate on the plane ,

$$D_x = X_d(k_1r^2 + k_2r^4 + \ldots), \quad D_y = Y_d(k_1r^2 + k_2r^4 + \ldots), \quad r = \sqrt{X_d^2 + Y_d^2}$$

where k_i s are distortion coefficients.

From Eqn. (1), (2) and (3), transformation from 3D World Coordinate to Camera Coordinate can be decrypted as:

$$X_d(1 + k_1r^2 + k_2r^4) = f\frac{r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}.$$
$$Y_d(1 + k_1r^2 + k_2r^4) = f\frac{r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

Step 4: Real image coordinate (X_d, Y_d) to computer image coordinate (X_f, Y_f) transformation:

$$X_f = X_d / d_x + C_x \qquad Y_f = Y_d / d_y + C_y \tag{4}$$

where (C_x, C_y) is row and column number of the center of computer frame memory, d_x is the center to center distance between adjacent sensor elements in X (scan line) direction, d_y is the center to center distance between adjacent sensor elements in Y (scan line) direction.

We use two-stage method for the calibration of our camera based on the method proposed by Tsai. In the first step, the parameters of the linear system can be solved by the perspective transform matrix. And in the second step, let the solved parameters be the initial guess, adding the coefficients of radial distortion (initially zeros), enter an iterative optimization procedure.

The calibration procedure for a projector is similar to that of a camera. We can think of a projector as a reverse pin-hole model of a camera. We assume a projector has a virtual “CCD” like a camera, and the projected pattern on the computer screen exists on this virtual “CCD”. We could regard the projected pattern as an imaging result of the projected object on the “CCD”. So, all the parameters used in camera calibration also apply to a projector.

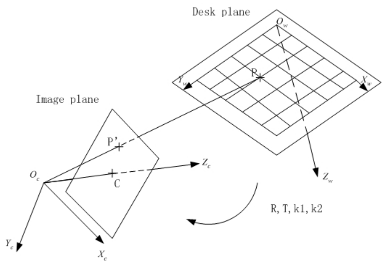


Fig. 3. Camera calibration system

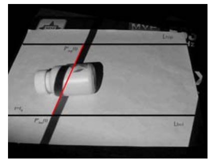


Fig. 4. Spatial and temporal shadow location

For calibrating a projector, we draw a checker board on the computer screen. Then project the checker image onto the desk plane on which we put the object. From another side, we use a camera (calibrated before) to take a photo of the projected object. From this photo, we can get 2D pixel coordinates of the grid corners. We can back-project these 2D pixels to 3D Word Coordinate System via the calibration parameters of the camera. On the other hand, we can also get 2D pixel coordinates of the corresponding grid corners on the checker board image. Now, we are able to feed the group of 2D and 3D coordinates into the Tsai's calibration program, and get the final parameters of the projector. All the calibration parameters are similar to those of a camera and will be utilized for the consequent 3D construction of an object.

4 Shadow Edge Localization

The traditional edge finding method only takes the space changes into consideration. The texture and the interreflection can cause edges. We only want to find the line shadow's edge. So we use a spatio-temporal approach to detect the shadow boundaries in the sequence of images $I(x, y, t)$.

At the image plane, we draw two horizontal lines L_{top} and L_{bot}. The Object is between the two lines (see Figure 4). The two major tasks to accomplish within this stage are:

- ♦ Find the shadow edge $P'_{top}(t)$ and $P'_{bot}(t)$ along the top and bottom line in the image plane at every time instant t, leading to the set of all shadow planes $S_s(t)$.
- ♦ Estimate the time $t_s(P')$ (the shadow time) at which the edge of the shadow passes throughany given pixel P' in the image.

Through those two tasks, two very different types of processing are required: spatial processing (to retrieve $S_s(t)$ at all times t) and temporal processing (to retrieve $t_s(P')$ at every pixel P' in the image). Both stages correspond to finding the shadow edge, but the search domains are different: first one operates on the spatial coordinates (image coordinates) and the other one on the temporal coordinate.

As the shadow is scanned across the scene, each pixel $P(x, y)$ sees its color going from the color under white light down to the color within the shadow and then back up to its initial value as the shadow goes away. Define the maximum color contrast $I_{contrast}$ for a given pixel *(x, y)* in the image as:

$$I_{contrast}(x,y) = \max_t\{\max_t\{I(x,y,t,v)\} - \min_t\{I(x,y,t,v)\}\} \quad (5)$$

where $v \in \{R, G, B\}$ is the color channel, *I(x, y, t, v)* means the color value in v channel at time t.

The pixels in shadows due to occlusions do not provide any relevant depth information. For this reason we restrict the processing to pixels that have enough swing between maximum and minimum color value. We only keep regions of the image where $I_{constract}(x, y, v)$ is larger than the pre-defined threshold.

We define the shadow edge time $t_s(x, y)$ at pixel *P(x, y)* as the time instant at which the pixel color value crosses for the first time the mean value between its maximum value and minimum value in channel v. This gives a different threshold value for every pixel. We call it the threshold $I_{thresh}(x , y)$:

$$I_{thresh}(x,y,v) = \frac{1}{2}(I_{\min}(x,y,v) + I_{\max}(x,y,v)),$$

where *v* comes from (4).

Define a shadow edge to be the zero crossing locations of the difference image $\Delta I(x, y, t)$ as follows:

$$\Delta I(x,y,t) = I\ (x,y,t,v) + I_{thresh}(x,y,v)$$

For P(x, y), if there is not $\Delta I(x, y, t)=0$ at all time instance, we can get $t_s(x,y)$,by interpolating linearly between t_0 and t_0-1 ($\Delta I(x, y, t_0)>0$, and $\Delta I(x, y, t_0-1)<0$). Shadow boundary localization can be alternatively performed in the spatial domain (in pixel coordinates *P(x, y)* fixing the time $t=t_0$) constant or in the temporal domain (in time coordinate t at a fixed pixel point *P(x ,y)* in the image) in a consistent way.

Because the line shadow scans forward from one side to another, here we assume the vertical line scan from left to right. The pixels on a horizontal line see the shadow edge in sequence. So there is the constrain: $t_s(x,y)>t_s(x+1,y)$. if $t_s(x,y)> t_s(x+i,y),i=1,\ldots,n$.

We can see this point *P(x, y)* as a bad point because of it's discontinuity in time space.

5 Triangulation

Once the shadow time $t_s(x, y)$ is estimated at a given pixel *P'*, the P'_{top} and P'_{bot} at time t_s can be found. The lines $O_cP'_{top}$ and $O_cP'_{top}$ intersect the desk plane S_d. We can get P_{top} and P_{bot} which can identify the corresponding shadow plane $S_s(t_s(x, y))$. Then, the 3D point *P* associated to *P'* is retrieved by intersecting $S_s(t_s(x, y))$ with the optical ray O_cP'. Notice that the shadow time $t_s(x, y)$ acts as an index to the shadow plane list $S_s(t)$. Since $t_s(x, y)$ is estimated at sub-frame accuracy, the final plane $S_s(t_s(x, y))$ actually results from linear interpolation between the two planes $S_s(t_0-1)$ and $S_s(t_0)$, if $t_0 -1< t_s(x, y)< t_0$ and t_0 is integer.

6 Implement Results

Fig. 5 shows the 3D points cloud after triangulation. We can see that the scanner can give satisfying result on the object made of paper, plastic, and metal.

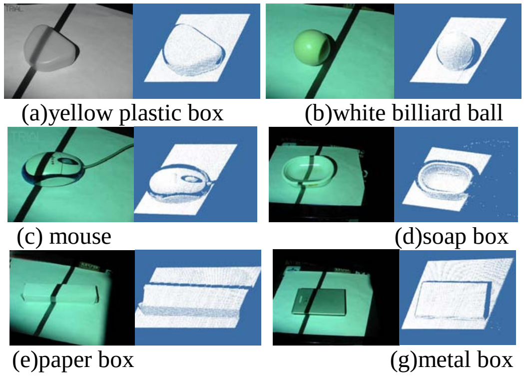


Fig. 5. Results after triangulation

7 3D Filter

The previous sections allow us to get 3D points cloud. Usually, the output is unsatisfactory. In our experiments, we observed that most areas are seriously affected by noise. Wrong triangulated points are scattered all over the object especially in the areas with low contrast. First, the points in and desk plane are discarded. Then we therefore designed two post-processing algorithms to annihilate the triangulated points that are most probably not part of the object.

7.1 Meshing Using Context Knowledge of the Mesh Topology

Assuming that the points belong to the surface of the object, we further assume that a point is valid if and only if it has two neighboring points to produce a triangle surface. Our experiments have demonstrated that using the connectivity of the points (that is the mesh topology), has improved the results. The algorithm is following:

```
Procedure FilterThroughMeshing(match_data)
for rows=0 to number of rows-1
  for column=0 to rowLength
    if (match at row happens before match at row+1)
      if (match upper left has less than 2 neighbors)
        remove match and its neighbors.
      else if (match lower left has less than 2 neighbors)
        remove match and its neighbors.
      end if
    end if
  end for
end for
```

with `match_data` a data structure implemented as a 1D array of lists. The 1D array of `match_data` represents the rows of the input image.

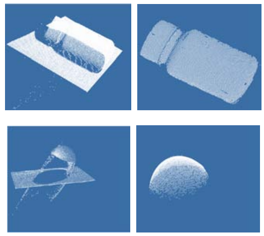


Fig. 6. Right: point cloud with noise Left :point cloud after using 3D filter

7.2 Face Orientation Culling

The previous procedure actually generates a triangle mesh. Using this mesh, the task now is to remove those triangular surfaces that are not visible to the camera and the projector. This is simply a backface culling algorithm that computes the scalar product of the triangle surface normal with the normalized direction of the line of sight to the center of this triangle. Thereafter, negative values are omitted. It actually suffices to consider only one of the triangle corners as the center. Furthermore, we leave out surfaces with normals that are almost orthogonal to the direction of the line of sight. This can be parameterized with the corresponding angle or scalar product value. In practice, depending on the shape of the object, an angle of up to 60 degrees between the surface normal and the normalized line of sight vector can be reasonable. Fig.6 shows the results after 3D filter.

8 Implement and Accuracy Analyses

The necessary hardware devices are a video projector, a digital camera and a standard PC. All these components are consumer level technology, allowing an affordable solution. We experimented with the HITACHI CP-HX1080 LCD-projector, which has a 32 bits color resolution and display patterns of *2592x1944* pixels. The selected camera is the OLYMPUS C5060WZ, that can provide images with sufficient resolution (up to 6 megapixels) and quality, as uncompressed raw images in TIFF format could be saved. PCs we use have Pentium 4 processors with 2.0 GHz, 512 MB RAM and nVidia graphics cards. Tab.1 shows the process time using our scanner on several samples.

Table 1. Process time of our scanner

objects	yellow plastic box	billiard ball	soap box	paper box	Mouse	metal box	Medical bottle
Times (seconds)	1.30	1.22	1.80	1.21	1.77	1.68	1.25

9 Accuracy Analysis

9.1 Calibration Accuracy Analysis

We define the following projection error formula:

$$DIPE=\sqrt{(X_f - X_f')^2 + (Y_f - Y_f')^2} ,$$

where *DIPE* means Distorted Image Plane Error, (X_f, Y_f) is the measured position of the point's image and $(X_{f'}, Y_{f'})$ is the position of the point's 3D coordinates (X_w, Y_w, Z_w) projected by the estimated camera calibration parameters.

$$UIPE = \sqrt{(X_{u2} - X_{u1})^2 + (Y_{u2} - Y_{u1})^2} \ ,$$

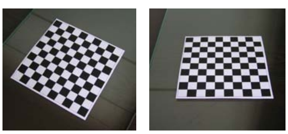
where *UIPE* means Undistorted Image Plane Error, (X_{u2},Y_{u2}) is calculated from the measured position of the point's image (X_f,Y_f), and (X_{u1},Y_{ul}) is calibrated from the 3D coordinates of the point (X_w,Y_w,Z_w).

$$OSE = \sqrt{(X_{c2} - X_{c1})^2 + (Y_{c2} - Y_{c1})^2} \ ,$$

where OSE means Object Space Error, (X_{c2}, Y_{c2}) is calculated from the measured position of the point's image (X_f, Y_f), and (X_{c1},Y_{cl}) is calculated from the 3D coordinates of the point (X_w, Y_w, Z_w).

Table 2. Error Data calculated from two demo photo I and II in fig. 6

Photo	Error Type	Mean	Standard Deviation	Max
I	DIPE (pix)	0.1226	0.0673	0.3614
	UIPE (pix)	0.1265	0.0701	0.3760
	OSE (mm)	0.0776	0.0423	0.2232
II	DIPE (pix)	0.3291	0.2206	1.0360
	UIPE (pix)	0.3476	0.2298	1.0458
	OSE (mm)	0.1648	0.1067	0.4728



(a) (b)

Fig. 7. Photo I (a) and Photo II (b)

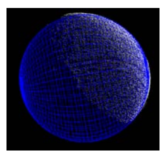


Fig. 8. Fitting result of reconstructed point cloud of the white billiard ball

9.2 Measurement Accurate Analysis

We fit the point data of the white billiard ball to a sphere. Then we visualized the fitting result shown in **Fig.** 8.

The formula we use for estimate relative error is:

$$Error = \frac{1}{N * R}\sqrt{\sum_{i=1}^{N}(R_i - R)^2} \ ,$$

where R_i are distances from measured points to center, R is the radius of the fitted sphere, and N is the number of points. The relative error of the point data is only 0.0031385%.

10 Conclusion and Future Work

We have presented a low cost 3D shape scanner which can measure the objects having specular reflections. The distortion of the line shadow on the object is used to get the object's 3D shape. The specific technical contributions include first a novel space-time edge finding algorithm that overcomes the interreflection influence. Second, we use the relation of the viewpoint and objects to filter the point cloud, which can filter most bad data. Third we use only a project, a digital camera and a PC to setup a low cost and small scale scanner, the method required little processing and image storage so that it can be implemented in real time. There are many important topics to explore in future work. First, the camera, the projector and the objects are fixed now. the pixel in the shadow area because of occlusions can not give any 3D information. A natural extensional extension of our current work is to make the projector, the camera or the objects movable. The alighment and merging techniques of 3D point data and uncalibrated vision method can be used in the future.

Acknowledgments. This work is supported by the National Natural Science Foundation of China under Grant Nos. 60373070.

References

1. Brown, D.C.: Calibration of close range cameras. In: Proc 12th congress Int. Soc. Photogrammetry, Ottawa, Canada (1972)
2. Tsai, R.Y.: A versatile camera calibration technique for high accuracy 3d machine vision metrology using off-the-shelf tv camera and lensed. IEEE J. Robotics Automat, RA 3(4), 323–344 (1987)
3. Tsai, R.Y.: An efficient and accurate camera calibration techniwue fpr 3d machine vision. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, pp. 364–374 (1986)
4. Li, H., Straub, R., Prautzsch, H.: IASTED International Conference on Visualization, Imaging and Image Processing, Fast subpixel accurate reconstruction using color structured light (2004)
5. Rocchini, C., Cignoni, P., Montani, C., Pingi, P., Scopigno, R.: A low cost 3D scanner based on structured light. Eruographics 2(3) (2001)
6. Rusinkiewicz, S., Hall-holt, O., Levoy, M.: Real-Time 3D model Acquisition, ACM SIGGRAPH, 21(3) (2002)
7. Curless, B., Seitz, S.: 3D photography, ACM SIGGRAPH, Course No. 19 (2000)
8. Pagès, J., Salvi, J., García, R., Matabosch, C.: Overview of coded light projection techniques for automatic 3D profiling. In: IEEE International Conference on Robotics and Automation, ICRA 2003 (2003)

Automatic Joints Extraction of Scanned Human Body

Yong Yu[1,2], Zhaoqi Wang[1], Shihong Xia[1], and Tianlu Mao[1,2]

[1] Institute of Computing Technology, Chinese Academy of Sciences,
100080 Beijing, China
[2] Graduate School of the Chinese Academy of Sciences, 100080 Beijing, China
{yyu,zqwang,xsh,ltm}@ict.ac.cn

Abstract. This paper presents an automatic method to extract joints accurately from scanned human body shape. Firstly, model is divided into slices by a group of horizontal parallel planes and primary landmarks are determinated according to the slices. Additionally, we propose a model intersection method based on direction of limbs. Model is separated into five segments including torso and four limbs and approximate direction of each segment can be calculated according to the primary landmarks. Five groups of parallel planes perpendicular to corresponding segments are employed to divide each segment into slices, and contours are calculated by intersection detection between planes and triangles. Finally, we propose a circularity function differentiation technique to extract key contours at joints from segments and the joints can be calculated according to centroids of the key contours. The experimental results demonstrate that the method has more advantages than the conventional ones, especially in algorithm's accuracy and robustness.

Keywords: Joints Extraction, Virtual Human, Scanned Human Body.

1 Introduction

Since the whole-body scanner was introduced into human body modeling, interests have been focused on automatic joints extraction from scanned human body models. With the development of three dimensional laser scanning technology, human body surface information accurately describing the human shape can be captured easily. Although the scanned human body data contain thousands of triangles and vertexes, the data have little semantic information and are difficult to be used in character animation, anthropometry and cloth simulation. For the purpose of effective use of the human body data, it is necessary to extract semantic information from the data, especially joints of the human body.

Nurre[2] presented a method to segment the human body in a standard posture (erect posture with arms and legs slightly apart) into six parts. This method divided human body into lots of horizontal slices and assigned each slice to corresponding part by analyzing the size and position of convex hulls of each slice. Ju[3] proposed an approach to extract joints from human body in the standard posture. In this approach, after horizontal slices were extracted by firing rays from the centers of the body to the surface, body was segmented into five lump parts according to the perimeter

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 286–293, 2007.

distributions of slices, and each part was further segmented based on curvatures of perimeter profiles of the part. Dekker[4], Buxton[5] and Ma[6] described techniques to segment human body by reentrant algorithm and to detect important joints such as shoulder, elbow, wrist, waist, hip, knee and ankle by analyzing the distributions of the slices' average radius. João[7] refined the techniques by detecting important joints according to local shape characteristics such as local maximum curvature. Werghi[1] and Xiao[9] proposed an approach to segment human body into parts by Morse theory and Reeb graph, but this approach did not extract the joints of human body which is very important for character animation and anthropometry. Lawson [11] proposed a method which discretized the model, computed its discrete medial surface, and created the skeletal structure. The method could even be applied to the animals, but it did not give an accurate result. Wang [10] proposed a method to extract joints by detecting feature lines and using the statistical anthropometric proportion. Yueqi Zhong[8] used the proportion of head length to body height as the benchmark in model segmentation and joints extraction.

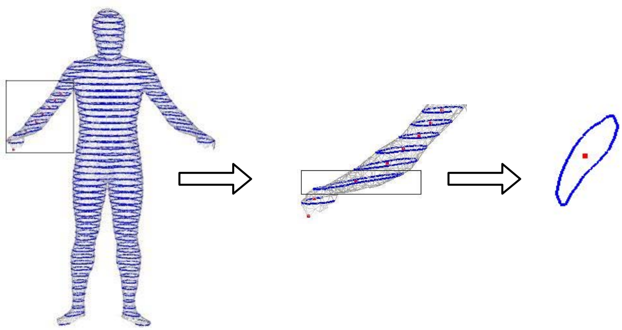


Fig. 1. Errors for intersecting with horizontal planes

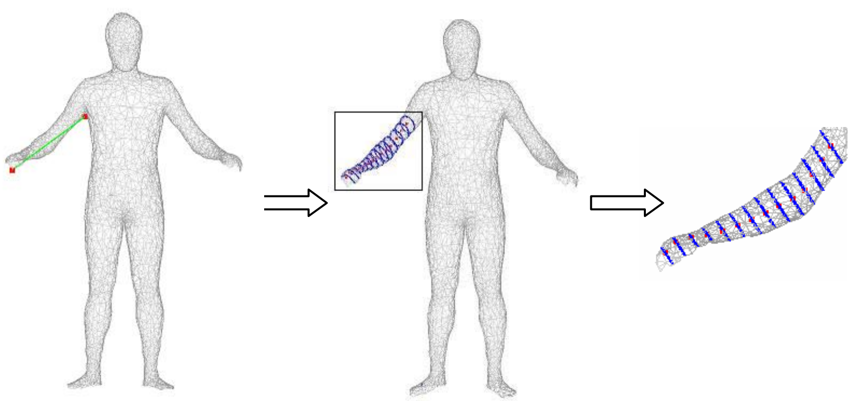


Fig. 2. Intersecting arm with parallel planes perpendicular to arm

These methods were proposed to extract joints from scanned human body automatically. But all of them have two limitations: 1) not robust. Model intersection with horizontal planes often brings errors, especially when human posture is slightly changed (Fig. 1); 2) not very accurate, because no prior anthropometric information is effectively utilized.

This paper proposes a novel joints extraction method to overcome the aforementioned limitations with three main steps. Firstly, model is divided into slices by horizontal planes to find primary landmarks[1]. Secondly, we propose a model intersection method based on direction of limbs to improve the model intersection process, so errors are eliminated when the human posture is slightly changed (Fig. 2). Finally, we propose a circularity function method to produce accurate results according to prior anthropometric assumption that the key contour near the joint is more irregular than the rest. We describe the method in the next section, and show our results and conclusions in section 3 and section 4 respectively.

2 Joints Extraction

All human bodies in this work are scanned in an erect posture with arms and legs slightly apart, and also lightly clothed, which allowed us to carry out 'touchless' measurements. The scanned human body model is represented as triangular meshes and each scanned model has about 5000 triangles and 2300 vertices (see Fig. 3).

In our work, the human model is treated as an articulated layered model. We use two layers: skeleton and skin. The structure and hierarchy of the skeleton that we wish to build are shown in Fig. 4.

In this section, parallel planes perpendicular to directions of limbs and torso is utilized to divide the human body model into slices, and the circularity function method is used to evaluate the characteristic of the slices for extracting joints from model.

2.1 Model Intersection

The previous methods produced contours by intersecting model with horizontal parallel planes, and extracted joints from the model by differentiating perimeters or average radii of a group of contours. However, model intersection in these methods often produced errors, when posture of the model was slightly changed. Fig. 1 shows the errors in intersecting model that a contour in Fig.1 even covers both palm and elbow because angle between forearm and horizontal plane is less than 20 degrees. Centroid of the contour that covers both palm and elbow in Fig.1 is usually considered as a joint in the previous methods, because the contour often has extremum of perimeter or average radius among all contours. The errors in that condition will not be reduced, unless a novel model intersection method is utilized.

To overcome the difficulty, we propose a novel method to intersect models correctly. Firstly, human models are divided into a group of slices by horizontal parallel planes. Unlike the previous method, these slices are not utilized to extract joints in this method, but merely to detect the primary landmarks, such as the top of the head, left and right armpits, crotch and ends of the arms and legs, by using the previous method [4].

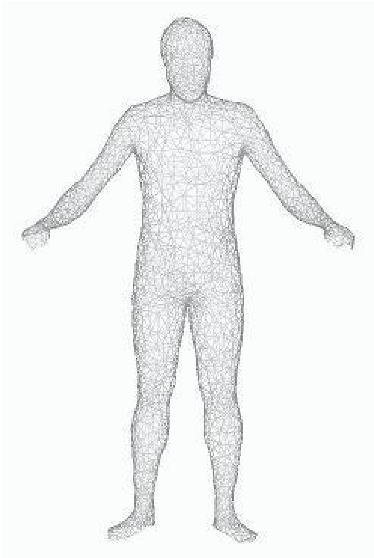


Fig. 3. Scanned human body

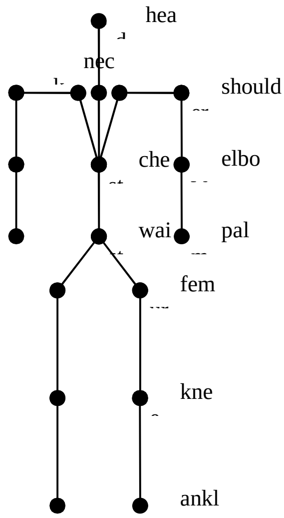


Fig. 4. Structure and hierarchy of skeleton

Secondly, directions of the left arm, the right arm, the left leg, the right leg and the torso are calculated. The arms can be segmented from the torso by detecting the transition slice that indicates the branch point at the armpits. Afterwards, the direction of the arms can be determinated approximately by the armpits and the ends of the arms. Similarly the legs can be segmented from the torso by detecting the crotch and the end of the legs, and their direction can also be calculated. The direction of the torso can be determinated by detecting the crotch and the armpits.

Thus, the human body model is composed of five segments including left arm, right arm, left leg, right leg and torso. Each segment is divided into a group of contours by intersecting the segments with parallel planes perpendicular to the direction of the corresponding segment. Fig. 2 shows result of arm intersection with planes perpendicular to the direction of the arm.

2.2 Joints Extraction

During the model intersection, the contours are produced to represent the human body model. The centroid of each contour can be calculated using formula 1, and every centroid $\overline{p_i}$ is a candidate joint. Thus, the problem of finding out the joint from the centroids in transformed into finding out the key contour at the joint from all contours of the segment.

$$\overline{p_i} = \frac{1}{n}\sum_{j=1}^{n} p_{ij} \qquad (1)$$

In Formula 1, *n* is the number of vertices in a contour, p_{ij} is the vertex *j* in contour *i*. $\overline{p}_i$ is the centroid of the contour *i*.

We propose a new method, namely circularity function method (CFM), to extract joints from the model. In other words, we are trying to distinguish the key contour at the true joint from the other contours.

An assumption is made for the circularity function method that the key contours at the joints are irregular because the bones and muscles clearly affect the shape of contours, on the contrary, the other contours are more regular and closer to the shape of circle because the connective tissue covers bones evenly.

The circularity function (shown in Formula 2) is utilized to estimate how close the contour is to circle.

$$g(M) = \frac{4\pi \cdot s(M)}{c^2(M)} \tag{2}$$

M is a contour of the model shape, *s* is the area of the contour, *c* is the perimeter of the contour and *g* is the circularity of the contour. When M is like a circle, *g* runs to 1. Contrarily, *g* runs to 0.

According to the assumption, the circularity function gains local minimum in the key contour whose centroid is a true joint because the key contour is more irregular than the others. Thus, the problem of joints extraction from scanned shape can be converted to a problem of resolving minimum of the circularity function of the contours in human body (Formula 3). In Formula 3, M_{joint} is the key contour that contains the joint.

$$M_{joint} = \arg\min(g(M)) = \arg\min(\frac{4\pi \cdot s(M)}{c^2(M)}) \tag{3}$$

Thus, the position of the joint can be calculated by Formula 4.

$$\overline{p}_{joint} = \frac{1}{n}\sum_{i=1}^{n} p_i, \qquad p_i \in M_{joint} \tag{4}$$

The circularity function curve of trunk and right leg is shown in Fig. 5. All the joints except in the shoulders and thighs can be extracted from the body model by the circularity function method, such as palms, elbows, knees, ankles, waist, chest and neck. Shoulders can be obtained from armpits, and thighs can be calculated based on waist and crotch.

3 Experimental Results

The scanned human body model in our experiment is represented as triangular meshes and each scanned model has about 5000 triangles and 2300 vertices. Twenty models are available for experiment of joints extraction, and some results are shown in Fig. 6.

Some previous methods[2][7] are implemented in this work to compare with our method.

Some parameters that can accurately reflect results of joints extraction and can be easily obtained from position of the joints are selected elaborately according to anthropometry, like length of leg, thigh, forearm and upper arm. Table 1 shows true value of the parameters and results of the previous methods and our method. Obviously, our method is more accurate and more robust than the previous methods.

Table 1. Results of the previous methods and our method

Average length (cm)	True value	Perimeter method [2]	Feature detection method [7]	Our method
Length of leg	40.6	37.4	41.3	40.5
Length of thigh	49.2	51.1	47.1	48.9
Length of forearm	25.6	22.8	25.0	25.6
Length of upper arm	31.6	32.3	32.1	31.1

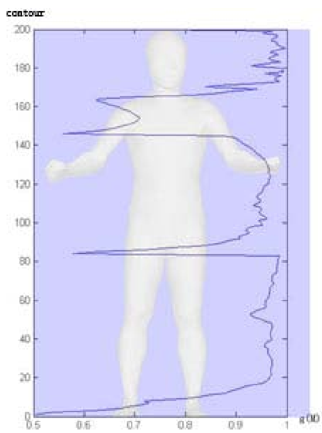


Fig. 5. Circularity function of trunk and right leg

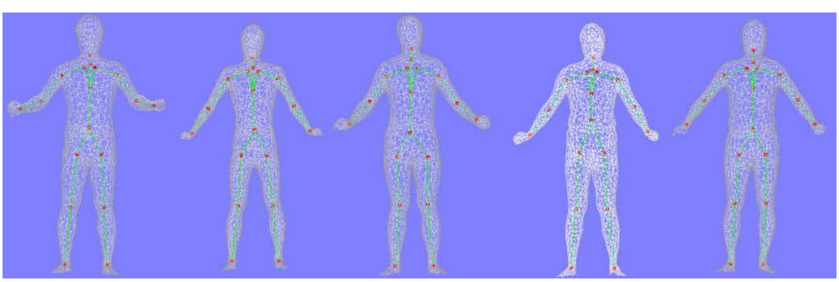


Fig. 6a. Joints and scanned human body

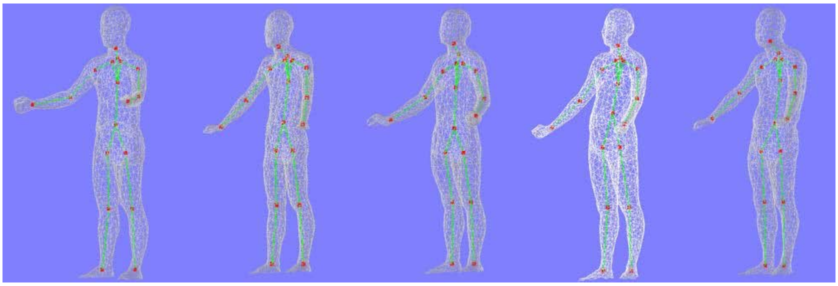


Fig. 6b. Joints and scanned human body

4 Conclusion

We have presented a novel algorithm to intersect scanned human body model and extract joints from it. This method uses a new model intersection technique and circularity function, avoiding errors by intersecting errors and improving the result of the joints extraction. Compared to the previous methods, our Algorithm has two advantages:

1) Robustness
Previous methods that intersect the model by horizontal planes can not cope with scanned human body in different postures. In this paper we overcome this problem by the new model intersection technique.
2) Accuracy
Previous methods that calculate the joints of the body by resolving minimum of the perimeter or average radius function of the contours do not make any sense in anthropometry, while this paper proposes circularity function to calculate the joints on the assumption of anthropometry. The experimental results have demonstrated the accuracy of the method.

Acknowledgement. This work is supported in part by NSFC-60403042, NSFC-60573162, NSFC- 60533070, NSFC-60603082, CISTC-2005DFA11060, BNSFC-4051004, BNSFC-4062032, 2006AA01Z336(863).

References

1. Yijun Xiao, W.N., Siebert, J.P.: A functional-based segmentation of human body scans in arbitrary postures. IEEE Transactions on Systems, Man, and Cybernetics, Part. B: Cybernetics 36(1), 153–165 (2006)
2. Nurre, J.: Locating Landmarks on Human Body Scan Data. In: Proceedings of the International Conference on Recent. In: Advances in 3-D Digital Imaging and Modeling (3DIM'97), Ottawa, Canada, pp. 289–295 (May 1997)
3. Yang, J.X., Naoufel, W., Paul, S.J.: Automatic segmentation of 3D human body scans. In: Proceeding of International Conference on Computer Graphics and Imaging 2000, Las Vegas, pp. 1–6 (2000)

4. Dekker, L., Douros, I., Buxton, B., Treleaven, P.: Building symbolic information for 3D human body modeling from range data. In: Proc.the Second International Conference on 3D Digital Imaging and Modeling, pp. 388–397, IEEE Computer Society (1999)
5. Buxton, B., Dekker, L., Douros, I., Vassilev, T.: Reconstruction and Interpretation of 3D Whole Body Surface Images. In: Numerisation 3D Scanning 2000 Conference Proceedings (April 2000)
6. Ma, Y.-Y., Zhang, H., Jiang, S.-W.: Realistic Modeling and Animation of Human Body Based on Scanned Data. Journal of Computer Science and Technology 19(4), 529–537 (2004)
7. Oliveira, J.F., Zhang, D., Spanlang, B., Buxton, B.F.: Animating Scanned Human Models. WSCG 2003 (2003)
8. Zhong, Y., Xu, B.: Automatic segmenting and measurement on scanned human body. International Journal of Clothing Science and Technology 18, 19–30 (2006)
9. Xiao, Y., Siebert, P., Werghi, N.: A Discrete Reeb Graph Approach for the Segmentation of Human Body Scans. 3-D Digital Imaging and Modeling, pp. 378–385 (2003)
10. Wang Charlie, C.L., Chang Terry, K.K., Yuan Matthew, M.F.: From laser-scanned data to feature human model: A system based on fuzzy logical concept. Computer-Aided Design 35(3), 241–253 (2003)
11. Wade, L., Parent, R.E.: Automated generation of control skeletons for use in animation. The Visual Computer 18(2), 97–110 (2002)

Wavelet Transform and Singular Value Decomposition of EEG Signal for Pattern Recognition of Complicated Hand Activities

Xiaodong Zhang, Weifeng Diao, and Zhiqiang Cheng

School of Mechanical Engineering,Xi'an Jiaotong University, Xi'an, Shaanxi Province 710049, China
xdzhang@mail.xjtu.edu.cn

Abstract. Electroencephalogram (EEG) is a useful bioelectrical signal in pattern recognition of hand activities because of its following characteristics: (1) patterns of EEG are different when doing diverse movements and mental tasks; (2) EEG can be real-timely or instantly extracted; (3) the measurement of EEG can reach an enough precision. A new approach for the pattern recognition of four complicated hand activities based on EEG is presented in this paper, in which each piece of raw data sequence for EEG signal is decomposed by wavelet transform (WT) to form a matrix, and the singular value of the matrix is extracted by singular value decomposition (SVD). And then the singular value, as the feature vector, is input to the artificial neural network (ANN) to discriminate the four hand activities including grasping a football, a small bar, a cylinder and a hard paper. Finally the research results show the correct classification rate of 89% was achieved by the approach mentioned above.

Keywords: Electroencephalogram, Pattern recognition, Wavelet transform, Singular value decomposition, Artificial neural network.

1 Introduction

EEG is a very useful signal produced by the electrical activity of brain. The signal contains a lot of information about the state of brain, which has been an important tool for diagnosing brain diseases, detecting brain functions, and analyzing brain activities. With the great achievement of modern science and technology, the researches about EEG are progressing rapidly. EEG is not only applied to clinical medicine, but also military medicine, aeromedicine, recognition of mental tasks, and psychology. In the latest years, owing to the progress of computer technology, modern electronics technology and microelectrode technology, EEG has the potential to become the information resource of precise control. Furthermore, the accurate pattern recognition technology of EEG is its precondition with no doubt.

For the recognition technology on EEG, a great deal researches have been done, and a series of valuable methods and technologies have been presented in the world. In terms of feature extraction and automatic recognition of EEG, the representative

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 294–303, 2007.

technologies should include time-frequency analysis, wavelet transform (WT), artificial neural network (ANN), adaptive filtering (AF), and blind source separation (BSS), etc.

With the deep development of the recognition technology above, some researches for EEG have been applied to the pattern recognition of hand movements. A Hilbert-transform based method for the predictions of left and right index finger movements from EEG measurements was reported in 1992, and its classification accuracy was 90%[1]. An approach for predicting the hand movements including grasping, opening and holding from the event-related EEG using neural network was presented in 1999, and an accuracy of 80% was reached [2]. A LVQ neural network was investigated in 2000 to classify single trial and repetition of hand movement, and the accuracy were 88% and 78%, respectively [3]. A system based on HMM was applied to discriminate brisk extension and flexions of the index finger, and movement of shoulder and fingers in 2001, the accuracy were 80% and 76% respectively [4]. An algorithm integrating both BP and ERD for classifying EEG of left and right hand movement was proposed in 2004, and the accuracy was 84%[5]. Y. Li, et al presented an algorithm to discriminate between the single-trial EEG of two different finger movement tasks, and the algorithm produced a classification accuracy of 92.1%[6].

For EEG signal applied in the real-time control of the dexterous robot hand finally, X. D. Zhang, et al proposed an approach for pattern recognition of hand activities based on EEG and Fuzzy Neural Network, and the classifying accuracy of hand movement from the other usual accompanying mental tasks was near 100%[7]. In this paper, an EEG based approach for pattern recognition of complicated hand activities using WT, SVD and ANN will be deeply investigated, and its high recognition accuracy will also be achieved.

2 Wavelet Transform of EEG Signal

WT is a new time-frequency analysis method for unstable signal. Differing from short time Fourier transform (STFT) and Wigner distribution, WT does not directly express signal in the time-frequency space but in the time-scale space. Every scale corresponds to a certain frequency range. So WT can be used to precisely locate the high frequency components and estimate the trend of the low frequency components. Now WT has been widely used in image processing, singularity detection, denoise filter and data compression. At the same time, WT shows a great potential in biomedical signal processing [8, 9].

2.1 Discrete Wavelet Transform

In computerized processing, continuous wavelet must be separated into discrete wavelet. The most common method is to discrete the scale factor into binary scales. We let $a = 2^j$. So we can define the dyadic wavelet transform (DWT) as follow.

$$\begin{aligned} WT(j,b) &= \langle x(t), \psi_{j,b}(t) \rangle \\ &= \frac{1}{\sqrt{2^j}} \int x(t)\psi^*\left(\frac{t-b}{2^j}\right)dt \end{aligned} \tag{1}$$

Suppose that the center frequency of mother wavelet $\psi(t)$ is Ω_0, and bandwidth is Δ_Ω. When $a=2^j$, the center frequency of $\psi_{j,b}(t)$ is as follows

$$(\Omega_j)_0=\Omega_0/2^j=2^{-j}\Omega_0 \tag{2}$$

And the bandwidth is

$$\Delta_{\Omega j}=2^{-j}\Delta_\Omega \tag{3}$$

We always hope that when the scale factor a increases from 2^j to 2^{j+1}, the analytic windows of $\psi_{j,b}(t)$ and $\psi_{j+1,b}(t)$, namely $[(\Omega_j)_0-\Delta_{\Omega j}, (\Omega_j)_0+\Delta_{\Omega j}]$ and $[(\Omega_{j+1})_0-\Delta\Omega_{j+1}, (\Omega_{j+1})_0+\Delta\Omega_{j+1}]$ can be connected together, so that the Fourier transform of $\psi_{j,b}(t)$ can cover the whole Ω axes. Obviously when $\Omega_0=3\Delta_\Omega$, the condition is satisfied.

In multi-level decomposition, the discrete signal $x(t),(t\in Z)$, is first split into the approximation a_1 and the detail d_1, i.e. $x(t)=a_1+d_1$. Then a1 is further split into a_2 and d_2. This procedure continues until the specified level is reached, for example, a_i and d_i ($i=1,2,\dots,n$) are the low-frequency component and high-frequency component of the *ith* level respectively.

2.2 Wavelet Decomposition of the EEG Signal

The frequency of human EEG (Δf) is between 1~60Hz[10]. According to the general feature of the EEG signal, there are five basic frequency components usually defined as follows.

(1) δ- delta wave with 0.5~4 Hz
(2) θ- theta wave with 4~8 Hz
(3) α- alpha wave with 8~14 Hz
(4) β- beta wave with 14~30 Hz
(5) γ- gamma wave with more than 30 Hz

So we can decide the number of decomposition level (i) by the above content and the sampling frequency (We can take fs=256Hz as an example).

According to the frequency relation between the approximation and detail as shown in Fig. 1, we can decide the decomposition level i=5. The useful components for recognition are as follow.

d2: 32~64Hz
d3: 16~32Hz
d4: 8~16Hz
d5: 4~8Hz
a5: 0~4Hz

Compared by experiments, Daubechies4 wavelet was chosen as the wavelet basis function, which is orthogonal and compacted and its order is 7. According to the knowledge of multi-resolution analysis and filter bank, Daubechies4 wavelet can be described by its scale function $\varphi(t)$ and wavelet function $\psi(t)$ as shown in Fig. 2. The first 1024 points of a piece of raw data were input to be decomposed with wavelet.

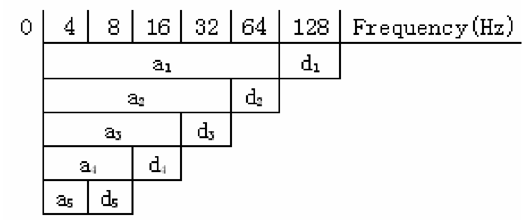
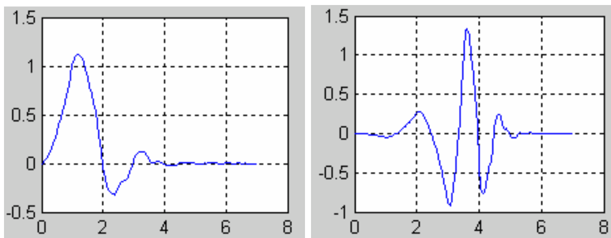


Fig. 1. Frequency relation between approximation and detail



(a) φ(t) (b) ψ(t)

Fig. 2. Daubechies4 wavelet

After that, a 5×1024 matrix (M_w) composed of five sequences, d_2, d_3, d_4, d_5 *and* a_5, was obtained for further processing.

As the same situation, the EEG signals of complicated grasping operations including grasping a football, grasping a small bar, grasping a cylinder and grasping a hard paper, should be respectively decomposed into 10 parts according to Fig. 1, but only d_2, d_3, d_4, d_5 and a_5 were five useful parts that include all the effective frequency bands of EEG, and denotes the signal in the frequency range of 32~64Hz, 16~32Hz, 8~16Hz, 4~8Hz and 0~4Hz respectively.

3 Singular Value Decomposition and Feature Extraction

According to Section II, the EEG data of complicated grasping operations above are decomposed by Daubechies4 wavelet, and we can get four 5×1024 matrix ($M_{wi}=[d_{2i}; d_{3i}; d_{4i}; d_{5i}; a_{5i}]$, $i=1,2,\ldots,4$). In order to pursue the simple and effective feature vector for the following recognition by ANN, the singular value of M_{wi} will be obtained by the singular value decomposition (SVD).

3.1 Singular Value Decomposition

Let A, B ∈Cm×n, if there are m order unitary matrix U and n order unitary matrix V, UHAV=B, then we say A and B are unitary equivalent. SVD is a kind of normalized form under the unitary equivalence.

If $A \in C_r^{m\times n}(r>0)$, the eigenvalues of AHA should have the following relation equation as

$$\lambda1 \geq \lambda2 \geq \ldots \geq \lambda r \geq \lambda r+1 = \ldots = \lambda n = 0 \tag{4}$$

then $\sigma_i = \sqrt{\lambda_i}$ ($i=1,2,\ldots,n$) are the singular values of A.

If $A \in C_r^{m\times n}(r>0)$, then there must be m order unitary matrix U and n order unitary matrix V, they satisfy (5).

$$U^H AV = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \tag{5}$$

where $\sum = diag(\sigma_1, \sigma_2, \ldots, \sigma_r)$, *and* $\sigma_i(i=1,2,\ldots,r)$ are the nonzero singular values of A. Equation (5) can be transformed to (6).

$$A = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^H \tag{6}$$

Equation (6) is called the SVD of A.

3.2 Feature Extraction of Complicated Grasping Operations

According to the method mentioned above, the singular value matrix of M_{wi} was obtained as follows.

$$M_{svdi} = [svd_{d2i}, svd_{d3i}, svd_{d4i}, svd_{d5i}, svd_{a5i}] \tag{7}$$

where $i=1,2,\ldots,4$. Obviously, for the EEG data of complicated grasping operations, the singular value matrix is a 4×5 matrix as follows.

$$M_{svd} = \begin{bmatrix} svd_{d21} & svd_{d31} & svd_{d41} & svd_{d51} & svd_{a51} \\ svd_{d22} & svd_{d32} & svd_{d42} & svd_{d52} & svd_{a52} \\ svd_{d23} & svd_{d33} & svd_{d43} & svd_{d53} & svd_{a53} \\ svd_{d24} & svd_{d34} & svd_{d44} & svd_{d54} & svd_{a54} \end{bmatrix} \tag{8}$$

If there is the following definition as

$$x_{\max} = \underset{i=1,44}{\overset{j=1,5}{Max}} \{m_{i,j}\} \tag{9}$$

$$x_0 = 10 x_{\max} / 9 \tag{10}$$

then, for the ith piece of EEG data, the normalized feature X can be obtained as follows.

$$X = [x_1, x_2, x_3, x_4, x_5] \tag{11}$$

$$X = [\, svd_{d2i} / x_0 , svd_{d3i} / x_0 , svd_{d4i} / x_0 , svd_{d5i} / x_0 , svd_{a5i} / x_0 \,] \tag{12}$$

4 Experimental Equipment and Recognition Model

4.1 Experimental Equipment

Fig.3 shows an experimental system set up by us, which mainly includes a subject, a PC, an EEG Amplifier named UE-16BZ and its matching parts. The EEG amplifier has a USB interface to connect to the PC, a high precision 16-bit A/D, and a related Impedance checking part to check the linking state between its electrodes and scalp. As shown in Fig.3, the amplifier has sixteen special EEG channels and other five bioelectric channels, which could be used to measure ECG, RESP, blood pressure according to actual demand. While some EEG signal is gathered, the amplifier will primarily accomplish signal gathering, signal amplifying and simple filtering.

Fig. 3. EEG experimental system

Our ultimate target is to put the pattern recognition of EEG into the practical humanoid control of artificial limb or robotic hand. So, it is a precondition how to complete the real-time processing of the EEG signals. Further more, as we know EEG signal at P3 and P4 point is sensitive to reflect hand operation nicely, we select EEG signal from P4 electrode as feature extraction source for the pattern recognition of the hand operations in the paper to simplify the feature extraction and guarantee its identifying precision.

4.2 Recognition Model

ANN is a good classifier widely applied in artificial intelligence, information processing, robotics and pattern recognition etc. Furthermore, back propagation (BP) network is a most common and mature kind of ANN for pattern recognition.

In this paper, a three layers BP network was adopted as shown in Fig.4, to complete the pattern recognition for four types of EEG signals.

According to feature extraction above, there are 5 neurons in the input layer for inputting the feature vector as in (11).

There are two neurons in the output layer, whose output can express the four patterns of EEG to be recognized as shown in Table 1, namely grasping a football, grasping a small bar, grasping a cylinder and grasping a hard paper, respectively.

Because there is no universal means to decide the number of neurons in the hidden layer, we had tried several numbers, and found that 8 is a suitable number of neurons in the hidden layer.

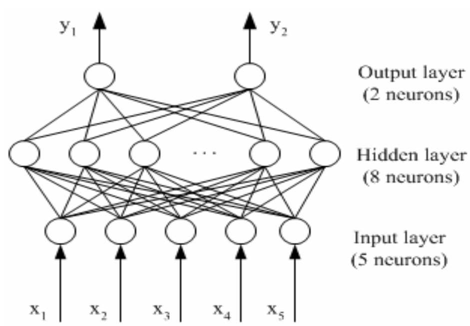


Fig. 4. Model of the BP network adopted

Table 1. Output and Corresponding EEG Pattern

Activity	y_1	y_2
Grasping a football	0	0
Grasping a small bar	0	1
Grasping a cylinder	1	0
Grasping a hard paper	1	1

5 Pattern Recognition of Complicated Hand Activities

According to the approach of the feature extraction and the recognition model of the BP network above, we took a lot of raw EEG data on complicated hand activities for extracting their feature, and randomly select the related learning samples 20 groups for training the BP network. On the other hand, 9 pieces of EEG data as the patterns to be recognized were extracted into the input feature vector and input to the BP network for testing its recognition accuracy.

5.1 Training of BP Network

All our recognition works were done based on MATLAB version 6.5, and the parameters of the BP network were set as follow.

Transfer function of the hidden layer: tansig
Transfer function of the output layer: logsig
Network training function: trainlm
Maximum number of epoch: 500
MSE performance goal: 0.01

Based on the training parameters, inputting the 20 groups of learning samples as described above, the BP network had finished its training with the mean square error of 0.00523432 after 62 epochs, and the whole training procedure was shown in Fig.5. The training result shows the BP network has a fast learning speed so that it has the potential for real-time processing.

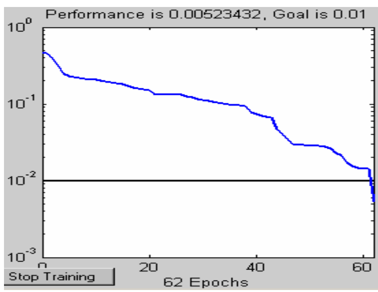


Fig. 5. Training result

5.2 Recognition Results

Table 2 shows the testing result for the recognition ability of the BP network. Through the analysis of comparing with Table 1, we can find the recognition possibility should be 100% for the two samples to be recognized of grasping a football or grasping a cylinder. Although there is a recognition result as y_1=0.7227 and y_2=0.9844, we can still think that the recognition possibility is nearly 100% for the three samples to be recognized of grasping a hard paper. However, for the two samples to be recognized of grasping small bar, the first recognition result as y_1=0.004 and y_2=1.000 is correct with no doubt, and the second recognition result as y_1=1.000 and y_2=1.000 is obviously wrong so that the hand activity of grasping a small bar is thought as grasping a hard paper, so its recognition possibility is 50%.

To sum up, 8 out of 9 testing samples were correctly recognized, so the correct classification rate of 89% has been achieved by the approach mentioned above.

Because the BP network as the recognition model adopts gradient descent (GD) algorithm, which often causes the network weights to converge on a local minimum

Table 2. Recognition Results of BP Network

Pattern input X to be recognized							Recognition result Y	
No.	Activity	x_1	x_2	x_3	x_4	x_5	y_1	y_2
1	Grasping a football	0.6569	0.1060	0.0631	0.0411	0.0240	0.0023	0.0000
2	Grasping a football	0.6735	0.1053	0.0628	0.0405	0.0220	0.0006	0.0006
3	Grasping a small bar	0.7824	0.1666	0.0742	0.0406	0.0236	0.0004	1.0000
4	Grasping a small bar	0.8327	0.1788	0.0765	0.0398	0.0192	1.0000	1.0000
5	Grasping a cylinder	0.6949	0.1369	0.0727	0.0418	0.0218	1.0000	0.0000
6	Grasping a cylinder	0.7192	0.1452	0.0771	0.0432	0.0205	0.9967	0.0000
7	Grasping a hard paper	0.6517	0.1053	0.0591	0.0330	0.0160	0.7227	0.9844
8	Grasping a hard paper	0.6701	0.1283	0.0581	0.0414	0.0182	1.0000	1.0000
9	Grasping a hard paper	0.9914	0.2441	0.0929	0.0552	0.0303	1.0000	1.0000

solution instead of a global minimum solution, the recognition system above inevitably has its drawback. Using different groups of training samples to train the BP network is an effective way to avoid the local minimum solution, and a bigger sample should be used further to test the reliability of the system.

6 Conclusions

WT, SVD and ANN were adopted to classify four types of hand activities, namely grasping a football, grasping a small bar, grasping a cylinder and grasping a hard paper in this paper. The research results show that there are four important conclusions obtained as follows.

(1) WT has a strong ability to highlight the characteristics of EEG signals both in time domain and in frequency domain, so it is suitable to measure the transient state of EEG signals.

(2) WT combining with SVD technology can be used for extracting the feature vector with good separability from the raw EEG data.

(3) Through providing proper feature vector and selecting some correct network parameters, the BP network has a correct classification rate of 89% in the pattern recognition of complicated hand activities based on EEG.

(4) The complicated activities of human hand can be discerned by analyzing the corresponding EEG signals.

Acknowledgments. Authors are very grateful for the support from National Natural Science Foundation of China (No.50375115) and its related NSFC-KOSEF Scientific Cooperation Program “Networked teleoperation of the dexterous robot hand based on EEG”.

References

[1] Medl, A., Flotzinger, D.: Hilbert-transform based predictions of hand movements from EEG measurements. Engineering in Medicine and Biology Society, 1992, vol. 14. In: Proceedings of the Annual International Conference of the IEEE, vol. 6 (October 29-November 1, 1992)

[2] Erfanian, A., Moghaddas, S.: Prediction of hand movement from the event-related EEG using neural networks, Engineering in Medicine and Biology, 1999. In: 21st Annual Conf. and the 1999 Annual Fall Meeting of the Biomedical Engineering Soc., 1999. Proceedings of the First Joint BMES/EMBS Conference, vol. 1, pp. 13–16 (October 1999)

[3] Pilla Jr, V.: Detection of movement-related desynchronization of the EEG using neural networks. Engineering in Medicine and Biology Society, 2000. In: Proceedings of the 22nd Annual International Conference of the IEEE, vol. 2, pp. 23–28 (July 2000)

[4] Stastny, J., Sovka, P.: EEG signal classification. Engineering in Medicine and Biology Society 2001. In: Proceedings of the 23rd Annual International Conference of the IEEE, vol. 2, pp. 25–28 (October 2001)

[5] Wang, Y.J., Zhang, Z.G.: BCI Competition 2003—Data Set IV: An Algorithm Based on CSSD and FDA for Classifying Single-Trial EEG. Biomedical Engineering, IEEE Transactions on, 51(6) (June 2004)

[6] Li, Y., Gao, X.R.: Classification of single-trial electroencephalogram during finger movement. Biomedical Engineering, IEEE Transactions on, vol. 51(6) (June 2004)

[7] Zhang, X.D., et al.: An approach for pattern recognition of hand activities based on EEG and fuzzy neural network. Journal of Mechanical Science and Technology 19(1), 87–96 (2005)

[8] Unser, M., Aldroubi, A.: A review of wavelets in biomedical applications. Proceedings of the IEEE 84(4), 626–638 (1996)

[9] Kalayci, T., Ozdamar, O.: Wavelet preprocessing for automated neural network detection of EEG spikes. Engineering in Medicine and Biology Magazine, IEEE 14(2), 160–166 (1995)

[10] Kiloh, L.G., McComos, A.J.: Clinical electroencephalography. Butterworths, Boston, London (1981)

Capturing 3D Human Motion from Monocular Images Using Orthogonal Locality Preserving Projection

Xu Zhao and Yuncai Liu

Shanghai Jiao Tong University, Shanghai 200240, China

Abstract. In this paper, we present an Orthogonal Locality Preserving Projection based (OLPP) approach to capture three-dimensional human motion from monocular images. From the motion capture data residing in high dimension space of human activities, we extract the motion base space in which human pose can be described essentially and concisely by more controllable way. This is actually a dimensionality reduction process completed in the framework of OLPP. And then, the structure of this space corresponding to special activity such as walking motion is explored with data clustering. Pose recovering is performed in the generative framework. For the single image, Gaussian mixture model is used to generate candidates of the 3D pose. The shape context is the common descriptor of image silhouette feature and synthetical feature of human model. We get the shortlist of 3D poses by measuring the shape contexts matching cost between image features and the synthetical features. In tracking situation, an AR model trained by the example sequence produces almost accurate pose predictions. Experiments demonstrate that the proposed approach works well.

Keywords: OLPP, Human Motion Analysis, Monocular Images.

1 Introduction

Capturing 3D human motion from 2D images is a significant problem in computer vision. For many image understanding applications, the 3D configurations of people in images provide usable semantic information about human activity. This is a challenging problem suffering from the obstacles conduced mainly by the complicated nature of 3D human motion and the information loss of 2D images. There are two main state-of-art approaches to deal with this problem [3]. *Discriminative methods* try to find the direct mapping from image feature space to pose state space by learning the mapping models from the training examples. This approach can supplies effective solution schemes for pose recovering problem if some additional issues can be well solved. However, the inherent one-more mapping from 2D image to 3D pose is difficult to learn accurately because the conditional state distributions are multimodal. The quantity and quality of training samples are also key factors, which can lead to some intractable problems to deal with. *Generative methods* follow the prediction-match-update philosophy. In the prediction step, the pose candidates are generated from the state prior distribution. The followed match step

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 304–313, 2007.

evaluates the pose-image similarity with some measurement. Finally, the optimal solution is found by the state update operation. Such approach has sound probabilistic support framework but generally computationally expensive because of the complex search over the high dimension state space. Moreover, prediction model and initialization are the bottlenecks of generative method especially for the tracking situation.

In this paper, we present a novel generative approach, by which we try to widen the bottlenecks mentioned above with lower computing expense. We represent the human poses by a 3D body model explicitly, whose configurations are expressed by the joint degrees of freedom (DOFs) of body parts. In our body model, there are more than fifty full body DOFs. This is a very large state space to search for the correct poses matching with the given images. Hence, the state space should be cut in order to avoid absurd poses. In general, the reasonable pose datum pool in some compact subspace of the full state space. We extract the subspace with the OLPP of motion capture data. In this concise subspace, there are some advantageous characteristics for pose estimation, which will be introduced detailedly in the followed sections. Based on the consistency of human motion, the structure of this subspace is explored with data clustering and thus we can divide the whole motion into several typical phases represented by the cluster centers. States prediction is a common difficulty of complicated non-linear problems for the absence of effective prediction model. We choose the Gaussian mixture model as state prediction model because this model can well approximates the multimodal pose distribution with the outcomes of data clustering. By the efficient shape contexts [7] matching, we evaluate the pose predictions and finally recover the 3D human pose. In the tracking situation, an autoregressive process guides the state prediction.

2 Related Work

There has been considerable prior work on capturing human motion [1-3]. However, this problem still hangs in doubt because it's ill conditioned in nature. For knowing how the human 3D pose is configured, more information are required than images can provide. Therefore, much work focus on using prior knowledge and experiential data. Explicit body model embody the most important prior knowledge about human pose and thus are widely used in human motion analysis [1]. Another class of important prior knowledge comes from the motion capture data. The combination of the both prior information causes favorable techniques for solving this problem.

Agarwal and Triggs [6] distill prior information of human motion from the hand-labeled training sequences using PCA and clustering on the base of a simple 2D human body model. This method presents a good tracking scheme but has no description about pose initialization.

Urtasun et al. [8,9] construct a differentiable objective function based on the PCA of motion capture data and then find the poses of all frames simultaneous by optimizing the function. Sidenbladh et al. [10,11] present the similar method in the framework of stochastic optimization. For a specific activity, such methods need many example sequences for computing the PCA and all of these sequences must keep same length and same phase by interpolating and aligning. Huazhong Ning et al.

[14] learn a motion model from semi-automatically acquired training examples that are aligned with correlation function. Unlike these methods, we extract the motion base space from only one example sequence of a specific activity using the lengthways OLPP and thus have no use for interpolating or aligning.

The methods mentioned above utilize the prior information in generative fashion. By contrast, discriminative approach [15-17] makes use of prior information by directly learning pose from image measurements. In [15], Agarwal and Triggs present several regression based mapping operators using shape context descriptor. Sminchisescu et al. [16] learn a multimodal state distribution from the training pairs based on the conditional Bayesian mixture of experts models. These methods can bring the interest of fast state inference after finishing the training. However, they are prone to fail when the small training database are used.

The styles of using prior information are multiform. Mori et al. [12] contain the prior information in the stored 2D image exemplars, on which the locations of the body joints are marked manually. By the shape contexts matching with the stored exemplars, the joint positions of the input images are estimated. And then, the 3D poses are reconstructed by the Taylor method [18].

Comparing with these methods, extracting the common characteristic of a special motion type from prior information is of particular interest to us. At the same time, we ensure the motion individuality of the input sequences in the generative framework with a low computational expense based on the efficient analysis of prior information.

3 OLPP-Based State Space Analysis

In this study, we represent the 3D configurations of human body as the joint angles vectors of body model. These vectors reside somewhere in the state space. The potential special interests motivate us to analyze the characteristics and structure of this space. Such interests involve mainly modeling the human activities effectively in the extracted base space and eliminating the curse of dimension. The state space analysis is performed on the base of OLPP.

3.1 Orthogonal Locality Preserving Projection (OLPP)

Orthogonal Locality Preserving Projection (OLPP) is a novel subspace learning algorithm presented by Deng Cai et al. [4]. This algorithm is built on the base of the Locality Preserving Projection (LPP) method [5] and primitively devised to solve the problem of face recognition. Actually, OLPP is an effective dimensionality reduction method falling into the category of manifold learning.

Considering the problem of representing all of the vectors in a set of n D-dimensional samples $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$ by n d-dimensional vectors $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n$, respectively, $D > d$. The objective function of LPP [5] is as follows:

$$\min \sum_{ij} \| \mathbf{y}_i - \mathbf{y}_j \|^2 S_{ij} \tag{1}$$

where S is a similarity matrix. A possible way of defining S is as follows:

$$S_{ij} = \begin{cases} \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / t), & \|\mathbf{x}_i - \mathbf{x}_j\|^2 < \varepsilon \\ 0, & otherwise \end{cases} \quad (2)$$

where ε is sufficiently small, and $\varepsilon > 0$. Here ε defines the radius of the local neighborhood. In other words, ε defines the locality. Therefore, minimizing the objective function is an attempt to ensure that, if $\mathbf{x}_i$ and $\mathbf{x}_j$ are close, then $\mathbf{y}_i$ and $\mathbf{y}_j$ are close as well. Finally, the basis functions of LPP are the eigenvectors associated with the smallest eigenvalues of the following generalized eigen-problem:

$$XLX^T\mathbf{w} = \lambda XDX^T\mathbf{w} \quad (3)$$

where $X = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$ and D is a diagonal matrix ; $D_{ii} = \sum_j S_{ji}$. $L = D - S$ is the Laplacian matrix and $\mathbf{w}$ is the transformation vector. The algorithmic procedure of OLPP is stated below.

1. *PCA projection:* projecting the high dimensionality points $\mathbf{x}_i$ into the PCA subspace by throwing away the components corresponding to zero eigenvalue. The transformation matrix of PCA is W_{PCA}.
2. *Constructing the adjacency graph:* Let G denote a graph with n nodes. The i-th node corresponds to $\mathbf{x}_i$. Putting an edge between nodes i and j if $\mathbf{x}_i$ and $\mathbf{x}_j$ are close, i.e. $\mathbf{x}_i$ is among p nearest neighbors of $\mathbf{x}_j$ or $\mathbf{x}_j$ is among p nearest neighbors of $\mathbf{x}_i$.
3. *Choosing the weights:* according to equation (2).
4. *Computing the orthogonal basis functions:* Let $\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_k$ be the orthogonal basis vectors, defining:

$$A^{(k-1)} = [\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_{k-1}] \quad (4)$$

$$B^{(k-1)} = [A^{(k-1)}]^T (XDX^T)^{-1} A^{(k-1)} \quad (5)$$

The orthogonal basis vectors $\mathbf{w}_1, \mathbf{w}_2, ...\mathbf{w}_k$ can be computed as follow.

– Compute $\mathbf{w}_1$ as the eigenvector of $(XDX^T)^{-1}XLX^T$ associated with the smallest eigen-value.

– Compute $\mathbf{w}_k$ as the eigenvector of

$$M^{(k)} = \left\{ I - (XDX^T)^{-1} A^{(k-1)} \left[B^{(k-1)}\right]^{-1} \left[A^{(k-1)}\right]^T \right\} \cdot (XDX^T)^{-1} XLX^T \quad (6)$$

associated with the smallest eigenvalue of $M^{(k)}$.

5. *OLPP Embedding:* Let $W_{OLPP} = [\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_l]$, the embedding is as follows.

$$\mathbf{x} \rightarrow \mathbf{y} = W^T \mathbf{x}, \quad W = W_{PCA} W_{OLPP} \tag{7}$$

In this paper, utilizing the orthogonality of the base functions, we reduce the dimensionality of state space and then reconstruct the data.

3.2 Pose Representation

We represent the human pose using the explicit body model. Our fundamental 3D skeleton model (see Figure. 1a) is composed of 34 articulated rigid sticks. There are 58 pose parameters in our model, including 55 joint angles of body parts and 3 global rotation angles. Therefore, each body pose can be viewed as a point in the 58D state space.

Figure. 1b show the 3D convolution surface [19] human model, which actually is an isosurface in a scalar field, defined by convolving the 3D body skeleton with a kernel function. Similarly, the 2D convolution curves of human body as shown in Figure. 1c are the isocurves generated by convolving the 2D projection skeleton. As the synthetical model features, the curves will match with the image silhouettes.

Fig. 1. (a) The 3D human skeleton model. (b) The 3D human convolution surface model. (c) The 2D convolution curves.

3.3 Extracting the Base Space

All of the human poses distribute in the 58D state space. The poses belong to a special activity, such as walking, running, handshaking, etc., generally crowd in a subspace of the full state space. We extract this subspace from the motion capture data obtained from the CMU database [20].

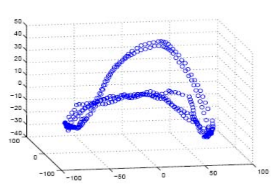


Fig. 2. The manifold of the walking sequences in 3D base space

According to the OLPP, any pose vector (t is the time tag) in a training sequence can be expressed as:

$$\mathbf{x}_t = \mathbf{y}_t W^T \tag{8}$$

where $\mathbf{x}_t$ and $\mathbf{y}_t \in \mathbb{R}^q$ are column vectors and W^T is the transformation matrix. $p(<q)$ is the dimension of the base space. Here, we take $p=5$ $p=5$, which means that recovering the human pose in the 5D base space only lose negligible information. In this way, we extract the base space covering a special human activity from a single training sequence. Actually, the training sequences belonging to a same motion type but performed by different subjects can produce similar outcomes. For example, our experiments demonstrate that the walking training sequences generate the similar manifold in the 3D base space as shown in Figure. 2. Thus, by extracting the base space, we represent the pose as a 5D vector in base space.

The interests of extracting the base space include not only the dimension reduction but also the advantages for analyzing. We have known that the special human motion type shows the special manifold in the base space. Essentially, this manifold indicates the common identity of the motion type. Therefore, our focus is transferred from the base space to the more local part: special manifolds, which actually are the point set presenting special geometry shape in the base space. We analyze the manifolds with the k-means clustering. Based on the activity continuity, the set of $\mathbf{y}_t$ can be partitioned into different connected subsets and every subset represents a special motion phase. Here, we choose the number of clustering as 4. Every clustering center is the key-frame of the motion sequence. Figure. 3 shows the 3D human poses corresponding to the clustering centers. In the followed tracking process, the clustering outcomes are used in the pose prediction model.

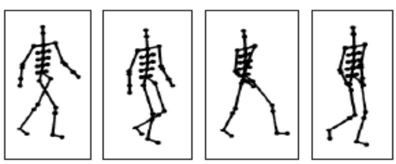


Fig. 3. The pose key frames in walking sequence

4 Image Matching Likelihood

In generative framework, pose recovering be formulated as a Bayesian posterior distribution inference:

$$p(\mathbf{y} \mid \mathbf{o}) \propto p(\mathbf{y})\, p(\mathbf{o} \mid \mathbf{y}) \tag{9}$$

where $\mathbf{o}$ represents the image observations. The likelihood function $p(\mathbf{o} \mid \mathbf{y})$ is used for evaluating every pose candidate generated by the prediction models.

We choose the image silhouettes as the observed image feature as shown in Figure.4.

Fig. 4. (a) Original image. (b) Image silhouette.

We describe the image silhouettes and the convolution curves using shape context descriptor [7], a robust and discriminative shape descriptor. The likelihood function is constructed by the shape contexts matching [13]. In the matching process, we first sample the edge points of the image silhouettes as the query shape. Next, the point set sampled from the convolution curves are as the known shapes. Before matching, the image shape and the candidate shape are normalized to same scale. We denote the image shape as $\mathbf{S}_{query}$ and the candidate pose shape as $\mathbf{S}_i$. The matching cost can be formulated as:

$$C_v(\mathbf{S}_{query}, \mathbf{S}_i) = \sum_{j=1}^{r} \chi^2 (SC_{query}^j, SC_i^*) \tag{10}$$

where SC is the shape context, r is the number of sample point in image shape, and $SC_i^* = \arg\min_u \chi^2 (SC_{query}^j, SC_i^u)$. Here, we use the χ^2 distance as the similarity measurement.

5 Tracking

For image sequences, tracking means that the pose estimation in current time step depends on the outputs of previous estimation. Thus, the most important part of tracking is dynamical model that indicates how the state evolves with time. Another intractable problem in tracking is the state initialization. In this section, we deal with the problems of tracking based on the outcomes of state space analyzing in generative framework.

5.1 Initialization

Initialization is the first step of tracking, aiming for finding the correct pose of the first frame in a given image sequence. We present an auto-initialization scheme based on the Gaussian mixture model. In the base space depicted in section 3, a pose $\mathbf{y}$ can be viewed as a 5D random vector that is generated from a multimodal distribution. This distribution can be formulated as:

$$p(y) = \sum_{i=1}^{c} \omega_i \cdot N(\mathbf{y}; \mathbf{y}_{ci}, \Sigma_i) \tag{11}$$

where, $c = 4$ is the number of pose clustering, $\{\omega_i : i = 1,2,3,4. \sum_{i=1}^{4} \omega_i = 1\}$ are the weights of single Gaussian distributions and $\{\sum_i : i = 1,2,3,4\}$ are the variances of these distributions which can be computed from the training sequence.

The procedure of *auto-initialization* is performed as follows:

1. Estimating of the global rotations by:
– Partitioning the global angle scopes into 8 bins (Relying on the robustness of the matching method, 8 bins is enough.).
– Generating N samples from each single Gaussian distribution in every bin. (In our experiments, N=3.)
– Performing shape contexts matching between the query image and convolution curves produced by the sample poses.
– Evaluating the bins according to the matching score. The bin containing the minimum cost score wins. By the way, recording the matching scores of every sample pose.
2. Determining the pose in the query image.
– Generating pose samples from the Gaussian mixture distribution as formulated in the Equation (11). The weights are determined as follows: (1) Taking out the minimum matching score of each Gaussian distribution from the winning bin (see step 1). (2) Obtaining the weights by normalizing the matching scores to [0,1].
– Evaluating these pose samples and determining the pose shortlist in which there are n samples with minimum matching scores.
– The final pose is the weighted sum of poses in shortlist.

5.2 Dynamical Model

Because of the complicated nature of human motion, it's difficult to obtain an analytical physical model for it. We prefer to seek the statistical dynamical model of human motion from the training data. Similar to the model introduced in [6], we learn a second order Auto-Regressive Process (ARP) for the time domain prediction of pose in the base space. In tracking situation, the probability distribution of pose in time t can be formulated as:

$$p(\mathbf{y}_t \mid \mathbf{o}) \propto p(\mathbf{o} \mid \mathbf{y}_t) p(\mathbf{y}_t \mid \mathbf{Y}_{t-1}) \tag{12}$$

in our model, $\mathbf{Y}_{t-1} = \{\mathbf{y}_{t-1}, \mathbf{y}_{t-2}\}$. And the prediction distribution $p(\mathbf{y}_t \mid \mathbf{Y}_{t-1})$ is modeled by the second order ARP:

$$\mathbf{y}_t = \mathbf{M}_1 \mathbf{y}_{t-1} + \mathbf{M}_2 \mathbf{y}_{t-2} + v_t \tag{13}$$

where the fifth order matrices $\mathbf{M}_{1,2}$ and the variances of Gaussian white noise v_t are learnt from the training sequences. These parameters are corrected in the process of pose recovering according to the estimated outcomes. Guided by the dynamical model, we find the correct poses using particle filter. The computational expense of our method is low because the special manifold that represents the common

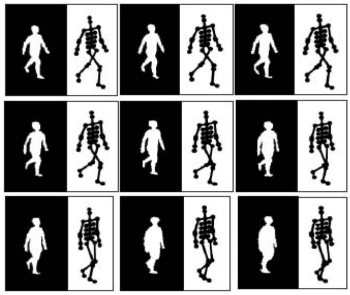


Fig. 5. The tracking results

characteristic of special motion type in the base space lead to the accurate dynamical model and therefore tracking can be proceeded with few particles. The results of tracking are shown in Figure.5. Experiments demonstrate that our method works well.

6 Conclusion

We have introduced a novel approach to tracking 3D human motion. This method extracts the compact base space from motion capture data that contain the prior information about human motion. Actually, in so doing, we extract the nature of a motion type and represent it by a compact way. Corresponding to a special motion type, a special manifold in base space indicates the common identity of this motion type. This can lead to the efficient estimation of human poses. We use the shape context matching to measure the similarity between the query image and the candidate poses. Experiments demonstrate that this is a robust and discriminative matching method. As the predict model, the Gaussian mixture model and the ARP model wok well in the process of tracking. In terms of future work, we will cover more types of human motions by including a wider range of training data. We plan to improve the matching method in order to reduce the computing expense further. And, the conception of base space will be extend to the recognition of human activity.

References

1. Aggarwal, J.K., Cai, Q.: Human Motion Analysis: A Review. Computer Vision and Image Understanding 73(3), 428–440 (1999)
2. Moeslund, T.B., Granum, E.: A Survey of Computer Vision-Based Human Motion Capture. Computer Vision and Image Understanding 81, 231–268 (2001)
3. Sminchisescu, C.: Optimization and Learning Algorithms for Visual Inference. In: ICCV 2005 tutorial (2005)
4. Cai, D., He, X., Han, J., Zhang, H.-J.: Orthogonal Laplacianfaces for Face Recognition. IEEE Transactions on Image Processing 15(11), 3608–3614 (2006)
5. He, X., Yan, S., Hu, Y., Niyogi, P., Zhang, H.-J.: Face recognition using laplacianfaces. IEEE Trans. on Pattern Analysis and Machine Intelligence, 27(3) (2005)

6. Agarwal, A., Triggs, B.: Tracking Articulated Motion Using a Mixture of Autoregressive Models. In: Proc. European Conf. Computer Vision (2004)
7. Belongie, S., Malik, J., Puzicha, J.: Shape Matching and Object Recognition Using Shape Contexts. IEEE Trans. Pattern Analysis and Machine Intelligence 24(4), 509–522 (2002)
8. Urtasun, R., Fleet, D.J., Fua, P.: Monocular 3D Tracking of The Golf Swing. In: Proc. IEEE CS Conf. Computer Vision and Pattern Recognition vol. 2, pp. 932–938 (2005)
9. Urtasun, R., Fua, P.: 3D Human Body Tracking Using Deterministic Temporal Motion Models. In: Proc. European Conf. Computer Vision, Prague, Czech Republic (May 2004)
10. Sidenbladh, H., Black, M., Sigal, L.: Implicit Probabilistic Models of Human Motion for Synthesis and Tracking. In: Proc. European Conf. Computer Vision, vol. 1 (2002)
11. Ormoneit, D., Sidenbladh, H., Black, M., Hastie, T.: Learning and Tracking Cyclic Human. In: Advances in Neural Information Processing Systems, vol. 13, pp. 894–900. The MIT Press, Cambridge (2001)
12. Mori, G., Malik, J.: Recovering 3D Human Body Configurations Using Shape Contexts. IEEE Trans. Pattern Analysis and Machine Intelligence 28(4), 1052–1062 (2006)
13. Mori, G., Belongie, S., Malik, J.: Efficient Shape Matching Using Shape Contexts. IEEE Trans. Pattern Analysis and Machine Intelligence 27(11), 1832–1837 (2005)
14. Ning, H., Tan, T., Wang, L., Hu, W.: People Tracking Based on Motion Model and Motion Constraints with Automatic Initialization. Pattern Recognition 37, 1423–1440 (2004)
15. Agarwal, A., Triggs, B.: Recovering 3D Human Pose from Monocular Images. IEEE Trans. Pattern Analysis and Machine Intelligence 28(1), 44–58 (2006)
16. Sminchisescu, C., Kanaujia, A., Li, Z., Metaxas, D.: Discriminative Density Propagation for 3D Human Motion Estimation. In: Proc. IEEE CS Conf. Computer Vision and Pattern Recognition. vol. 1(1), pp. 390–397 (2005)
17. Rosales, R., Sclaroff, S.: Learning Body Pose Via Specialized Maps. In: NIPS (2002)
18. Taylor, C.J.: Reconstruction of Articulated Objects from Point Correspondences in a Single Uncalibrated Image. Computer Vision and Image Understanding 80, 349–363 (2000)
19. Jin, X., Tai, C.-L.: Convolution Surfaces for Arcs and Quadratic Curves with a Varying Kernel. The Visual Computer 18, 530–546 (2002)
20. CMU database: http://mocap.cs.cmu.edu/

Human Motion Simulation and Action Corpus

Gang Zheng[1], Wanqing Li[1], Philip Ogunbona[1],
Liju Dong[1,2], and Igor Kharitonenko[1]

[1] School of Computer Science and Software Engineering, University of Wollongong, Australia
[2] College of Information Engineering, Shenyang University, P.R. of China
{gz207,wanqing,philipo,liju,igor}@uow.edu.au

Abstract. Acquisition of large scale good quality training samples is becoming a major issue in machine learning based human motion analysis. This paper presents a method to simulate continuous gross human body motion with the intention to establish a human motion corpus for learning and recognition. The simulation is achieved by a temporal-spatial-temporal decomposition of human motion into actions, joint actions and actionlets based on the human kinematic model. The actionlet models the primitive moving phase of a joint and represents the muscle movement governed by kinesiological principles. Joint actions and body actions are constructed from actionlets through constrained concatenation and synchronization. Methods for concatenation and synchronization are proposed in this paper. An action corpus with small number of action vocabularies is created to verify the feasibility of the proposed method.

Keywords: Human Motion, Actions, Simulation, Motion Editing.

1 Introduction

Automatic recognition of human motion has been one of the most active research topics in computer vision in the past decades [1,2]. Recent research [1,2,3,4,5] has demonstrated that machine learning approach has the potential to lead to a generic solution. However, the success of machine learning approach hinges on the quality and amount of training samples. Although limited individual action samples maybe obtained using motion capture devices [6,7,8,9,10,11], acquisition of large amount of training samples remains challenging considering the cost and amount of work required to obtaining such training samples. In this paper, we propose to generate large amount of quality motion data through simulation for the purpose of learning and recognition.

Simulation of human motion has been previously studied mainly from the perspective of graphical animation [6,7,8] and qualitative and quantitative biomechanical studies [12,13] with a focus on individual actions and small part of the body. For instance, the commercial software *Anybody* [14] and *Jack Software* [15] are used for biomechanical and ergonomic study and design of products. The software enables users to position biomechanically accurate digital humans of

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 314–322, 2007.

various sizes in virtual environments, assign them tasks and analyze their performance. Both softwares employ complex kinetics analysis and empirical models [16,17,18] and are able to simulate individual actions in a natural and realistic manner. However, they are not designed for generating large amount of motion data, especially the variations of individual actions.

This paper presents a method to simulate gross human body motion with the intention to establish a human motion corpus for learning and recognition. We propose to simulate human motion by hierarchically decomposing it into actions, joint actions and primitive actions, referred to as *actionlets*. The actionlet models the primitive moving phase of a joint and represents the muscle movement governed by kinesiological principles. Joint actions and body actions are constructed from actionlets through constrained concatenation and synchronization.

The rest of the paper is organized as follow. Section 2 presents the kinematic model and the hierarchical decomposition of human motion into the primitive movement, actionlets. Section 3 describes how a complex human motion can be generally simulated through hierarchical concatenation and synchronization of the actionlets. An approach to establish a motion corpus based on the proposed simulation scheme is given in Section 4. Section 5 demonstrates some simulation reults and conluding remarks are made in Section 6.

2 Kinematic Model and Motion Decomposition

Human body is often viewed as an articulated system of rigid links or segments connected by joints. Fig. 1(a) shows a kinematic model with 15-joint, 22-DOFs (Degree Of Freedom). The model is tailored to represent the postures of human body, which is descriptive enough for gross human motion. A detailed model may be adopted with more DOFs and segments such as hands and feet.

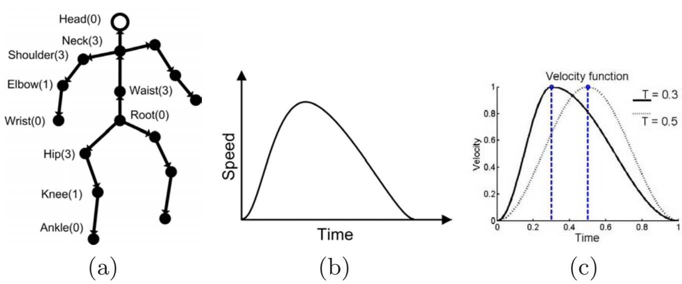


Fig. 1. The adopted kinematic model and velocity curves. (a) Kinematic model and main joints, (b) A typical velocity curve based on kinesiological study [19], (c) Two simulated velocity curves using a concatenation of two half cycle *sine* function (Eq. 1).

Human motion can be considered as a continuous change of his/her body posture and, therefore, represented as a sequence of temporally sampled postures. A posture at a given time is decided by the spatial configuration of the segments of the human body. The co-existance of spatial and temporal nature has made human motion much more complicated to describe and analyse than speech. In [20], we proposed to decompose *human motion* into a series of temporally concatenated actions (temporal decomposition). Each action is formed by a set of coordinated *joint actions* (spatial decomposition). A joint action is further decomposed into a sequence/sequences of *actionlets*(temporal decomposition), where each actionlet represents a cycle of acceleration and deceleration of a joint according to kinesiological study [19,21]. Fig. 1(b) shows a typical velocity curve [19] of an actionlet. This curve may be emulated by the following function

$$v(t) = \begin{cases} \frac{1}{2}(1 + \sin(\frac{\pi}{T}(t - \frac{T}{2}))); & 0 \leq t \leq T \\ \frac{1}{2}(1 + \sin(\frac{\pi}{1-T}(t - (T - \frac{1-T}{2})))); & T \leq t \leq 1, \end{cases} \tag{1}$$

where T is a normalized variable that controls the shape of the function. Fig. 1(c) shows the emulated curves with different $T = 0.3$ and $T = 0.5$ respectively.

In order to represent the posture in a way that is independent of the body size [20], Euler angle is adopted in our research to represent the spatial configuration of the joints of the kenematic model. Let $\boldsymbol{Y} = (y_0, ..., y_{21})$, where $y_0, ..., y_{21}$ are all Euler angles corresponding to 22 DOFs. An action that happens between the period, $t \in [T_{begin}, T_{end}]$, can be represented as

$$\boldsymbol{Y(t)} = (y_0(t), ..., y_{21}(t)); t \in [T_{begin},\ T_{end}]. \tag{2}$$

where $y_i(t), t \in [T_{begin}, T_{end}]$ represents the time series of the Euler angles of the $i'th$ DOF. $y_i(t)$ may consists of N_i cycle of accelarion and deceleration or silence (static periods). In other words, $y_i(t)$ is a concatenation of N_i actionlets. Therefore,

$$Y(t) = \begin{cases} concat(f_{[0,0]}(\tau_{0,0}), f_{[0,1]}(\tau_{0,1}), \cdots, f_{[0,N_0-1]}(\tau_{0,N_0-1})) \\ \vdots \\ concat(f_{[21,0]}(\tau_{21,0}), f_{[21,1]}(\tau_{21,1}), \cdots, f_{[21,N_{21}-1]}(\tau_{21,N_{21}-1})) \end{cases} \tag{3}$$

where $concat(\cdot)$ represents concatenation, and N_i represents the number of actionlet in the $i'th$ DOF. $f_{[i,j]}(.)$ is the Euler angle series of the $j'th$ actionlets in the $i'th$ DOF, which can be specified by users or obtained from motion capture data. $\tau_{i,j}$ is the time variable corresponding to each $f_{[i,j]}(.)$. The velocity, temporal changing rate of the Eulaer angle, of any actionlet follows the basic principle of muscle movement as shown in Fig. 1(b,c).

Let $v(t)$ be a velocity function of any $y_i(t)$ that consists of one actionlet. The condition of actionlet is

$$v(t_{begin}) = v(t_{end}) = 0 \tag{4}$$

$$v(t) \neq 0, t_{begin} \leq t < t_{end}, \tag{5}$$

where t_{end} is the exact next zero velocity point to the t_{begin}, so that any kind of action concatenation can be abstracted into concatenation between actionlets.

According to [22], any muscle acts in cycles of contracting, keeping still, strectching and keeping still. In the angular velocity space, it means a cycle of acceleration and deceleration. Soderberg [19] gave a typical angular velocity curve cooresponding to one joint action (Fig. 1(b)) that can be approximated by half cycle of *sine* function as shown in Fig. 1(c).

3 Simulation

With the proposed decomposition, simulation of human motion amounts to the composition of a sequence of actions using actionlets. Any actionlet has to satisfy all possible physical constraints which the corresponding joint may be subject to. The two major composition operations that are required to build a specific type of actions (e.g. running) from actionlets are *concatenation* and *synchronization*. The reality of the simulated motion depends on how well these operations perform.

Concatenation is employed to build a joint action from actionlets or to form a motion from a sequence of actions. We apply the continuity of the Euler angles and their first order derivatives (velocity) to all DOFs. When two actionlets or actions do not meet the concatenation criteria, either the two actionlets or actions will be blended or a transition actionlet or action is inserted between them.

Considering two actionlet sequences SA and TA, where SA has two actionlets $sa1$ and $sa2$, TA has one actoinlet $ta1$. Suppose $[t_{sa1}, t_{sa2}]$ is the time scope of $sa1$, $[t_{sa2}, t_{sa3}]$ is the time scope of $sa2$, and $[t_{ta1}, t_{ta2}]$ is the time scope of $ta1$. The concatenation between SA and TA can be generally represented as an insert operation, i.e. to 'cut' a piece of action $[t_{begin}, t_{end}]$ from source action SA and 'paste' it into target action TA at the time t_{target}, where $t_{sa1} < t_{begin} < t_{sa2} < t_{end} < t_{sa3}$ and $t_{ta1} < t_{target} < t_{ta2}$. The result action has to be maintained smooth.

Suppose v_a is the velocity function corresponding to each actionlet, a. The process of the insertion operation is summaried in table 1 and table 2, where the target actionlet is changed to two actionlets $ta1'$ and $ta2'$, and

$$t'_{ta1} = t_{ta1},$$

$$t'_{ta2} = t_{target} + (t_{sa2} - t_{begin}),$$

$$t'_{ta3} = t_{ta2} + (t_{end} - t_{begin}).$$

Here, the source actionlet is 'pasted' into the target actionlet. Similar process is employed when there are more than one actionlet in the time scope $[t_{begin}, t_{end}]$. This algorithm can maintain the smoothness of target's anguler displacement because the velocity values at the ends of target actionlets keep zero.

Synchronization is required when an action involves a number of joints. The joints have to act in a coordinated manner in order for the action to be realistic and meaningful. A typical example is "walking" where the arms and legs and the two joints (keen and hip) on the legs have to be synchronised properly.

Table 1. The source and target actionlets before concatenation

Actionlet	Begin time	End time	Velocity function
$sa1$	t_{sa1}	t_{sa2}	v_{sa1}
$sa2$	t_{sa2}	t_{sa3}	v_{sa2}
$ta1$	t_{ta1}	t_{ta2}	v_{ta1}

Table 2. The source and target actionlets after concatenation

Actionlet	Begin time	End time	Velocity function
$ta1'$	t'_{ta1}	t'_{ta2}	v_{ta1}
$ta2'$	t'_{ta2}	t'_{ta3}	v_{sa2}

Let $\wp_g(t)$ be a joint action group, then

$$\wp_g(t) \equiv \{y_k(t) \,| k \in \Im\} \tag{6}$$

where $\Im$ is a set of valid joints, and $y_k(t)$ corresponding to the $k'th$ DOF showed in Eq. 2.

Assume $\wp_p(t)$ and $\wp_q(t)$ are two groups of joint actions, we consider both $\wp_p(t)$ and $\wp_q$ as periodic functions, the synchronization between them can be represented as

$$\wp_q(t) = \wp_p(sh * t + \phi) \tag{7}$$

where sh denotes the temporal scale coefficient and ϕ is a phase variable.

4 Action Corpus

The major objective of the proposed simulation is to provide an easy and low-cost scheme to generate enough quality motion data, action corpus, for learning and recognition. With the hierarical decomposition of the human motion into the primitive actionlets, the process to establish such an action corpus becomes to define a set of resuable actionlets and construct actions and sequence of actions from the actionlets. This is a reverse process of decomposition as illustrated in Fig. 4. To make the construcion process easy, we introduce a number of intermediate reusable components: *joint actions*, *limb actions* and *joint group action.*

1. *Joint Action.* In the kinemati model as showed in Fig. 1(a), joints may have more than 1 DOF. *Joint action* is defined as a group of coupled DOF functions $y_i(t)$ within a specified time scope for a particular joint. For example, the joint action on the left shoulder joint includes $y_3(t)$, $y_4(t)$ and $y_5(t)$ and its dimentionality is 3.

2. *Limb Action.* We define a *limb action* as a group of two or more phisically connected joint actions in a specified time scope. For example, the left arm includes left elbow and left shoulder and the limb action of the left arm includes $y_3(t)$, $y_4(t)$, $y_5(t)$ and $y_7(t)$, where $y_7(t)$ corresponds to the left elbow's DOF function. This limb action is 4 dimension.
3. *Joint Group Action.* A joint group action is defined as a group of any coombination of joint actions and limb actions. In other words, it could be any combination of DOF functions within $y_i(t)$, where $0 \leq i \leq 21$, for there are 22 DOF functions.

The definition of *joint actions*, *limb actions* and *joint group actions* connects the real world's description of human actions and the underliying mathematical approximation models, the action models and the actionlet models. They provide us with a set of flexible and resuable *action units* to describe and simulate a variety of human motions. With a small number of resualble action units together with the concatenation and synchrononization methods described above, virtually, unlimited types of human motion can be simulated.

Due to the demand for large number of samples of a particular action for learning, we introduce a *mutation* process in the simulation to generate variations of a particular action. The mutation process eliminates the need for specifying an action for each sample and allows us to easily generate samples with variations for the same action. There are two ways to mutate the samples. One is to choose same type of actionlets but with different velocity functions from the actionlet database. Another is to perturb the parameters of the velocity functions as shown in Fig 1(c).

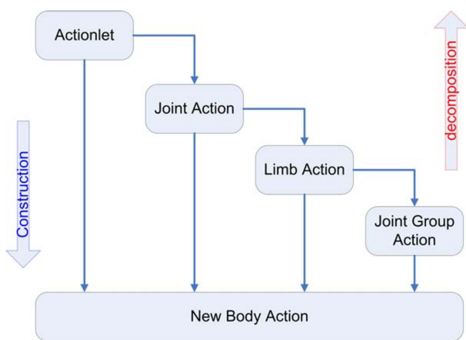


Fig. 2. Decomposition and Construction

5 Results and Visualization

A simulation system has been implemenetd with a number of essential motion editing functions and a set of parametrized actionlets. Humanoid animation (H-Anim), part of W3D/VRML ISO standard, is adoped for visualization to

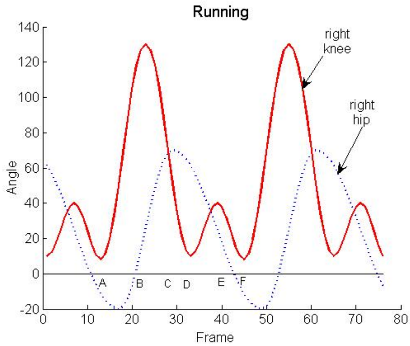


Fig. 3. Angle curve of two joints for running

check the reality of the simulated actions. Fig. 4 shows a simulated 'running' action. Fig. 5 shows the angle changes of the two joints: right knee and right hip. Corresponding to the six indicated positions, six static postures were cut from a 3-D browser as shown in Fig. 4.

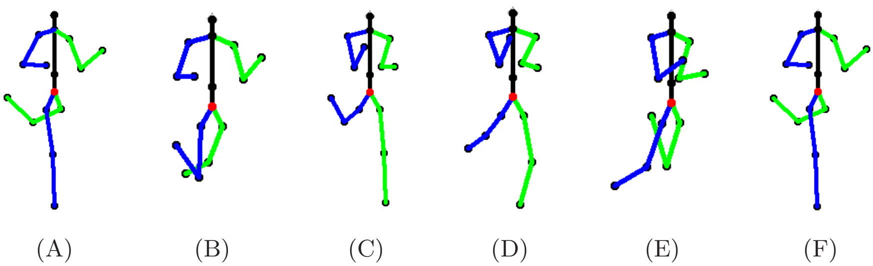
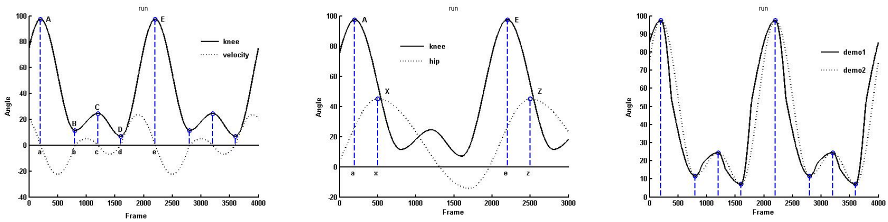


Fig. 4. Running action. A, B, C, D, E, and F corresponding to the six positions marked in Fig. 5

Fig. 5(a) shows one *joint action* (knee) while running at the temporal resolution of 3000 frames per second. The action was composed of two types of actionlets AB (or CD) and BC (or DE). Fig. 5(b) shows the synchronisation between joint "knee" and "hip" for running. *Actionlets* AB, BC, CD, DE construct the *joint action* AE for "knee" and *joint action* XZ for hip. There is a phase $|A - X|$ difference between these two *joint actions*, but the two *joint actions* have the same period, that is $|E - A| = |Z - X|$. All *action unit* mentioned above are stored in one database that can be reused to generate other actions. After the mutation process, a large variety of the 'running' action were generated. Fig. 5(c) shows the mutation of the two joints in action 'running'.



(a) (b) (c)

Fig. 5. Simulation of action "running". (a) Two types of actionlets for joint "knee" (AB and BC). (b) Synchronization of "knee"' and "hip" (c) Mutation result of "knee", actionlets were modelled by Eq. 1 with $T = 0.3$ (solid line) and $T = 0.5$ (dotted line).

6 Conclusion

In this paper, we improved the action model proposed in [20], and resolved the concatenation issue in the simulation process. Using the method, two or more actions can be concatenated together without complicated kinematic analysis, and keep good nature-looking. The system is based on novel scheme to decompose human motion into actions, joint actions and actionlets. This decomposition makes feasible a large scale simulation of human body motion in which not only a virtually unlimited types of actions can be simulated from a finite set of actionlets, but also an unlimited variations of one action can be generated.

References

1. Moeslund, T.B., Granum, E.: A survey of computer vision-based human motion capture. Computer Vision and Image Understanding 81, 231–268 (2001)
2. Moeslund, T.B., Hilton, A., Kruger, V.: A survey of advanced in vision-based human motion capture and analysis. Computer Vision and Image Understanding 104, 90–126 (2006)
3. Hu, W., Tan, T., Wang, L., Maybank, S.: A survey on visual surveillance of object motion and behaviors. IEEE Transactions on Systems, Man. and Cybernetics, Part. C 34(3), 334–352 (2004)
4. Zhang, X., Liu, Y., Huang, T.S.: Motion analysis of articulated objects from monocular images. IEEE Transactions on Pattern Analysis and Machine Intelligence 28(4), 625–636 (2006)
5. Green, R.D., Guan, L.: Quantifying and recognizing human movement patterns from monocular video images–Part I: A new framework for modelling human motion. IEEE Trans Circuits and Systems for Video Technology 14(2), 179–190 (2004)
6. Boulic, R., Magnenat-Thalmann, N., Thalmann, D.: A global human walking model with real-time kinematic personification. International Journal of Computer Graphics 6(6), 344–358 (1990)
7. Fang, A.C., Pollard, N.S.: Efficient synthesis of physically valid human motion. Proceedings of ACM Transactions on Graphics 22(3), 417–426 (2003)

8. Xiao, Z., Yang, X., Zhang, J.J.: Motion data correction and extrapolation using physical constraints. IEEE Proceedings of the 9th International Conference on Information Visualisation (IV 05) 00, 571–576 (2005)
9. Faraway, J.: Human animation using nonparametric regression. Department of Statistics University of Michigan, Tech. Rep. 382 (2002)
10. Ijspeert, A., Nakanishi, J., Schaal, S.: Movement imitation with nonlinear dynamical systems in humanoid robots. In: Proceedings of the IEEE International Conference on Robotics and Automation, Washington, DC, May 2002, vol. 2 , pp. 1398–1403 (2002)
11. Pettre, J., Simeon, T., Laumond, J.: Planning human walk in virtual environments. IEEE / RSJ International Conference on Intelligent Robots and Systems (IROS'02) vol. 3, pp. 3048–3053 (September 2002)
12. Chang, C.-C., Brownb, D.R., Bloswickb, D.S., Hsiang, S.M.: Biomechanical simulation of manual lifting using spacetime optimization. Journal of Biomechanics 34(4), 527–532 (2001)
13. Park, W., Chaffin, D.B., Martin, B.J., Faraway, J.J.: A computer algorithm for representing spatial-temporal structure of human motion and a motion generalization method. Journal of Biomechanics 33, 2321–2329 (2005)
14. [Online]. Available: `http://www.ugs.com/products/tecnomatix/human_performance/jack/`
15. [Online] Available: `http://www.anybodytech.com`
16. Wrotek, P., Jenkins, O.C., McGuire, M.: Dynamo: dynamic, data-driven character control with adjustable balance. In: sandbox '06: Proceedings of the 2006 ACM SIGGRAPH symposium on Videogames, pp. 61–70. ACM Press, New York, USA (2006)
17. Kry, P.G., Pai, D.K.: Interaction capture and synthesis. In: SIGGRAPH '06: ACM SIGGRAPH 2006 Papers, pp. 872–880. ACM Press, New York, USA (2006)
18. Xia, S., Qiu, X., Wang, Z.: A novel framework for athlete training based on interactive motion editing and silhouette analysis. In: VRST '05: Proceedings of the ACM symposium on Virtual reality software and technology, pp. 56–58. ACM Press, New York, USA (2005)
19. Soderberg, G.L.: Kinesiology application to pathological motion, 1st edn. Williams & Wilkins (1986)
20. Zheng, G., Li, W., Ogunbona, P., Dong, L., Kharitonenko, I.: Simulation of human motion for learning and recognition. In: AI 2006. LNCS (LNAI), vol. 4304, pp. 1168–1172. Springer, Heidelberg (2006)
21. Tözeren, A.: Human Body Dynamics. Springer, Heidelberg (1999)
22. Zatsiorsky, V.M.: Kinematics of Human Motion, 1st edn. Human Kinetics (1998)

This page intentionally blank

Part II

Building and Applying Virtual Humans

User Experience Quality: A Conceptual Framework for Goal Setting and Measurement

Russell Beauregard and Philip Corriveau

Intel Corporation, Channel Platforms Group
5200 NE Elam Young Pkwy
Hillsboro, OR 97124, USA
{russell.beauregard,philip.j.corriveau}@intel.com

Abstract. Although the term 'user experience' has become ubiquitous, variations in its conceptualization can make design objectives unclear. This paper proposes a simple framework for conceptualizing the components of user experience in order to communicate with UX stakeholders and advance goal setting and measurement in applied settings. A deeper understanding of the components of experience provide a greater ability to set strategic direction for the user experience, guide design goals, and assess user experience outcomes. In educating stakeholders on a more complete view of user experience, UCD practitioners have the opportunity to play a key role in planning the level of user experience quality for the product user experience and influencing where user experience studies will have the most impact on products.

Keywords: UX, user experience assessment, experience quality, consumer experience, perceptual quality, customer emotions.

1 Introduction

Over the past decade, the term 'user experience' has become a buzzword within industry and in the HCI community (Jordan, 2002; Khalid & Helander, 2006; Hassenzahl & Tractinsky, 2006). From a professional practice perspective, user experience (UX) is often used to represent a wider approach than usability by going beyond usability to include aesthetics, hedonics, contextual, and temporal variables (Forlizzi & Battarbee, 2004). From a business perspective, having a UX strategy is increasingly being recognized as a means of controlling an aspect of the product or service value proposition. Delivering a better quality UX may be part of the differentiation strategy or potentially required to keep up with the UX offered by competitors. In competitive markets, differentiating on UX has potential to win new customers, increase market share over competition, or be used to relieve pricing pressures associated with technology commoditization.

2 Theory

Recent theoretical models of UX show that it is a complex construct which is why it is difficult to define succinctly (Swallow, Blythe, & Wright, 2005). Difficulty to

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 325–332, 2007.

define UX may be partially due to artificial dichotomies being drawn between usability constructs (such as simplicity and task effectiveness) and aspects of UX design (such as fun, joy, aesthetics, emotions…etc). These are sometimes framed as incompatible. Take for example game development. Enjoyable games must have challenge and even complexity in order to maintain interest (see Hassenzahl, Beu, and Burmerster, 2001) while a traditional usability approach is more concerned with making tasks easier or more efficient for users. This notion is supported by traditional definitions such as usability defined by ISO 9241-11, “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.” Usability defined this way lacks the kind of psychological constructs that are typically highlighted in discussions of UX. Differences in perspective among product designers and/or other product decision makers may lead to conflicts in their approach to design and/or less clear design goals and strategies.

In an attempt to help reconcile views between the roles of usability and UX, Jordan (2002) proposed a needs abstraction model useful for high level conceptualization of UX. Jordan suggested that products should engage people at three separate levels: functionality; usability, and UX. At the lowest level in the hierarchy is functionality which is basically what the product can do. To meet functionality needs, the product must reliably do what it claims to do at a basic level. Next, usability is how easily the target user segment interacts with the product to complete their activities in a give context. Users should find it intuitive and easy to accomplish intended their objectives using the product. At the top of the hierarchy is UX. UX is defined in terms by how the user relates to the product and higher level aspirations associated with use over time. Since each of these areas are conceptually distinct it is natural that each area of investigation requires its own set of methodologies that may be highly distinct from each other (McNamara & Kirakowski, 2005). Research questions and methodologies appropriate for usability, therefore may not be the same methods appropriate for UX.

In the HCI and emerging UX literature base, there are still few conceptual models that support practical UX goal setting. Given the restricted scope of usability approaches, these are not typically sufficient for larger UX concerns. Although usability goal setting and assessment methods are often critically important, they do not typically address higher level concerns that have become widely recognized as part of the UX literature. Since UX has been an umbrella term for many constructs, part of the challenge is to define and model UX in an intuitive way that supports goal setting and measurement.

The conceptual framework proposed in the next section of this paper highlights the importance of several key psychological constructs applicable to a wide range of products. This was designed to be easy to communicate and simple enough for use to educate high level decision makers within organizations and at the same time draw the appropriate attention to relevant constructs for the purpose of design goal setting and validation of a products UX quality. This has been successfully used to provide a basis for UX goal setting and assessment exercises (Beauregard, Younkin, Corriveau, Doherty, and Salskov, 2007).

3 Interaction Based Framework of UX

UX is defined here simply as the emotions, attitudes, thoughts, behaviors, and perceptions of users across the usage lifecycle. Figure 1 shows the components of UX and suggests a conceptual relationship between these constructs. UX arises from an interaction between a user and product within a context. This framework shows the psychological nature of UX by highlighting that many of the key categories of UX constructs are cognitive in nature. As such, UX is mainly accessible through self report, behavioral observation, and other proxies (e.g., physiological proxies) of cognitive processes. The construct categories (see detailed definitions of perceptions, emotions, thoughts, and attitudes below) are suggestive of useful distinctions for both design and potential tools for assessment.

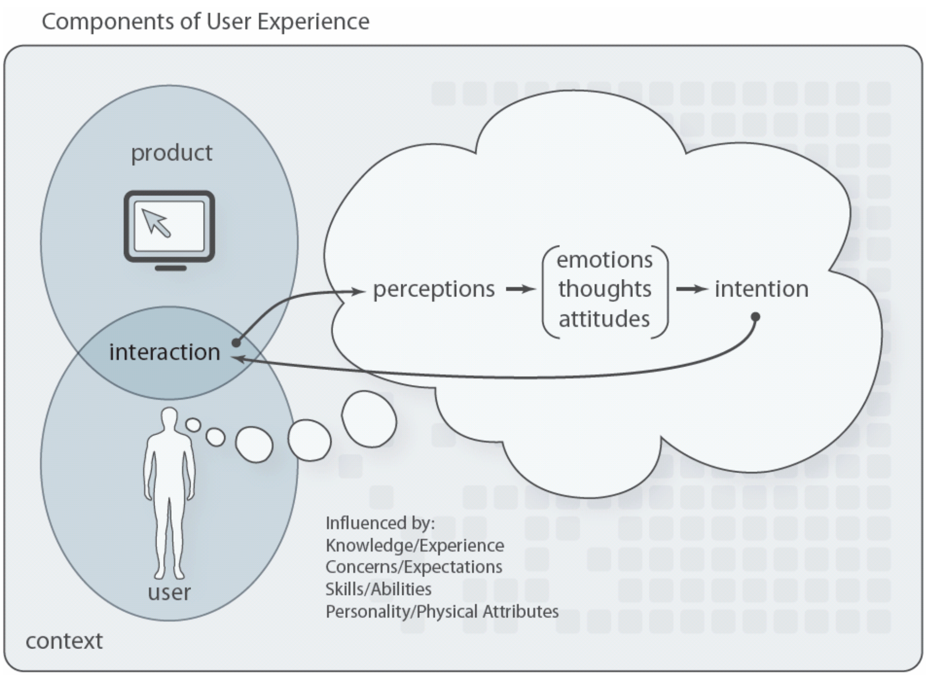


Fig. 1. Conceptual Framework

Essentially, this framework suggests that an interaction with a product or concept (perceived or noticed through one or more of the five senses) gives rise to emotions, thoughts, and attitudes. These, in turn, influence intentions and interactions (including observable behavior) with the product across time. How the interaction unfolds and each component of the UX itself is influenced by a larger set of individual differences including knowledge/experience, concerns/expectations, skills/abilities, personality and physical attributes. This cycle all occurs within a larger context, the surrounding backdrop that is made up of other cultural, social, and technological influences on experience.

Defining UX in terms of key psychological constructs provides a way to talk about the different components, set design goals, and create a strategy regarding measurement. The supporting descriptions below, based on the psychology and HCI literatures, are for the purposes of drawing useful distinctions between these terms as they relate to UX.

Perceptions: Perception is the process of acquiring and interpreting sensory information. End-user perceptual experience is an often ignored aspect of UX assessment. Focusing on intake and subjective assessment, psycho-visual and psycho-acoustics studies assess human perceptual variables related to product use such as such as video quality, audio quality, acoustical and thermal performance. Perceptual targets must comprehend that user expectations change over time as technology changes. For example, as consumers shift from standard-definition to high-definition television resolutions, the anticipation of high picture quality increases and this must be reflected in perceptual requirements. See Beauregard et al. (2007) and Corriveau (2006) for examples of applied perceptual validation techniques as an aspect of UX.

Emotions: Emotions are feelings or subjective states of consciousness that are critical to learning, trust, and assessment of what's desirable. Products may evoke positive or negative feelings which affect purchasing behavior, how much the technology is used, and what consumers say about the product to others. Despite the assumption that people make decisions logically, research shows that decisions are highly dependent on the emotional states of end users (Schwartz, 2004). Both quantitative and qualitative assessment techniques exist to understand the potential impact of emotions and factor this into planning and design. From a quantitative approach, Desmet (2003) describes several broad emotion categories and has created measures to assess specific emotions related to products. In applied settings, semantic differentials and likert-type questionnaires are often used for targeted emotions. Examples of commonly targeted emotions include pleasant surprise, desire, fascination, interest, and joy. Degree of negative emotional states such as boredom, disgust, frustration, sadness, fear, and anger are also often studied. In addition to quantitative techniques, case studies provide opportunities for rich qualitative data that can provide a deep and unique understanding of context (Strauss & Corbin, 1990; Hassenzahl, Beu, and Burmester, 2001).

Attitudes: Attitudes are judgments toward a target typically associated with value, good/bad, or helpful/harmful. Attitudes are a function of expectations and past experiences (Hassenzahl & Tractinsky, 2006). Examples of attitudes, showing the importance of these measures for UX, include satisfaction, perceived value, judged comparisons of products or past experiences, judgments of the degree of control over technology. Additional research is needed in this area particularly addressing known problems. For valid assessment, careful design of the study methodology is needed if results are to be generalized. Often interviews geared to assess attitudes have quantitative components (e.g., likert type scales), although the most important information is from deeper contextual interviewing techniques. Attitude measurements such as commonly used in usability studies or surveys have questionable validity when attributes are not entirely clear (surveys) or when the situation or contexts are not representative (lab studies or context simulations). Assessments of attitudes are also particularly sensitive to distortions in situations where there are non-realistic consequences to user behavior. Given these limitations and the importance the role of attitudes in UX, methodology research is needed in this area. Specifically, using the framework of UX as a guide, use of ethnographic and contextual inquiry methods, such as the "grounded theory technique" (case study described by Swallow, Blythe, & Wright, 2005) may provide useful data to better triangulate attitudes variables measured by more quantitative or hybrid studies.

Thoughts: Thoughts are mental processes that allows humans to model what they experience and to plan behavior. Understanding the mental models held by users before, during, and after interaction with a product can provide clear design opportunities and highlight roadblocks across stages of the use the usage lifecycle. Product goals can explicitly target and assess proper alignment between the user mental models and the mental models afforded by contact with the product.

Overall, each of the high level components of UX can provide useful categories in assessing needs, developing designs, and assessing the UX itself. These are proposed as closely interrelated threads. The UX directly influences behavior including further interaction with the product and context and continually loops back to the interaction to further effect perceptions, emotions, thoughts, attitudes, and behavior. As the "usage lifecycle" part of the UX definition suggests, this feedback process continues throughout the full usage lifecycle starting from initial discovery through purchase, out-of-box, usage, maintenance, upgrades, and disposal (e.g., Vredenburg, 2002). In this way, UX not only extends a temporal thread (anticipated experience through reflected experience over time) but also highlights the necessity to coordinate UX efforts across organizational functions to include everything the user sees, touches, and hears regarding the product, service, and/or brand.

The ability to accurately assess many of these aspects of UX for business purposes has been steadily advancing, but overall, is still in early stages. The proposed UX framework shows a high level structure for constructs of interest. Specific constructs within these categories will vary according to the product or service being assessed. A deeper understanding of what UX means provides a greater ability to set strategic direction for the UX, guide design goals, and assess UX outcomes. It is suggested here that setting UX goals on these or a subset of variables begins the right conversations that can link the strategy of the top decision makers directly to the project team and designers.

4 Setting UX Goals

Leading companies that consider UX as a core part of their business have been using UX quality measures as formal decision gates in their development processes. Based on a series of informal benchmarking interviews, companies including IBM, Microsoft, British Telecom, Google, and Yahoo each use some form of UX quality assessment data as part of ongoing assessments and go/no go decisions regarding product release. These and other companies, including Philips, Ford, and Proctor &

Gamble have indicated that they routinely assess the UX of products throughout the development process. The methods used for assessing UX in these companies tend to be part of a larger user-centered innovation effort.

From a goal setting and measurement perspective, the components of UX can be scoped and prioritized according to the type of product, comparative usages, target user segment characteristics, context of use, and business goals. These factors may each influence what components of UX are to be highlighted. For instance, given the context of use and the competitive landscape, in some cases (such as entertainment product development), emotional appeal will have a higher priority than in other cases such as in business applications. As relative priorities are set regarding the product vision, business strategy, and the UX constructs to be targeted, specific goals can be set.

Setting UX quality requirements should occur early in the product planning stage and well before detailed usability requirements or detailed use case development. It is important set these intended levels of UX quality early so that checkpoints can be inserted to ensure the final product will elicit the intended perceptions, emotions, thoughts, and attitudes from the target market segments. UX quality goals communicate both to the development teams and to top management the clear targets regarding how good the UX associated with the usages must be. Here, we describe a brief outline of three broad steps in setting UX quality goals.

The first step in setting UX quality goals involves identification and prioritization of the relevant UX dimensions. Market research, needs-finding processes (such as market segmentation, ethnographic research, and usage model definition) define the nature of the product opportunity (Anderson, Bramlett, Gao, & Marsh 2007). These need to be described and rank ordered in terms of the vision for the high-level key features and usages. Particular attention and priority should be given to the features and usages that are end-user noticeable, will be included in the marketing messaging, and differentiate the system from others that will be on the market. In addition, any usages involving perceptual quality (such as acoustics, video quality, audio quality, or even tactile thermal properties) can be called out as relevant according to the end-user value propositions being targeted.

The second step is targeting specific (and measurable) UX dimensions for each of the key usages and features. The UX conceptual framework provides guidance regarding the categories of constructs to target. This involves assessing what emotions, attitudes, perceptions, and thoughts are being targeted for the planned end-user value propositions. Selecting the proper dimensions to target and how to best measure them is where a background in psychological assessment, psychometrics, and/or ethnographic methods is essential. The variables, study design, and measurement techniques selected should be based on branding/marketing strategies as well as practical and experimental design considerations.

The third step is working with the UX owners to assign guidelines and cutoffs (quantitative or conceptual guidelines) for each of the key features with respect to the variables being measured. To do this, competitive analysis benchmarking data or prior baseline data can be used. If no prior UX data are available, then judgment based on experience with similar types of systems can be used to start with. The main objective is to set explicit goals for UX quality well in advance of product development so that these goals can serve as clear targets and bring appropriate attention to the UX throughout the development cycle. By highlighting what should be the UX outcomes to development teams and the accountable stakeholders, strategies and resources can be channeled to ensure user-centered design processes are prioritized appropriately with other business demands.

After goals have been set, measurements to assess the state of the UX can be planned. We refer to these checkpoints as critical UX milestones in the program timeline. At these milestones, decision makers and sponsoring executives can now have UX data to weigh tradeoffs that may affect both the future UX plans and other business outcomes. Common questions that UX quality assessment can help answer include: How good are the different aspects of UX for the target market? What levels of perceptual qualities will consumers notice and value? How does the end-user value proposition change when ecosystem partnerships or key functionality changes? How will the product compare to other product experiences?

Discussing and setting UX quality goals is particularly relevant in companies undergoing transitions toward establishing user-centered processes. It increases emphasis and visibility of the design function, key usage models being developed, and serves to communicate clearly the priority of UX relative to other objectives.

5 Conclusion

UX has become a hot topic and a key potential differentiator in competitive markets. This paper proposed a simple applied framework for conceptualizing the components of UX in order to communicate with UX stakeholders and advance goal setting and measurement within applied settings. The framework involves psychological constructs of UX that go beyond traditional validation, assessments of usability, and narrow conceptions of UX. The UX framework highlighted the roll of human perceptions, emotions, thoughts, attitudes, and behaviors resulting from the interaction with all aspects of the product over time.

Setting explicit UX goals guided by this framework and setting milestones at key gates in the development lifecycle, helps align team members with the business strategy regarding UX and affords attention in the right places regarding UX objectives. Focusing attention on the UX to management and executive sponsors not only generates appropriate support for design related functions but also helps to get the team pulling in the same direction when tradeoffs need to be made.

UX data collected throughout product development can be used to inform status compared to UX goals and to directly influence current and future product direction. innovation. By increasing visibility through explicit UX goal setting and measurement across stages of development, not only is UX emphasized as an important organizational objective, but strategies and resources can be better channeled to ensure user-centered design processes are prioritized appropriately relative to other business demands. In the future, UX studies based on behavioral research and psychology assessment techniques will go well beyond the classic satisfaction surveys and usability studies. As UX continues to become recognized as critical value propositions for customers, these techniques hold promise to provide deeper insight and better control into the design of successful products.

Acknowledgements. Thanks for the many contributions of our colleagues that made this effort possible. Specifically, we thank Paul Sorenson, Mike Payne, Genevieve Bell, David Sward and Amy Briggs for insightful conversations and input that helped determine the direction of the framework. Additional thanks go to Anne McClard, Cory Booth, Colin Bay, and Delbert Marsh for their expertise in data collection. We are also grateful to Brian Bramlett, Paul Osterlund, and Erik Simmons who have helped define the constructs relative to other efforts within Intel. Finally, we thank Herman D'hooge and Genevieve Bell who continue to successfully sponsor and champion these concepts at an executive level.

References

Anderson, G., Bramlett, B.W., Gao, J., Palmer, R., Marsh, D.: Intel Usage-to-Platform Requirements Process. Intel Technology Journal, 11(1) (2007)

Beauregard, R., Younkin, A., Corriveau, P., Doherty, R., Salskov, E.: User Experience Assessment. Intel Technology Journal, 11(1) (2007)

Corriveau, P.: Digital Video Image Quality and Perceptual Coding. In: Wu, H.R., Rao, K.R. (eds.): Video Quality Testing, pp. 125–153, Taylor & Francis, Boca Raton (2006)

Desmet, P.A.: A Multilayered Model of Product Emotions. The Design Journal 6(2), 4–13 (2003)

Forlizzi, J., Battarbee, K.: Understanding experience in interactive systems. In: Proceedings of the 2004 conference on Designing Interactive Systems: Processes, practices, methods, and techniques 2004, pp. 261–268 (2004)

Hassenzahl, M., Beu, A., Burmester, M.: Engineering Joy. IEEE Software [on ACM Portal] (2001)

Hassenzahl, M., Tractinsky, N.: User experience – a research agenda. Behaviour & Information Technology 25(2), 91–97 (2006)

Jordan, P.W.: Designing Pleasurable Products: An introduction to the new human factors. Taylor & Francis, Philadelphia, PA (2002)

Khalid, H.M., Helander, M.G.: Customer Emotional Needs in Product Design. Concurrent Engineering: Research & Applications 14(3), 197–206, 10p, 2 diagrams (2006)

McNamara, N., Kirakowski, J.: Defining Usability: Quality of Use or Quality of Experience. In: 2005 IEEE International Professional Communication Conference Proceedings, 2005, pp. 200–204 (2005)

Schwartz, N.: Metacognitive Experiences in Consumer Judgment and Decision Making. Journal of Consumer Psychology 14(4), 332–248 (2004)

Strauss, A., Corbin, J.: Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Sage Publications, Newbury park, CA (1990)

Swallow, D., Blythe, M., Wright, P.: Grounding experience: relating theory and method to evaluate the user experience of smartphones. In: ACM International Conference Proceeding Series, 2005, vol. 132, pp. 91–98 (2005)

Vredenburg, K.: Designing the Total User Experience at IBM: An examination of case studies, research findings, and advanced methods. International Journal of Human-Computer Interaction 14(3/4), 275–278 (2002)

Integrating Perception, Cognition and Action for Digital Human Modeling

Daniel W. Carruth, Mark D. Thomas, Bryan Robbins, and Alex Morais

Center for Advanced Vehicular Systems
Mississippi State University
P.O. Box 5405, Mississippi State, MS 39762
{dwc2,mthomas,bryanr,amorais}@cavs.msstate.edu

Abstract. Computational cognitive models are used to validate psychological theories of cognition, to formally describe how complex tasks are performed, and to predict human performance on novel tasks. Most cognitive models have very limited models of how the body interacts with the environment. The present research examines a simple human-machine interaction task and a simple object manipulation task. A cognitive model is integrated with a human avatar within a virtual environment in order to model both tasks.

Keywords: Digital Human Modeling, Cognitive Models, Perception, Cognition, Action.

1 Introduction

Computational modeling and simulation such as CAD/CAE and FEA have become invaluable tools for manufacturing and design. Radical new designs can be virtually constructed and mathematical models of materials and physics can be used to accurately simulate the performance characteristics of these new designs without expending large amounts of time or money for prototyping and testing. Ergonomic tools exist that allow simulation of the human component for analysis. However, these tools currently require significant input from the user and can often generate unlikely results. For each alternative design for a single task, a designer must specify the postures (manually or through simulation) for each step of the task. If the task could be specified more generally and the digital human model could execute the task on multiple designs, analysis of designs could proceed much more quickly.

Cognitive modeling may provide a mechanism that would allow a generic specification of how to perform a task. In addition, cognitive modeling provides an analysis of the role of perceptual, cognitive, and motor factors in the usability of physical designs. The current research integrates a cognitive architecture with a simple, human avatar in a virtual environment and applies it to two basic tasks: human-machine interaction and object manipulation.

1.1 ACT-R Cognitive Architecture

ACT-R is a computational architecture that implements a formal specification of psychological theories of cognition, perception, and action [1]. Within the

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 333–342, 2007.

architecture, individual models of human performance can be constructed to describe or to predict human behavior on many different tasks. ACT-R has been successfully used to investigate psychological phenomenon, to study human-computer interaction, and to develop intelligent tutoring systems.

The ACT-R cognitive architecture is a modular system with a production system as its core. The construction of a model within the architecture involves the specification of condition-action rules that represent the modeler's hypothesis of how the task is performed. The execution of the model involves matching the condition-action rules against the current state of the architecture's modules. Matching rules are then compared based on a utility function and the matching rule with the highest utility is executed updating the other modules. The functions of the other modules include goal-tracking, declarative memory, vision, audition, vocalization, and action. The architecture can be extended by modifying existing modules or creating entirely new modules.

2 Extending ACT-R for Upper-Body Task Modeling

In order to apply ACT-R to real-world tasks such as product assembly, vehicle maintenance, or human-machine interaction, the standard implementation of the cognitive architecture must be extended. The extensions in this research are limited to the perception and motor modules. The perception and motor modules contained in the standard implementation of ACT-R are largely based on the EPIC architecture developed by Meyer and Kieras [4].

2.1 Environment

The standard implementation of ACT-R supports interaction with user interfaces such as command line or GUI systems. This interaction is accomplished by generating information suitable for ACT-R's perception modules from the environment and translating ACT-R's action module's output into messages to the software environment. Since our goal is to model real-world tasks, our environment is a dynamic, virtual environment created in commercial off-the-shelf software. ACT-R is embodied by an animated human avatar within the virtual environment. Every 17 ms, the virtual environment analyzes the scene and generates a table of all of the visual and auditory features that the human avatar may be able to perceive. These tables are submitted to the appropriate modules in ACT-R where they are processed in the same way as the standard implementation processes information from user interfaces. Messages generated from our motor module are sent to the environment which activates stored animations, inverse kinematic simulations, and/or advanced models of human posture prediction.

2.2 Vision

The ACT-R vision module implements a feature-based theory of visual perception. The visual array is made up of features (i.e. color, size, shape, etc.) at particular locations in the array. In order to perceive an object in the environment, visual attention must be shifted to a location in the visual array and the features at that location must be integrated into a representation of the object at that location.

Visual attention is shifted to locations in one of two ways: a voluntary shift to a known location or an involuntary shift to a significant change in the features of the visual array. In ACT-R's vision module, involuntary shifts are made when the model is not currently attending to a visual object and a significant change occurs in the visual array (i.e. a new feature appearing). Voluntary shifts generally require two steps. In the first step, a model requests a search for a specified set of features be performed on the visual array. This very efficient search makes available a location that matches the specified features. Then, the model can shift attention to the location to encode the location's features into a visual object. Encoded visual objects are then available to the production system for processing.

Extensions. The standard implementation of ACT-R does not specifically handle motion as a feature. The vision module was extended to better support dynamic, virtual environments with many moving objects. These extensions impacted three aspects of the vision module: search of the visual array, involuntary shifts of attention, and encoding of movement information. In order to support motion in visual search, motion magnitude and motion direction features can be used to specify a search of the visual array. A model can search the visual array for a fast object moving amongst slow objects by specifying a search for motion magnitude greater than the speed of the slow objects. Alternatively, a model can search for an object moving left amongst objects moving right. If a motion feature changes significantly, then visual attention may be involuntarily drawn to the location of the change. For example, if an object suddenly starts moving, visual attention will involuntarily shift towards it. Motion magnitude and motion direction is then encoded and made available to the production system as part of the visual object.

Spatial Information. In addition to motion, spatial information is needed for interactions with the virtual environment. The exact nature of human spatial encoding, representation, and manipulation is a matter of some debate [6] [7]. In our current model of spatial representation, egocentric spatial relationships can be encoded and stored in declarative memory based on the visual objects encoded by the vision module. An egocentric spatial relationship encodes the observer's bearing to and distance from an object in the visual array. Object-to-object relationships can be encoded based on two egocentric spatial relationships. These relationships can encode a view-independent, allocentric representation of the environment. The current models generate, store, and use both types of representations. Other spatial extensions to ACT-R are currently being investigated by researchers [2] [3].

2.3 Motor

The motor module of ACT-R models the initiation and programming of motor actions. The module also estimates the time required to execute motor actions. The standard ACT-R module contains movement styles primarily suited for human-computer interactions including keyboard presses and mouse movements. Our extensions to the motor module include the addition of movement styles that drive animations for a human avatar within our virtual environment. The movements that our current model supports are walking, pushing buttons, reaching for and manipulating objects.

3 Empirical Data

In order to validate our extensions to the ACT-R vision and action modules and to test our integration with the virtual environment, human performance data was collected for multiple tasks including a human-machine interaction task and an object manipulation task.

3.1 Participants

Data was collected for 12 male participants (Mean age = 19.08 years) recruited from the Mississippi State University student population. Participants were screened for major health and vision problems. Two participants appeared to have below average visual acuity but were allowed to participate in the study.

Participants were asked to wear a Mobile Eye portable eye tracker manufactured by ASL. The Mobile Eye is a lightweight (76 grams) eye tracker that records (30 Hz) to a digital tape recorder worn on the waist. Computer software analyzes the recorded eye and scene videos for analysis. Three participants were dropped from data analysis because of technical difficulties with the eye tracking equipment. The whole-body motion of participants was also captured using an Animazoo Gypsy suit. The whole-body data is not currently relevant to the modeling effort.

3.2 The Vending Machine Task (Human-Machine Interaction)

Many tasks of interest to digital human modelers are industrial tasks that require some interaction between a human operator and a machine. In order to investigate a simple and common interaction between humans and machines, participants were asked to purchase a 20 oz. bottled drink from a vending machine. Each participant was given 10 dimes and directed to the location of the vending machine. At the vending machine, participants deposited the coins and selected a drink of their choice. The participants point of gaze and motions were recorded during the task.

As participants approach the vending machine, their focus is primarily on the coin slot. Similar to results reported by Pelz and Canosa [5], participants appear to look ahead at locations that will be necessary later in the operation of the machine. For example, during approach, participants would glance to the receptacle at the bottom of the machine where they would ultimately retrieve their drink. Also, while reaching from the coin slot to their hand for another dime, participants would often browse the button labels presumably to consider their drink selection. Throughout the task, eye movements (see Figure 1) primarily focused on the coin slot. Participants also verified their deposits by checking the display.

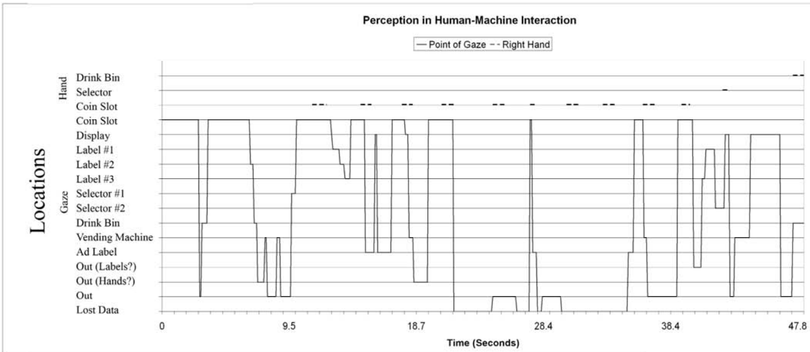


Fig. 1. Example eye and right hand movements

Participants had almost no errors during their interaction with the vending machine. One participant dropped a dime while depositing a coin in the coin slot. During one participant's trial, the vending machine did not accept one of the dimes. The participant used the coin return and successfully completed the task on the second attempt. In the human-machine interaction task, there are two types of errors: human errors and machine errors. In this task, human error could be a physical error (i.e. dropping a dime), a decision making error (i.e. choosing the wrong drink), or a visual encoding error (i.e. not noticing that a drink was available). The human operator also has to respond to machine errors including deposit errors (i.e. a coin getting stuck), out of stock errors, and unresponsive buttons.

3.3 The Block Modeling Task (Object Manipulation)

Many real-world tasks also involve the assembly and/or disassembly of parts. Participants were shown three models constructed from 9 of 14 children's blocks. Figure 2 shows the three models (M1, M2, and M3) and lists the 14 blocks available for constructing the models. Participants were asked to construct a copy of the example models by selecting the correct blocks from the 14 available blocks and placing them in the correct positions.

All participants performed the task quickly and successfully with few errors. Initially, participants assess the overall shape of the example model. Following this first assessment of the model, participants appear to employ one of two strategies. In the single-block strategy, participants select a single piece from the example model, search for a matching piece amongst the available blocks, and then orient and place the block in the appropriate position relative to the already placed blocks.

In the pattern strategy, participants appeared to identify patterns of blocks that make up the example model. The participants would then find all of the blocks making up the pattern and construct the pattern without referencing the example until the pattern was completed. For example, in *M1*, the top three blocks of the model (2 small blocks and the large triangle) could be encoded together and constructed at the beginning of the construction attempt. Once the pattern is completed, it might be set

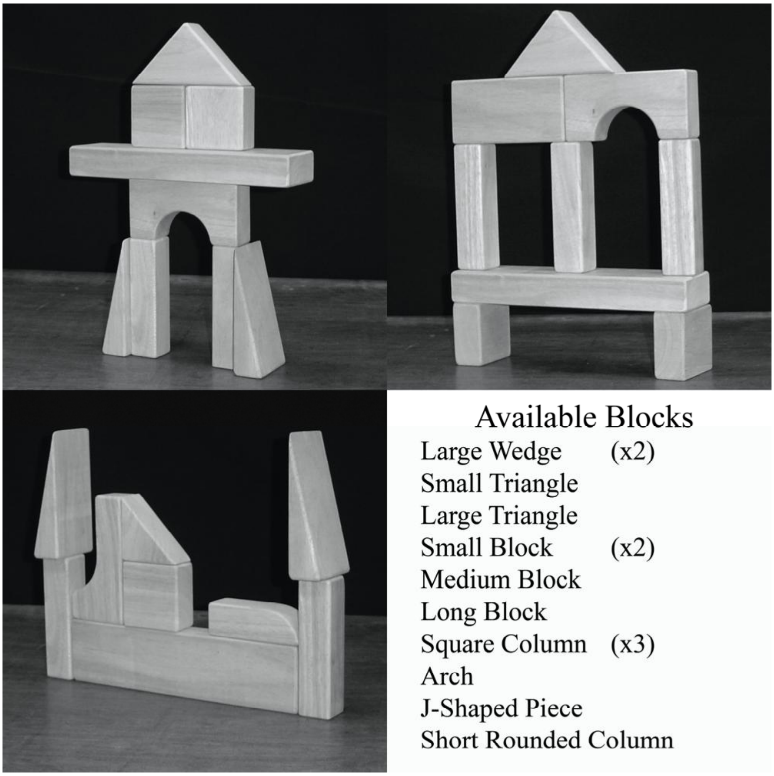


Fig. 2. Three models (*M1* in the top left, *M2* in the top right, and *M3* in the bottom left) constructed from children's blocks. Each model consists of 9 blocks drawn from a total of 14 blocks (listed in the bottom right).

aside until other construction was completed or placed immediately. When placed, participants were observed moving the pattern as a whole (i.e. lifting the three pieces at once) and disassembling and reassembling the pattern in parts.

Typical errors included orientation errors (i.e. in *M3* attempting to place the wedges facing in rather than facing out) or size errors (i.e. in *M1* or *M2* attempting to use the small triangle rather than the large triangle).

4 Modeling the Empirical Data

Based on the results of the study of human participants, a model of task performance was constructed for the two tasks within the ACT-R architecture using our extensions.

4.1 The Vending Machine Task (Human-Machine Interaction)

The vending machine task requires participants to perform a common interaction with a machine. In order to model this interaction, a virtual environment must be constructed with the appropriate objects and a cognitive model must be developed within the selected architecture.

The Environment. The virtual environment was developed within Virtools, commercial off-the-shelf software designed for rapid prototyping of dynamic, 3D environments. Models of the objects relevant to the task were initially created in CATIA then imported into the Virtools environment. The Virtools environment communicates with the ACT-R cognitive architecture via a TCP/IP network connection.

Due to the level of interaction, the environment model of the vending machine requires considerable detail. Each of the components that a participant may interact with was modeled as a distinct object making up the whole of the machine. Each of the modeled objects is tagged with the symbolic information that can be visually encoded. For example, the coin slot, selector, and vending machine body are tagged with their VALUE slot equal to "SLOT", "BUTTON", and "MACHINE." The labels next to the selectors are tagged with their VALUE slot equal to the drink name represented by the label (i.e. "Water," "Lemonade," etc.).

The Model. Once the environment is modeled and imported into the virtual environment, a cognitive model for the specific task can be created. It should be possible to create a single model that would perform both of the current tasks and many other tasks but, for the current research, separate models have been created for simplicity.

Based on the human performance data, four basic tasks were identified: deposit coins, determine a desired drink, select the desired drink, and retrieve the drink. The first two tasks are somewhat independent and are often interleaved during actual human performance. Selecting the desired drink cannot be done until the coins have all been deposited and the desired drink has been identified. Likewise, the drink cannot be retrieved from the drink bin until a selection is made.

4.2 The Model Assembly Task (Object Manipulation)

The object manipulation task requires participants to encode the shape, size, orientation, and spatial relationship of the blocks that make up the model. Participants must then use the stored information to construct a matching model from available parts. In order to model this task, the virtual environment must contain all of the parts that make up the block models.

The Environment. The object manipulation task required 15 virtual object prototypes: the 14 blocks used in the models and the workbench. All of the object prototypes were constructed using CATIA and imported into the Virtools environment. Two instances of the 14 blocks were created in the environment: the blocks available to the digital human (*source pile*) and the blocks used in the example model.

As in the vending machine environment, each object was tagged with symbolic information that could be encoded visually. In this environment, there are two classes of objects. The workbench is encoded using the basic visual-object tag and only basic information is available to the model. This includes location, size (in the visual array), distance, color, motion, etc. The blocks all share a block tag derived from the visual-object tag with additional parameters including: shape, size, and orientation.

When the model attends to one of the blocks, the block's shape, size, and orientation are encoded. For some parts of the model assembly, the model must encode the stability of the block. The model accomplishes this by monitoring the motion values of the basic visual-object tag. If a support block or a top block is moving, the model can perceive the possible instability and attempt to adjust the blocks. This is important for ensuring that bottom blocks successfully support top blocks. The model also encodes spatial relationships between the blocks. This encoding is critical for determining the correct position to place blocks.

The Model. The model for object manipulation is more complicated than the model for human-machine interaction. The object manipulation task includes encoding of block information and spatial relationships between blocks as well as additional movement styles to orient and place the blocks. However, participants tend to perform the block task in a clear, serial fashion without the interleaving of two sub-tasks as seen in the vending machine task.

Based on human performance data, 6 components of the task were identified: model assessment, target selection, search for a block, orientation of a block, placement of a block, and target re-assessment. The model begins the task by finding the example and examining its makeup.

Participants appeared to use one of two strategies for constructing a copy of the example. Some participants worked from the bottom up and selected a single target block. Other participants appeared to identify patterns, then select blocks to build the pattern. Once a pattern was built, participants moved on to the next piece or pattern. In the initial assessment, the actor may identify pattern components or single components. After the assessment, the actor selects one of the identified components as the target. If a pattern is available, the actor may select the pattern even if it is not the base. If a pattern is constructed, the construction of the pattern is completed before moving to another part of the example model. For example, if the 2-blocks-1-triangle pattern at the top of M1 is selected, the actor will construct the pattern without referring back to the example model. Once the pattern is completed, it is treated as a single object to be added to the rest of the actor's construction. If no pattern was identified, the model will select a single base object.

The visual features (i.e. shape, size, etc.) of the target block are encoded to guide the actor's search of the source pile. The actor shifts attention to the source pile and searches for a block matching its internal representation of the target block. As attention shifts to possible matches in the source pile, the features of the currently attended block are compared to the stored features of the target block. Once a target block was found, participants would sometimes refer back to the target block. Multiple interpretations of this reference are possible. Participants could be verifying their selection. They could have failed to encode orientation or position when the target was selected. This target re-assessment can occur after almost every step in the construction process when the actor needs more information. In the pattern strategy, participants may perform this re-assessment less as they may build an entire pattern before looking back to the main model.

Placing a block requires orienting the model correctly (and potentially re-assessing the target) and placing the model correctly (and potentially re-assessing the target). During placement, features such as motion are visually monitored to ensure the stability of the construction.

The model continues in this pattern until the example model has been successfully copied.

5 Discussion

The current research applies an initial integration of cognitive modeling and digital human modeling to two basic tasks. While the tasks currently being modeled are relatively simple, the fundamental cognitive processes may be applied to more complex tasks. For example, the perception and motor components of the operation of complex machinery are not significantly different from those used to accomplish the vending machine task. The operator encodes the features of the machine including display information and operators (buttons/levers/etc.). The operator performs a task based on if-then rules stored in procedural memory. The operator executes motions to act on the machinery. In the assembly task, the cognitive and perceptual components are similar. The motor capabilities in assembly tasks are particularly important and the use of a more effective digital human model than a simple animated avatar is required.

By extending ACT-R's sensory and action systems and integrating the architecture with a virtual environment, the possibilities for cognitive modeling may be greatly expanded. Future extensions may be made to include advanced computational models of human posture and motion. Once validated, these integrated digital human models will be powerful tools allowing designers to investigate both the cognitive and the ergonomic aspects of workspaces and products at the earliest stages of design.

Acknowledgments. This research was performed at the Center for Advanced Vehicular Systems at Mississippi State University.

References

1. Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, D., Lebiere, C., Qin, Y.: An Integrated Theory of the Mind. Psychological Review 111, 1036–1060 (2004)
2. Harrison, A.M., Schunn, C.D.: ACT-R/S: Look Ma, no cognitive map! In: Proceedings of the Fifth International Conference on Cognitive Modeling, Bamberg Germany, 2003, pp. 129–134 (2003)
3. Johnson, T.R., Wang, H., Zhang, J.: An ACT-R model of human object-location memory. In: Proceedings of the 25th Annual Meeting of the Cognitive Science Society, Boston, MA, 2003, p. 1361 (2003)

4. Meyer, D.E., Kieras, D.E.: A Computational Theory of Executive Control Processes and Human Multiple-Task Performance: Part 1. Basic Mechanisms. Psychological Review 104, 3–65 (1997)
5. Pelz, J.B., Canosa, R.: Oculomotor Behavior and Perceptual Strategies. Vision Research 41, 3587–3596 (2001)
6. Shelton, A.L., McNamara, T.P.: Systems of spatial reference in human memory. Cognitive Psychology 43, 274–301 (2001)
7. Tversky, B.: Structures of mental spaces: How people think about space. Environment and Behavior 35, 66–80 (2003)

The Usability of Metaphors with Different Degree of Abstract in Interface Design

Ming-Chuen Chuang and Inglen Lo

National Chiao-Tung University, Institute of Applied Arts,
Taiwan
cming@cc.nctu.edu.tw, inglenlo.aa89g@.nctu.edu.tw

Abstract. Recently, more and more devices everywhere are getting "smarter" with a multi-modal hierarchical menu and form interface. One of the main points of the menu or interface design is to provide users with ease-to-use operation environment. This make them not only learn efficiently but also feel fun (interested) in the process of learning. However, there is no one concept of design suit everyone because the needs and purposes of users are much different from individuals. To satisfy them, the varied design concepts have been suggested to fit for their distinct perceptions and experiences. Consequently, new assessment, called usability, is also required to estimate whether the design concepts are good or not. Therefore, this study attempted to investigate into the usability of 3D interface design. For that, 3 types of main menu of the mobile phone's interface metaphor design were developed as stimuli with different degree of abstract in this study. Then, a four-phase experiment was conducted to explore the usability evaluation of 3 types of metaphorical interface design with different degree of abstract, including: (1) to investigate users' opinions on a mobile phone's interface design; (2) to verify whether the simulated graphics and interactions corresponding to the metaphors intended (pilot study); (3) to measure the usability of 3 types of metaphorical interface design simulated in this study; (4) to compare the preference for any one of the 3 types of metaphorical interface design. The experimental procedures and the results of the analysis would be interpreted respectively according to different phases. Additionally, the degree of abstract in the metaphorical interface design was defined by the average ratings in phase 1: metaphor 3 were regarded as abstract interface design and metaphor 1 and metaphor 2 were regarded as concrete interface designs, but the degree of concrete in metaphor 1 was stronger than in metaphor 2.

Keywords: Metaphor, interface design, usability, SUS score, preference.

1 Introduction

This is a feasibility study focused on 3D interface metaphors of mobile devices. There were three reasons explained why the cellular phone's interface was decided as typical stimuli in this study: (1) the cellular phone is the most popular one of mobile product; (2) varied but short lifestyle of mobile phone product is caused by users'

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 343–352, 2007.

fickle in affection; (3) there is a tendency towards 3D interface of products but limited in applying to mobile phones. Also, despite plenty of design researches into mobile phone products, more of them focus on examining the essentials of exterior design. Consequently, the reformation of products' appearance design is becoming faster and faster except their interface design. Later, some researchers found that and made studies of the icon design towards Kansei impressions but limited in present mobile phone products.

Thereore, there are three purposes of this study: (1) to investigate users' opinions on a mobile phone's interface design; (2) to verify whether the simulated graphics and interactions corresponding to the metaphors intended (pilot study); (3) to measure the usability of 3 types of metaphorical interface design simulated in this study; (4) to compare the preference for any one of the 3 types of metaphorical interface design.

1.1 Design Principles of User Interface Design

The user interface is a medium for communication between users and computational devices (systems). As a result, the appraisal of usability is directly depended on whether the user interface design is good or bad. Therefore, there are various principles generalized from lots of researches into the user interface design so far. For examples, Microsoft (1995) defined 5 bases of interface design, including consistent, learnable, intuitive, extensible and attractive; Sutcliffe (1983) suggested 6 fundamentals to the software interface designs, information structure, consistency, compatibility, adaptability, economy and guidance not control included; Norman (2000) addressed 5 principles of good interface design as follows: affordance, good conceptual model, good mapping, feedback and visibility. Based on the opinion of user's experience, Veryzer (1999) recognized four essential design properties as the characteristics of interface designs: operation, understanding, structure and value. Furthermore, 5 frameworks of interface designs towards usability problems were advised by Nielsen (1993), including easy to use, efficient to use, few error, subjectively pleasing, easy to remember. To conclude, that ease-to-learn of operational environment to make users learn effectively is the essential principle of user interface design.

Recently, graphical user interface (GUI) is praised for ideal interactive environment in common. It mainly contains five elements: screen, windows, icons, menus and point devices. Sanders and McCormick (2002) defined them as : display, operational environment, control panel, integration of control and display, and information structure. However, there are also potential problems with GUI, such as lack of real touch and immediate responses, misunderstandings of symbolic cognition from the different contexts of cultures, overloading derived from too many invisible multi-layers, less interests of learning resulted from trivial operating sequences, etc. Therefore, designers need some guidelines, procedural advice and computational support in using concepts for their design problems (Carroll & Rosson, 1994). In response to the needs, metaphors are widely applied to visualize and specify the structure of GUI design (Kobara, 1991).

1.2 Metaphor

In the early 80s, Lakoff and Johnson (1980) offered a new vision of cognitive processes as the capacity of projecting from a given (well known) domain to a new (less known) domain. This capacity of mappings between domains is immediately recognized as "metaphor". Later, they (1987) developed their ideas of metaphor into a modern approach to cognition, known as experientialism. One of the main concepts of experientialism is the image-schema which is abstract pattern derived from our bodily experience and other (everyday) interactions with the external world. Briefly, the essence of user interface design is an appropriate metaphor which is intuitively related to their cognition, experience and knowledge (Erickson, 1990); the core of a metaphor is the understanding and experiencing one kind of thing in terms of another. After that, metaphors as a design source are often applied in designing the user interfaces to provide the underlying images, terms and concepts that make communication possible at all (Marcus, 1993). Also, the metaphors should be an intuitive access to all functionality. That means the user should be not only familiar with the metaphor domain but also able to perform the mapping between the metaphor domain and the application domain.

In conclusion, for designers, a good metaphor is necessarily to look after both sides of functionality (information structure) and visualization (aesthetics and consistency, etc.); for users, a good metaphor of interface design is necessarily able to help themselves get applicable instructions in procedures to reduce the chances of mistakes happened.

1.3 Usability Measurements: System Usability Scale (SUS) and Interview

Initially, usability was defined as the degree of efficiency and effectiveness of use within a specified range of users, tasks, tools, and environment (Bennet, 1984; Shackel, 1984), which results in over 87% of usability researches, especially in HCI domain, focus more on measuring objective estimations of effect on goals (Nielsen and Lavy, 1994). To a certain extent, the product performances based on the subjective measurements are actually better than others but uncertainly make users be satisfied with the product itself. In fact, user's satisfaction is considered as one of emotional expression.

To sum up, the emotional reactions accompany objective evaluations but are unconsciously operated to influence the degree of satisfaction (Zajonc, 1980). So, the subjective emotional issues should no be excluded from usability problems. In other words, usability should simultaneously take both objective measurements and subjective assessments into account (Norman, 2002, Fig 1). Therefore, the ISO9241-11 (1998) re-describes usability as "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". Further, "effectiveness" means the ability of users to complete tasks using the system, and the quality of the output of those tasks; "efficiency" means the level of resource consumed in performing tasks; "satisfaction" means users' subjective reactions to using the system.

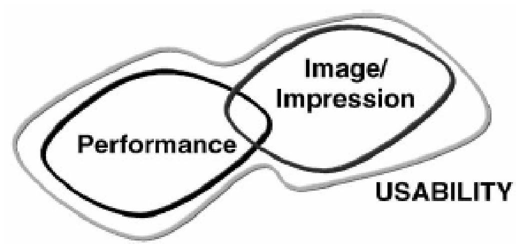


Fig. 1. The concept of Norman's usability

In general, the demands of evaluating usability of systems within an industrial context are often neither cost-effective nor practical to perform a full-blown context analysis and selection of suitable questions. But, users could be very frustrated if they were presented with a long questionnaire. It was very possible that they would not complete it and there would be insufficient data to assess subjective reactions to system usability. Nevertheless, what we need is a general indication of the overall level of usability of a system compared to its competitors or its predecessors. That means the measure had not only to be capable of being administered "quickly" and "simply" but also to be "reliable" enough to be used to make comparisons of user performance changes from version to version of a product.

In response to these requirements, the System Usability Scale (SUS) was developed. It is a simple, ten-item Likert scale giving a global view of subjective assessments of usability (cited from John Brooke). Generally, it is used after the respondent using the system (interface) being evaluated but before any discussions takes place. Due to the numbers for individual items are not meaningful on their own, SUS scores should be calculated according to following steps: first sum the score contributions from each item (each item's score contribution will range from 0 to 4) and then multiply the sum of the scores by 2.5 to obtain the overall value of SU. That means SUS scores have a range of 0 to 100.

Sometimes, interviews are also regarded as usability evaluation approaches. In Rubin's model (Rubin, 1994), interviews are not only used in the beginning of development stage to design the questionnaire but also used in the last stage of the evaluation stage of usability testing to clarify user responses and to collect additional information. Interviews are of two types: Structured and Open-ended (USINACTS Usability, 2004). Generally, structured interviews can be designed rigorously with a predefined set of questions to avoid biases and usually provide more reliable and quantifiable data than open-ended interviews. Later, the structured interviews were also applied in the last phase of the usability evaluation in this study.

2 Method

Initially, this was a cooperative project to propose applicable metaphors based on 3-D graphic to the interface of mobile phone product. After a succession of discussions and improvements, 3 types of metaphorical interface design with different degree of abstract were selected, simulated and constructed as Fig 2, Fig 3 and Fig 4 below. In response to the purposes of this study above, three phases of empirical experiments were proceeded as follows:

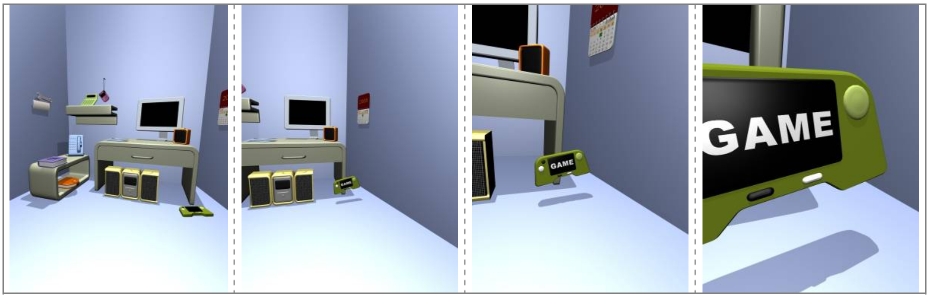


Fig. 2. Metaphor 1 - mobile phone's interface as personal room

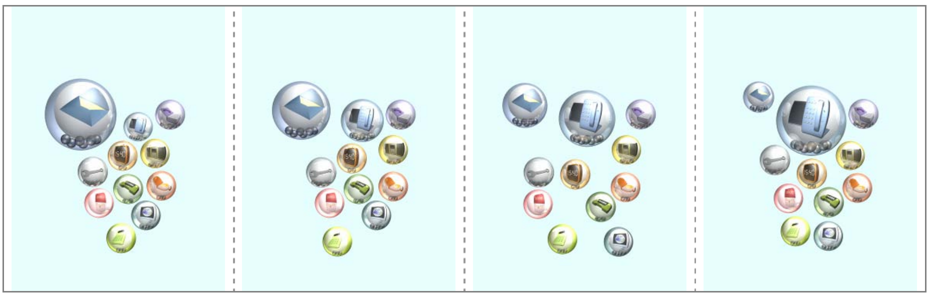


Fig. 3. Metaphor 2 - information structures as transparent bubbles

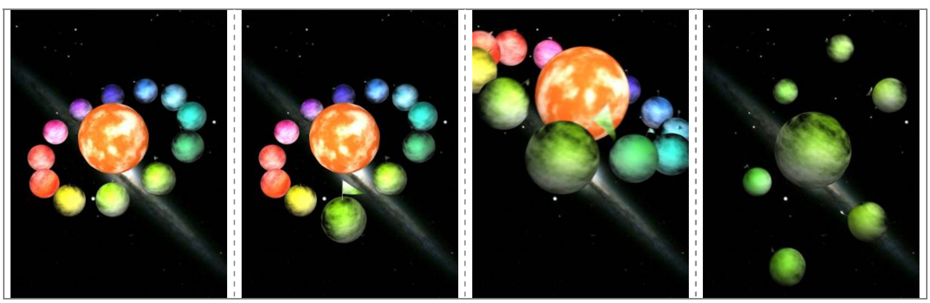


Fig. 4. Metaphor 3 – search for functionalities as an exploration of interplanetary

2.1 Phase 1: User's Response to Mobile Phone Products

Seventy-three people participated in this questionnaire survey, 43 of whom were students, 25 of whom were workers, and 5 of whom were on active duty of military service. There were 27 female and 46 male participants ranging in age from 19 to 35. The average age was 25.05 (SD = 3.01). On average, they have had 3.66 personal mobile phone products (SD = 1.25), ranging from 2 to 8.

In order to gather user's opinion on mobile phone products, participants were asked to respond to 3 similar questions with 14 possible answers prepared in advance

(13 of them were specific and 1 of them were any of other possibilities, shown in Table 1). The 3 questions were describe as: (1) which 3 of factors were the most important elements to evaluate whether a product of mobile phone is good or bad; (2) which 3 of factors were attractive essentials to individual to want to have a product on an impulse; (3) which 3 of factors were the most important issues if individual wanted to buy a new one. Also, participants were instructed to mark the factors in order according to the extent of importance (1 meant the most important factor, 2 was next, and 3 was next again).

Table 1. 14 possible answers to individual's opinions on mobile phone product

01. appearance	02. functionality	03. interface manipulation	04. graphic design of interface	05. size
06. weight	07. brand	08. price	09. quality	10. place of production
11. assessment (from others)	12. product name (or slogan)	13. limit of quantity	14. others ……	

2.2 Phase 2: (Pilot study) Examinations of Correspondence Between the Simulated Interface Designs and Metaphorical Concepts

Four volunteers were recruited to examine whether the simulated graphics and interactions corresponding to the metaphors intended, 3 of whom were graduated students and 1 of whom was worker. There were 2 female and 2 male volunteers ranging in age from 24 to 31. On average, they have had 4.25 personal mobile phone products, ranging from 3 to 6. To verify the correspondence between the visualization of 3 types of interface design and original metaphor concepts, open-ended interviews were applied in this phase. This aimed to estimate how much time volunteers could correctly guess the objects of metaphors. They were also asked to interpret the helpful cues to guess. Here, volunteers were allowed only to see the static graphics of the metaphorical interface design in response to all the questions.

2.3 Phase 3: The Usability Evaluations of 3 Types of Metaphorical Interface Design

Eight subjects were recruited to measure the usability of 3 types of metaphorical interface design with SUS questionnaire, 6 of whom were students and 2 of whom were workers. There were 5 female and 3 male subjects ranging in age from 15 to 27. On average, they have had 3.38 personal mobile phone products, ranging from 1 to 7. In this phase, without restrictions on time, subjects were permitted to manipulate the constructed interfaces by order and instructed to perform 2 tasks in each interface respectively. After operating and performing the tasks on an interface, they were requested to evaluate the usability of three different metaphorical interface designs with SUS questionnaire. Then, the SUS scores were calculated to indicate whether it was a good metaphorical interface design.

2.4 Phase 4: User's Preference for the Metaphorical Interface Design

Twelve subjects were enrolled in this phase, 9 of whom were students and 3 of whom were workers. There were 7 female and 5 male subjects ranging in age from 15 to 31. On average, they have had 3.67 personal mobile phone products, ranging from 1 to 7. Here, after operating and performing the tasks on each metaphorical interface design, subjects were requested to evaluate personal impressions on the 3 types of metaphorical interface design with SD questionnaire and compare individual's preference for any one of them.

3 Results and Discussion

The results of different experimental analysis in each phase were described respectively as follows:

3.1 Phase 1: User's Response to Mobile Phone Products

To interpret the results more systematically and easily, the observations were analyzed through the frequency of descriptive statistics. Then, a score (a cumulative number, abbreviated to "sum"; the highest score = 219 = 73 x 3) was obtained by adding across the answers (frequencies) in the same question but different orders.

In response to question 1 (to evaluate whether a product of mobile phone is good or bad), "appearance" was considered as the most important factor (sum = 54), next was "functionality" (sum = 39), next was "quality" (sum = 27), next was "interface manipulation" (sum = 25), and next were "brand" and "price" (the same as sum = 24). Effects on other factors were ignored because their sum total was less than 20.

In response to question 2 (to find out the attractive essentials of a mobile phone product), "appearance" was obviously regarded as the most important factor (sum = 65), next was "functionality" (sum = 27), and next was "price" (sum = 22). Effects on other factors were ignored because their sum total was less than 20.

In response to question 3 (to clarify which factors really determine a user's decision to buy), "appearance" was still regarded as the most important factor (sum = 57), next was "price" (sum = 48), and next was "functionality" (sum = 36), next were "brand" (sum = 21), and next was "interface manipulation" (sum = 20). Effects on other factors were ignored because their sum total was less than 20.

Furthermore, the average rating of questionnaire items on a 9 point scale ranging from "strongly agree to abstract" to "strongly agree to concrete" reflected the degree of abstract in metaphorical interface designs. Consequently, that the metaphorical interface designs scored 3.5, 0.75, and -2.25 in turn meant metaphor 3 were regarded as abstract interface design and metaphor 1 and metaphor 2 were regarded as concrete interface designs, but the degree of concrete in metaphor 1 was stronger than in metaphor 2.

3.2 Phase 2: (Pilot study) Examinations of Correspondence Between the Simulated Interface Designs and Metaphorical Concepts

There were no significant differences in making a guess at the metaphor of visualization and in performing completion times found among different metaphorical interface designs. However, large differences in surprise and interest ratings were evidently resulted from volunteer's personal attitude towards mobile phone products. Therefore, further results and discussions might be appropriately illustrated with the outcomes in next phases.

3.3 Phase 3: The Usability Evaluations of 3 Types of Metaphorical Interface Design

In order to appraise whether it was good metaphor at interface design, the usability evaluations of 3 metaphorical interface designs were performed by means of SUS questionnaire. Then, SUS scores were calculated and illustrated with Fig 5 below.

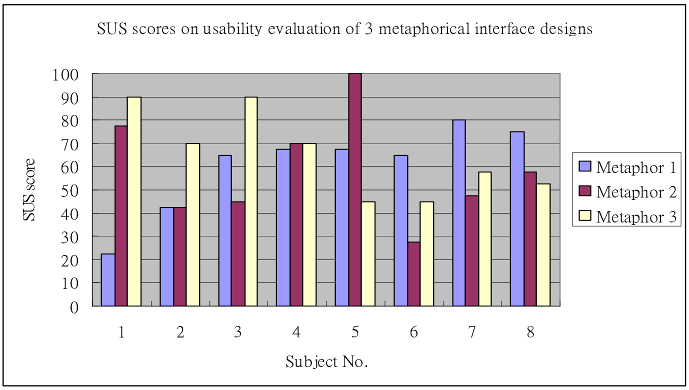


Fig. 5. SUS scores on usability evaluation of 3 metaphorical interface designs

As shown in Fig 5, it was indicated that there were large dissimilarity of cognitions between different participants (especially in the results of metaphor 2). That meant what was someone's favor might be disliked by another. Fortunately, the interview method was simultaneously combined in this phase. The discrepancies in the results of the scores were easily ascribed to subject's attitude towards mobile phone products. Despite some of them treat the products as only communicated tools, for example, another ones regarded them as representations of personality. Besides, the SUS scores were deeply affected by personal preferences, too. So, it was certainly needed to compare the metaphorical interface designs based on user's subjective responses.

3.4 Phase 4: User's Preference for the Metaphorical Interface Design

In this phase, subjects were requested to make a favorite choice between coupled selections of metaphorical interface design. Consequently, in a couple of metaphor 1 and metaphor 2, 8 subjects chosen metaphor 1 as their preference; in a couple of metaphor 2 and metaphor 3, 7subjects chosen metaphor 2 as their preference; in a couple of metaphor 3 and metaphor 1, 10 subjects chosen metaphor 1 as their preference. To conclude, metaphor 1 was more popular than other.

In addition, there might be a good idea to analyze and gain further useful information on the results of the coupled comparison by means of quantity analysis methodologies, such as through MDS analysis to indicate the potential mental space of human beings. Within time limitation in this study, only the result of metaphor 1 was taken for example and shown in Fig 6 below.

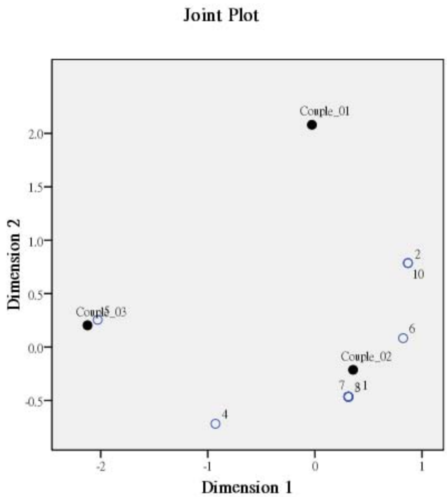


Fig. 6. The result of MDS analysis of metaphor 1

4 Conclusion

Although, the System Usability Scale (SUS) is a commonly used, freely distributed, and reliable questionnaire consisting of 10 items, it was also verified that some of subjects (non-native English speakers) failed to understand the word “cumbersome” in Item 8 of the SUS (“I found the system to be very cumbersome to use”) without instructor’s further explanation (Finstad, 2006). Also, in this study, it was found out that more of subjects might be not used to seek for further help to any questions spontaneously but appreciate if the instructor would voluntarily give them more information. So, it was suggested that the descriptions of the questionnaires should be confirmed with no doubt before being proceeded. Moreover, it was also noticed that subject’s attitude towards mobile phone products brought large differences in the results of both usability evaluation and subjective impressions appraisal. Because of SUS estimations and interviews being substitute for quantity measurement in this study, only 4-12 subjects were recruited to perform the assessment tasks. In consequence, it was difficultly aimed at generalize any principle of metaphorical interface design in common. So, that might be a good point to improve this study in further works.

References

1. Benyon, D., Palanque, P.: Critical Issues in User Interface Systems Engineering. Springer, New York (1996)
2. Brooke, J.: SUS: A Quick and Dirty Usability Scale. In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, I.L. (eds.) Usability Evaluation in Industry, Taylor & Francis, London (1996) Also see http://www.usabilitynet.org/trump/documents/Suschapt.doc
3. Finstad, K.: The System Usability Scale and Non-Native English Speakers. Journal of usability studies 1(4), 185–188 (2006)
4. Imaz, M., Benyon, D.: How Stories Capture Interactions. In: Proceedings of Human-Computer Interaction – INTERACT '99, Edinburgh, Scotland, September 1-3, 1999, pp. 321–328. IOS Press, Amsterdam (1999)
5. Kissel, G.V.: The effect of computer experience on subjective and objective software usability measures. In: Conference companion on Human factors in computing systems 1995 (1995)
6. Nielsen, J.: Usability Engineering. Academic Press, Boston (1993)
7. Nielsen, J.: Why You Only Need to Test With 5 Users. Alertbox, (March 19, 2000) (accessed on 06 February 2007) http://www.useit.com/alertbox/20000319.html
8. Norman, D.A.: The design of everyday things. Doubleday, New York (1990)
9. Norman, D.A.: Emotion & design: attractive things work better. Interactions 9(4), 36–42 (2002)
10. Shneiderman, B., plaisant, C.: Designing the user interface: strategies for effective human-computer interaction, 4th edn. Addison Wesley, London (2005)
11. Vatrapu, R., Perez-Quinones, M.A.: Culture and Usability Evaluation: The Effects of Culture in Structured Interviews. Journal of usability studies 1(4), 156–170 (2006)

Applications of the Visible Korean Human

Min Suk Chung and Jin Seo Park

Department of Anatomy, Ajou University School of Medicine, Suwon, Korea
dissect@ajou.ac.kr

Abstract. Visible Korean Human (VKH) consisting of magnetic resonance, computer tomography, anatomic, and segmented images was created. In the VKH, several techniques were developed and numerous data were acquired. The VKH techniques majorly contributed to the generation of advanced segmented images, Visible Living Human, and Visible Mouse. Also, a software for sectional anatomy, three dimensional images for virtual dissection and virtual endoscopy, was developed based on the VKH data distributed worldwide. The VKH technique and data are expected to promote development of other serially sectioned images and software, which are helpful in medical education and clinical practice.

Keywords: Visible Korean Human, Magnetic resonance images, Anatomic images, Segmented images, Three dimensional images.

1 Introduction

Visible Human Project was the first trial ever made to obtain serially sectioned images of cadaver's whole body. The data derived from Visible Human Project have contributed largely in the medical image field [16]. Furthermore, technique used for the Visible Human Project has been modified in Korea for Visible Korean Human (VKH) [9, 10, 13] and in China for Chinese Visible Human [18]. By using the improved technique for VKH such as magnetic resonance (MR) scanning, computerized tomography (CT) scanning, serial sectioning, photographing, and segmenting, the VKH team acquired better data consisting of MR images, CT images, anatomic images, and segmented images. The improved VKH technique was introduced through article [9, 10, 13]; the VKH data were distributed worldwide. The objective of this report is to promote new trials by other researchers to create other serially sectioned images applying VKH technique and three dimensional (3D) images with VKH data, which will be greatly helpful in medical education and useful in clinical practice. To achieve this objective, this report describes the ideas and experiences in applying the VKH technique and data to generate new image data contributing in medical image field.

2 Materials, Methods, and Results

2.1 Application of VKH Technique for Detail Segmented Images of VKH

Three years ago, 13 structures (skin, bones, liver, lungs, kidneys, urinary bladder, heart, cerebrum, cerebellum, brainstem, colon, bronchi, and arteries) in the anatomic

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 353–362, 2007.

images were outlined to obtain basic segmented images [9]. However, these basic segmented images were insufficient to produce various 3D images; therefore, advanced segmented images of the many more structures were decided to be made, in order to complement and replace the basic segmented images.

By using segmentation technique on Adobe Photoshop™ (version 7.0) [9], important structures identifiable in anatomic images were segmented as follows: 104 structures of head and neck including brain components, 15 structures of heart, and 84 structures of left upper limb as well as 114 structures of left lower limb including each bone, muscle, nerve, and artery. Few segmented structures such as skin were used as they stand, and other segmented structures such as bone were classified into each bone; other unsegmented structures such as muscles were newly outlined. According to color difference of each structure in gross examination, segmentation was performed automatically, semiautomatically, or manually on Adobe Photoshop. Through stacking the segmented images, coronal and sagittal segmented images were made in order to verify segmentation. As a result, advanced segmented images of 317 structures were produced (Fig. 1).

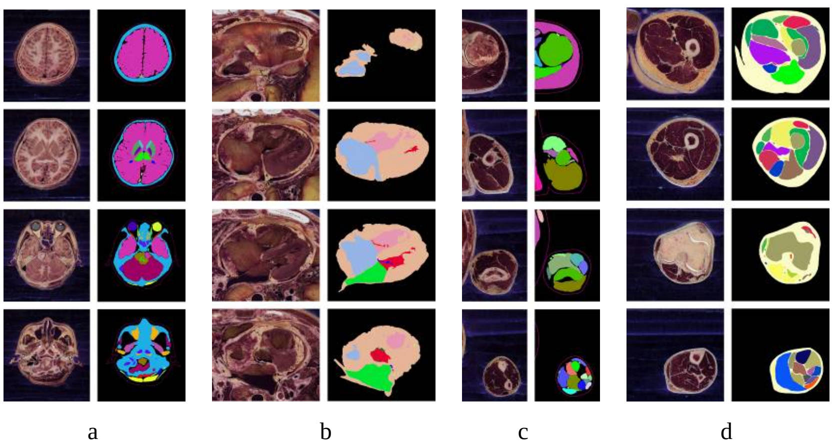
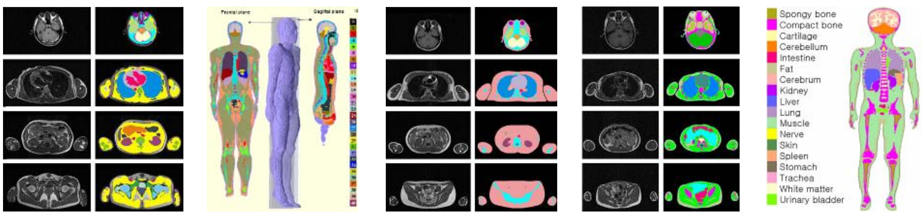


Fig. 1. Anatomic and segmented images of head (a), heart (b), left upper limb (c), and left lower limb (d)

2.2 Application of VKH Technique for Visible Living Human

Besides the VKH, Visible Living Human was newly planned to produce whole body MR images of living humans. While the Visible Living Human does not include anatomic images, it includes MR images of living human, whose body conditions are much better than those of cadaver. In order to carry out the Visible Living Human, MR images of a male adult, a female adult, and a male child were acquired and image processing was done as follows.

By utilizing MR scanning technique of the VKH, whole bodies of living humans were MR scanned. The whole body of an adult can not be MR scanned at once; through the experience from VKH, the first MR series from head to knees and the second MR series from knees to toes were scanned separately; subsequently, both MR series were combined and aligned. Living humans' organs move contrary to cadaver, thus new technique was utilized for the Visible Living Human: To minimize bowel movement, the volunteers had been starved for 12 hours prior to MR scanning; to compensate heart movement, electrocardiography sensor was attached; to compensate lung movement, lung movement sensor was worn by the volunteers. As a result, 613 MR images of male adult, 557 images of female adult, and 384 MR images of male child were acquired at 3 mm intervals (Fig. 2a,c,d) [8, 9].



a b c d e

Fig. 2. MR and segmented images of male adult (a) [8], female adult (c) [9], and male child (d). 3D images of male adult (b) [6] and male child (e) with the color bar which indicates the segmented structure.

Through the same segmentation technique, anatomic structures in MR images were segmented. The Adobe Photoshop was adequate for segmentation not only in anatomic images (24 bits color) but also in MR images (8 bits gray). Yet MR images did not show definite anatomic structures, so that more knowledge of anatomists and radiologists was necessary for segmentation process. In the Visible Living Human, segmentation of the anatomic structures was performed, whose absorptance of electromagnetic wave is quite different: 47 structures in male adult, 19 structures in female adult (segmentation in process), and 33 structures in male child (Fig. 2a,c,d) [6, 8, 9].

The segmented images of male adult were stacked and volume-reconstructed to produce 3D images of 47 structures. Electromagnetic wave was exposed on the 3D images in various ways to calculate influences of the electromagnetic wave on internal structures. In the same manner, the exposing simulation is being performed using the 3D images of the male child (Fig. 2b,e) [6]. In addition, the segmented images of male adult were stacked and surface-reconstructed to produce 3D images. A software on which the 3D images can be selected and rotated was produced for medical education [8].

2.3 Application of VKH Technique for Visible Mouse

Mouse anatomy is a fundamental knowledge for researchers who perform biomedical experiments with mice. So far the mouse anatomy has been educated through two dimensional images such as atlas, thus being difficult to be comprehended, especially the stereoscopic morphology of the mouse. In order to overcome this difficulty, it is necessary to produce 3D images made of Visible Mouse [12].

Through serial sectioning and photographing technique of the VKH, anatomic images of a mouse were obtained. The cryomacrotome with only 1 um error was designed for serial sectioning of any organism smaller than human. Using the cryomacrotome, a mouse was serially sectioned at 0.5 mm intervals to make sectioned surfaces. Every sectioned surface was photographed using a digital camera to produce 437 anatomic images. Distance between the sectioned surface and the digital camera could be adjusted closely to generate anatomic images with 0.1 mm pixel size. During photographing, the strobe light was flashed in the same condition to produce anatomic images with consistent brightness (Fig. 3) [10, 12].

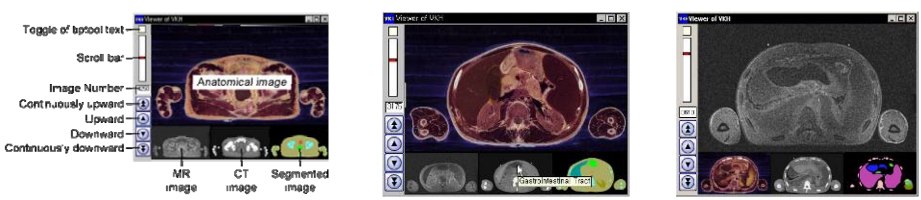
a b

Fig. 3. Anatomic images, segmented images (a) [12], and 3D images (b) of the mouse with skin and internal structures

By practising the same segmentation technique, 14 mouse structures in the anatomic images were segmented. The rodent structures were so small that exhaustive knowledge of mouse anatomy referred to mouse atlas was necessary for the segmentation process [5, 12]. From the segmented images, contours of the mouse structures were stacked on Alias^{TM} Maya (version 7.0); stacked contours were volume-reconstructed on Autodesk^{TM} AutoCAD (version 2007) (Fig. 3).

2.4 Application of VKH Data for Sectional Anatomy Education

Sectional anatomy is the course to learn anatomic structures on the sectional planes. The sectional anatomy has become much more important especially because medical students have to interpret MR and CT images. Because of the growing importance of sectional anatomy, browsing software of the VKH data was created. Raw data of the browsing software were 1,702 sets of MR, CT, anatomic, and segmented images



a b c

Fig. 4. Browsing software showing four images and graphic user interface (a), name of structure in CT image (b), and enlarged MR image (c) [11]

(1 mm intervals) of the whole body. On the browsing software, a set of four images corresponding to one another was displayed; different sets of images could be selected conveniently using graphic user interface; names of the segmented structures in all images were also displayed; among a set of four images, any image could be enlarged. The browsing software can be downloaded free of charge from our web site (anatomy.co.kr) (Fig. 4) [11].

2.5 Application of VKH Data for Virtual Dissection (Volume-Reconstruction)

As the traditional method in educating anatomy, cadaver dissection is exceptionally important; but the cadaver dissection with time and place restriction can not be performed commonly. The problem can be compensated by virtual dissection. For the virtual dissection, volume-reconstruction was carried out primarily to produce 3D images out of the VKH data because volume-reconstruction enables users to section 3D images of structures with the same color as real-life human structures [14].

Virtual dissection software of whole body was created at Inha University, Korea. For this research, basic segmented images of 13 structures and corresponding anatomic images were used as materials. As the preprocess, intervals and pixel size of the anatomic and segmented images were increased from 0.2 mm to 1.0 mm because original image files were too large to be processed on a personal computer. Anatomic and segmented images were stacked and volume-reconstructed on Microsoft™ Visual C++ (version 6.0) to produce 3D images. On programmed virtual dissection software, the 3D images with real color could be sectioned, selected, and rotated (Fig. 5a) [10].

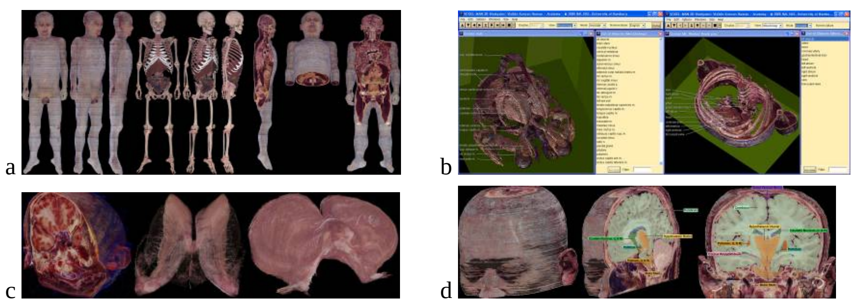


Fig. 5. 3D images by volume-reconstruction at Inha University, Korea (a) [10], University Medical Center Hamburg-Eppendorf, Germany (b), State University of New York at Stony Brook, US (c), and Texas Tech University, US (d)

Virtual dissection software of head and neck was produced at University Medical Center Hamburg-Eppendorf, Germany. For this research, advanced segmented images of 104 structures in head and neck were used as materials; Voxel-Man system was used for segmentation refinement and for volume-reconstruction [14]. On the virtual dissection software, the 3D images with real color could be stripped in sequence; the 3D images could be selected to display by names of structures; the 3D images could also be annotated. In the same manner, virtual dissection software of thorax including heart components was created (Fig. 5b). Additional virtual dissection software of head is being created at other institutes such as State University of New York at Stony

Brook (Fig. 5c), Texas Tech University (Fig. 5d), Stanford University Medical Media & Information Technologies, Pittsburgh Supercomputing Center, and Upperairway Company, US, as well as in MAÂT3D, France.

2.6 Application of VKH Data for Virtual Dissection (Surface-Reconstruction)

After the trial of volume-reconstruction, surface-reconstruction was tried with segmented images of the VKH to produce 3D images of structures with outstandingly small file size. The surface-reconstructed 3D images can be selected to display, rotate, and transform themselves in real time. In addition, the 3D images can be easily distributed through the Internet [17].

3D images of urogenital tract were made by surface-reconstruction at University Paris V René Descartes, France. Contours of 42 structures including urogenital tract, its neighboring bones, arteries, and skin were stacked and surface-reconstructed using SURFdriver software to produce 3D images. The 3D images could be manipulated individually or in-group with the 3D images' transparency adjusted (Fig. 6a) [17]. Additionally by surface-reconstruction, 3D images of skull and 3D images of lower limb bones were made at University Malaya, Malaysia (Fig. 6b) and at Konrad-Zuse-Zentrum für Information Stechnik, Germany (Fig. 6c), respectively.

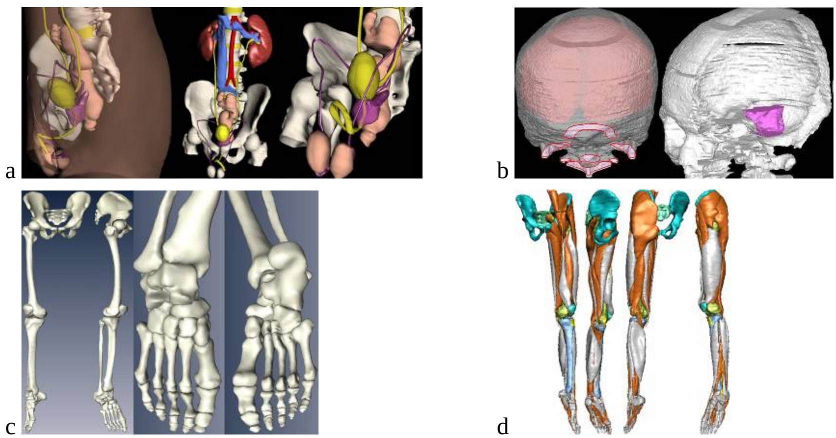


Fig. 6. 3D images by surface-reconstruction of University Paris V René Descartes, France (a) [17], University Malaya, Malaysia (b), Konrad-Zuse-Zentrum für Information Stechnik, Germany (c), and Ajou University, Korea (d) [15]

We tried to produce 3D images of left lower limb by surface-reconstruction on Maya and Rhino, which are both popular commercial software. Contours of 114 structures in left lower limb were stacked on Maya; gaps between contours were filled with non-uniform rational B-spline (NURBS) surfaces on Rhino; all NURBS surfaces were converted into polygon ones on Rhino; the surfaces were corrected to complete the 3D images on Maya. In this manner, surface-reconstruction can be done on the popular and commonly used software to produce 3D images in the Maya file format, thus be widely used by other researchers (Fig. 6d) [15].

2.7 Application of Vkh Data for Virtual Endoscopy

Virtual colonoscopy of a patient is becoming popular in diagnosing colon cancer. However, the virtual colonoscopy being based on the CT images, colon wall can not be displayed in real color. In the VKH, lumen of colon had been segmented. The anatomic and segmented images were stacked and volume-reconstructed to produce 3D image of colon wall's luminal surface, which held real color. Based on the 3D image, virtual colonoscopy was performed at Inha University, Korea. The virtual colonoscopy with real color was similar to real colonoscopy, so that educational effect could be enhanced. In the same manner virtual bronchoscopy and virtual arterioscopy were performed with the VKH data (Fig. 7) [13].

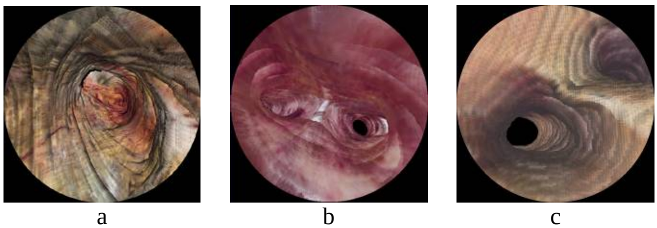


Fig. 7. Virtual colonoscopy (a), virtual bronchoscopy (b), and virtual arterioscopy (c) [13]

2.8 Application of Vkh Data for Radiation Simulation

Just as Visible Living Human data were used for exposing simulation of electromagnetic wave, VKH data were used for radiation simulation at Hanyang University, Korea. The raw data from VKH were the segmented images of 22 structures at 2 mm intervals. The segmented images were stacked and volume-reconstructed to produce 3D images. Then the 3D images were exposed by broad parallel photon beams in various ways to calculate effects of radiation on each structure. These results could be used to prevent workmen in radiation-polluted environment from receiving radiation damage [4].

3 Discussion

Various applications performed with the VKH technique and data have been described. In discussion, further potential applications to be performed in the future are presented.

Segmented images of the VKH will be acquired to completion. For the VKH, it required a day and three months only to obtain MR images and anatomic images, respectively. Then it required three years to obtain segmented images of 317 structures (Fig. 1); nevertheless, segmentation has not been finished yet [9, 10, 13]. It will take three more years to finish segmentation in whole body. Three years later, the segmented images accompanied by corresponding anatomic images will be distributed worldwide free of charge. It is anticipated that other researchers reform some incorrect segmented images and make more detailed segmented images for their own purposes. In any case, the segmented images, which will be finished by authors, are expected to save precious time and effort of other researchers.

Female data of the VKH will be acquired. The female data, which are later than male data, need to be upgraded as follows. The female cadaver for the VKH needs to be of young age, good body contour, and with few pathologic findings. Anatomic images of whole body need to have 0.1 mm intervals, 0.1 mm pixel size, and 48 bits color depth. It seems that 0.1 mm voxel size of anatomic images is the last trial in this research, which belongs to gross anatomy. Additionally, the female data are desirable to be followed by the child data, fetus data, and embryo data. The systematic data according to sex and developmental age will be the main contents in the medical image library [10, 13].

Visible Head data including 7.0 Tesla MR images will be made. State-of-the-art 7.0 Tesla MR machine has been developed by Professor Zang-Hee Cho in Neuroscience Research Institute of Gachon University, Korea [3]. By using the 7.0 Tesla MR machine, a cadaver's head were MR scanned recently. Surprisingly, several structures of brain (for example, cerebral arteries) appeared better in the MR images than in the anatomic images of VKH while the other structures appeared better in the anatomic images (Fig. 8). The cadaver's head will be serially sectioned to obtain anatomic images in correspondence to the MR images. These Visible Head data are expected to be milestone images in neuroscience field.

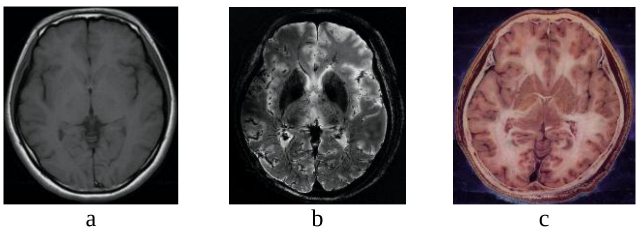


Fig. 8. 1.5 Tesla MR image with T1 (a), 7.0 Tesla MR image with T2 (b), and anatomic image (c)

Visible Patient data will be acquired by registration of the VKH data to a patient's data. For example, MR images of a patient are produced. Anatomic images of the VKH are transformed, thus be corresponding to the MR images of the patient. For the transformation, high level of registration technique and segmented images of VKH will be utilized. The anatomic images registered to the patient have color information and high resolution. Therefore, anatomic images can be used in production of realistic 3D images, which will be the basis for preoperative virtual surgery of the patient [7].

Pregnant female data of the Visible Living Human will be made. MR images of male adult, female adult, and male child have been scanned (Fig. 2) [8, 9]. Likewise, MR images of pregnant female will be scanned, including fetus images. The pregnant female data will be used for the exposing simulation of electromagnetic wave too [6].

Visible Rat data and Visible Dog data will be made. Like the Visible Mouse (Fig. 3) [12], Visible Rat and Visible Dog can be performed for biomedical experiment and veterinarian education by using the same cryomacrotome of VKH. In order to make better images than other Visible Rat [1] and other Visible Dog [2], micro MR machine and micro CT machine will be used for the rat; high level of serial sectioning and detail segmentation will be tried for both the rat and the dog.

In this report, our ideas from experiences to apply the VKH technique and data have been introduced. By applying the VKH technique, other images such as Visible Living Human (Fig. 2) [6, 8]. Visible Mouse (Fig. 3) [12] can be created; by applying the VKH data, sectional anatomy (Fig. 4) [11], virtual dissection (Figs. 5, 6) [9, 10, 17], virtual endoscopy (Fig. 7) [13], radiation simulation [4], and Visible Patient can be performed. Based on this report and the distributed VKH data, other researchers are expected to make better images and more progressive applications for use at medical education and clinical practice.

References

1. Bai, X., Yu, L., Liu, Q., Zhang, J., Li, A., Han, D., Luo, Q., Gong, H.: A high-resolution anatomical rat atlas. J. Anat. 209, 707–708 (2006)
2. Böttcher, P., Maierl, J.: Macroscopic cryosectioning. A simple new method for producing digital, three dimensional database in veterinary anatomy. Anat. Histol. Embryol. 28, 97–102 (1999)
3. Cho, Z.H., Jones, J.P., Singh, M.: Foundations of medical imaging. Wiley, New York (1993)
4. Choi, S.H., Lee, C.S., Cho, S.K., Chung, M.S., Na, S.H., Kim, C.H.: Construction of a high-definition ‘Reference Korean’ voxel phantom for organ and tissue radiation dose calculation. In: Proceedings WC 2006 Conf 2006, pp. 4061–4064 (2006)
5. Iwaki, T., Yamashita, H., Hayakawa, T.: A color atlas of sectional anatomy of the mouse. Maruzen Co. LTD, Tokyo (2001)
6. Lee, A.K., Choi, W.Y., Chung, M.S., Choi, H.D., Choi, J.I.: Development of Korean male body model for computational dosimetry. ETRI J 28, 107–110 (2006)
7. Li, L., Liu, Y.X., Song, Z.J.: Three-dimensional reconstruction of registered and fused Chinese Visible Human and patient MRI Images. Clin. Anat. 19, 225–231 (2006)
8. Lee, Y.S., Chung, M.S., Park, J.S., Hwang, S.B., Cho, J.H.: Three dimensional MRI and software for studying normal anatomical structures of an entire body. J. Korean Soc. Magn. Reson. Med. 9, 117–133 (2005)
9. Park, J.S., Chung, M.S., Hwang, S.B., Lee, Y.S., Har, D.H.: Technical report on semiautomatic segmentation using the Adobe Photoshop. J. Digit. Imaging 18, 333–343 (2005)
10. Park, J.S., Chung, M.S., Hwang, S.B., Lee, Y.S., Har, D.H., Park, H.S.: Visible Korean Human. Improved serially sectioned images of the entire body. IEEE Trans. Med. Imaging 24, 352–360 (2005)
11. Park, J.S., Chung, M.S., Choe, H., Byun, H.Y., Hwang, J., Shin, B.S., Park, H.S.: Serially sectioned images of the whole body (Sixth report: Browsing software of the serially sectioned images for learning sectional anatomy). Korean J. Anat. 39, 35–45 (2006)
12. Park, J.S., Chung, M.S., Hwang, S.B.: Serially sectioned and segmented images of the mouse for learning mouse anatomy. Korean J. Anat. 39, 305–312 (2006)
13. Park, J.S., Chung, M.S., Hwang, S.B., Shin, B.S., Park, H.S.: Visible Korean Human: Its techniques and applications. Clin. Anat. 19, 216–224 (2006)

14. Pommert, A., Höhne, K.H., Pflesser, B., Richter, E., Riemer, M., Schiemann, T., Schubert, R., Schumacher, U., Tiede, U.: Creating a high-resolution spatial/symbolic model of the inner organs based on the Visible Human. Med. Image. Anal 5, 221–228 (2001)
15. Shin, D.S., Chung, M.S., Park, J.S.: Technique of the semi-automatic surface reconstruction of Visible Korean Human data on the commercial software. In: Proceedings SPIE Medical Imaging 2007 Conf (in press)
16. Spitzer, V.M., Ackerman, M.J., Scherzinger, A.L., Whitlock, D.G.: The Visible Human Male: A technical report. J. Am. Med. Inform. Assoc. 3, 118–130 (1996)
17. Uhl, J.F., Park, J.S., Chung, M.S., Delmas, V.: Three-dimensional reconstruction of urogenital tract from Visible Korean Human. Anat. Rec. A Discov. Mol. Cell. Evol. Biol. 288, 893–899 (2006)
18. Zhang, S.X., Heng, P.A., Liu, Z.J.: Chinese Visible Human Project. Clin. Anat. 19, 204–215 (2006)

Role of Humans in Complexity of a System-of-Systems

Daniel DeLaurentis

School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN, USA
ddelaure@purdue.edu

Abstract. This paper pursues three primary objectives. First, a brief introduction to system-of-systems is presented in order to establish a foundation for exploration of the role of human system modeling in this context. Second, the sources of complexity related to human participation in a system-of-systems are described and categorized. Finally, special attention is placed upon how this complexity might be better managed by greater involvement of modeling of human behavior and decision-making. The ultimate objective of the research thrust is to better enable success in the various system-of-systems that exist in society.

Keywords: system-of-systems, complexity, human behaviors, connectivity.

1 Introduction

A system-of-systems (SoS) consist of multiple, heterogeneous, distributed, occasionally independently operating systems embedded in networks at multiple levels that evolve over time. While the *moniker* may be recent, the *notion* of a "system-of-systems" is not new. There have been and still are numerous examples of collections of systems that rely upon the interaction of multiple, but independently operating, systems. Ground and air transportation, energy, and defense are high profile examples. These existed before the phrase "system-of-systems" entered common use, have been studied extensively through several fields of inquiry, but rarely have been examined as a distinct problem class. In recent years, the formal study of systems-of-systems has increased in importance, driven largely by the defense and aerospace industries, where the government's procurement approach has changed. Where government customers once issued detailed requirements for a specific platform system, they now ask instead for a broad set of capabilities that are needed over a significant time span. As a result, the system developers and contractors must determine the appropriate mix of systems and related interconnections to provide these capabilities.

As a result of the defense system heritage, perhaps the most active group discussing SoS is the systems engineering community. Several contributions in the recent literature from this community indicate a growing recognition that systems engineering processes are not complete for the new challenges posed by the development of SoS [1]. Thus, researchers are looking for new characterizations. Rouse, for example, describes implications of complexity on systems engineering approaches [2]. Sage and Cuppan present a working collection of traits for SoS that points to the possibility of a "new federalism" as a construct for dealing with the variety of levels of cohesion in

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 363–371, 2007.

SoS organization [3]. We describe later in this paper an additional view with an angle toward structuring complexity in a SoS.

While defense-related applications have driven the recent emphasis on systems-of-systems, many of society's needs across diverse domains are met through systems-of-systems. Unfortunately, in many instances, an inflexibility in response to disruption (artificial or natural) or increased demand for service is observed. Supply chains acquire excessive inventory and capacity if manufacturers and distributors ignore opportunities to collaborate on demand forecasts. Healthcare networks experience breakdowns of information flows and lose continuity of care as patients migrate among autonomous hospitals, out-patient clinics, nursing homes, etc. Electric power blackouts illustrate consequences of emergent behaviors; modest breakdowns in some component systems of energy grids mushroom into major disruptions. Well-publicized transportation delays from over-capacity operations at major airports prevent passengers and cargo from reaching their destinations in a timely, productive manner.

We postulate that a common thread through all these challenges is the complexity that arises from their nature as a system-of-systems. Further, much of the effort towards development of effective approaches for system-of-systems is coming from the engineered systems community. "SoS solutions" are still conceived primarily as the appropriately organized mix of artificial systems. Less attention is being paid to the role of human behavior in these applications and to research areas such as the socio-technical systems approach, which addresses the complexities and uncertainties that result when human and technical systems interact.

Therefore, the goals of this paper are to describe the system-of-systems problem class, identify primary sources of complexity within this class, and highlight particular roles of humans in this context. It is hoped that this paper will spur interdisciplinary endeavors that link research threads from the human factors community (particularly digital human modeling) with those addressing system-of-systems problems from a complex systems engineering and design perspective. While the limited space in this paper does not allow for detailed presentation of design methodologies or analysis results from particular applications, several references will be given throughout to select methods and results.

2 Characterizing a System-of-Systems

A set of distinguishing traits for SoS problems have been proposed by Maier [4]. Maier's criteria are operational independence, managerial independence, geographic distribution, evolutionary behavior, and emergent behavior. The first three primarily describe the problem boundary and mechanics while the latter two describe overall behavior. Emergent behavior—the manifestation of behavior that develops out of complex interactions of component systems that are not present for the systems in isolation presents a particular challenge because it is unpredictable, is often non-intuitive, and can be manifested in a good manner (e.g., a new capability) or bad manner (e.g., a critical failure). A primary challenge in performing System-of-Systems Engineering (SoSE) with greater effectiveness is to understand the mechanism of emergent behavior, develop cues to detect it, and create a methodology for managing it intelligently. The one-sentence characterization for an SoS offered at the beginning of the introduction to this paper attempts to encapsulate these traits as well as additional aspects such as heterogeneity of component systems and multi-level structure.

This latter aspect especially indicates a need to characterize the structure of an SoS. For this, a lexicon has been formed [5]. Since many SoS problems have some interconnectedness in a hierarchical manner, the lexicon enumerates these interacting *levels* of various components with Greek letters. Starting from α, the levels proceed upward, leading to γ or δ or higher depending on the complexity of the problem and the depth of analysis desired. Further, a β-level constituent represents a network of α-level entities; a γ-level constituent represents a network of β-level entities, and so on.

In addition to these hierarchical labels representing levels of complexity, a set of scope dimensions are defined. These dimensions, or *categories*, highlight the trans-domain aspects of SoS. Specifically, not all entities within the levels are similar in their basic character; they can be made of entities or systems of different characteristics and representations. Thus, a framework is needed to classify the different components of each level. Further, these components can be categorized primarily into *Resources*, *Operations*, *Economics*, and *Policy*. Each of these categories independently comprises the previously described levels, thereby completing the SoS lexicon. The categorization of the levels lends clarity in dealing with the different facets of the problem while maintaining the lucidity provided by the levels. The relationship between the categories and the levels forms a pyramid, which is depicted in Fig. 1.

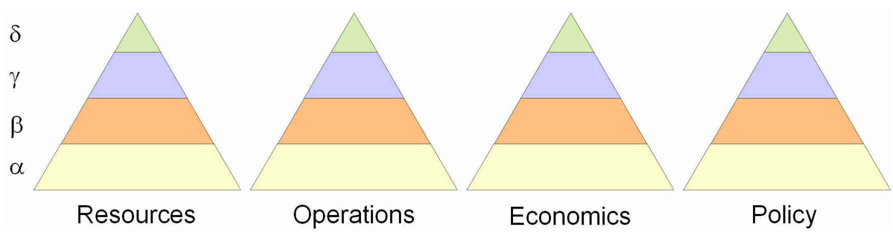


Fig. 1. Scope categories and hierarchy of levels: an unfolded pyramid unifying the lexicon

While a systematic representation of scope and structure is crucial, the ability to characterize SoS problems in the analysis domain is the next required step. Further, the SoS of interest must be defined in a way that exposes appropriate level(s) and categorical scope dimensions. Towards these objectives, a taxonomy has been proposed [6] consisting of a three-dimensional space characterized by system type, autonomy, and connectivity, illustrated in Fig. 2 below. Analysis methods must be appropriate for the **system types** *(S-axis)* that constitute the SoS. Some SoSs consist predominantly of technological systems – independently operable mechanical (hardware) or computational (software) artifacts. Technological systems have no purposeful intent; i.e., these resources must be operated by, programmed by, or activated by a human. Other SoSs consist predominantly of humans and human enterprise systems—a person or a collection of people with a definitive set of values. The second SoS dimension is the **degree of control** *(A-axis)* over the entities by an authority or the autonomy granted to the entities. This relates to Maier's discussion of operational independence and managerial independence of systems within an SoS. Emphasizing the importance of control / autonomy, others refer to a collection of systems with operational, but limited managerial, independence as a "system-of-systems" and a collection of systems with little central authority as a "federation of systems" [3]. Finally, systems involved in a system-of-systems are interrelated and **connected** *(C-axis)* with other (but likely not all) systems in the SoS. These interrelationships and communication links form a network. A key focus for design methods research in an SoS context lies in analysis and exploitation of interdependencies (i.e., network topology) in addition to the attributes of systems in isolation. These dimensions serve as a taxonomy to guide the formulation and analysis of the SoS design problem. A particular SoS can therefore be considered as a "point" or "region" in the three-dimensional space formed by the aforementioned dimensions as axes. Based on its location in this space, and other indicators of particular problem structure, the approach and methodology necessary for analysis and design can be more intelligently selected.

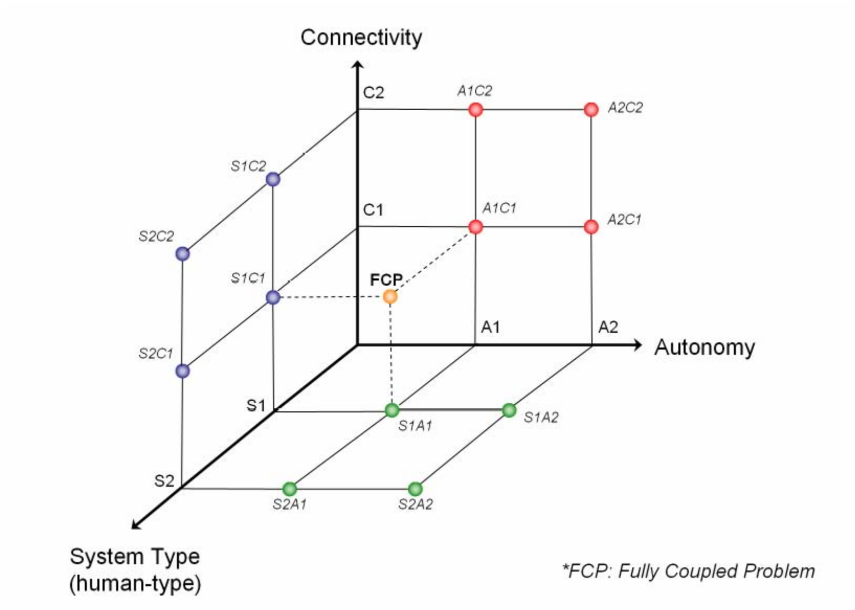


Fig. 2. Three key dimensions for system-of-systems

3 Role of Humans in SoS Complexity

3.1 General Sources of Complexity

Views of complexity include internal system complexity, external complexity, computational (algorithmic) complexity, etc. While complexity can be viewed in different manners, it is safe to say that complexity is always a comparative measure. A given system, at a particular scale and using a particular measure, can be more or less complex than another system examined at the same scale using the same measure. One particular view holds that complexity measures the amount of information necessary to describe a system [7]. More complex systems require greater information to define them. Complexity in an SoS stems primarily from the heterogeneity of its constituent systems, dynamic and uncertain connectivity that arises with operation at multiple levels, and the role of humans who bring "self-organization" and differing perspectives on operational context within an SoS. In this vein, sources of complexity in an SoS can be gleaned from their characteristics and placement in the taxonomy discussed. Further, in the face of these sources, representing the structure of organization is key to managing complexity in SoS.

The following notional example, depicted in Fig 3 and using the previously introduced lexicon, illustrates the manifestation of several of these sources of complexity in a system-of-systems. The α-level comprises the most basic components of the SoS. The result of the interactions at an α-level are felt at the corresponding β-level. Hence, emergence is evidenced at the β-level and higher. In addition, there is evolution at play in any given β-level during any interval of time. As shown in Fig. 3, the different α-level components are nodes in a β-level network and the connectivity may shift over time (t_1, t_2 and t_3). There is an undirected relation between 1 and 2 at time t_1 while there is only a one-way relation between them at time t_2. Also, new entities appear and existing entities depart over time. Due to these changes among different α-level constituents, the performance of the β-level entities is altered by both evolution and emergence. New configurations lead to altered interactions between the α-level constituents (emergence) and this emergence in turn affects the course of the future make-up of the β-level (evolution). Thus, evolution and emergence are mutually influential and are manifested at each level.

Relative complexity between two SoS may differ for each level. One SoS may have simplistic organizational structure (lower complexity) at β-level but its α-level systems are more complex than those of another SoS. This multilevel aspect is especially important from a computational complexity perspective. Integration of high fidelity analysis models across multiple layers of abstraction is impractical, and a more refined tact that is selective in which information is appropriate is required. This is well explained by Herbert Simon, who asserts that, "Resemblance in behavior of systems without identity of the inner systems is particularly feasible if the aspects in which we are interested arise out of the *organization* of the parts, independently of all but a few properties of the individual components" [8].

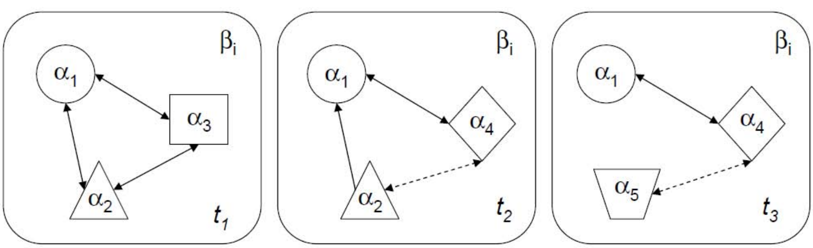


Fig. 3. Notional example of a system-of-systems

3.2 Role of Humans

When the taxonomy indicates that a particular SoS has heterogeneous system type which may imply significant role of human behavior (e.g. S1-S2 along S-axis in Fig. 2), an SoS-oriented approach will be most challenged by complexity sources peculiar to these human behaviors. This is also when such an approach is most needed in order to uncover the complicated interactions between entities that emerge at the higher levels of abstraction. Along these lines, the recently released study conducted by the U.S. Air Force Scientific Advisor Board entitled "System-of-Systems Engineering for Air Force Capability Development" [9] identified the role of the human as critical for the successful implementation of SoS in the field. Quoting from their summary:

> "Whenever the Air Force generates a system-of-systems, interaction among the systems often includes human-to-human interactions. If the machine-to-machine aspect of SoS is weak, then it falls upon the humans to achieve the interaction. This can, and often does, create a very challenging environment for the human; sometimes leading to missed opportunities or serious mistakes. The lack of sound Human System Interface designs can exacerbate this. Coordinated situation awareness is difficult to manage if the individual systems miss or convey confusing or conflicting information to their operators.
>
> Assuming there are sound human systems interfaces for individual systems, the Air Force can greatly reduce the burden on the human-to-human coordination if effective inter-system interoperation is established. It is the objective of our study to suggest ways for improving intersystem interoperation at the hardware and software level while keeping in mind that sometimes the human-to-human interaction is the appropriate interface. An effective system-of-systems will promote collaborative decision-making and shared situation awareness amongst the human operators. This occurs by addressing the need for common, consistent human-system interfaces."

To more explicitly tie together the concepts of human behavior and sources of complexity in an SoS, we introduce a tangible example. The information presented in Fig. 4 describes air transportation as an SoS using the lexicon introduced earlier in this paper. This figure is adopted from our work in seeking to forge robust architectures that could transform air transportation from its current strained state to one that better scales to shifts in demand and disturbance.

Human values and the organizations in which they participate influence the SoS through application of self-interests and perspectives (via operational and managerial independence). Dynamics from this influence take place at multiple levels and under multiple time-scales. For example, humans can operate and implement policies for α-level systems. In air transportation, these roles include pilots, maintenance crew, inspectors, etc. Further, actions at these lower levels of hierarchy tend to evolve on a short time scale; they typically are near-term actions of a practical variety in response to the situation at hand and are operational in nature. Complexity may arise here from ambiguous instructions, differing skill levels in distributed teams [10], goals and

		← System of Systems Dimensions →			
	Level	**Resources**	**Operations**	**Economics**	**Policy**
Base Level	α ($\vartheta\,10^6$)	Aircraft, Tower	Pilot/Crew Deployment, Maintenance Schedules	Fuel Price, Investments	Type Certification, Flight Procedures
↑ Network of Networks ↓	β ($\vartheta\,10^4$)	Airport	Airline	Fuel Market, Labor/Union Costs	Airport Traffic Mgmt, Noise Policies
	γ ($\vartheta\,10^2$)	Air Transportation System	Commercial Air Operations	Airline Industry	Air/Ground Safety, Accessibility
	δ ($\vartheta\,10^1$)	National Transportation System	Operators of Total National Transportation System	Overall Transportation Forecasts/Market	National Transportation Policies
	ε ($\vartheta\,10^0$)	Global Transportation System	Global Operators in the World Transportation System	WTO, Global Marketplace	Global Transportation System Policies

Fig. 4. Air transportation system-of-systems description

strategies and detection [11] and diagnosis of system failures [12] in a network context. Humans can also participate as managers of α-levels systems, or parts of organizations at even higher levels. In air transportation, these roles include airlines, air traffic regulatory bodies (e.g., Federal Aviation Administration, FAA), labor unions, etc. Actions of these human-based entities tend toward the longer-term influence based upon considered assessment of organizational goals; these decisions may establish well-intentioned policies that unintentionally restrict the existence of promising solutions.

The interactions of humans and human enterprises at various levels of the SoS exhibit the same kind of self-organizing behaviors described in the previous section on complexity. In order for human engineered/operated/managed systems interact, they must form “convergence protocols” [9] which enable interoperability. Driving the need for these protocols is the multiplicity of perspectives which arise from “institutional frames of reference and associated operational context scope that system users and system designers implicitly assume in their behavior and their designs” [13]. These perspectives form a backdrop for the operational assumptions adopted for the given system and enable (or prohibit) its integration with other systems in the system-of-systems. When systems come together (to exchange mass, energy, or information), the perspectives under which they were instantiated may conflict and require human intervention to resolve. These so-called “trigger events” force a reorganization of the convergence protocols, assumptions under which the systems operate, the context in which they interact, or some combination of these. This reorganization is nothing more than a manifestation of the self-organization of a complex system. Humans effect this restructuring of the SoS to resolve this interoperability challenge, but require additional information (about unstated assumptions, perspective, and so forth) in order to respond effectively. As a result, this increase in information also increases the complexity of the overall SoS.

Capturing these interactions with humans at various levels of organization in the SoS has been undertaken by some researchers in the SoS community in an effort to improve the understanding of SoS dynamics. For example, in the air transportation domain, traveler preference models (an α-level behavior) have been modeled by Lewe et. al [14] and Trani [15]. Further, models of β-level entities have been developed, such as MITRE's JetWise model of airlines [16], and higher level dynamics in DeLaurentis et. al [17]. Some of these studies have used agent-based modeling (ABM) to mimic the human factors and influences on the system-of-systems. ABM employs a collection of autonomous decision-making entities called agents imbued with simple rules of behavior that direct their interaction with each other and their environment. The mathematical representation of agent rules is often quite simple, but the resultant system-wide behavior is often more complicated, unexpected, and thus instructive [18]. One major limitation of this method involves the simple-rule-based models that represent human behaviors and organizations. ABM is a modeling tool for the study of Complex Adaptive Systems [19], which represents a problem class with many of the same dynamic behaviors of SoS problems (e.g., emergence). For air transportation, Donohue has specifically called for a CAS approach [20]. Recent advancements in analyzing complex networks also provide useful tools for addressing the global connectivity of a given SoS problem without reliance on low level interactions [21].

However, a need remains for an increased ability to understand and replicate the complex human factors involved in a human-technical system-of-systems. This includes more than simply the behaviors of individuals in the SoS, but also the representation and scope assumptions issues discussed in the previous paragraphs.

4 Summary

A brief introduction to system-of-systems problems was presented, especially crafted as an entrée into dialogue with researchers in human system modeling. The sources of complexity related to human participation in an SoS were briefly introduced, couched in a lexicon and taxonomy for the problem class. We observe that these sources reside at multiple levels and across different scope dimensions. In particular, each human participant brings a unique perspective, and thus interpretation, to information in an SoS. A challenge for effective design in a system-of-systems context is to make these perspectives transparent during interactions. Finally, a better understanding of the mixing between short and long-term time scales for human influence on an SoS must be obtained.

We recommend collaborative explorations between engineered system architects and designers with human system modeling experts in order to increase the effectiveness of system-of-systems that may involve significant human presence in their constitution as described via the proposed taxonomy.

References

1. Keating, C., Rogers, R., Unal, R., Dryer, D., Sousa-Poza, A., Safford, R., Peterson, W., Rabadi, G.: System-of-Systems Engineering. Engineering Management 15(3), 36–45 (2003)
2. Rouse, W.: Engineering Complex Systems: Implications for Research in Systems Engineering. IEEE Transactions on Systems, Man, and Cybernetics—Part. C: Applications and Reviews 33(2), 154–156 (2003)

3. Sage, A.P., Cuppan, C.D.: On the Systems Engineering and Management of Systems of Systems and Federations of Systems. Information, Knowledge, Systems Management 2(4), 325–345 (2001)
4. Maier, M.W.: Architecting Principles for System-of-Systems. Systems Engineering 1(4), 267–284 (1998)
5. DeLaurentis, D.A., Calloway, R.K.: A System-of-Systems Perspective for Public Policy Decisions. Review of Policy Research 21(6), 829–837 (2004)
6. DeLaurentis, D., Crossley, W.: A Taxonomy-based Perspective for Systems of Systems Design Methods. In: Proceedings of IEEE System, Man, & Cybernetics Conference, Hawaii, 10–12 October 2005, Paper 0-7803-9298-1/05 (2005)
7. Bar-Yam, Y.: Dynamics of Complex Systems. Westview Press (Perseus Books Group), Boulder, CO (1997)
8. Simon, H.: Sciences of the Artificial, 3rd edn. MIT Press, Cambridge (1996)
9. Saunders, T., et al.: System-of-Systems Engineering for Air Force Capability Development. United States Air Force Scientific Advisory Board, Report SAB-TR-05-04, Washington, DC (July 2005)
10. Caldwell, B.: Analysis and Modeling of Information Flow and Distributed Expertise in Space-related Operations. Acta Astronautica 56, 996–1004 (2005)
11. Patrick, J., James, N., Ahmed, A.: Human Processes of Control: Tracing the Goals and Strategies of Control Room Teams. Ergonomics 49, 12–13, 1395–1414 (2006)
12. Rasmussen, J., Rouse, W.B. (eds.): Human Detection and Diagnosis of Systems Failures. Plenum Press, New York (1981)
13. Polzer, H., DeLaurentis, D., Fry, D.: Multiplicity of Perspectives, Context Scope, and Context Shifting Events. In: Proc. IEEE SMC Society Second Annual International Conference on System-of-systems Engineering, San Antonio, TX , April 16–18, 2007 (2007)
14. Lewe, J., DeLaurentis, D., Mavris, D., Schrage, D.: Modeling Abstraction and Hypothesis of a Transportation Architecture. Journal of Air Transportation 11(3) (2006)
15. Baik, H., Trani, A.A.: A Transportation Systems Analysis Model (TSAM) to Study the Impact of the Small Aircraft Transportation System (SATS). In: 23rd International Conference of the System Dynamics Society, Boston, MA, July 17-21, 2005 (2005)
16. Niedringhaus, W.P.: The Jet:Wise Model of National Air Space System Evolution. Simulation 80(1), 45–58 (2004)
17. DeLaurentis, D., Han, E.-P., Kotegawa, T.: Establishment of a Network-based Simulation of Future Air Transportation Concepts. In: 6th AIAA Aviation Technology, Integration and Operations Conference (ATIO), Wichita, Kansas, 25–27 September 2006, AIAA 2006-7719 (2006)
18. Bonabeau, E.: Agent-based Modeling: Methods and Techniques for Simulating Human Systems. Proc. Natl. Acad. Sci. 99(3), 7280–7287 (2002) (accessed on 6 Feb 2006 from) http://www.pnas.org/cgi/reprint/99/suppl 3/7280
19. Waldrop, M.M.: Complexity: The Emerging Science at the Edge of Order and Chaos, 1st edn. Simon & Schuster Inc., NY (1992)
20. Donohue, G.: Air Transportation is a Complex Adaptive System: Not an Aircraft Design. In: AIAA International Air and Space Symposium and Exposition: The Next 100 Years, Dayton, Ohio, 14–17 July 2003, AIAA 2003-2668 (2003)
21. Barabási, A.-L., Albert, R.: Emergence of Scaling in Random Network. Science no. 286, 509–512 (1999)

Digital Human Modeling for Product Lifecycle Management

H. Onan Demirel[1] and Vincent G. Duffy[1,2]

[1] School of Industrial Engineering
[2] School of Agriculture and Biological Engineering
Purdue University
47907 West Lafayette, IN, USA
hdemirel@purdue.edu, duffy@purdue.edu

Abstract. This paper presents the current and probable future applications of Digital Human Modeling (DHM) in Product Lifecycle Management (PLM) implementations. A Formula 1 race car and a marine vessel were developed by the integration of PLM and DHM software packages. Dassault Systemes' CATIA V5 PLM solution package was used for CAD/CAE design/analysis and UGS Tecnomatix JACK software was utilized for visual/mathematical ergonomics and human analysis. Literature review for future work, [1] and [2], was discussed to investigate the potentials of DHM in PLM for simulation of a blast motion in Navy vessels. The results and observations indicated that integration of DHM and PLM packages have potentials to improve the product development efforts and offer an enhanced approach for simulation of complex systems where there is human-machine integration.

Keywords: Digital Human Modeling, Advance Manufacturing, Product Development, Human Factors and Ergonomics, Product Lifecycle Management, Engineering Design, Motion Capture, Simulation & Modeling.

1 Introduction

Use of DHM applications and integration of those with PLM solutions have noticeable advantageous in product development and digital manufacturing. Number of design changes and cost of rapid prototyping can be minimized by DHM and PLM integration [3]. CIMdata's research conclude that on average, organizations using digital manufacturing technologies can reduce lead time to market by 30%, the number of design changes by 65% and time spent in the manufacturing planning process by 40%. Production throughput can be increased by 15% and overall production costs can be cut by 13% [4].

1.1 Human Element in PLM

PLM is a strategic business approach which incorporates managerial and engineering aspects of product definition by integrating people, process and information. PLM uses product definition information to manage the entire lifecycle of the product from

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 372–381, 2007.

concept creation to end of life by composing a strong collaboration across different segments in the enterprise. PLM applies a consistent set of business solutions in support of the innovative creation, manufacturing, management, distribution and service of products [5]. Complex product development strategies should consider the human-machine integration and ergonomics aspects of the concept product before it's ever launched. One key product development strategy studied in this paper is the utilization of DHM for PLM applications. Many industries have the same challenges where the human element is not being effectively considered in the early stages of the design, assembly, and maintenance of products. However the use of DHM tools can increase the quality by minimizing the redundant design changes and the improve safety of products by eliminating ergonomics related problems [6,7]. An example, at Ford Motor Co. involved the JACK human simulation solution to evaluate ergonomics and worker safety in installing a new satellite digital antenna radio system. This analysis occurred early, at the product design review stage. This reduced the late design modification efforts and helped assess performance and ergonomics based problems prior to prototyping. [8].

1.2 Need of DHM for PLM

DHM uses digital humans as representations of the workers inserted into a simulation or virtual environment to facilitate the prediction of performance and/or safety. Also, it includes visualizations of the human with the math and science in the background [9]. DHM helps organizations design safer and efficient products while optimizing the productivity and cost [10]. Applications that incorporate DHM have the potential to enable engineers to integrate ergonomics and human factors engineering principles earlier in the design process [11]. One of the advantages of the DHM applications is that Motion Capture tools can be used to drive the DHM and facilitate reduction of injuries & comfort prediction through virtual interactive design of work stations and new products [12]. This method allows manufacturers and designers to predict potential risk before production begins [13].

Traditionally DHM applications have been utilized by manufacturing and design industry. One of the first digital human modeling applications was implemented by the U.S. military for cockpit design. Virtual drivers were used to assess the safety and the performance of the vehicles, and the need of expensive physical mockups were minimized. The largest application areas of digital human models are vehicle design, auto and off-road equipment manufacturing [14]. More recently, technological developments and advancement in the CAE tools expands the application areas of DHM. Studies show that one can find current/probable applications of DHM in transportation design, work environment simulation, training and simulation of healthcare personnel and cognitive human modeling.

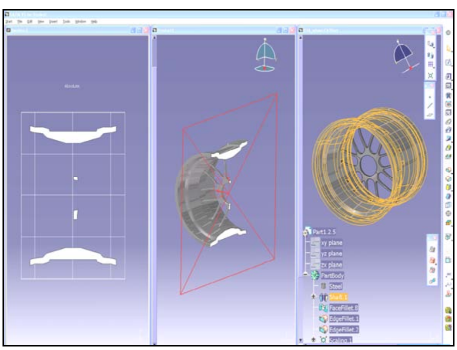
Design and development of advance products or systems is an iterative process where designs are needed to be modified, changed or reverse engineered due to customer, work or manufacturing demands and constraints. Also, the essential components of a product development; people, departments, equipment and systems can not perform work needed in isolation. There is a definite need for integrating those essential components during design, analysis, production, marketing and service of the products. Integration of DHM and PLM tools may resolve the challenges of product development from design to service, and provides control of product development process. This integration also increases the engineering design and analysis capabilities, improves the product ergonomics, enables human-machine virtual reality applications and provides cost and time savings.

2 Method

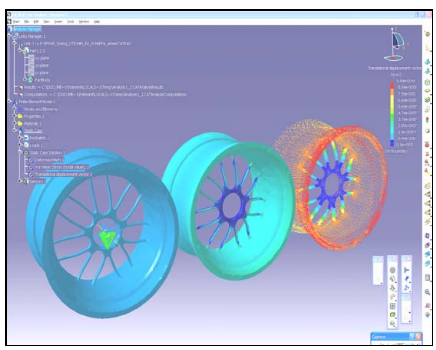
This study examined two product development applications by integration of DHM and PLM tools. A Formula 1 race car and a marine vessel were designed by the integration of PLM and DHM software packages. Dassault Systemes' CATIA V5 PLM solution package was used for CAD/CAE design/analysis and JACK DHM software was utilized for visual/mathematical ergonomics and human analysis [15,16]. Authors' design methodology and DHM strategies of new product development process are explained in detail and figures from design and analysis phases are provided.

2.1 Formula 1 Race Car Design and Driver Comfort Assessment

The utilization of DHM for PLM is evaluated by the integration of UGS Tecnomatix Jack, one specific piece of DHM software by UGS, with Dassault Systemes' CATIA V5 PLM software in development of a sample Formula 1 race car. The complete CAD model is designed by using CATIA V5 software. (see Fig.2) The CAD model is composed of 40 constraint-based individual sub-models. Each sub-model is also composed of numerous mechanical parts which are created and then assembled by using CATIA V5's Mechanical Design workbench. Utmost care is taken during constraint based solid modeling in order to construct a valid model with transferability characteristics. Constraint based solid modeling approach brings the advantage of integrating the created model with other CAE software (Fluent, Ramsis, MADYMO) Also, Sectioning and Clash Analysis are applied to visually inspect the parts in assembly. (see Fig.1a.) During CAD model development mechanical integrity of several sub-models are tested by CATIA V5's Generative Structural Analysis.



1a



1b

Fig. 1a. Constraint based CAD solid modeling **Fig. 1b.** Generative structural analysis

A static analysis case is developed to test the boundary conditions by assigning different scenarios (forces, accelerations, etc.) to check whether the part withstands the effects of applied forces without losing its mechanical integrity. (see Fig. 1b.)

Ergonomics Design and Analysis workbench of the CATIA V5 is utilized to optimize the cockpit space and comfort of the driver. The digital human model is inserted into the Formula 1 race car from built-in anthropometric library and a suitable driver posture is assigned by using Human Posture Analysis workbench. Visual checks (vision, reach envelope, clash detection…) are applied to ensure that driver position is visually legal. (No penetration between human model and race car,

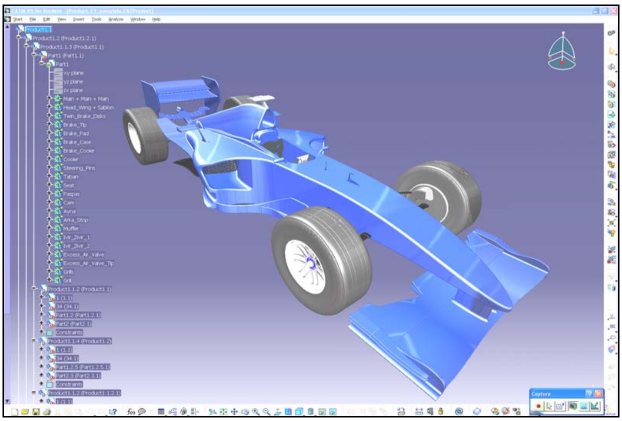
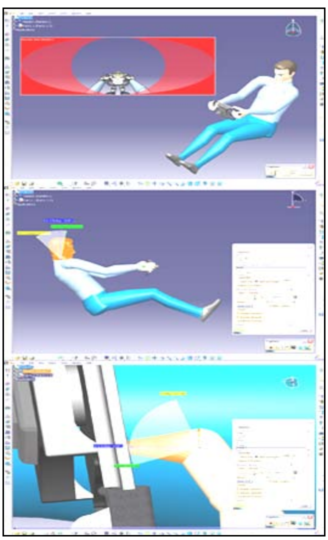
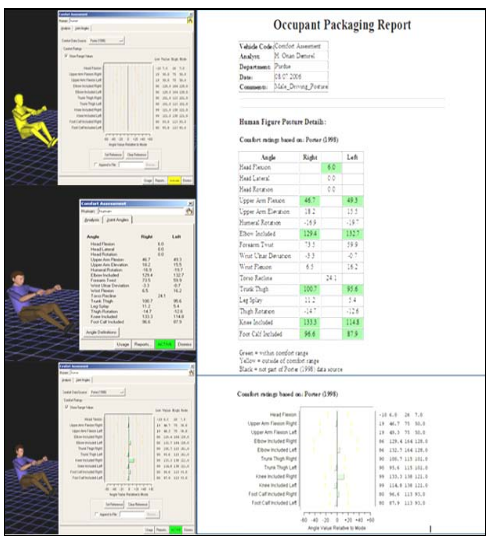


Fig. 2. Final CAD model with Real Time Rendering applied



3a



3b

Fig. 3a. DHM Virtual Design

Fig. 3b. Driver Comfort Analysis

no visual awkward body posture…) (see Fig. 3a.) The given posture and the car then exported to Jack 5 software. The assigned driving posture in CATIA V5 environment is mimicked in Jack 5 software by using built-in Jack human models. The mimicked posture is then analyzed for driver comfort by using Comfort Assessment module under Occupant Packaging Toolkit menu. Comfort Assessment module determines if the assigned posture is within the comfort range based on 3D joint angles and overall position analysis. The Comfort Assessment analysis output highlights if any of the body parts are within/outside the comfort range. (see Fig. 3b.)

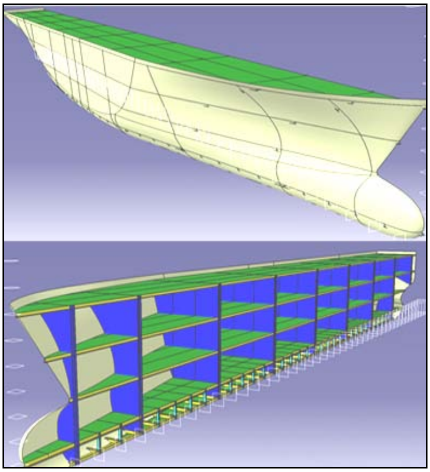
Several cockpit features are changed and awkward body positions found are modified until the comfort of the human model and structural integrity of CAD model is matched to the design objectives. This dual validation method, virtual and mathematical, provides a more extensive analysis and offers a better resolution to accommodate optimum driving standards and comfort. After completion of assembly, the complete CAD model is rendered by Real Time Rendering workbench, which enhances the product development attempts by adding almost-real visualization effects (shadows, reflections, etc.) to the model. (see Fig. 2.)

2.2 DHM in Marine Vessel Design and Simulation

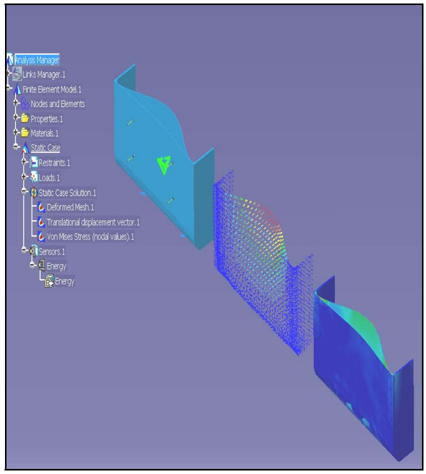
The product development methodology discussed in Formula 1 race car design can be utilized to resolve the design and analysis challenges in marine applications. The paper discussed [2] indicates that integration of DHM applications may improve the design, analysis and simulation needs of marine vessels.

The paper [2] also indicates that the underwater shock produced by an underwater explosion remains one of the biggest threats to ships and shipboard personnel due to extreme high accelerations in very short duration. Authors used a lumped parameter model which is a simplified mechanical system that divides the human body (sitting upright in a chair) in different segments. The human body is subjected to a vertical shock and the lump parameter model is used to identify the parts of the human body where the impact forces exceed the critical limits. Also, the influences of structural damping and stiffness of the peak loads acting on the human body can be investigated by using this lumped parameter model. [2]

A parallel design and analysis approach similar to the Formula 1 race car is studied to show the potential of DHM and PLM integration with advanced CAE applications in naval vessel design and analysis. Similar to the Formula-1 race car CAD modeling procedure, a ship model is developed from scratch by using CATIA V5. The Generative Shape Design workbench is utilized to sketch the external skin (External hull in CATIA V5 tutorial is utilized) of a naval vessel. After that the external skin is used as the boundary of the deck and transversal bulkhead plates of the base unit. Then, a simplified inner structure of the ship is developed by replicating the base unit by utilizing Structure Object Detailed Design workbench. (see Fig. 4a.) Longitudinal deck plates, transversal bulkhead plates and beams are designed and stiffened to form the skeletal structure of the ship. Mechanical integrity of the external hull is tested by CATIA V5's Generative Structural Analysis. A static analysis case is developed to observe the what-if effects of applied distributed forces on a non-stiffened external hull. (see Fig. 4b.)



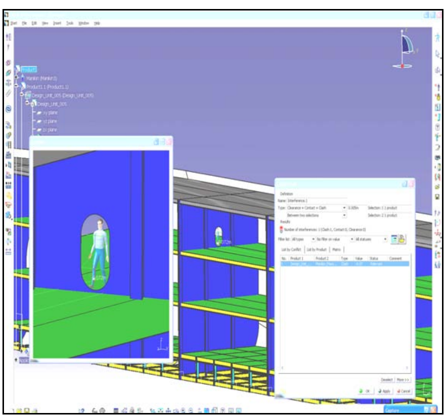
4a



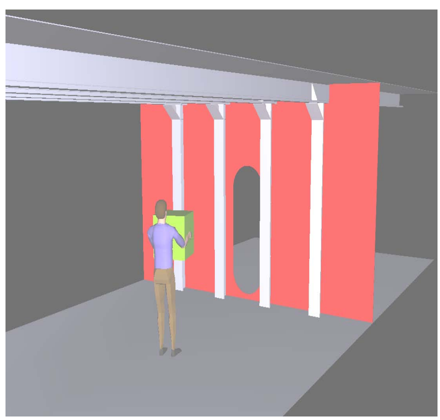
4b

Fig. 4a. Constraint based CAD solid modeling **Fig. 4b.** Generative structural analysis

Ergonomics Design and Analysis workbench of the CATIA V5 is utilized simultaneously with CAD construction tools to visually confirm posture, position and visual ratio of the inserted human model respect to ship structure. A door panel is created on one of the transversal bulkhead plates by carefully assessing the inserted human model's anthropometric information. (height, arms span width, etc.) Location of the door (door height from deck plate) and door size is visually fixed in regard to the inserted human model. After the door panel is created, a walking posture is assigned to human model and visual checks (vision, reach envelope, clash detection, etc.) applied to ensure that human model is visually legal. (No penetration between



5a



5b

Fig. 5a. CATIA Clash Test **Fig. 5b.** JACK PUT task simulation human model and panels, no visual awkward body posture) (see Fig. 5a.) The given posture and the design unit are then exported to JACK software. The assigned walking posture in CATIA V5 is mimicked in JACK by using built-in JACK human models. The mimicked posture assigned to JACK model is then inserted into Task Simulation Builder module to simulate a walking task through the door panel. The assigned walking simulation is used to check if the door height and width is blocking the passage. Furthermore, a PUT task simulation with 25kg cubic load (0.5m x 0.5m x 0.5m) is assigned to human model to verify if the walking passage is still valid with assigned load. (see Fig. 5b.)

3 Results

CATIA V5 multi-task design and analysis interface gives a great flexibility to user by incorporating complex CAD modeling, generative structural tests and human visual performance tools on a single software package. Throughout the design and analysis development, CAD models are exported to JACK software to assess Driver Comfort Analysis. Depending on comfort analysis, several dimensions of the driver's cockpit is changed to accommodate optimum driving standards and comfort. It was observed that this dual file transfer/editing protocol available between CATIA V5 and JACK integration increased the speed and the quality of product development while reducing time needed for virtual prototyping and redundant changes. The analysis of critical body joints and clash detection between human and race car improved the safety and comfort of the driver. The Occupant Packaging Report shows that the current cockpit design meets the ergonomics, safety and design objectives. In addition, the advance Real-time Rendering workbench of CATIA V5 also enhanced the product development efforts by adding an aesthetic touch to the CAD model.

The literature review and the expertise developed during Formula 1 race car design shows parallel potentials in development of marine vessels by using DHM and PLM tools. The lumped parameter model (multi-degrees of freedom) used by Z.Zong and K.Y.Lam [2] is an improved way over SDF (single-degree-of-freedom) to asses the effects of blast motions in marine vehicles. The simplified mechanical system shows a good approximation of body parts under the stress of underwater shock motions. However further research and development is needed to investigate design, analysis and simulation of complex systems with the aid of advance CAE software applications which incorporate powerful visual and mathematical analysis capabilities. Also, CATIA V5's Structure Detailed System/Object Design module is a powerful tool in ship structure design and user friendly interface eases the design and analysis efforts. Integrating this PLM solution package with dynamic DHM packages may enhance the capabilities of current DHM and PLM implementations.

4 Conclusion

PLM approach proved its strength and capabilities by resolving broad engineering and business challenges. DHM have significant impact on the product design and it is able to shorten design-to-built process time and cost [17]. Also, DHM applications help engineers to improve ergonomics, safety and human factors in design and product development [15]. The literature review, Formula 1 race car and marine vessel design practices show that integration of DHM and PLM can be a common solution for complex product development challenges. Advance vehicles can be designed and checked for human safety and comfort prior to prototyping and production. Design of complex systems and their interaction with the humans can be simulated through DHM and PLM integration. Overall, DHM and its integration with PLM have potential to improve the product development challenges and provide control of the entire process of designing and analyzing a product before it's ever launched.

Formula-1 race car and marine vessel design discussed in this paper illustrates a product development methodology through concept sketching, solid modeling, mechanical analysis, manufacturing, and marketing under a single frame. One of the most important aspects of the product development challenges studied in this paper is that up-to-date engineering techniques are harmonized with aesthetics skills to demonstrate a versatile marketing opportunities for a profitable product. This advanced integration needs further development and research efforts however studies and applications motivate that the integration between DHM, PLM and CAE tools would be a probable solution for design/analysis and simulation of complex vehicles/systems.

5 Future Work

Observations and results presented in this paper were motivated by the need to develop advanced simulation and design methodologies for navy ships. Despite the advantageous of mathematical models discussed in the paper by Z.Zong and K.Y.Lam [2], still more studies are needed to incorporate visual/virtual design components to mathematical models. Previous studies discussed rely on effectiveness of crash test dummies and accelerometers installed in the navy ships and data is collected by live-firing tests. This time consuming tests are raising the questions about safety, cost effectiveness and environmental issues. The elimination of live firing tests by computer simulations have potential to improve the safety, minimize the environmental and financial concerns and reduce the time needed for analysis phase of the design. One possible solution for above challenges can be the integration of DHM and PLM tools with advanced CAE applications which have the capabilities of dynamic analysis of humans in mechanical systems.

Products that are designed and analyzed in this paper rely on the integration between CATIA V5 and UGS Tecnomatix JACK. Both software applications and developed products showed potential improvements in design and analysis of complex systems where human-machine integration is present. However, the demands and constraints of reviewed naval project [2] require a comprehensive approach in utilization of DHM and PLM tools. In order to simulate an underwater shock and to investigate its threats to ship and shipboard personnel, an advanced integration of DHM, PLM and CAE applications with the finite element and dynamic human-machine analysis capabilities are needed.

One of the industry wide available dynamic human-machine simulation software is the MADYMO, developed by Tass Inc. MADYMO is the worldwide standard for occupant safety analysis featuring generic multi-body and finite element capability, a full range of predictive and efficient dummy models and tools to support the restraint optimization process [18]. There are available CAD solid models, Finite Element and Fluid Dynamics models/analysis to observe the effects of the blast motion on naval ships however additional efforts are needed to combine these models with DHM tools to have a comprehensive dynamic human-machine simulation. Integration of MADYMO with available PLM and DHM tools may make this simulation possible. The MADYMO software can be used to simulate complex dynamic response of humans and mechanical systems subjected to impact/blast forces. The paper [1] supported by U.S. Navy demonstrates the capabilities of MADYMO's prediction of response of human in blast-induced blunt trauma by replicating an experimental anthropometric crash test dummies placed in an office space where they exposed to blast forces. Results show that the MADYMO has ability to demonstrate the effects of the blast motions applied and it properly validates human-machine interaction virtually and mathematically [1]. Even though the experiment environment is different (building vs. marine vessel); still there are considerable similarities in simulation development. The ship by itself can be considered as a steel reinforced building where each individual deck beams corresponds to stories. Also, the transverse beam and longitudinal frame of the ship can be considered as columns and arches of a building. Also the biggest application similarity would be each office in building can be thought as compartment in a ship.

Further research and development is needed to resolve the challenges discussed above. A more extensive integration of DHM, PLM and CAE solution packages, (includes dynamic human analysis tools) are demanded to satisfy the comprehensive requirements of complex product design, analysis and simulation needs. However, expertise and motivation gained throughout this study and the literature review demonstrate that technological improvements in CAE tools may enhance this integration and extend the application capabilities of DHM and PLM tools.

References

1. Bogosian, D., Avanessian, H.D., Ashbery, W.: Analytical and Experimental Studies to Predict Response of Humans to Blast-Induced Blunt Trauma 2001. In: Presented at 10th International symposium of Munitions eith Structures, San Diego, California 2001 (2001)
2. Zong, Z., Lam, K.Y.: Biodynamic Response of Shipboard Sitting Subject to Ship Shock Motion. Journal of Biomechanics 35–43 (2002)
3. Chaffin, D.B.: On simulating human reach motions for ergonomics analyses. Human Factors and Ergonomics in Manufacturing 12, 235–247 (2002)
4. Teresko, J.: Digital Manufacturing: Not Only Manufacturing. Industry Week (November 2006) http://www.industryweek.com/ ReadArticle.aspx?ArticleID=12821
5. CIMdata, What is PLM? CIMdata (2006) http://www.cimdata.com/PLM/
6. Duffy, V.G. (ed.): IE 590D Applied Ergonomics. Spring 2006. Purdue University (November 2006) http://web.ics.purdue.edu/ duffy/index590D
7. Duffy, V.G. (ed.): IE 656 Digital Human Modeling for Design. Spring 2006, Purdue University (November 2006) //web.ics.purdue.edu/ duffy/index656
8. Teresko, J.: Process Planning: Human Simulation. Industry Week (November 2006) http://www.industryweek.com/ReadArticle.aspx?ArticleID=12822

9. Sundin, A., Ortengren, R.: Digital human modeling for CAE applications. In: Salvandy, G. (ed.): Handbook of Human Factors and Ergonomics, pp. 1053–1078. John Wiley & Sons, New York (2006)
10. Yang, J., Abdel-Malek, K., Nebel, K.: Restrained and unrestrained driver reach barriers, SAE Technical Paper 2004-01-2199
11. Chaffin, D.B., Faraway, J., Reed, M.B., Woolley, C., Rabourn, R.: From modeling to improved ergonomics in designed systems, 2005, University of Michigan HUMOSIM Laboratory, p. 27 (2005)
12. Li, K., Duffy, V.G., Zheng, L.: Universal accessibility assessments through Virtual Interactive Design. International Journal of Human Factors Modeling and Simulation (in press, 2006)
13. Duffy, V.G.: Using the Virtual Build methodology for computer-aided ergonomics and safety. Human Factors and Ergonomics in Manufacturing (in press, 2006)
14. Reed, M.P.: Research: Digital Human Modeling University of Michigan (2006) http://mreed.umtri.umich.edu/mreed/research_dhm.html
15. UGS Tecnomatix Jack. UGS (November 2006) http://www.ugs.com/products/Tecnomatix/human_performance/jack/
16. Catia V5 Dassault Systemes (October 2006) www.3ds.com
17. Chaffin, D.B.: Digital Human Modeling for Vehicle and Workplace Design. Society of Automotive Engineers, Warrendale, PA (2001)
18. Tass TNO-MADYMO. MADYMO (January 2007) http://www.madymo.com/cms/index.php

Bridging the Gap: Exploring Interactions Between Digital Human Models and Cognitive Models

Robert G. Feyen

Department of Mechanical and Industrial Engineering
University of Minnesota – Duluth
rfeyen@d.umn.edu

Abstract. For years now, most researchers modeling physical and cognitive behavior have focused on one area or the other, dividing human performance into "neck up" and "neck down." But the current state of the art in both areas has advanced to the point that researchers should begin considering how the two areas interact to produce behaviors. In light of this, some common terms are defined so researchers working in different disciplines and application areas can understand each other better. Second, a crude "roadmap" is presented to suggest areas of interaction where researchers developing digital human form and other physical performance models might be able to collaborate with researchers developing cognitive models of human performance in order to advance the "state-of-the-art" in replicating and predicting human performance.

Keywords: Human performance, cognitive models, digital human models.

"The modeling of cognition and action by individuals and groups is possibly the most difficult task humans have yet undertaken. Developments in this area are still in their infancy. Yet, important progress has been and will continue to be made."[1]

1 Introduction

Indeed, considerable progress is being made in our ability to model humans and their behavior. With the improvements in computing technology over the past three decades, we have been fortunate to have researchers able to take advantage of the increases in processing speeds, graphics capabilities, and computational power in order to generate digital human form models with increasing levels of realism. The same improvements have allowed other researchers to specify computational models that predict a human's physical performance and cognitive function with surprising levels of accuracy.

However, the claim made above that work in this area is still in its infancy may be more apt than ever. Those familiar with child development might observe that our digital human form models have less cognitive abilities than an infant and, in large part, are still developing the neuromuscular patterns for coordinating and executing limb and other body motions. They might also observe that our cognitive models still have difficulty perceiving much of the environment around them and figuring out

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 382–391, 2007.

what to do in response without us helping them out by telling them what's going on around them and what they should do. Finally, they would likely observe that these models typically are developed and used separately from one another.

They might then argue that, since physical and cognitive development in an infant are intertwined, perhaps it would make sense to stop dividing human performance into "neck up" and "neck down" (a distinction made by many in human factors and ergonomics between cognitive and physical performance) and instead begin considering how the two interact to produce behaviors. However, as many things in research and life, this is far easier said than done. But the argument has strong merit; thus, the two main goals of this opinion paper are to (1) define some common terms so researchers working in different disciplines and application areas can understand each other better and (2) present a crude "roadmap" that suggests areas of interaction where researchers developing digital human form and other physical performance models and researchers developing cognitive models of human performance might be able to collaborate in order to advance the "state-of-the-art" in replicating and predicting human performance.

2 Defining Common Terminology

For over a century now, researchers have been working to understand the nature of the various interactions between man and machine while engineers and designers have been attempting to apply this understanding to the design of products and systems. But to do this effectively, the engineers and designers must be able to readily access and utilize this information in the earliest stages of the design process. Designers already have access to tools that can predict hardware and software performance; but to consider human-machine interaction in the conceptual stages of a design, they need tools that can predict how the human will perform using task-based metrics associated with the speed, actions, timing, and errors that may arise while undertaking a task.

Such tools often are called "human performance models" or HPMs. HPMs can be found throughout a diverse range of disciplines and all attempt to represent or describe some aspect of human performance. Disciplines include psychology, cognitive science, engineering, artificial intelligence, computer science, biomechanics, physiology, and medicine. HPMs have been used in the design of complex systems ranging from industrial processes and nuclear power plants to surface and air transportation systems; they have also been used extensively in the field of human-computer interaction. Types of human performance modeled by these approaches include perceptual functions such as signal detection, visual scanning, object recognition, and motion estimation as well as models of various cognitive functions including problem solving, memory recall, training, skill acquisition, and errors. Other approaches consider a person's anthropometric and biomechanical characteristics while yet other approaches provide insight into the workload - both mental and physical - experienced by a person. Some approaches distinguish between individual operators and group operations (e.g., multi-operator systems such as found in passenger jet cockpits) while others distinguish between performance on a single task and performance in situations where several tasks must be completed at the same time. This diverse range of applications and disciplines makes discussing tools for predicting human performance difficult since the same terminology often has different meanings, depending on who is using the tool.

To avoid confusion, the terms "approach", "framework" (or, "architecture"), and "model" should be defined separately since they are often used interchangeably in the HPM literature and commonly confused.

- An *approach* encompasses the methodology used to specify the task, environment (including equipment), and human in a consistent manner and the means by which this information is incorporated into the modeling efforts to generate measures of performance.
- The *framework* is a general structure that contains elements and parameters common to any scenario to be modeled (e.g., information processors and their default service times). The task, environment, and human information as specified by the approach indicate which elements of the framework are required and which parameters must be altered when modeling a specific system scenario.
- The resulting structural elements and the set of parameters representing a specific system scenario are called the *model*.

Another area of confusion pertains to the output of human performance approaches: these approaches yield *engineering* models, not *psychological* models. In the psychological literature, a "human performance model" is narrowly focused, with the term often used synonymously with a theory of performance. Typically, a human performance model in the psychological literature is independent of the system or task contexts in order to be applicable in a wide range of settings. A few of these models tend to be analytic and limited in scope with an exact solution for a given set of parameters (e.g., Fitt’s law) while the majority of others are qualitative models (e.g., principles of stimulus-response compatibility). However, their primary purpose is to evaluate and develop the theory underlying the model. When a model does not agree with data obtained or behaviors observed from the real world, the model and theory are either discarded or reworked until agreement is reached. This differs from the engineering approaches in which, regardless of the characteristics of a particular model, a model has value as an engineering or design tool as long as the model can aid decision making by yielding simulated performance data that approximates real performance data within acceptable limits.

Yet another complicating issue is defining what one means by “model”. For example, researchers in one discipline may conceive of models that differ dramatically from what researchers in another discipline might call a model. However, Baron *et al* [2] and Elkind *et al* [3] have proposed useful classifications to help bridge that gap. In particular, they identified eight (8) pairs of characteristics that can be used to define the purpose of a given model:

1. A *predictive* model is capable of accurate performance predictions without requiring extensive validation data. On the other hand, the parameters of a *descriptive* model must be adjusted in order to fit the model to existing performance data. Obviously, a predictive model is more useful to a decision-maker and, if accurate, also classifies as a descriptive model.
2. Primarily focused on system performance, a *top-down* model starts with a set of system goals, which are then decomposed into progressively smaller sub-goals and

processes until a level is reached at which the resulting processes are deemed as "primitives" (which cannot be broken down any further). A *bottom-up* model starts with a predefined set of primitives and links various primitives together to form an overall system. This approach concentrates primarily on the process that dictates the system's performance, although the performance itself is also considered.

3. *Simulation* models imitate human performance and, by manipulating the inputs and outputs, allow interactions between the human, the system, and the environment to be simulated and analyzed. *Analytic* models represent human performance mathematically, providing concise descriptions that in some cases may be referred to as "laws" of human behavior.
4. A *process* model attempts to predict the human output by representing the specific human processes used to accomplish the task. A *performance (or, output)* model attempts to predict performance without any consideration of the human processes involved and places no requirements on the structure or validity of the inner workings of the model. In many situations, output models are sufficient for the purposes of analysis; but, since process models are applicable to a wider range of tasks and conditions, they are the more powerful of the two.
5. A *descriptive* model describes the current human behavior, performance, or process while a *prescriptive* model (also known as normative or rational action) suggests the human behavior or process that will lead to optimal performance.
6. A model may have either *depth* or *breadth*, depending on the amount of detail desired and the level at which behavior is being modeled.
7. Elkind *et al* [3] define a model focused on a single task as *limited* whereas *comprehensive* refers to a model capable of addressing a variety of different tasks. However, this definition can be expanded to include the range of human performance issues addressed by an HPMA. By necessity, early attempts at formulating HPMAs addressed only limited aspects of human performance (e.g., visual scanning or strength capabilities) since modeling multiple aspects was too complex and time-consuming to evaluate. But continued research combined with rapid advances in computer technology has rekindled efforts to integrate numerous aspects of human behavior into a single, coherent modeling framework. For example, recent comprehensive approaches such as SOAR [4], ACT-R [5], MIDAS [6], EPIC [7] and the QN-MHP [8] are capable of predicting certain aspects of performance accurately across an increasingly wider range of tasks and human behaviors.
8. While *qualitative* models based on verbal theories and empirical data can be useful in analyzing human performance, they are not especially useful for performance prediction. Some authors refer to certain *quantitative* models as "formal models", defined as those arising from one of two distinctly different types of approach: those based on mathematical theories and those based on computational techniques.[1]

[1] In March 2001, the Human Factors journal published a call for participation regarding a special issue to be titled "Quantitative Formal Models of Human Performance." In the call, the co-editors (Wayne Gray and Michael Byrne) state "…'Formal models' are construed to include both computational and mathematical models…"

Even within these latter two approaches to quantitative models, different conceptual frameworks are used to elicit measures of human performance. For example, linear statistical and control theoretic frameworks can be identified separately within the mathematical category while information processing, task networking and knowledge-based frameworks fall under the computational umbrella [9, 10]. This delineation is not absolute as recent advances in many frameworks borrow features or techniques from various other frameworks. This trend towards integration undoubtedly will continue as no single approach is likely to provide a comprehensive general model capable of describing or predicting human behavior in various task settings (see Gluck and Pew [11] for an excellent review of the state-of-the-art in integrated approaches).

3 Interactions Between Physical and Cognitive Models

For this same reason, physical and cognitive models of human performance also need to be integrated. Approaches such as HOS [12] and MIDAS [13] as well as work by researchers at Sandia National Laboratories [14] represent initial efforts to integrate physical and cognitive models into a single approach by including independent "modules" (micromodels containing algorithms based on human experimental literature) that can be used as necessary to predict an aspect of performance. One criticism is that performance predicted by these "plug and play" modules actually may not be independent; that is, interactions between different aspects of performance may produce behaviors and performance that the modular approach would have difficulty predicting accurately.

But is this criticism valid? The answer is – no one seems to know. Published research on the interactions *between* various models is scarce, if not nonexistent. What interactions need to be studied? Space does not permit a comprehensive discussion of this, but three major types of interaction warrant further attention – physical, physiological, and emotional. But before discussing these, three external influences impacting these interactions of cognitive and physical performance should be noted: the activities of the person, the equipment available to the person for carrying out these activities, and the environment in which the person acts.

3.1 External Influences Impacting Interactions

The first external influence is the person's *activities*. This encompasses more than the elements of a specific task that a person may be attempting to complete. Models of task performance often fail to capture the full range of activities undertaken by a human, primarily because these tasks occur in a broader context which may include elements unrelated to the task at hand. The source of these "other" activities may be rooted in the person (e.g., physiological needs requiring some response or prospective memory leading to performance of another task), the situational context (e.g., task deadlines or other externally defined performance requirements, safety implications, local events), or the environment (e.g., workplace layout, climate, ecology). Further, these activities impact the person's physiological system, having effects that may accumulate over time (both short term and long term) and trigger a physiological response that must be addressed by an activity that interrupts the current task.

The second external influence is the *equipment available* for use by the person to carry out these or other activities. Once again, not all of the available equipment may be used by a person in a given location or time frame to complete the task at hand, but may simply be present. However, such equipment has potential to produce distractions or interruptions leading to activities unrelated to the immediate task at hand. For example, telephones, e-mail and instant messaging alerts, instrument displays, and alarm monitors can produce responses that detract from the performance of the task at hand. Secondly, many activities can be completed satisfactorily using different types of available equipment, requiring the person to choose between these options and the resulting strategies in order to complete the chosen activity. For example, imagine the various ways in which people can write reminders to themselves. At any given time, the method selected often depends on the type of writing surface available (sticky notes, scrap paper, one's hand), type of writing utensils (pens, pencils, markers), computer or phone access (send oneself an e-mail or voicemail), available surfaces for attaching notes (e.g., bulletin boards, whiteboards, walls), etc.

The person's *environment* acts as the third external influence. The workplace itself may hinder or facilitate activities, depending on housekeeping practices as well as the organization and location of equipment and objects that must be seen, heard, or manipulated. Activities taking place in warm, dry conditions may require additional steps or entirely different methods if carried out in cold, wet conditions. As with activities, physiological responses to environmental conditions often detract from task performance and may lead to interruptions as the person carries out an activity unrelated to the task at hand in order to address a basic physiological need. Other environmental conditions such as noise, poor surface conditions, vibration, rain and sunlight often produce responses extraneous to those needed to complete a given task or lead to errors in the performance of the task.

These three influences (activity, equipment, and environment) should not be surprising to anyone working in the area of either digital human (physical) or cognitive modeling. But what becomes apparent in this discussion is the strong bidirectional relationship between the physical and cognitive aspects of human performance. At all times, a person's actions and physical presence in the environment and perceived proximity to nearby objects along with the person's internal physiological responses generates information used by the cognitive system to determine what to do next. Taking this diverse information into account with the situational context, the cognitive system attempts to select a strategy likely to be feasible, successful, and appropriate *for that moment in time*. For example, importance of an activity, the available time, the time required to complete an activity, and the individual's appraisal of their ability to complete an activity successfully (i.e., self-efficacy) are significant factors impacting which activity or set of activities a person chooses to do at any given time [15, 16, 17, 18]. However, once a strategy is selected, the cognitive system must also dictate how the physical system will respond while accounting for the person's functional limitations, equipment capabilities, and environmental constraints. Once initiated, this response impacts the status of all activities currently underway, alters the environment, and affects the body's physiology and its position relative to any available equipment. And the cycle repeats. Notably, this iterative balancing act implies that not all activities carried out by an individual will be rational, at least not in the narrow context of achieving a single task-oriented goal as assumed in many cognitive modeling approaches. Thus, to improve our ability to model human performance, we need a better understanding of how these various sources of information influence the choice of activities, strategies and responses – and vice versa.

3.2 Interactions Between Models

3.2.1 Physical Interactions

Performance monitoring. From the digital human modeling standpoint, perhaps the most straightforward area of interaction to explore is between a person's physical and perceptual limitations and their response selection. For example, a number of digital human modeling approaches already include tools to identify constraints on performance such as range of motion, posture and strength capability assessment. When placed in a virtual environment, most digital human form models can be used to analyze reach and clearance envelopes as well as determine the field of view at a given moment. Some of these modeling approaches can generate realistic motion paths of the limbs and the whole body; many can also detect collisions between the person and an object. However, other than the modeling approaches mentioned at the outset of this section (e.g., HOS, MIDAS), the digital human models rarely, if ever, utilize any of this information internally to monitor physical safety and alter its performance in real-time. For example, what will most digital human models do if they encounter a load exceeding strength capabilities? Nothing different; the model will perform as scripted even if the task requirements exceed physical capabilities. Rather, safety and performance issues are identified mostly post-hoc.

Information Acquisition. On the other side of this interaction is the need of the cognitive system to acquire information in support of current activities. Most cognitive models assume that the information needed to complete the task being modeled is readily accessible, either directly or using some standard sequence of motor actions along with perfect knowledge of the location of the information source. But the role of search and eye movements in acquiring information is not a trivial issue. Add any need for physical actions other than eye movements to assist in the search process (e.g., postural or position changes, rearrangement of objects in the environment, etc.) and the impact is magnified further. Or, if information sources are present, but cannot be reached or inadequate clearance is present for a person to access them, the overall task of information acquisition will fail. In fact, Gray and Boehm-Davis [19] argue that a person is capable of selecting between different strategies for completing a task step on the basis of differences as small as 150 milliseconds. Further, they show that subtle differences in the location of information sources can bring about changes in which strategy is selected, resulting in significant effects on task performance. By utilizing the ability of the digital human model to track body position and posture relative to equipment in the environment as well as the body movements required to access information sources, one can account for the interaction between body position and information access. In many scenarios, this would reduce the variation between predicted and actual performance measures.

Errors. Another area of interaction to explore between digital human and cognitive models is the impact of errors. Given that most digital human models are scripted for performance, movements tend to be targeted and executed perfectly. Yet, as we all know, perfect performance rarely occurs in real life. And when the wrong button is pressed or an object is knocked off the desk, etc., we typically have to do something to recover from this error. This interaction between the imprecision of human movement and the need for cognitive strategies to recover from any resulting errors has a major impact on our ability to predict true human performance accurately. From the cognitive side, errors in navigation, decision making and information acquisition often add layers of complexity to the physical requirements of a task which, among other things, can lead to even more errors, higher likelihoods of failure, longer completion times, fatigue, and even injuries.

3.2.2 Physiological Interactions

Another major area of interactions is the effect of physiological influences on cognitive and physical performance. Human behavior is rife with apparently extraneous movements as itches are scratched, noses are blown, and postures are changed to improve comfort in response to physiological triggers. These unintentional behaviors produce variations in body positions and activity sequences that impact overall task performance. Increased heart rate and respiration act as indicators for localized and whole body fatigue; but whereas both physical and cognitive task performance are known to degrade with fatigue and can be accounted for in performance models, shouldn't models also account for any physical actions counter to the cognitive performance goals that become necessary to prevent physical harm and aid recovery? Temperature effects on both physical and cognitive performance are well-documented; but once again, physical actions counter to the cognitive performance goals may also be necessary. To complicate this even further, what strategies, if any, do people use for deciding when and what physiological actions to interject in a task?

3.2.3 Emotional Interactions

The most difficult interactions to consider will involve the influence of emotion on physical and cognitive performance. Anecdotally, strong negative emotions such as anger, fear and frustration often lead to actions that degrade task performance (e.g., performing too quickly, incorrect strategy selection, quitting tasks before completion or performing actions unrelated to any task at hand). Methods such as facial gesture recognition and physiological monitoring hold promise for identifying general emotions during human task performance (e.g., augmented cognition methods as described in Schmorrow [20]), but ethical issues involved in provoking these strong negative emotions in a controlled laboratory setting will need to be addressed before models accounting for the interaction of these emotions and task performance can be developed.

4 Conclusion

Ultimately, we are looking for human models that, among other things, can adapt to the environment, get tired, make mistakes and become frustrated or angry - just like we do. Clearly, a lot of work is yet to be done. The empirical studies and modeling efforts required will be time-consuming enterprises and, as noted at the outset, are among the most difficult tasks a researcher could choose to undertake. But, modeling human performance - whether the physical form, the cognitive function, or some other aspect – is rapidly approaching a point where efforts should not (and perhaps cannot) be undertaken without collaborative efforts to put the results in the proper human context and to understand the inevitable interactions. Despite these challenges, this is a fascinating and gratifying field in which to work and, through all of these efforts - collaborative and otherwise - we should be able to watch our current human models grow from infancy and perhaps even reach adulthood.

References

1. Pew, R.W., Mavor, A.: Modeling human and organizational behavior: Applications to military simulations. National Academy Press, Washington DC (1998)
2. Baron, S., Kruser, D., Huey, B. (eds.): Quantitative modeling of human performance in complex, dynamic systems. National Academy Press, Washington, D.C (1990)
3. Elkind, J.I., Card, S.K., Hochberg, J., Huey, B.M. (eds.): Human performance models for computer-aided engineering. Academic Press, Inc., San Diego (1990)
4. Newell, A.: Unified theories of cognition. Harvard University Press, Cambridge, MA (1990)
5. Anderson, J.R., Lebiere, C. (eds.): Atomic components of thought. Lawrence Erlbaum, Hillsdale, NJ (1998)
6. Laughery, K., Corker, K.: Computer modeling and simulation. In: Salvendy, G. (ed.) Handbook of Human Factors and Ergonomics, 2nd edn., Wiley, New York (1997)
7. Kieras, D.E., Meyer, D.E.: An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. Human-Computer Interaction 12(4), 391–438 (1997)
8. Liu, Y., Feyen, R., Tsimhoni, O.: Queuing network-model human processor (QN-MHP): A computational architecture for multi-task performance in human-machine systems. ACM Transactions on Computer-Human Interaction 13(1), 37–70 (2006)
9. Sheridan, T.: Humans and automation: System design and research issues. HFES Publications, Santa Monica, CA (2002)
10. Baron., et al. (1990)
11. Gluck, K.A., Pew, R.W.: Modeling human behavior with integrated architectures: Comparison, evaluation, and validation. Lawrence Erlbaum, Hillsdale, NJ (2005)
12. Lane, N.E., Strieb, M.I., Glenn, T.A., Wherry, P.A.: The human operator simulator: An overview. In: Moraal, J., Kraiss, K.F. (eds.) Manned System Design: Methods, Equipment, and Applications, Plenum Press, New York (1981)
13. Laughery and Corker (1997)
14. Bernard, M., Xavier, P., Wolfenbarger, P., Hart, D., Waymire, R., Glickman, M., Gardner, M.: Psychologically plausible cognitive models for simulating interactive human behaviors. In: Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society, 2005, pp. 1205–1210 (2005)

15. Van Eerde, W., Thierry, H.: Vroom's expectancy models and work-related criteria: A meta-analysis. Journal of Applied Psychology 81(5), 575–586 (1996)
16. Vroom, V.: Work and motivation. Wiley, New York (1964)
17. Kernan, M.C., Lord, R.G.: Effects of valence, expectancies, and goal performance discrepancies in single and multiple goal environments. Journal of Applied Psychology 75(2), 194–203 (1990)
18. Klein, G.A.: Recognition-primed decisions. Advances in man-machine systems research 5, 47–92 (1989)
19. Gray, W.D., Boehm-Davis, D.A.: Milliseconds matter: an introduction to microstrategies and to their use in describing and predicting interactive behavior. Journal of Experiment Psychology: Applied 6(4), 322–335 (2000)
20. Schmorrow, D.: Foundations of augmented cognition. Lawrence Erlbaum, Hillsdale, NJ (2005)

Translating User Experience to Requirements

Ji Gao, Glen Anderson, Brian W. Bramlett, Ryan Palmer, and Delbert Marsh II

5200 NE Elam Young Pkwy,
Hillsboro, OR 97124, USA
{jade.gao,glen.j.anderson,brian.bramlett,ryan.palmer,
delbert.k.marsh.ii}@intel.com

Abstract. In this paper we introduce the Usage-to-Platform Requirements (U2PR) process in the context of a platform product lifecycle at Intel that involves challenges beyond the development of an individual product. We describe the types of research necessary to inform the process, the methods we have evolved in developing use cases and usage requirements, and the practice used to communicate the requirements to the right audiences. The U2PR process provides a practical approach to fill the gap between descriptions of marketing opportunities and actionable, technology capability definitions and experience-quality requirements. It demonstrates how one can apply a general user-centered design process to strategic planning and design.

Keywords: platform, user experience, usage model, use case, requirement.

1 Introduction

A platform, as understood within Intel, is a set of products that one or more companies deliver to the market as an integrated solution. One of the principal aspects of Intel's transformation from a microprocessor company into a platform company is the building of a comprehensive customer focus to complement technology excellence. Delivery of platform value depends on standardization, and an ecosystem of products and services that support and enable the delivery. Because open, evolving platforms are not limited to a single product type, or even a product line, they are especially challenging to design. Adding to the challenge, the same platform may be delivered through a range of business models in multiple countries, may need to accommodate multiple types of users and uses, and may be required to change rapidly to accommodate changes in technology. In order to manage these challenges, the business, usage, and technology perspectives have to be integrated as a coherent system, referred to as Intel's "Three-Circle Model" [1]. The Usage to Platform Requirements (U2PR) plays an important role in linking all three aspects together.

Prior to the implementation of U2PR process, platform planners and technology architects had difficulty interpreting the usage models identified by Intel marketing teams. In particular, the usage model descriptions did not include the detail necessary to generate actionable engineering requirements. A high-level usage model such as "consuming media anywhere and anytime within the home" had a good deal of thought and research behind it, and the planners also had anticipated what

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 392–398, 2007.

technologies would likely be employed to provide such an experience; however, handing off the concept to architects with only this level of detail had several problems. The U2PR process was created to specify the usage details and translate it to the usage requirements that could help technology architects design the platform appropriately to deliver the intended user experience. It found its first implementation in the work in 2004 aimed at the Enterprise and Home platforms planned for 2007 and now has been adopted for multiple platforms across Intel.

1.1 U2PR in the Platform Product Lifecycle

The focus of usage-oriented roles evolves over the course of the platform product lifecycle as knowledge and decisions shape the platform. For most of platform or product companies, there are four major phases in the Platform Product Lifecycle (PPLC) characterized as discover, design, develop, and deploy, as shown in Figure 1.

The discover phase is focused on identifying opportunities and establishing strategy. Emerging usage trend-analysis and synthesis from research identify usage models and architectures. The design phase concentrates on transforming strategy into actionable and effective plans. The bulk of U2PR efforts occur during this phase, as key use cases determine specific platform requirements and feasibility. In the develop phase, engineering consumes platform requirements while generating product specifications and implementation details. Development teams must also determine whether the products under development meet U2PR targets for experience quality and usability. The deploy phase deals with getting products built, delivered, and supported. Experience design can encompass the entire "usage lifecycle," ranging from the initial search for a product to the out-of-box experience to ongoing maintenance. U2PR is a key component of delivering on the promises intended during early planning, and of realizing the value of the platform. The following sections will introduce the U2PR process and the scope of work in greater detail in the context of PPLC, as demonstrated in Figure 2.

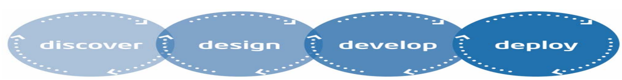


Fig. 1. Platform Product Lifecycle (PPLC) phases

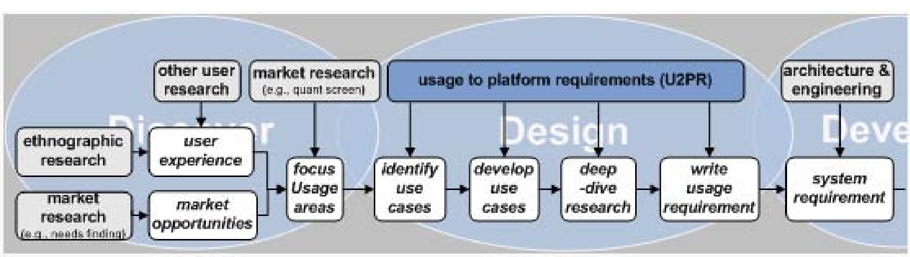


Fig. 2. Role of U2PR in the Platform Product Lifecycle (PPLC)

2 Before U2PR – Identifying Usage Models

Before the U2PR process begins, a process called usage modeling gives Intel strategists a tool for depicting how a particular technology will be perceived and adopted within the context of people's lives. Usage models enable a diverse group of technologists, strategists, and researchers within the company to begin to understand a product's impact on people before the product is actually built. Such information is gathered by market research, ethnographic studies, and review of existing products, etc. The development of usage models ensures a balance between business, technology, and user goals. Once a usage model has been identified as an opportunity for Intel, it is up to U2PR professionals to dig deeper by depicting customer experience, ease-of-use, system interactions, and value propositions, as shown in Figure 2. In preparation for U2PR work, U2PR professionals must have an adequate understanding of the potential business impact and current technology solutions and a thorough understanding of user needs.

3 U2PR Process

The core steps within U2PR are identifying the use cases, developing the use cases, conducting deep-dive research, and writing usage requirements, as shown in Figure 2.

3.1 Identifying Use Cases

As stated earlier, a usage model describes the interaction between a user and a system and demonstrates the benefit of the system to the user within a specific context. A use case defines a more specific sequence of interactions between one or more actors and the system. More than one use case is usually needed to fully define a usage model.

In the example usage model cited earlier, "consuming media anywhere and anytime within the home," the U2PR professional would develop several key use cases to specify this usage model. For example - user selects and streams music from a PC to a device in the home; user adds music to a favorite playlist from any device on the home network; user manages playlists and streams to any device on the home network. It is important to identify a key set of use cases that cover the user types, device types, and media types that would be critical to consider for this usage model.

For platform companies such as Intel, multiple use cases are typically contained with a usage model, and usage requirements enable one or more use cases. While the platform company itself may not develop everything necessary to implement all of the use cases, it is still important to consider the usage model as a whole. This allows the platform company to identify needs and potential gaps for products provided by other companies in the ecosystem. Products to fulfill those needs and gaps can usually be categorized as hardware, software, services, or standards.

3.2 Developing Use Cases

In developing the details of use cases the main objective is to communicate the intended usage experience to others involved in developing a product. The primary vehicle for this communication is the use-case document. Use cases can take many different forms, which are intended for various purposes. Table 1 depicts the format and description of the elements that comprise Intel use cases as implemented by the User Centered Design Group.

Table 1. Use case elements

Identification
Use Case Title: Short name for reference purposes–the title should accurately represent its contents to avoid misinterpretation or confusion.
Use Case Tag: Unique identifier. Allows for traceability and common reference.
Summary: Brief description to highlight the main activity.
Context
Actors: People/systems who play a part in the use case. They may include formally defined personas (see discussion below).
Pre-conditions: Any pre-requisites or triggers for the use case.
Post-conditions: The end condition, the outcome of the use case.
Body
Sample Scenario(s): Narrative story, depicting a realistic situation, to demonstrate how the actors interact to satisfy the objective of the use case and the context in which the activity occurs. This is invaluable for communicating the intended user experience to architects and others.
Basic course of Events: Task flow visually representing the flow of information and interaction between the actors
Conceptualization: Screen shot, conceptual mockup, storyboard, or other graphical representation of usages to help the audience visualize the intended user experience.
Simultaneous Usage: Usages that may occur at the same time on the same system.
Alternate Paths: Narrative story depicting situations in which the interaction between the actors has deviated from the normal flow. It is important to capture the cause for deviation, and distinctions and considerations from the normal situation.
Exceptions: Listing of the instances when the use case would NOT be triggered.
Requirements Management
Priority: Rating to indicate the importance this use case holds with respect to meeting the usage model objectives.
Author: Name of the individual who wrote/developed use case for tracking and reference purposes for any questions regarding the specifics of the use case.
Usage Owner/Originator: Name of the individual who conceptualized the usage for tracking/reference purposes for any questions regarding the origins of the use case.

In developing use cases, it may be valuable to take an approach that Cockburn terms “Black Box Use Cases” [2]. With this approach the use case does not address specific technologies but focuses instead on the user and the desired usage experience, leaving the implementation open for alternatives and design innovation. At other times a technology may be central to defining how a use case unfolds. In these cases, the U2PR professionals may take a pragmatic approach that includes a particular technology, which may involve some limitations to the user experience, as part of the use case.

3.3 Deep-Dive Research

One approach that has been used to inform the U2PR work is referred to as “deep-dive” research. This approach is a combination of human factors and design research methodologies, which are commonly used in contextual inquiry and task analysis. The purpose of deep-dive research is to better understand user needs in the context of identified usage models and therefore to validate use cases and capture important requirements. As Figure 2 indicates, the deep-dive takes place at a later stage of PPLC. The findings from deep-dive research may require that the use cases be modified.

U2PR professionals conduct deep-dive research via a series of face-to-face, semi-structured interviews with individuals who are identified as representative of users that have shown interests in a particular usage. The first objective of the research is to understand the pain points in participants’ current user experience by encouraging users to talk about their daily work, frustrations, and even walking through typical tasks. The second objective is to gather user feedback on proposed technology solutions that enable the usage to determine what features are needed and how the platform should be designed to ensure the delivery of a desired user experience. The interviewer introduces the participants to a solution through conceptualizations (e.g., storyboards) and records their responses to pre-designed questions and elicits ratings of the appeal and relevance of various attributes of the proposed solution.

The essential output of ‘deep-dive’ is detailed information that allows usage models to be developed into a set of usage requirements, as explained in the following section. The deep-dive and the followed brainstorm and ideation activities would be helpful discovering new usages. A wide variety of methodologies may be applied to ideation and are covered in depth in Kelley [3]. Though low-fidelity prototypes and story boards that show example user interfaces may be useful in this type of research, deep-dive research generally precedes product design.

3.4 Writing Good Requirements

The requirements coming out of the ‘deep-dive’ research is referred to as usage requirements, which directly describe the human experience of the system. The technical requirements come from other sources, such as technical specifications and standards. The main tenets of effective usage requirement include complete, correct, clear, concise, consistent, coherent, connected, feasible, prioritized, and verifiable [4]. Going beyond these basic tenets, the most important consideration when constructing usage requirements is to at all times develop them with the users in mind. What are the users’ objectives? What interactions are reasonable to expect of them in the context of the activity that is taking place? Usage requirements must lay the necessary framework for the experience to satisfy the user.

Requirement writers should compose in a language that is easily interpreted by systems architects or software developers. The usage requirements communicate affirmative statements that clearly articulate what should be developed to improve the experience people have with the product. Each requirement also contains a rationale for why it needs to be considered when developing the technology to support the usage. The rationale points to primary and secondary research and data supporting the requirement. As with use cases, the focus is on the fundamental need to be met, not a specific technological implementation. An example of a usage requirement is as follows:

Name: Secondary Recording Status Display
Description: The PC shall present recording status when the PC is in the process of recording a TV show or capturing digital content, even if the main display is turned off.
Rationale: TV recording devices typically show a status light on the bezel to show recording in progress, allowing the user to see status without turning on a connected screen.
Priority: High – essential for competitive offering

Usage requirements can be categorized as functional and non-functional. The example above is a functional requirement, since it describes a particular function the system must perform. A non-functional requirement defines a quality of the system, such as how good the perceived video quality must be. Both types are outputs of the U2PR process.

4 After U2PR – Communicating Usage Requirements

The first key element to ensure the requirements are well consumed is getting requirements to the right people. Usage requirements are communicated both formally and informally to a diverse team of engineers, architects, software developers, planners, and designers involved with product development across the platform. To ensure user-oriented concerns are considered along with technology-centric approaches to platform architecture, user requirements are communicated with the rationale derived from research and human factors expertise. Mapping usage requirements to key technology capabilities allows architects to reference the end-user value that the technology enables. Reliance on more tangible implementations of the user experience is important. Aligned with this, U2PR professionals communicate requirements in the form of compelling storyboards, videos, vignettes, and concept prototypes. Visual conceptualizations also provide a tool for eliciting feedback on future usages that do not have an existing frame of reference. In addition, U2PR professionals use personas [5] in the context of a use case to show how a user's values and experience play a role in the user's interaction with a system. The persona gives product teams an opportunity to take a walk in target users' shoes and experience their goals and behaviors.

The U2PR at Intel ends with a set of usage requirements that feed an architectural analysis process. In some organizations, the same group that does a U2PR-like process may do this work as well, but for the scope of this paper, U2PR ends with the communication of the usage requirements. However, the user experience torch should be carried through prototype evaluation and post-shipment, through a “user experience quality” measurement program.

5 Conclusion

Companies that develop platforms composed of multiple products face special challenges in filling the gap between marketing plans and platform architecture. This creates key roles for U2PR professionals who can communicate across these product groups and organizations, and it requires the corporate to invest resources to enable the creation of a coherent platform. Another challenge is getting U2PR professionals in the correct mindset to define a platform, instead of only one product that will present the user interface. It requires the close work with ecosystem partners to ensure success in the final platform. Those challenges go beyond those inherent in individual product development. The U2PR process fills the gap using a combination of research, human factors methodologies, use case and requirements processes, and visual conceptualization.

Acknowledgements. The authors thank Paul Sorenson for recognizing the need for a U2PR process at Intel, for establishing the program, and for ongoing leadership and guidance. We thank Erik Simmons for helping to define the process and thank Herman D’hooge for fostering a corporate culture that recognizes the need to bring the user perspective into technology development.

References

1. Simmons, E.: The Usage Model: Describing Product Usage during Design and Development, IEEE Software, vol. 23, no. 3, May/Jun, (2006) pp. 34–41
2. Cockburn, A.: Writing Effective Use Cases, Addison-Wesley, Boston, MA, ISBN 0-201-70225-8, (2001)
3. Kelley, T.: The Art of Innovation, Doubleday, New York, NY, (2001)
4. Hooks, I. and Farry, K.: Customer-Centered Products, Creating Successful Products through Smart Requirements Management, Amacom, ISBN 0-8144-0568-1, (2001).
5. Cooper, A.: The Inmates Are Running the Asylum: Why High Tech Products Drive Us Crazy and How to Restore the Sanity, SAMS Publishing, Indianapolis. IN, (1999)

Involving Engineers in User Research and User Experience Design of ICT for China

Chaoyu Huang and Huogao He

2nd. Floor, Block 4, 555 Dong Chuan Road
Shanghai Zizhu Science Park
Shanghai, PRC, 200241
Chaoyu.huang@intel.com, kenny.he@intel.com

Abstract. Chinese culture and consumer patterns are dramatically different from those in the US and Europe. That greatly impacts the types of products that are relevant to the Chinese market, the product development life cycle and methods by which these products are defined and developed for information and communication technologies (ICT). To address those unique differences, the User Centered Design (UCD) research team at Intel China has developed and refined techniques for involving engineering staff in the early stages of product design, namely user research and experience design. This early involvement has many advantages and improves the overall effectiveness of the product development team. This article describes the role of the engineers in the early phases of the user centered process, and the learnings and challenges that come from this approach. Real-world case studies are used to illustrate the methodologies.

Keyword: ICT, user experience design, user centered design, user centered research, ethnography, brainstorming, field work.

1 Introduction

China, with its huge population and rapidly growing economic development has drawn the attention of the world as a huge new market and an opportunity for growth for companies that are able to capture some of that market. Many reports indicate that China is currently the second largest computer market not far behind the United States. ICT products like computers and mobile phones are quickly becoming an integral part of daily life for many Chinese.

As China opens its doors and markets to the world, major differences in culture, lifestyle and behavioral patterns render many proven Western and mature market business and commercial practices ineffective for China. Numerous reports and studies attempt to demystify China, but many of the new entrants in the Chinese market are often unsure of how to tackle this market. The market environment characterized by rapid growth, fierce competition, China unique considerations and fast changing consumer tastes also demands agility and a short product life cycle for product developers. Thousands of new products are flowing into the market everyday with most of them quickly moving out of fashion. That means that user research plays an even

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 399–408, 2007.

more important role in new product design and development: it requires an integrated understanding of market opportunity, the business ecosystem and user needs.

User research based on ethnographic methodologies have become increasingly popular in the field of human computer interaction (HCI) to inform product definition and span the product development life cycle from understanding user needs, to formulating relevant user requirements, defining user experience that responds to the need, product and technology development as well as the evaluation of new products and services.

The research practice of ethnography in China, and especially its use in an industrial context, has a relatively short history compared with the United States and Europe. The methods originated from the Western society are mostly based on Western culture and models of thought. So we are facing challenges in adapting these research methods for user research aimed at the product development for Chinese market. [4]

One big challenge is that academic ethnographic research in its purest form is very time intensive. “Typically, ethnography will take place over a period of several months with at least the same amount of time spent in analysis and interpretations of the observations [1]”. But it is almost impossible to spend months or even weeks in the field gathering data in the fast moving commercial world especially in China. Pressures of time to market, people and budgets force trade-offs to be made in the way research is conducted. [6] Those constrains are greatly impacting research design by keeping the focus on the real goal: to generate relevant and reasonably accurate product design and engineering requirements that can be developed into commercial products. Rather than exploring human condition in great detail, commercial application has to zoom in on the most important activities as opposed to create full context ethnographic research.

And Even more important is that user research really needs to be conducted by professionals with anthropology, psychology or sociology backgrounds, so one question that needs to be addressed is how to narrow the gap of interpretation of the research findings and user experience design to engineers and designers who are expected to act upon those findings. It’s always a big challenge for researchers to translate their findings into engineering product or technology design requirements. There are many good methodologies developed in the HCI field involving the use of personas, story boards, use cases and usage scenarios, and models to name a few. But we are always looking for other techniques that are most effective at helping engineers better understand the meaning of the research to lessen the gap between the different functional disciplines on a product development team.

In short, what’s needed is a research methodology that is cost-effective, provides a good basis for communication, is easy to act on by the different functions and really helps inform product design. This paper describes the methods we used in our research work in China for an ICT product developed in 2004. These methods specifically involved engineers in the early stages of product definition. As will be explained, they were quite active even in the user centered research work.

2 Why We Need Engineers to Be Involved in User Research

Generally, engineering participation in the product definition process creates a deep sense of co-ownership and buy-in into the product ideas by the technologists involved. Engineers who participated in the user experience definition would also be expected to fully participate in user research work. By involving engineers in the user research phase a project can really achieve the effect that “when the creative process is conducted as a truly integrated, mixed-discipline collaboration versus an *engineering project*, the resulting product concept that emerges is usually far superior and makes better use of technologies.”[2]

From user researchers’ side, they are usually not well-versed in technology. It’s especially obviously in China as we mentioned that the use of HCI user research in China has not well-established yet. So it can be a barrier for Chinese local researchers to catch crucial insights to inspire the technology innovation during user research phase. The mix of technology expertise into research work to offset local researchers’ weakness on such knowledge is necessary. Engineers’ technology background can help research in identifying key questions that contribute to technology development. The presence of engineers and designers, helps appropriately focus the field research before and during the field work which in turn can reduce the total research time and cost.

Engineers, of which there is no shortage in China, but overall tend to be less comfortable and skilled at interpersonal communication as the traditional culture teaches people to be modest, show respect for others and focus on one’s own responsibility. People are sometimes too modest and polite to readily share ideas in open brainstorm sessions. Critique of someone else’s ideas is often perceived as un-professional, even personally offensive. However innovation work relies heavily on the team’s interaction rather than individuals working independently. The involvement of engineers in user research phase is essential to build an environment that encourages wide open communication, that allows people to be open minded and where organizational authority between participants can be temporarily put aside.

Earlier involvement of engineers helps them see and understand the real world user needs and environment directly. That helps engineers better appreciate ideas and concepts created in response by user experience designers that they might otherwise consider as too wild and too difficult to implement. They can start thinking from the users’ perspective instead of getting secondary research data reported by the researchers. It creates empathy among the engineers for the user and their problems. That helps overcome the stereotypical behavior where “engineers and technologists often continue to take a one-sided technology view of what the world needs.”

The advantages of earlier involvement of engineers is obviously now. When placing engineers among the real users and the environment in which the product will have to operate, they can have very effective interaction with user researchers in driving them to articulate additional questions that will help them shape product design and resolve engineering tradeoffs. And it provides a good basis for better communication between researchers and engineers in product development that helps shorten the total development time and deliver even fancy product to pursue users’ value.

3 How to Involve Engineers in User Centered Research Work

To involve engineering people in product definition has a long history in Intel. While by the user research practice in China, engineers are even involved in earlier stage: the user research phase. It helps to be explained by looking into Intel's UCD process. (Fig. 1) Within Intel, the UCD process can be broken down into roughly four phases [2]:

The first phase is *to understand people by user research.* Generally speaking, we capture insights from observation of users, identify user's goals and needs, discover value proposition and context constraints which are used to inform and inspire the definition team.

Then is the second phase, *define the user experience*. The goal of this phase is to define the external user-observable behavior of the product to be developed.

The third phase is to *translate into engineering requirements*. Here, high level *external* product requirements are translated into quantifiable and concise technical engineering requirements that are directly actionable by engineers and later testable.

And the last phases, develop, *integrate, productize*. The real engineering work and technology development happens in this stage.

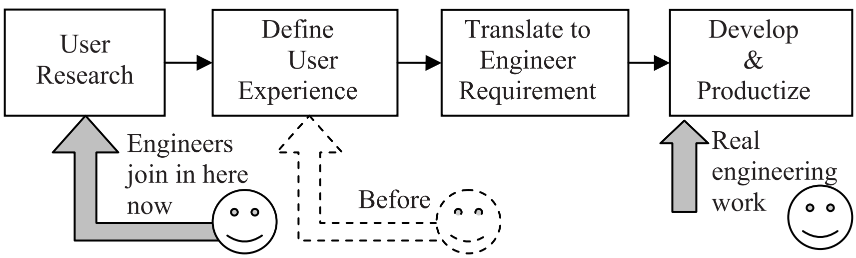


Fig. 1. User Centered Design process and the phase to involve Engineers

Of course, taking engineers into the field can be a bit of a challenge to most researchers, as engineers lack the research knowledge such as proper interviewing techniques, how to interact with the study subjects, and so on. They are usually overly focused on technology specifics as opposed to for example understanding more subtle social interactions and deeper motivations that should also be a major part of good field work. We will next introduce the process and some techniques that help engineers to be fit better with the field environment.

3.1 Phase 1: Homework

This stage allows the engineers to ramp up in thinking from the user viewpoint. We ask the engineers to execute some simple tasks that the trained researchers designed for them. All the activities are for a single task, a clear objective, offer a low time duration and are easy to complete. It's a very effective approach for engineers to start understanding people as people by acting as a user, thinking as a user or observing a user's activities.

3.1.1 Short Story Telling

This technique is used most frequently in our practices. It gives several questions for engineers for which they have to find the answers from target users. These target users don't need to be as accurate as the real target user we have defined; it could be anyone having a related experience with the implied technology. The users to interview are easy to find and talk with. It could be their family members, friends, neighbors, or, while not ideal, they could even use themselves as the target users. The questions they are given include three categories:

1. Story of a bad experience of technology use. This is a very short descriptive question that to which it should be easy to generate a story like answer. These questions help the engineer determine a user's unmet needs or broken expectations with current technology.
2. Good experiences with using technology. These questions help explore what the user values in technology, what technology means to the user, and how it impacts their life.
3. Ideal product or technology that could improve the usage they described earlier. These questions are aimed at determining the ideal desired experience for the user.

3.1.2 Map of a Typical Day in the Life

We provide pictures, magazines or sticks with symbols and text to the engineers. They are then asked to use these materials or anything else they can readily find to create a map of their typical day involving the use of technology products. They can even take photos of what they do as they use technology products. The map has to include key elements such as time, place, what they are doing, what technology they are using and how they feel about it. [8]

3.1.3 Photo Journal

In a similar technique we give engineers a topic and ask them to take photos to best describe it. For example they may be asked to take photos of how people use a certain technology product, take photos of their favorite technology product or the product they used most frequently.

From creating the stories or map of their typical day of life engineer is starting to think what technology means to the person using it and also how they use it. We are not targeting on finding real user needs and context through these parts to shape an actual product, but they are an effective introduction to user research work for engineers.

3.2 Workshop Discussion

After team member finished their homework, all the results are taken into a workshop for discussion. This is aiming at providing an environment that encourage people open discussion, sharing information and generate idea systematically. In the workshop session team members show and tell the stories of their work one by one, and others ask short questions after each 2-5 minute presentation. With this sharing of data, everybody get more stories of technology use and user problems. This is followed by a brainstorming session that explores several elements: user needs, possible solutions, areas that need additional understanding and would require follow-on user research. The discussion can become very intense as everybody really fully engages in the creative process. The final part of the workshop is to categorize or clump the brainstorming results and listing the key areas where further research in needed. (Fig. 2)

Fig. 2. Category user need in workshop **Fig. 3.** Conduct user observation with engineers

After the workshop, the team ends up with a relatively crisp requirements list for designers and engineers resulting from the research work, and more research questions for researcher to get answered. Based on the latter, researchers can form a research plan and protocol for field work. Similarly, engineers can start technology design based on the user-level requirements list. Their direct involvement in the creation of that list makes it easy for them to also appreciate the underlying motivation for those requirements. The advantage is clear: every functional discipline walk out with a clear direction for their next steps and project time saved as everybody are working of a shared understanding of who the target user is, what user problems are to be addressed, why the requirements have been selected. It all happens in parallel and at the same time; nobody has to wait for others to deliver results for their next step.

3.3 Boot Camp

Boot camp is a basic training preparation for people prior to joining the field work. The research protocol is discussed and finalized, non-researchers are trained on research principles and ground rules such that they will behave properly during the field work. We also use techniques for make impressions of these research principles to engineers.

3.3.1 Field Scenario Role Play

Two trainers act as informant and researcher respectively, they show appropriate behavior and also inappropriate behavior by acting out specific scenarios that simulate how real scenarios can happen in the field.

Sometimes we also ask engineers to act as the researcher while the trainer acts as the informant and go through a short part of research guideline, the rest of the audience will help to identify the problems during the simulated “research field work”, and the trainer will provide comments after that.

As the research plan and protocol are designed for maximum effectiveness, every person on the field research team will be assigned a role and responsibility in the field during boot camp. Then everybody including the engineers is ready to head out.

3.4 Structured Field Work

We call it *structured* field work because engineers do not participate in all aspects of the field work the way trained and experienced user researchers do. Researchers still take responsibility for most of the field work as usual for obvious reasons. Engineers only participate in the sessions scheduled near the end of all the field work sessions. These sessions are designed to be one to two days in length and include visiting key sites and interviewing key informants that have been carefully selected by the researchers from the earlier field work sessions. That selection is expected to "make the most" of the field time and to improve the field data collection with both abundance of data and the chance that interesting events and objectives will be observed.

Several activities occur in the structured field work sessions include: structured interviews, activity walkthroughs, and contextual inquiry. The “map of a typical day” technique we mentioned earlier is frequently used here to ask the informant to describe their typical day. We also ask people to show us their technology products and describe how they’re actually used; even ask them to show us how they use their desktop computer to do specific tasks such as download an MP3 file to their digital audio player.

By using those research techniques in structured field work, researchers present their findings to the engineers in the real end user environment. (Fig. 3)This natural setting provides for a much richer representation of the research findings and makes it easier for the engineers to makes sense of the situation. And discrepancies and gaps in understanding of the research insights can be readily be illustrated and resolved. In the same time the impact of having the engineers observe and interact with the informants at that location allows them to generate key insights for themselves that often directly drive aspects of the design and engineering requirements. Often the engineers are more likely to pay attention to some the more technical aspects of the context.

After completing each field work session, a short discussion should be held to improve the protocol and justify the research activities before the team starts the next field session. This is repeated iteratively and the research results are consumed in the concept brainstorming process. After a few iterations, there are no additional unanswered research questions and the resulting concepts can be considered stable. (Fig. 4)

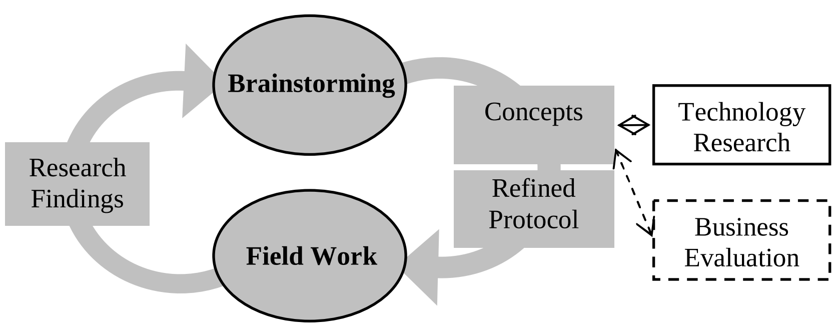


Fig. 4. Structure of engineers participated field work

It's interesting to highlight why this method is well suited for Chinese culture. Chinese people are usually not likely to express their opinions without fully understand the real purpose. They seem closed the first time you meet with them, and they will not likely be forthcoming in answering sensitive questions. As people become more familiar with each other, they also become very warm-hearted and willing to reveal much more about themselves and their feelings. So a good informant replies heavily on the personal relationship that is built up during the research activity. Once that relationship and trust is in place, research findings become much more interesting and insightful to the point that we consider these informants as key informants [3], and they are going to share more and more important ideas when keeping a good relationship with them.

Our results with engineer participation in field research is very encouraging, they are constantly surprised by the abundant information they collect while in the field and really recognize that what they used to think about the user is quite different from what they observed and heard from users by being there. This instills a profoundly different attitude in the engineers that positively impacts their further development work and they also learn to appreciate much more the value and importance of other research results in which they did not themselves participate. All this helps foster a good working relationship within the cross-functional development team, since everyone on the team is focused on a common goal.

4 Case Study

We'll next take a look at a project named "gaming engine". The objective is to design a product for China internet café to improve young people's gaming experience.

Engineers were asked to search for answers to three questions at home:

- Please tell us a story of your most exciting/ happy experience with gaming
- Please tell us a story to describe an unsatisfied area about any of the existing gaming machine that you are using (can be: PC, gaming console, mobile, or other)
- Please tell us a story about your dream of the ideal gaming experience

Engineers tend to take their work very seriously. They took a lot of interesting stories into discussion. One mentioned that it's really exciting that his friends in the dorm built their LAN together and play against each other for LAN-based action gaming in the university. Another one mentioned that the ideal gaming machine should have a function like "love seat" in cinema, so couples or boy and girl friends could play together.

During the workshop discussion, teams were asked to present their story of good and bad experiences. All the stories were pasted on the wall and broken into two categories: good experiences and bad experiences. The team then discusses these two categories respectively. In the good experience category, we are trying to highlight what satisfied the users; what is valuable to them. In the bad experience category, we find out users' unmet needs and capture them as to keywords with paraphrasing.

After highlighting user value and unmet needs, the team members were divided into two teams to design a "gaming concept" based on the output of the team's discussion, and present the concept back to the team with sketches and a brief description. After the two concepts have been presented, the team holds another round of discussion to improve the concepts and bring up new questions for user research as well as areas for technology study as a next step.

The user research work was designed as a combination of home and site visits, with in-depth interviews and focus groups. There's no doubt that engineers have no problem being involved in observing in-depth interview and focus group as they are in the safe zone behind the one-way mirror. But also in home and site visits, engineers tend to follow the process and protocol and contribute many good ideas to team.

We found that engineers are very surprised by the research findings they get from personally participating in the field work. For example, before the field work, performance (speed) was always assumed to be the biggest concern for young gamers, so our efforts focused on performance almost exclusively. The research surfaced other and different needs that pointed at the importance of other aspects that are key to successful commercial solutions. Although performance is still very important, meeting the social need around gaming is driving a new and different set of requirements. We found that the personal presence in the gamer community is very important. To go out with friends for LAN-based gaming is a very common social activity for them and feedback from online gamers is that they are sometimes seeking a kind of social relationship that is different from the real world. The social purpose has driven the team to think and develop solutions that enhance group communication for gamers.

With the collaboration of engineers in user research and user experience design, the development team went through the work more smoothly and everybody was passionate to devote their talents to the product that users really want.

5 Summary

With the reality of constraints such as timing, people and budget of product development projects, the method of involving engineers into the earlier stages of user centered product innovation provide an approach that shortens the development life cycle and fosters improved interaction between different functional disciplines involved in development.

The method is developed and refined in the context of ICT product design in China, but it could easily be adapted to any geography by adjusting for local culture differences.

Research technologies can also be introduced during research that can help engineers participate productively in user research stage besides the methods mentioned in this article. For example: by capturing video documentary of the session or by capturing an audio record, or take still pictures. [7] The data can then alter be filtered and condensed by trained researchers. Researchers should trade-off the methods used based on the resources available to them.

Acknowledgements. The authors thank Intel's User-Centered Design team and People And Practices Research (PAPR) group for their help in identifying appropriate research techniques and for sharing valuable best practices that have kept our development methods up to date. In addition, we thank Intel Channel Platforms Group (CPG) Asia platform definition team for their ongoing support for innovative user-centered research and other talent that has helped improve research feasibility in practice.

References

1. Bentley, R., Hughes, J.A., Randall, D., Rodden, T., Sawyer, P., Shapiro, D., Sommerville, I.: Ethnographically-informed system design for air traffic control. In: Proccedings of the Conference on Computer Supported Cooperative Work, Tornoto, pp. 123–129, November 1992 (1992)
2. D'Hooge, H.: User-Centered Research Fuels Technology Innovation at Intel. Technology @Intel Magazine, 3-10 April 2005 (2005)
3. Millen, D.R.: Rapid Ethnography: Time Deepening Strategies for HCI Field Research. In: Proceedings of the conference on Designing interactive systems: processes, practices, methods, and techniques. New York, United States. pp. 280–286 (2000)
4. Ann, E.: Founder and Director Kaizor Innovation. Cultural Differences affecting Ethnographic Research Methods in China http://www.kaizor.com/web_doc
5. Spradley; Paperback. The ethnographic interview
6. Dray, S.M., Siegel, D.A.: Penny-wise, pound-wise: making Smart Trade-offs In Planning Usability Studies. Interactions (May, June, 1999)
7. Buur, J., Binder, T., Brandt, E.: Taking Video beyond 'Hard Data' in User Centred Design. In: Proposed for Participatory Design Conference, New York, 2000, 7 (2000)
8. Gaver, B., Dunne, T., Pacenti, E.: Cultural Probes. Interactions (January, February 1999)

Shape Analysis of Human Brain with Cognitive Disorders

Tianzi Jiang[1,2], Feng Shi[2], Wanlin Zhu[2,3], Shuyu Li[4], and Xiaobo Li[2]

[1] Institute of Biomedical Informatics, School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China
[2] National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China
[3] Department of Mathematics, Zhejiang University, Hangzhou 310027, China
[4] Department of bioengineering, Beihang University, Beijing 100083, China
jiangtz@nlpr.ia.ac.cn

Abstract. In this paper, we present some of our current studies on how human brain structures are influenced by cognitive disorders occurred from various neurological and psychiatric diseases based on magnetic resonance imaging (MRI). We first give a brief introduction about computational neuroanatomy, which is the basis of these studies. In Section 2, several novel methods on segmentations of brain tissue and anatomical substructures were presented. Section 3 presented some studies on brain image registration, which plays a core role in computational neuroanatomy. Shape analysis of substructures, cerebral cortical thickness and complexity was presented in Section 4. Finally, some prospects and future research directions in this field are also given.

1 Introduction

Computational neuroanatomy here aims at computationally demanding quantitative neuroanatomic analyses, and computational modeling of brain structure and spatial organization based on such quantitative data. It has played an essential role in the study of cognitive disorders, especially on relationships between anatomical abnormalities and cognitive disorders [1]. The basic research topics in this field include brain tissue segmentation of MR images, intra- and inter-modality image registration, automatic lesion detection and segmentation, brain structure segmentation, registration and shape analysis and so on.

In recent years, more and more studies show their interest in these directions and also achieve satisfying results [1][3]. In this paper, we will present some advances of our current studies on detection of the anatomical abnormalities of human brain with neurological and psychiatric diseases based on magnetic resonance imaging (MRI). First, several novel methods on segmentations of brain tissue and anatomical substructures, brain image registration, and computation of cerebral cortical thickness and complexity will be presented. Then, we will present some interesting findings on anatomical abnormalities when these methods were applied to human brain with various cognitive disorders, including Alzheimer’s diseases, Schizophrenia, Attention Deficit Hyperactivity Disorder, early blind and deaf, compared with matched normal controls. Finally, some prospects and future research directions in this field are also given.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 409–414, 2007.

2 Segmentation of Brain Image

Image segmentation is a process of separating an image into several disjoint regions in which characteristics are similar such as intensity, color, texture, or other attributes. In the field of brain image analysis, we are mainly concerning with brain tissue segmentation and brain substructure segmentation, which are also the mainly preprocessing step in many medical research and clinical applications.

Brain Tissue Segmentation: In MR images, intensity inhomogeneity is a very common phenomenon which can change the absolute intensity for a given voxel in different locations. So it becomes a major obstacle to any automatic methods for MR image segmentation and makes it difficult to obtain accurate segmentation results. In order to address this issue, we proposed a novel method called Multi-Context Fuzzy Clustering (MCFC) based on a local image model for classifying 3D MR data into tissues of white matter, gray matter, and cerebral spinal fluid automatically [3]. Experimental results on both simulated volumetric MR data and real MR images showed that the MCFC outperforms the classic method of fuzzy c-means (FCM) as well as other segmentation methods that deal with intensity inhomogeneity.

Another related work is pixon-based adaptive scale method for image segmentation. Markov random fields (MRF)-based methods are of great importance in image segmentation, for their ability to model a prior belief about the continuity of image features such as region labels, textures, edges. However, the main disadvantage of MRF-based methods is that the objective function associated with most nontrivial MRF problems is extremely nonconvex, which makes the corresponding minimization problem very time consuming. We combined a pixon-based image model with a Markov random field (MRF) model under a Bayesian framework [4]. The anisotropic diffusion equation was successfully used to form the pixons in our new pixon scheme. Experimental results demonstrated that the proposed method performs fairly well and computational costs decrease dramatically compared with the pixel-based MRF algorithm.

Brain Substructure Segmentation: We also proposed another variational based segmentation algorithm. The originality of formulation was on the use of J-divergence (symmetrized Kullback-Leibler divergence) for the dissimilarity measure between local and global regions. The intensity of a local region was assumed to obey Gaussian distribution. Thus, two features mean and variance of the distribution of every voxel were used to ensure the robustness of the algorithm when the noise appeared. J-divergence was then employed to measure the distance between two distributions [5]. The proposed method was verified on synthetic and real medical images and experimental results indicated that it had the ability to segment brain substructure robustly.

Accurate segmentation of brain structures such as the brain ventricles is needed for some clinic applications. In recent years, the active-models-based segmentation methods have been extensively studied and widely employed in medical image segmentation and have achieved considerable success. However, the current techniques are going to be trapped in undesired minimum due to the image noise and pseudoedges. We proposed a parallel genetic algorithm-based active model method and applied it to segment the lateral ventricles from magnetic resonance brain images

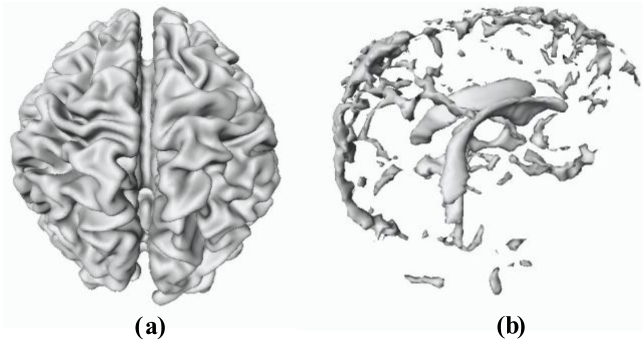


Fig. 1. Brain cortex segmentation results. (a) Inner surface of cerebral cortex. (b) CSF surface

[6]. The proposed method was demonstrated successfully to overcome numerical instability and was capable of generating an accurate and robust anatomic descriptor for complex objects in the human brain as lateral ventricles.

3 Registration of Brain Image

Image registration plays major role in computational neuroanatomy. In terms of satisfying the technical requirements of robustness and accuracy with minimal user interaction, rigid registration has been applied to many applications including multimodality and inter-subject registration. However, due to the nonlinear morphometric variability between subjects. The requirement of both global and local registration accuracy asks for non-rigid registration. The method of non-rigid medical image registration usually include physics-based and geometry-based. We have made our effort on both of them.

Physics-Based Method: A non-rigid Medical Image Registration by Viscoelastic Model was presented in [7], by assuming the local shape variations were satisfied the property of Maxwell model of viscoelasticity, the deformable fields were constrained by the corresponding partial differential equations. Applications of the proposed method to synthetic images and inter-subject registration of brain anatomical structure images illustrate the high efficiency and accuracy.

Geometry-Based Method: We also proposed an efficient registration framework which has feature of multiresolution and multigrid [8]. In contrast to the existing registration algorithms, Free-Form Deformations based NURBS (Nonuniform Rational B Spline) are used to acquire nonrigid transformation. This can provide a competitive alternative to Free-Form Deformations based B spline on flexibility and accuracy. Subdivision of NURBS is extended to 3D and is used in hierarchical optimization to speed up the registration and avoid local minima. The performance of this method is numerically evaluated on simulated images and real images. Compared with the registration method using uniform Free-Form Deformations based B spline, this method can successfully register images with improved performance.

Fig. 2. Registration results on MRI images. Contours of the reference image (white line) were overlaid on the resulting images as a reference of registration accuracy. (a) Reference image, (b) result image using affine registration, (c) result image using FFD registration, (d) result image using NFFD registration.

Deformation Based Morphometry: Deformation based morphometry (DBM) derived a voxel-wise estimation of regional tissue volume change from the deformation field required to warp subject to the template image. By using this technique, anatomic differences in boys with Attention-Deficit/Hyperactivity Disorder (ADHD) was characterized [9]. The statistical results reveal some pronounced volume alterations in the brains of ADHD, which confirm that there are widespread abnormalities in volume of boys suffering by ADHD.

4 Shape Analysis

Cognitive disorders can cause the variations of anatomical shape in human brain. Statistical Shape Analysis (SSA) is a powerful tool for noninvasive studies of pathophysiology and diagnosis of brain diseases. It is another key component of computational neuroanatomy. The population-based shape analysis not only reveals the difference between the healthy and diseased subjects, but also provides the dynamic variations of the patients' brain structures over time.

Hippocampus Shape Analysis: We proposed a new method which incorporated shape-based landmarks into parameterization of banana-like 3D brain structures to address this problem [10]. The triangulated surface of the object was firstly obtained and two landmarks were extracted from the mesh, i.e. the ends of the banana-like object. Then the surface was parameterized by creating a continuous and bijective mapping from the surface to a spherical surface based on a heat conduction model. The correspondence of shapes was achieved by mapping the two landmarks to the north and south poles of the sphere. The proposed method was applied in a Alzheimer's disease (AD) study [11]. The machine learning methods were used to characterize shape variations in AD based on these surface-based shape measures. Correct prediction rate were above 80% in bilateral hippocampi with leave-one-out cross validation.

Cortical Morphometry Analysis: We proposed a 3-phase variational segmentation model for extraction of inner and outer surfaces of cerebral cortex [6]. As the brain

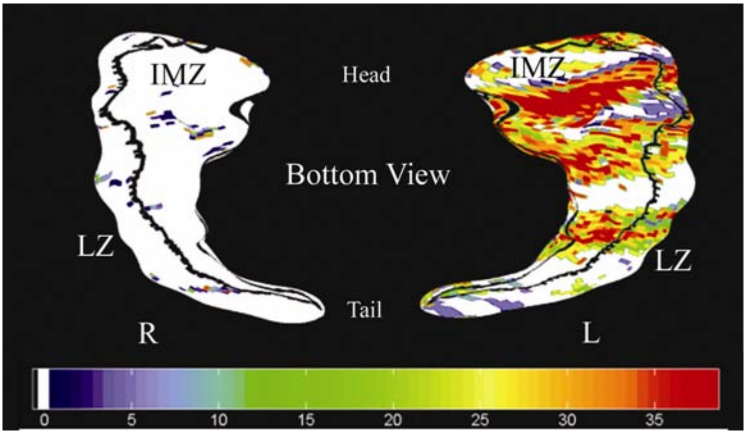


Fig. 3. Visualization of surface shape variations of hippocampi in AD compared to healthy controls

tissue can be decomposed into white matter, grey matter and cerebrospinal fluid, any two tissues have boundaries. Thus competition should be introduced for any two regions. The proposed 3-phase model based on our J-divergence active contour model and designed specifically for three regions segmentation. It has been successfully applied to cerebral cortex reconstruction. Cortical thickness analysis has also been performed for finding abnormal pattern of brain structure between normal controls and early blind patients. The T-average was used as cortical thickness definition. After statistics analysis, the final results showed us that normal people had more thinner cortical thickness for part of occipital region. This may be caused by absent usage of this brain region during development of brain.

Besides cortical thickness, there is another widely used measurement of cortical morphometry. Cortical complexity, which is usually employed to describe the degree of cortical convolution or gyrification, is often assessed by using fractal dimension. We developed a voxel based method to compute information dimension of cortical pial surface and applied to subjects with attention-deficit/hyperactivity disorder. A left-greater-than-right prefrontal cortical convolution complexity pattern was found in both groups.

5 Conclusions and Future Directions

In this paper, we have presented some representative techniques to detect the anatomical abnormalities of human brain with neurological and psychiatric diseases. We have been applying various modern neuroimaging techniques to combat the neurological and psychiatric diseases, especially Alzheimer's Diseases and Attention Deficit Hyperactivity Disorder. Our long-term goal is to find the early markers for the neurological and psychiatric diseases based on not only neuroimages and but also genome datasets. It would be very interesting to identify the genetic basis of the anatomical and functional abnormalities of human brain with neurological and psychiatric diseases. Anatomical and functional abnormalities of human brain with neurological and psychiatric disorders are very promising endophenotypes for these disorders. In fact, several publications have been available and a new field - imaging genomics, named by Hariri and Weinberger, has emerged [13]. It is at its infant stage and we expect a lot of important progress can be made in the future.

Acknowledgements. This work was partially supported by the Natural Science Foundation of China, Grant 30425004 and 60121302, and the National Key Basic Research and Development Program (973), Grant No. 2004CB318107.

References

1. Csernansky, J.G., Wang, L., Joshi, S.C., Ratnanather, J.T., Miller, M.I.: Computational anatomy and neuropsychiatric disease: probabilistic assessment of variation and statistical inference of group difference, hemispheric asymmetry, and time-dependent change. NeuroImage 23 (Suppl. 1), 56–68 (2004)
2. Thompson, P.M., Toga, A.W.: A Framework for Computational Anatomy. Computing and Visualization in Science 5, 13–34 (2002)
3. Ashburner, J.: Computational neuroanatomy. Ph.D. thesis, University College London (2000)
4. Zhu, C.Z., Jiang, T.Z.: Multi-context Fuzzy Clustering for Separation of Brain Tissues in MR Images. NeuroImage 18, 685–696 (2003)
5. Yang, F.G., Jiang, T.Z.: Pixon-Based Image Segmentation With Markov Random Fields. IEEE Trans Imag. Proc. 12, 1552–1559 (2003)
6. Zhu, W.L., Jiang, T.Z., Li., X.B.: Segmentation of Brain MR Images Using J-Divergence Based Active Contour Models. In: Parametric and Geometric Deformable Models: An application in Biomaterials and Medical Imagery, vol. II, Springer, Heidelberg (in press, 2007)
7. Fan, Y., Jiang, T.Z., Evans, D.J.: Volumetric Segmentation of Brain Images Using Parallel Genetic Algorithm. IEEE Trans Med. Imaging 21, 904–909 (2002)
8. Tang, S.Y., Jiang, T.Z.: Non-rigid Registration of Medical Image by Linear Singular Blending Techniques. Pattern Recogn Lett. 25, 399–405 (2004)
9. Wang, J.Z., Jiang, T.Z.: Nonrigid Registration of Brain MRI Using NURBS, vol. Pattern Recognition Letters 28(2), 214–223 (2007)
10. Wang, J.Z., Jiang, T.Z., Cao, Q.J., Wang, Y.F.: Characterizing Anatomical Differences in Boys with Attention Deficit Hyperactivity Disorder Using Deformation Based Morphometry. American Journal of Neuroradiology 28(3), 543–547 (2007)
11. Zhu, L.T., Jiang, T.Z.: Parameterization of 3D Brain Structures for Statistical Shape Analysis. In: proceedings of SPIE Medical Imaging'04, California, USA , 2004, pp. 14–17 (2004)
12. Li, S.Y., Shi, F., Jiang, T.Z. et al.: Hippocampal shape analysis of Alzheimer's disease based on machine learning methods, American Journal of Neuroradiology (in press, 2007)
13. Hariri, A.R., Weinberger, D.R.: Imaging genomics. Br Med Bull. 65, 259–270 (2003)

Finding Origin Points for New Coordinate System Suitable for Sign Animation

Tomohiro Kuroda[1,2], Kazuya Okamoto[2], Ryo Kitauchi[3],
Tadamasa Takemura[1,2], Naoki Ohboshi[3], and Hiroyuki Yoshihara[1,2]

[1] Department of Medical Informatics, Kyoto University Hospital,
54 Kawahara-cho, Shogo-in, Sakyo-ku, Kyoto, 606-8507, Japan
[2] Graduate School of Informatics, Kyoto University,
36-1, Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-8501, Japan
[3] School of Science and Engineering, Kinki University,
3-4-1, Kowakae, Higashi-Osaka, 577-8502, Japan
{Tomohiro.Kuroda,kazuya,takemura}@kuhp.kyoto-u.ac.jp,
stern@info.kindai.ac.jp, lob@kuhp.kyoto-u.ac.jp

Abstract. This paper proposes new coordinate system suitable for denoting sign language motion. As the proposed coordinate system consists of polar coordinate systems whose origins are certain points of human body, postures shown on the system can be proportional for avatars with any possible shape and fit with existing subjective sign notation systems. This paper extracted coordinate origins from Japanese-Japanese Sign Language Dictionary via morphological analysis. Selected 85 points are successfully mapped on H-ANIM standard humanoid avatar.

Keywords: Japanese Sign Language, Natural Language Processing, Sign Language Dictionary, Sign Animation.

1 Introduction

Sign linguistic engineering is a group of research to develop communication aid (called sign information system) for the Deaf and the Hearing Impaired, who have communication barrier in social lives, using information technology [1]. Fig. 1 shows the general view of sign information systems. Any sign information systems, such as virtual reality based sign language telephones, automatic translation system between sign languages and phonetic languages, and even sign language dictionary, share same basic structure. Most of sign information systems use three-dimensional computer graphic model of human beings called avatar to show signs. To obtain signs, though several systems using image recognition technologies, most of them are utilizing motion capture systems consists of data-gloves and position and orientation sensors either magnetic-field-based, ultrasound-based, and image-based. The authors also developing sign language telephone named S-TEL [2] and consequent systems based on motion capture system consists of two data-gloves and three position and orientation sensors.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 415–422, 2007.

Most of sign animations handles motion as sets of rotation angles of each joint and generate animation applying the data to an avatar model, just as same as MPEG-4 H-ANIM standard [3]. However, a certain body motion parameter to show a certain sign on a certain avatar cannot produce the same sign when it applied on another avatar with different body shape. Thus, the sign linguistic system needs to prepare an avatar, which holds the same body shape as a certain person to produce proper signs. Although authors developed a system to morph a standard avatar to fit a certain person through image recognition [4], it cannot give users full selection of avatar's outlook to have favorable view. Thus, to give flexibility and usability on sign information systems, a new motion coding system, which is independent from body shape and easily transferable from or to standard motion data such as a captured motion, data via a certain motion capture system or H-ANIM.

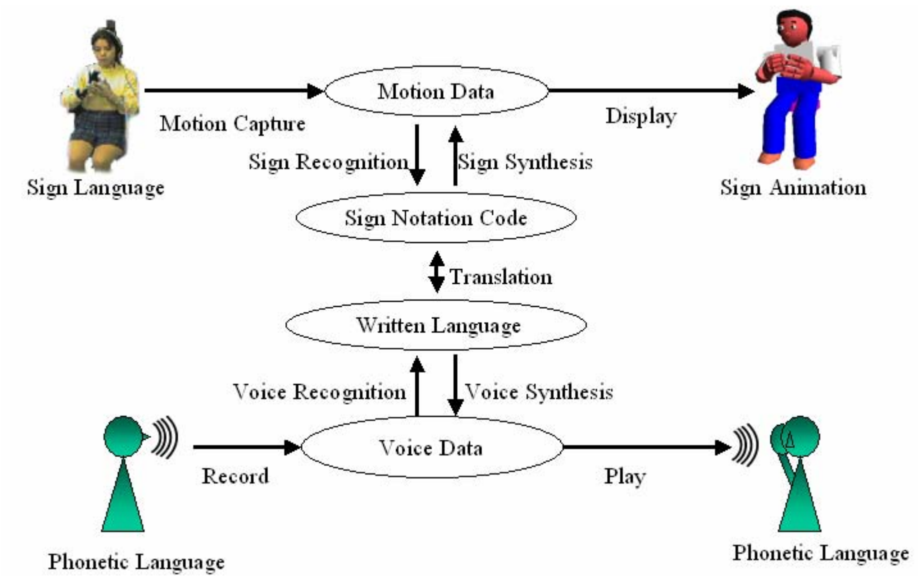


Fig. 1. General Overview of sign information systems

This paper proposes new coordinate system suitable for sign animation. Notations of textbooks or dictionaries of sign language present a position and orientation of hands in a certain posture in relation to a certain point of human body.

2 Foregoing Coding Systems

Foregoing researches on sign language developed several types of sign coding systems. Most of the sign notation system follows framework of Stokoe [5], which claims signs are consists of "Cheremes", that is, hand posture (Dez), position (Tab) and motion (Sig). Thus, most of the notation system denotes signs as a series of combination of hand posture, orientation and position, although some of them denote motions as shape of trajectories. Sometimes several non-manual signals, such as facial expressions, are added on them.

HamNoSys [6], one of the most well-known sign notation system, also denote signs as a combination of hand posture, orientation and position. HamNoSys denote position of hands as combination of upper/middle/bottom of the shoulder and left and right of body, although it adds several more codes on update of version 4.0, the resolution is still quite rough.

SignWriting [7], another one of the most well-known sign notation system, denotes hand posture and position, orientation and speed. As SignWriting is designed for easy denoting and reading of sign language and is derived from dance notation system, the notation is quite pictorial and subjective. Thus, automatic conversion from or to motion capture data seems quite difficult.

Kurokawa [8] developed a motion coding system for nonverbal interface and applied the coding system to denote sign language. However, resolution of the coding system, again, quite rough to synthesis signs animation from given code as the system designed mainly for motion recognition.

As discussed above, foregoing sign notation systems are based on subjective coding and its resolution are not detailed enough, new coding system, which mediates such subjective coding system and motion data, such as FACS [9] for facial expression, is essential for sign language information systems. As hand posture can be mapped via simple vector quantization as authors presented on previous conference and implemented on StrinGlove® [10], the coding system to denote hand position and orientation can be sufficient.

3 Designing New Coordinate System

3.1 Basic Approach

As discussed above, the target of this research is to develop hand posture/orientation coding system to mediate raw captured or kinematics data and existing subjective sign notation codes.

Careful observation of foregoing sign notation systems, especially new location symbols added on HamNoSys in version 4.0 [11], gives interesting insight of sign posture coordination. Figure 2 shows several new added location symbols. As shown in the figure, new version of HamNoSys gives locations of nominal hand by several body parts such as lip, ear lobe, nose and arms. On the other hand, quick review over several textbooks of sign languages also gives positions of hands or fingers by referring certain body parts. Therefore, human being seems to position hands relative to a certain body parts, such as lip, nose, chest, elbow, etc. Thus, coding system that let us denote relation between hands and other body parts can fit subjective insight of human beings as well as existing sign notation code.

On the other hand, to enable automatic conversion from or to “raw” motion data, such as captured motion data or generated kinematics data, the coding system must not be sparse or rough. The coding system should have enough resolution to show its precise position in a certain coordinate.

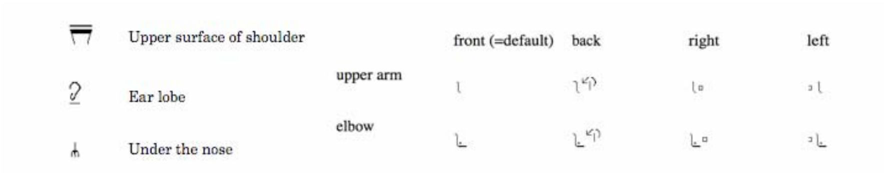


Fig. 2. New Location Symbols on HamNoSys 4.0 and its usage

The authors propose to develop new coordinate system suitable for sign motion rather than coding system. The coordinate system consists of sets of polar coordinate system whose origins are certain points (the reference points) on human body, such as "inner corner of right eyelid" or "eye". As the reference points are proportional to shape of body, the proposed system can proportion position and orientation of hands to various avatar models including cartoon-like model with big head as shown in Fig. 3.

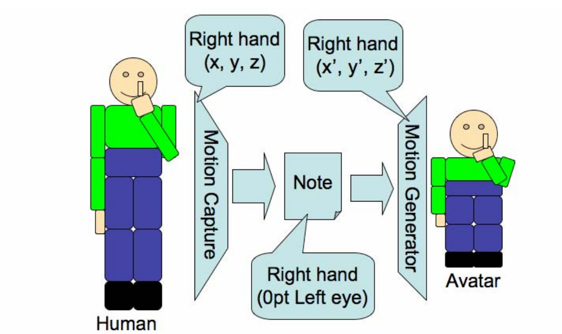


Fig. 3. Basic Concept of the Proposed Coordinate System

This coordinate system allows automatic conversion from "raw" motion data by calculating out distance of hand from several reference points and to "raw" motion data by mapping them on certain avatar. Additionally, a certain motion denoted on subjective notation code can be mapped onto the coordinate by giving distance on a certain rule.

3.2 Picking Up the Reference Points

To obtain reference points, this paper analyzes notations of Japanese – Japanese sign language dictionary issued by the Deaf association of Japan [12]. The dictionary has 8320 headwords (Japanese) and one headword has a pair of gestures (Japanese signs) and notation of each sign gestures in Japanese as shown in Fig. 4. In total, the dictionary has 16640 sign notations.

The authors scanned all pages of the dictionary and obtained notations via optical character reader. The obtained notations are analyzed through ChaSen, a morphological parser of the Japanese Language developed by NAIST Computer Linguistic Laboratory [13], which equips medical corpus developed for natural

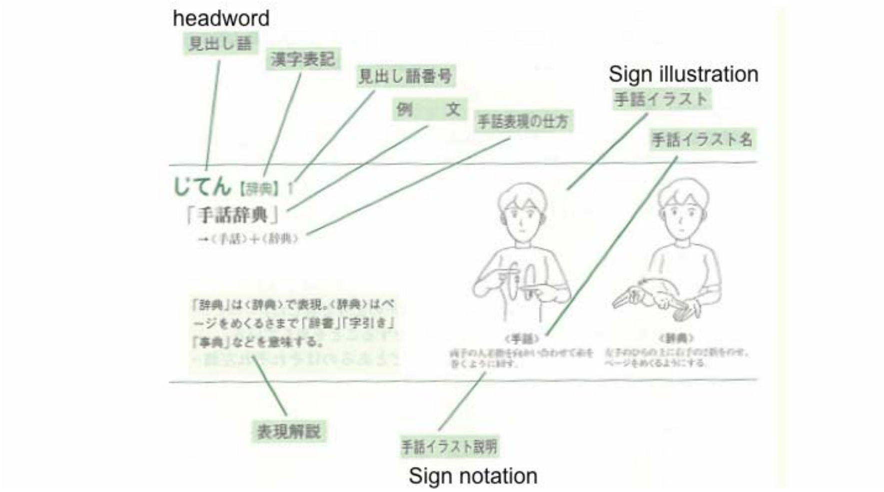


Fig. 4. Explanatory Notes of Japanese-Japanese Sign Dictionary [12]

language processing in medicine by the authors [14] to obtain proper anatomical nouns. Analysis extracted 43 candidates of reference points which can be successfully mapped on either a certain point of cinematic model or reference model of MPEG4 H-ANIM standard [15].

3.3 Verification of Hypothesis

The reference points are extracted based on a certain hypothesis, which a sign notation may gives positions of hands or fingers by referring certain body parts, obtained through quick review of sign textbooks. This section inspects the hypothesis through detailed analysis on sign documents.

As the sign notations obtained from the Japanese-Japanese sign language dictionary contains certain amount of error due to error of the optical character reader, this paper performs simple automatic correction before detailed analysis. As shown in Fig. 4, the dictionary uses one or two signs to denote one Japanese headword, and, as

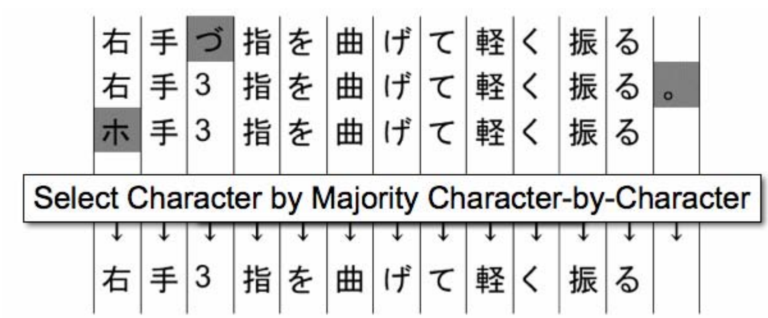


Fig. 5. Simple decision-by-majority algorithm

a whole, 3235 signs are used to denote 8320 Japanese words. Therefore, most of the signs appear repeatedly. This paper compares multiple sign notations of a certain sign and decides character to pick by majority as shown in Fig. 5. This simple algorithm produces 2022 error-free sentences.

Selected 2022 sentences analyzed again via ChaSen with medical corpus. Through careful review on the result of morphological analysis, sentences with following features are marked as sentences supporting the hypothesis.

1. Sentence contains compound noun, which consists of body parts and postpositional modifier denoting position connected by possessive particle "NO", such as "yubi no saki (tip of finger)".
2. Sentences contains compound noun includes body parts followed by elative ("kara"), illative ("e") or inessive ("de") particles, such as "kuchimoto kara (from mouth)".

Tab. 1 shows the result. 301 sentences can be categorized as pattern 1 and 379 sentences can be categorized as pattern2. Thus, about 35% of the sentences denoted as the manner to support hypothesis.

Table 1. Analytical result

	Sentence with body parts and postpositional modifier of position	Sentence with body parts followed with elative/illative/inessive particles	Others	Total
Number of signs	301	379	1342	2022
Times to appear in the dictionary	1768	2507	7309	11665

The result shows rest 65% of sentences doesn't have points to refer. However, it doesn't means these sentences doesn't fit with the hypothesis. In natural sign conversation, most of signs are shown in front of signer's chest, called neutral position, as it is most natural position to keep hands while moving. Quick review of rest 65% indicates most of the signs denoted without position are the signs shown at neutral position. So, more detailed analysis including idea of neutral position is required to verify the hypothesis.

4 Discussion

The analysis of Japanese-Japanese sign language gives 43 body positions, which can be mapped on H-ANIM standard avatar. Thus the coordinate system may denote given sign motion in a certain manner.

On the other hand, text analysis on source sign notation obtained from Japanese-Japanese sign language didn't clearly support the underlying hypothesis, as majority of source sign notations does not contains referring position. However, most of sign notations tacitly indicate that their target signs are shown around neutral position. For example, most of sign notations include information of motions, such as “move right hand upward” without denoting starting point, where the neutral position. Therefore, more detailed analysis is required to inspect whether the hypothesis is supported or not.

Although this paper proposed and mapped new coordinate system logically suitable for signs, sign language or human posture can be quite complicated. For example, to show “eye”, a signer can indicate his/her eye from any direction of eye if the fingertip of index is 0pt (or small amount) from targeting eye. Thus, to define motion of sign posture, to define dominant part of hand is indispensable. However, as the dominant part can be defined as nearest point to the reference point, or dependent on hand posture itself, automatic conversion from “raw” data into the proposed coordinate may be realized to a certain extent.

5 Conclusion

This paper proposed new coordinate system suitable for sign animation, and defined possible origin of coordinates through morphological analysis of Japanese Sign Language Dictionary. The investigation clears that the proposed coordinate can proportional to avatar and be applicable for most of sign expressions. However, analysis on source sign notations cannot clearly support or decline the underlying hypothesis. More detailed analysis to verify the hypothesis is still open question. The authors are developing coordinate conversion to evaluate availability of the proposed coordinate system.

Acknowledgements. This research is partly supported by Grant-in-Aid for Young Scientists (A) (17680008) from The Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

1. Nagashima, Y., Kanda, K.: Present State and Issue of Sign Linguistic Engineering. In: Proc. HCI Int. vol. 3, pp. 387–391 (2001)
2. Kuroda, T., Sato, K., Chihara, K.: S-TEL an Avatar Based Sign Language Telecommunication System. Int, J. Virtual Reality 3(4), 21–27 (1998)
3. Information Technology – Coding of Audio-Visual Objects- Part2: Visual. ISO/IEC 14496-2 (2001)
4. Kuroda, T., Makari, W., Manabe, Y., Chihara, K.: Transformation of Signing Avatar to Fit User’s Figure. Moroccan J. Cont. Comp. Sci. and Sig. Proc. 3(1) (2001)
5. Stokoe, W.C., Casterline, D., Croneberg, C.: A Dictionary of American Sign Language on Linguistic Principles. Gallaudet University Press, Washington DC (1965)
6. Hanke, T.: HamNoSys; Representing Sign Language Data In Language Resources and Language Processing Contexts. In: Proc. LREC, pp. 1–6 (2004)
7. Sutton, V.: The SignWriting Literacy Project. In: Proc. Cognition AREA Conf. (1998)
8. Kurokawa, T.: Gesture Coding and a Gesture Dictionary for a Nonverbal Interface. IEICE Trans. Fund. Elec. 75, 112–121 (1992)

9. Ekman, P., Friesen, W.V.: Facial Action Coding System: a Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto (1978)
10. Kuroda, T., Tabata, Y., Goto, A., Ikuta, H., Murakami, M.: Consumer Price Data-Glove for Sign Language Recognition. In: Proc. ICDVRAT, pp. 253–258 (2004)
11. Schmaling, C., Hanke, T.: HamNoSys 4.0. (2004) http://www.sign-lang.uni-hamburg.de-/Projekte/HamNoSys/HNS4.0/englisch/HNS4.pdf
12. Yonekawa, A., (eds.): Japanese-Japanese Sign Language Dictionary. Japanese Federation of the Deaf, Tokyo (1997)
13. Matsumoto, Y., Kitauchi, A., Yamashita, T., Hirano, Y., Matsuda, H., Takaoka, K., Asahara, M.: Moriphological Analysis System ChaSen Manual. Nara Institute of Science and Technology (2001)
14. Takemura, T., Ashida, S.: A Concept of Medical Corpus for Natural Language Processing in Medicine. In: Proc. Int. Work. Enterprise Networking Comp. in Health Care Industry (2001)
15. Kuroda, T., Kazuya, O., Takemura, T., Nagase, K., Yoshihara, H.: New Coordinate System Suitable for Sign Animation. Int. J. Disability and Human Development 5(2), 113–117 (2006)

User Experience Modeling and Enhancement for Virtual Environments That Employ Wide-Field Displays

James J.W. Lin[1] and Donald E. Parker[2]

[1] Siemens Corporate Research, 755 College Road East, Princeton, NJ, USA
[2] Human Interface Technology Laboratory, University of Washington, Seattle, WA, USA
jameslin@siemens.com, deparker@u.washington.edu

Abstract. User experience in virtual environments including presence, enjoyment, and Simulator Sickness (SS) was modeled based on the effects of field-of-view (FOV), stereopsis, visual motion frequency, interactivity, and predictability of motion orientation. We developed an instrument to assess the user experience using multivariate statistics and Item Response Theory. Results indicated that (1) presence was increased with a large FOV, stereo display, visual motion in low frequency ranges (.03 Hz), and high levels of interactivity; (2) more SS was reported with increasing FOV, stereo display, .05-.08 Hz visual motion frequency, lack of interactivity and predictability to visual motion; (3) enjoyment was increased with visual motion in low frequency ranges (.03 Hz) and high levels of interactivity. The resulting response surface model visualizes the complex relationships between presence, enjoyment, and SS. Overall, increasing interactivity was found to be the most profound way to enhance user experience in virtual environments.

Keywords: user experience modeling, virtual environment, presence, simulator sickness, enjoyment, interactivity.

1 Introduction

The assessment of user experience in virtual environments (VEs) has been primarily focused on evaluating the sense of presence reported by users while in VEs. Virtual presence is defined as being the extent to which participants feel they are in a place away from their actual physical setting while experiencing a computer generated simulation [1]. This is considered to be a product of a VE's immersive properties [2]. Previous studies suggested that VEs that produce an increased sense of immersion or effectively direct users' attention towards certain tasks would produce higher levels of presence. It is generally noted that experiencing presence requires focus on a meaningfully coherent set of stimuli in the VE to the exclusion of unrelated stimuli in the physical location. A number approaches have been demonstrated to be effective for the enhancement of virtual presence, such as increasing FOVs of VE displays and visually inducing self-motion perception of VE users.

Applying these approaches for presence enhancement in VEs, unfortunately, could also increase the likelihood for their users to develop symptoms similarly to those in motion sickness. The phenomenon of VE induced sickness, or simulator sickness

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 423–433, 2007.

(SS), has affected a great number of the users that it has become a focal point of discussion and research during the past two decades [3][4]. Referring to former studies in motion sickness, the sensory conflict theory has been widely used as the explanation for the development of SS symptoms in VEs. According to the sensory conflict theory, SS is thought to be the result of conflicting motion and orientation signals from the visual and vestibular senses [5][6], as in a VE the self-motion perception induced by the visual cues is expected to be deviated from that suggested by the cues from the vestibular receptors. While a user experiences SS in a VE, s/he could develop a number of symptoms such as increasing salivation, sweating, nausea, fatigue, dizzy, headache, eyestrain, blurred vision, difficulty focusing, and difficulty concentrating. Although the symptoms could vary upon individuals, a prolong accumulation of SS severity could often make the users feel uncomfortable and even withdraw from a VE exposure. Consequently, this side-effect could contribute negatively to user experience, and may also impact the sense of presence or enjoyment in a VE. For instance, the development of presence could be interrupted because of difficulty focusing on the visual scenes or tasks, whereas the level of enjoyment could sink due to general physical discomfort.

The use of VE technology is at the core of many games and other entertainment platforms. A major goal of the entertainment-based VEs is to deliver the experiences to the user in a more engaging and fun way. Unlike presence and SS, degree of enjoyment users experience in a VE has not yet been widely studied. Various factors that influence degree of presence and SS in a VE system may affect degree of enjoyment in similar or different ways. In addition, user experience regarding enjoyment might be impacted directly by the level of presence and/or SS a user experienced. Nichols et al. [7] studied relationships among presence, enjoyment, and SS during exposure to a VE. They hypothesized that subjects who enjoyed participating in the VE would report higher presence and lower SS. The results from this study suggested that presence added to the enjoyment of the VE and that there was no significant correlation between SS and enjoyment. However, another study described in a doctoral thesis found a negative correlation between SS and enjoyment [8]. Unpacking their hypothesis, it seemed to imply enjoyment as a mediator between VE attributes and presence or SS. Their results, nevertheless, seemed to reveal that enjoyment would be affected by presence or SS. This indicates that enjoyment could be associated with presence and SS, however, the underlying model and possible causal relationships among them need to be examined.

Furthermore, although degree of enjoyment is likely to be associated with the presence or SS that users experience in a VE, enjoyment may not just simply be a product of presence and/or SS. Various factors of a VE system could contribute to enjoyment independently of their influence on presence and SS. Planteb et al. [9] studied if virtual reality technology enhances the psychological benefits of exercise in a laboratory setting. Their results suggested that pairing VE with exercise enhances enjoyment, and reduces tiredness compared with VE or exercise alone. This study implied that coupling certain tasks in a VE was likely to increase level of enjoyment. In a VE requiring performance of certain tasks or activities, participants needed more mental focus. Lee et al. [10] examined the influence of mental focus on performance and enjoyment on students' course work. They addressed the role of mental focus in predicting enjoyment and performance and found that mental focus was positively related to both enjoyment and performance. Considering participants' exposure to a VE, they could just passively sit and do nothing but observe, or they could more actively perform certain tasks such as driving through a route in the VE. Passively observing versus actively driving in a VE involved different degrees of mental focus, and consequently led to different degrees of enjoyment during the VE experience. This led to a very interesting avenue for our research – exploring the effects of interactivity on enjoyment in a VE.

In summary, the sense of presence and SS have been considered as crucial aspects of users' responses in VEs; enjoyment should also be a crucial aspect to study since it could be particularly of interest to the VE-based entertainment applications. While studying user experience in a VE, we consider presence, SS, and enjoyment should be assessed simultaneously, in particular that numerous inter-correlations among presence, SS and enjoyment could exist. In addition, many studies have looked into the approaches of tuning VE variables (e.g. FOV of VE displays) to enhance the sense of presence. It is likely that factors that enhance presence might also impact the level of enjoyment and/or SS. It is therefore important to study both of the more positive aspects, i.e. presence and enjoyment, and the aftereffect, i.e. the SS, together when investigating the approaches of using different variables to enhance user experience in a VE. The objective of this research was to gain a deeper understanding of those relationships among presence, enjoyment, and SS, and seek to model user experience in a VE in a more systematic and holistic fashion. Furthermore, through the development of the model we hope to discover the approach of best tuning the VE variables, which will ultimately enhance user experience as a whole. In terms of the positive aspects of user VE experience, we hoped to find approaches to enhance presence and enjoyment and diminish SS; whereas in terms of aftereffect of user VE experience, we aimed to find approaches that alleviate SS while at the same time preserving presence and enjoyment.

2 Development of Assessment Tool

As discussed above, the levels of presence, enjoyment, and SS are the primary aspects for assessing user experience in VE in this research. Regarding the measurement of SS, the Simulator Sickness Questionnaire (SSQ) developed by Kennedy provides information to distinguish the multiple dimensions of the sickness including nausea, disorientation and oculomotor disturbance [11]. SSQ has been generally used as the instrument to monitor or assess the SS symptoms participants may develop in VEs and motion simulators. Our research used SSQ or Revised SSQ (RSSQ) [12], which is a variation of SSQ, to assess the user experience concerning SS. Regarding the measurement of presence and enjoyment, as there are no universally accepted measure techniques for presence and enjoyment in VEs, by referring to previous researches and consolidating the instruments from other studies, we developed a refined scale – the Enjoyment, Engagement, and Immersion Scale (E^2I Scale) to assess user experiences in VEs.

2.1 The E^2I Scale

Engagement and immersion are two elements that contribute to the idea of presence. Following Witmer and Singer's [2] approach to the concept of presence, engagement indicates the degree to which the subjects' attention is directed to the VE, similar to when one is engrossed in a novel or movie. Immersion is defined as the experience that one is wrapped in a surround, similar to being inside an elevator. Enjoyment is the feeling of pleasure or contentment during the VE experience. The Enjoyment, Engagement and Immersion Scale, or the E^2I Scale, was developed to assess enjoyment, engagement, and immersion experienced by participants in a VE.

Development of the scale began with the creation of Likert-scale (7-point) questions. Each question was considered an item in the construction of the scale. To examine engagement and immersion, 9 items were developed and modified based on the well-known presence questionnaires described by Singer and Witmer [13] and by Slater et. al [14]. These items were constructed based on the ideas of engagement (EG) and immersion (IM) associated with the concept of presence. These 9 items and their associated ideas were as follows. (Crayoland is the name of the VE used in the experiments.)

Q1. *How much did looking at Crayoland (the VE) involve you; i.e., how much did the visual scene attract your attention? (EG)*
Q2. *To what extent did events such as noise occurring outside Crayoland distract your attention from Crayoland? (EG)*
Q3. *How compelling was your sense of objects moving through space? (EG, IM)*
Q4. *How consistent were experiences in the virtual environment; i.e., to what extent did you feel as though you were actually moving through Crayoland? (EG, IM)*
Q5. *How completely were you able to actively survey or search the environment using vision? (EG, IM)*
Q6. *Were you involved in the memory task to the extent that you lost track of time? (EG)*
Q7. *How much did you have a sense of "being there" in the virtual environment? (IM)*
Q8. *During the time of the experience, which was strongest on the whole, your sense of being in the driving simulator room or in Crayoland? (IM)*
QMem. *A set of questions to assess structure of memory. (EG, IM)*

The enjoyment component of the E^2I Scale was based on the attributes of pleasure (PL) and satisfaction (SF) during a VE experience. The 5 enjoyment items were as follows. The item "QTime" was added to the questionnaire as the 15th item for the later experiments.

Q9. *To what degree did you feel unhappy when the experience was over? (PL)*
Q10. *How much did you enjoy yourself during the experience? (PL, SF)*
Q11. *Would you like to repeat the experience you just had? (SF)*
Q12. *How interesting was your experience in Crayoland? (PL)*
Q13. *How much would you be willing to pay to have a similar experience? (PL)*
QTime: *Circle the amount of time you estimate the experience you just had took.*

A unique feature of the E^2I Scale design was the memory test (QMem) embedded in the questionnaire. Previous research has suggested that the degree to which a subject memorizes the features in a VE may be an indicator of how "present" the subject is, and could therefore be useful in measuring presence. The memory structure for a VE was suggested to include the six dimensions [15] – types of objects, shapes, colors, relative locations, relative sizes, and event sequences. In our study, each of the six dimensions of the subjects' memory structure of the VE was assessed by a specific type of questions concerning the attributes of the dimension [16].

2.2 The Refinement of E^{2}I Scale

Evaluation of the E^{2}I Scale involved (1) using factor analysis to detect the underlying structure and extracting factors, (2) selecting and weighing items for each factor based on their contributions using reliability tests and some multivariate approaches, (3) examining the item qualities including discrimination power and location for each factor, and finally (4) developing overall index for each factor as a subscale based on the analysis of Item Response Theory (IRT). Two subscales – the Presence and Enjoyment Subscales were constructed in the E^{2}I Scale. Based on the responses to the items, standardized Z scores provided an index to indicate the levels of presence or enjoyment.

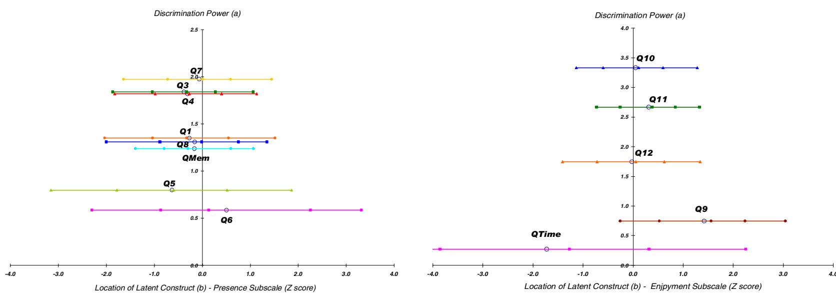


Fig. 1. The discrimination powers against the locations and thresholds from the ITR analyses for the 8 items in Presence Subscale (left panel) and Enjoyment Subscale (right panel)

Figure 1 (left) illustrates the qualities of the items in the Presence Subscale from the analysis of IRT, based on the discrimination power axis relative to the location axis for each item. The figure indicates that items Q7, Q3, and Q4 had greater discrimination powers than the others. Interestingly, both Q3 and Q4 were items related to vection (visually induced self-motion perception). This seems to suggest that degree of vection should be a significant indicator for assessing the degree of presence that a user experiences in a VE. The item Q7, which directly asked about the degrees to which a subject had a sense of “being there” in a VE, also appeared to have a greater capability for discriminating responses on presence, based on the response patterns from the data. On the other hand, the Item Q6 seemed to be less sensitive for discriminating presence. Item Q6, which asks about the degree to which subjects “lost track of time”, had a higher location on the Presence Latent Construct. This suggests that as users approach a higher presence level they gradually lose their ability to keep track of time. The implications of the results provided a deeper understanding and suggested possible approaches for refining the items in the Presence Subscale.

The qualities of the items in the Enjoyment Subscale are illustrated in Figure 1 (right), which similarly indicates that item Q10 had greater discrimination powers than the other items. Item Q10, which directly asked about the degree of enjoyment experienced in the VE, had the greatest capability for discriminating levels of enjoyment. Item Q9, which at the end of each trial asked about the feeling of “unhappiness” due to the VE experience terminating, exhibited a higher location on the Enjoyment Latent Construct. This implies that the user needs to approach a higher level of enjoyment in order to feel unhappy when the VE experience has terminated. The QTime item does not appear to be sensitive enough to reflect the difference of subjects' enjoyment experience. Part of the reason may be because subjects had differing estimates of how much time they spent in the VE, despite the fact that the stimuli were identical. Subjects may also have had difficulty in specifying the number of seconds in a given time period without being presented with a reference time period. In other words, it could be easier for the subjects if there is a time period being presented as a baseline, and they could do direct magnitude estimation of the target time period. The IRT results suggest the future refinement of this item. Again, these analyses provide a deeper understanding and suggest possible approaches for refining the items in the Enjoyment Subscale.

3 Experiment Settings

A Real Drive driving simulator (Illusion Technologies International, Inc.) including a full-size Saturn car (General Motors Company), three 800 x 600 pixel Sony Superdata Multiscan VPH-1252Q projectors, and 3 230 x 175 cm screens was used to display the VE for the experiment. The virtual world (Crayoland) was generated by the CAVE software library (developed at the EVL, University of Illinois, Chicago) using a Silicon Graphics Onyx2 system. Crayoland is a cartoon world that includes a cabin, pond, flowerbeds, and a forest. Participants sat on the driver's seat of the car and wore the CrystalEyes stereo glasses (StereoGraphics Inc.) while moving through the Crayoland VE along a trajectory that included left and right turns (yaw movements) as well as forward translation. A number of variables including image characteristics (such as FOV and stereoscopic depth cues), visual motion frequency, interactivity, and visual cues or intervention for motion prediction were manipulated in a series of 6 experiments as listed in Table 1. The effects of the variables on presence, enjoyment, and SS were measured in all of 6 experiments using E^2I and SSR/RSSQ. The experiment environment is illustrated in Figure 2 and the details of the experiments can be found at [16][17][18][19][20].

Table 1. A list of independent variables applied in the 6 experiments included in this study

Exp. ID	Independent Variable	# Conditions	# Participants
4.1	FOV	4	10 (5W, 5M)
4.2	Stereo	4	10 (4W, 6M)
4.3	Visual Motion Frequency	6	20 (11W, 9M)
5	Interactivity	4	12 (5W, 7M)
6	Visual Cues for Motion Prediction	3	12 (7W, 5M)
7	Virtual Guiding Avatar (VGA)	4	12 (8W, 4M)

4 Results and Discussion

The results from the Experiments 4.1, 4.2, 4.3, and 5 showed that manipulation of those VE variables (i.e. FOV, stereo, visual motion frequency, and level of interactivity) can significantly contribute to the positive aspects of user experience,

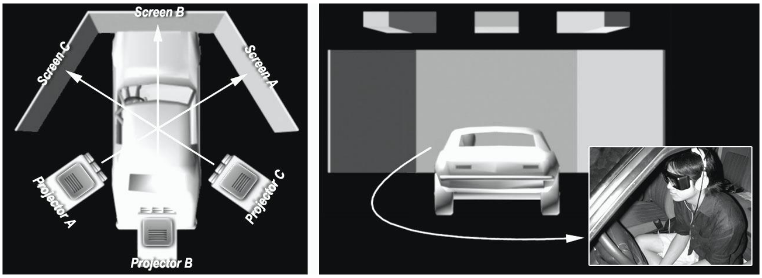


Fig. 2. Equipment setting. The left panel shows a bird's eye view. The right panel shows a back view of the set up as well as a close-up view of subject's posture in the experiment.

including presence and enjoyment. These factors, however, also have significant negative impact on user experience. In addition, new approaches to alleviating SS were developed and then tested in Experiments 6 & 7. These approaches using visual interventions in the VE scenes were demonstrated to be effective for SS reduction. The impact of the use of these approaches on presence and enjoyment was measured during the experiments, in order to evaluate the potential side-effects of using these visual intervention techniques on the positive aspects of user experience. Several findings derived from these experiments can be applicable as design implications for enhancing user experience in immersive display systems.

4.1 Enhancing Presence and Enjoyment

Figure 3 (left) illustrates that the sense of presence in a VE is significantly associated with FOV, stereopsis, visual motion frequency, and levels of interactivity. It suggests that a VE with the following four attributes is likely to make a more compelling user experience as a result of higher levels of presence: (1) large FOV (i.e, 180°), (2) visual scene presented in stereo, (3) visual scene moving at a very low frequency (i.e., 0.03 Hz, and (4) an active mode of interaction with the visual scene. Figure 3 (right) indicates that the degree of enjoyment in a VE is significantly associated with visual motion frequency, levels of interactivity, and use of visual interventions. It suggests that a VE with the following three attributes is likely to make a more interesting user

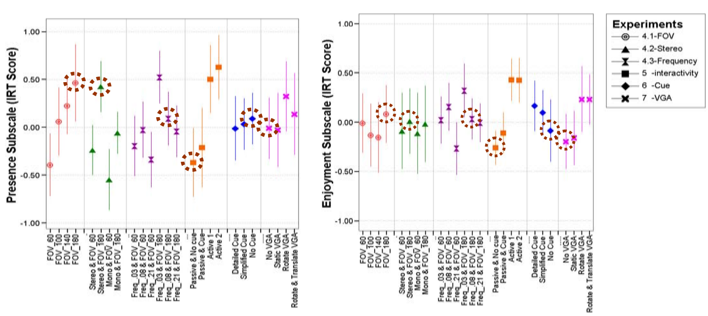


Fig. 3. Responses on Presence (left) and Enjoyment (right) Subscales

experience, resulting from a higher degree of enjoyment: (1) a visual scene moving at a very low frequency (such as 0.03 Hz), (2) an active mode of interaction with the visual scene, and (3) the use of visual interventions to alleviate SS.

Manipulating theses factors would vary the resulting levels of presence or enjoyment in a VE. A practical question to look into is which factors have large effects than the others. Across the 6 studies, a common condition bearing similar properties was included in each experiment, as identified in circles in Figure 3. These properties included a 180° FOV, stereo graphics, 0.05~0.08 Hz continuously oscillating visual motion, subjects in passive mode, and no visual intervention for SS alleviation applied. Considering these common conditions into the same baseline level, the results imply that levels of interactivity play a key role in enhancing both presence and enjoyment. It has become clear that when participants have active control of a VE, the degree of presence they experience as well as the enjoyment level could be higher than when the other factors were manipulated. In addition, FOV also makes a significant contribution in levering the degree of presence. However, when the FOV reaches its ceiling (full-coverage FOV: about +/- 110° from the center of a visual field), providing active control of scene motion will have a great influence on participant’s sense of presence in a VE.

4.2 Reducing the Aftereffect – SS

Figure 4 (left) indicates that SS induced in a VE is significantly associated with FOV, stereopsis, visual motion frequency, levels of interactivity, and the ability to predict upcoming motion. It suggests that SS could be reduced in an immersive display system bearing the following attributes: (1) a small FOV (such as 60°), (2) visual scene not presented in stereo, (3) visual scene moving at a very low frequency (such as 0.03 Hz), (4) allowing VE users to be in an active mode interacting with the visual scene, and (5) using visual interventions to provide predictive cues for upcoming motion. Similarly, taking the common conditions into account, the overall trend again suggests that levels of interactivity play a significant role concerning SS induction in a VE. When users have active control of scene motion, the degree of SS reduction is larger than when varying the levels of the other factors. Since having control over the scene motion implies having the theoretical capability of being able to predict upcoming visual motion, two kinds of visual interventions were developed to alleviate SS: unobtrusive prediction cues and the VGA. However, these procedures were not as effective as directly providing the active control of scene motion. The figure also indicates that decreasing FOV or having low frequency visual scene motion would also reduce the degree of SS.

4.3 Enhancing User Experience as a Whole

As noted above, the three primary aspects of user experience in a VE – presence, enjoyment, and SS could intimately link and may be correlated with one another. In order to seek for the approaches to enhance user experience as a whole, it is important to use a more holistic approach to further demonstrate the relationships among the three aspects in relation to these VE variables used in the experiments. Using a response surface based on a nonparametric smoothing density estimation of

Enjoyment Subscale scores by SSQ and Presence Subscales scores, the resulting response surface model, as illustrated in Figure 4 (right), visualizes these complex relationships. It suggests four properties that concern correlations between presence, enjoyment, and SS: (1) a positive correlation between presence and enjoyment, (2) a negative correlation between SS and enjoyment, (3) no specific correlation between presence and SS, and (4) either a decline in presence or a growth in SS resulting in a decrease in enjoyment.

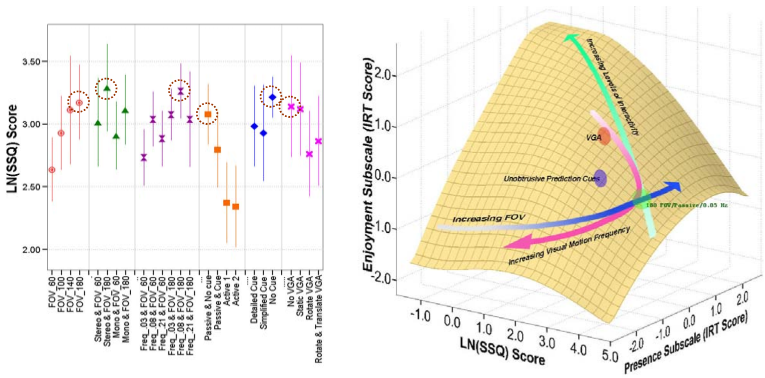


Fig. 4. Responses on SSQ (Left). The right panel shows a response surface model summarizing the overall trend of the effects of different factors on presence, enjoyment and SS.

Based on this construct, we summarize the following practical design implications for enhancing user experience in VEs:

- Increasing level of interactivity in a VE is the most profound way of enhancing the overall user experience of all factors examined in this dissertation. It leads to a high degree of presence and enjoyment as well as a low degree of SS.
- Visual motion frequency influences user experience in a complex way. If possible, it is useful to decrease the frequency when the visual motion is around .05 to .08 Hz. As the frequency of visual motion is above the .05 to .08 Hz rage, the higher the frequency, the lower the resulting presence and SS will be.
- Visual interventions for motion predictions may enhance user experience. Both approaches in this study effectively reduced SS. A VGA type of intervention may be more effective in enhancing user experienced and might increase presence.
- FOV has a significant impact on both presence and SS. However, varying the FOV does not have a great effect on enjoyment. As indicates in the figure, increasing FOV leads to an increase in both presence and SS. SS causes the overall user experience to decrease. In Figure 3 (right), the relatively flat effect on enjoyment across different FOV also reflects this finding, where enjoyment level remained similar when FOV was the major factor being manipulated.

The model presented in Figure 4 (right) clearly summarizes the findings across the different studies described in this paper. The effects of FOV, visual motion frequency, interactivity, and visual interventions on presence, enjoyment and SS are visualized.

This model represents an overall trend in the findings. Further studies are needed to investigate the effects regarding detailed variations within each factor as well as the interactions among these factors.

References

1. Barfield, W., Sheridan, T., Zeltzer, D., Slater, M.: Presence and performance with in virtual environments. In: Barfield, W., Furness, T.A. (eds.) Virtual Environments and Advanced Interface Design, pp. 473–513. Oxford University Press, Oxford (1995)
2. Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: a presence questionnaire. Presence: Teleoperators and Virtual Environments 7, 225–240 (1998)
3. Cobb, S.V.G., Nichols, S., Ramsey, A., Wilson, J.R.: Virtual reality induced symptoms and effects (VRISE). Presence: Teleoperators and Virtual Environments 8, 169–186 (1999)
4. Kennedy, R.S., Stanney, K.M.: Aftereffects of virtual environment exposure: Psychometric issues. In: Smith, M., Salvendy, G., Koubek, R. (eds.) Design of computing systems: Social and ergonomic considerations, pp. 897–900. Elsevier Science Publishers, Amsterdam (1997)
5. Reason, J.T., Brand, J.J.: Motion Sickness. Academic Press, London (1975)
6. Griffin, M.J.: Handbook of Human Vibration. Academic Press, London (1990)
7. Nichols, S., Haldane, C., Wilson, J.R.: Measurement of presence and its consequences in virtual environments. International Journal of Human-Computer Studies 52, 471–491 (2000)
8. Nichols, S.: Virtual reality induced symptoms and effects (VRISE): Methodological and theoretical issue. Doctoral thesis, University of Nottingham, Nottingham, U.K (1999)
9. Plante, T.G., Aldridge, A., Bogden, R., Hanelin, C.: Might virtual reality promote the mood benefits of exercise. Computers in Human Behavior 19, 495–509 (2003)
10. Lee, F.K., Sheldon, K.M., Turban, D.B.: Personality and the goal-striving process: The influence of achievement goal patterns, goal level, and mental focus on performance and enjoyment. Journal of Applied Psychology 88, 256–265 (2003)
11. Kennedy, R.S., Lane, N., Berbaum, K., Lilienthal, M.: Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. International Journal of Aviation Psychology 3, 203–220 (1993)
12. Kim, D.H.: Development of method for quantification and analysis of simulator sickness in a driving simulation environment. Doctoral dissertation, Hanyang University, Seoul, South Korea (1999)
13. Singer, M.J., Witmer, B.G.: Presence measures for virtual environments: Background & development. Draft ARI Research Note, U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA (1996)
14. Slater, M., Usoh, M., Steed, A.: Depth of presence in virtual environments. Presence: Teleoperators and Virtual Environments 3, 130–144 (1994)
15. Usoh, M., Catena, E., Arman, S., Slater, M.: Using Presence Questionnaires in Reality. Presence: Teleoperators and Virtual Environment 9, 497–503 (2000)
16. Lin, J.J.W., Duh, H.B.L., Parker, D.E., Abi-Rached, H., Furness, T.A.: Effects of Field of View on Presence, Enjoyment, Memory, and Simulator Sickness in a Virtual Environment. In: Proceedings of IEEE Virtual Reality 2002, pp. 164–171 (2002)

17. Lin, J.J.W.: Enhancement of User-Experiences in Virtual Environments that Employ Wide-Field Displays. Dissertation for Doctor of Philosophy, University of Washington, Seattle, USA (2004)
18. Lin, J.J.W., Razzaque, S., Parker, D.E.: Effects of Simulated Motion Frequency in Virtual Environments. In: the Proceedings of the International Conference on Theoretical Issues in Ergonomics Science (TIES 2005) (2005)
19. Lin, J.J.W., Parker, D.E., Lahav, M., Furness, T.A.: Unobtrusive Turns Prediction Cues Reduced Simulator Sickness during Passive Motion in a Driving Simulator. Ergonomics 48, 608–624 (2005)
20. Lin, J.J.W., Abi-Rached, H., Lahav, M.: Virtual Guiding Avatar: An Effective Procedure to Reduce Simulator Sickness in Virtual Environments. Proceedings of ACM SIGCHI, 719–726 (2004)

AIPlayer: A Platform of Intelligent Simulation of Virtual Human in Virtual Environment

JianGuo Liu[1,2], YanSheng Lu[1], and JiuYun Chen[3]

[1] Institute of Computer, Huazhong University of Science and Technology,
430074 Wuhan, China
[2]Automobile School,Wuhan University of Technology,
430070 Wuhan, China
[3] Management School,Wuhan University of Technology,
430070 Wuhan, China
ljg424@163.com,
chenjiuyun@163.com

Abstract. The research of intelligent simulation of virtual human in virtual environment is an interesting direction in the domain of AI after the research of virtual human's behavior simulation. This paper presents a new research platform named AIPlayer in order to promote this work. Virtual environment is come close to the real world. Some characteristic in virtual environment can realistically represent the real world, at the same time it can provide some advantages to science research . This virtual environment may be MMORPG or virtual reality. The motive of the research is to simulate a virtual human who can autonomous live in the virtual environment. We named this research as intelligent simulation of virtual human in virtual environment. First, this paper argues the significance of this research in theory and application, then analyses the demand of the AIPlayer, the characteristic of its living environment, the aim of AIPlayer and some correlative research domain may be involved, then it describes the architecture of AIPlayer and the core of AIPlayer: BSOAA(based smart object agent architecture), at last a prototype of AIPlayer in virtual environment is introduced.

1 Introduction

Since 1950 while the paper of Turing's "Can machine think?"[1]has been published, people continuously inspired for this goal, however, even at present, when the computer operating speed is much higher than those days, can machine think? This question still caused the researchers to be puzzled. Although in some certain special domain such as game, some specific domain expert system and so on, we have obtained the significant achievement, but when we change the vision to the simple daily life, we are disappointed, some people even started to suspect, the machine really can think? Although some searchers already studied the robot which can simulate the person to stand and walk by his two legs or to dance, to go up and down the staircase or even to ride the bicycle, but in replying the question whether the machine can

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 434–442, 2007.

think, similarly we lacked the enough evidence, although started from the Turing, we believed one day the machine may think[2].

Looking back the history of research on artificial intelligence, we can clearly find that according to the difference of research object and appliance scope, we divide it into three phases. The first phase is to take the machinery theorem prove as object and use it in the math axiom testifying; The second phase is to create some expert system fitted to some special scopes; The third phase is to take the realism world as study object[3].

Even nowadays we take realism world as study object, the researchers deeply study in some traditional scope. These domains involve chess, I-go, expert system, the construction of commonsense knowledge base, different chatting system etc.

The study of artificial intelligence comes into an odd circle: on the one hand we would like to bring the realism world into study scope; on the other hand we have to hunt for some means in traditional scope which can be used in realism world such as the methods of presentation knowledge, machine learning. The most important reason is that: the realism world is too complex to be studied.

This paper put forward a new research object which carrying on the research of the artificial intelligence in a virtual digital environment. It means that creating a virtual person who can be independent live in this virtual environment. In the meantime the really person will take part in this environment. This kind of virtual environment can be a MMORPG or perhaps virtual reality. Our target is to imitate a person, who can independently live in this environment. Possibly this is the most close to artificial intelligence research. In order to build a research platform, the author constructs AIPlayer.

The AIPlayer is a virtual human intelligence research platform, it can completely control a virtual role existing in a virtual environment, living independently and own an ability to communicate with the other roles in the virtual world.

Some scholars of the abroad carried on similar research. For example in California University research group carried on Gamebots in 2002, attempting to build an independent virtual player system in a virtual world. The virtual world is UT 2003. it is a first person action game[4][5].

The arrangement of this paper is that section 2 carries on AIPlayer demand analysis, including constitute of the virtual environment and the advantages of intelligence simulation, the target of the platform; section 3 introduces the AIPlayer's realization technology including a prototype of the core named BSOAA, and the AIPlayer total structure, section 4 summaries this paper.

2 AIPlayer Demands Analysis

2.1 AIPlayer Existing Environment Characteristics

(1)AIPlayer lives in the virtual world close to the mankind reality world, this similarities involve the society, cooperation and competition. A virtual social background, the demand living in the virtual society is including Common sense, basic ability, the ability of communication with others and the ability of study. This kind of context environment and common sense is limited and controlled, but to the artificial

intelligence it is really very beneficial through controlling the problem scale, so we can broaden the virtual human's intelligence gradually from easy to complex.
(2)Simulation of human’s daily life. Compared with the real world, AIPlayer’s environment is simple because it is abstract and imitation of reality life.
(3) A controllable and visual environment. The virtual environment is different from the real world because it is consist of many controllable and visual items which are convenient for redearch.

The virtual environment is including story background, story rules, virtual objects, and other virtual players and so on. The virtual role controlled by the true player will complete various mission in the virtual environment such as acquiring various product, carrying out the true person's amusement purpose.

We will consider the difference between AIPlayer and the two kinds of simulation in the virtual environment: the artificial intelligence in the game[6] and the Bot which can control the role in the game[7].

Different from the artificial intelligence in general game, the AIPlayer is independent of the game. While the artificial intelligence used in the general game, mostly is aim to raise the intelligence of various role in the game, making its behavior more like a true player. By applying the artificial intelligence technique in the game, we can construct this environment more like the real world. But the AIPlayer’s target is to control the role, product and other objects in this virtual world in order to acquire the feeling similar to the mankind.

Comparing with the Bot which can control the role in the game, the AIPlayer has two basic differences:

(1)the AIPlayer’s behavior is more like human being’s: observing environment by the video and sound, responding by the mouse and the keyboard to control a role. Most Bots acquire the environment information and control game role through the interface provided by the game or directly crack the data segment, so as to get all states of the game.

No matter what adopting legal methods controlling the game through interface provided by the game developer, for instance, GameBobs project is connected with the game through interface which provides by UT2003 development group, thus Game Bots can completely control the game or illegal method cracking data segment about the game world, thus obtaining the entire game world condition understanding, for instance current majority network game automatically hanging machine software. The simulation world in the viewpoint of Bots is completely open, it is accurate and unmistakable, but AIPlayer obtains the simulation world condition information is incompletely, possibly makes mistakes, the ability of processing incomplete knowledge and the indefinite knowledge is AIPlayer and the general Bots basic difference.
(2) AIPlayer is different in the game goal with general Bots. Living in the simulation game world, to general Bots it does not have any special significance, but to AIPlayer, living in the game world, has the explicit goal. This kind of goal may be comes from AIPlayer in, also may be comes from exterior: Assigns instruction or goal. How to describe this kind goal coming from the AIPlayer interior is one of our research contents.

2.2 Target of AIPlayer

The Ultimate Target of AIPlayer. The requirements of exploring a virtual human who can independently survive in the virtual game environment are shown as following.

(1) acts according to own hobby, the emotion choice hypothesized world;
(2) initiatively gains about this hypothesized world knowledge;
(3) like the humanity players, only can use the video frequency the way to observe the game picture, distinguishes the game advancement through the sound, according to plays the picture and the sound content, makes the corresponding decision-making;
(4) like the humanity players, only can controls the hypothesized role in the game through the keyboard and the mouse;
(5) like the humanity players, may be initiative, have the goal to play the game;
(6) can do cooperation and the competition with other player;
(7) like the humanity players, may finally leave the hypothesized world.

Obviously the AIPlayer's ultimate target is too complex to realize in current, it needs a long study process, although the author believes that this target certainly can be realized. Based on the present situation in the intelligent research as well as the computer hardware, the author suggests the basic target of AIPlayer.

The Basic Target of AIPlayer

(1) firstly, assigns the hypothesized world;
(2) obtains the knowledge of the hypothesized world under expert's instruction;
(3)through the video frequency observe game, makes corresponding decision-making, temporarily does not consider the sound information obtaining;
(4) controls the game through the keyboard and the mouse the hypothesized role;
(5) beforehand assigns in the hypothesized world goal;
(6) and other plays the family to carry on limited the cooperation andthe competition which controls;
(7) leaves the hypothesized world instruction to produce by exterior.

The Correlative Domain of AIPlayer. As the author pointed out above, the artificial intelligence research has faced a strange problem. On the one hand, people want to bring the realistic world into the research scope; on the other hand, people have to look for various methods applied to the realistic world in the traditional research realm, including the representation methods of knowledge…etc. The gap between the traditional research domain and realistic world is huge, therefore, it seems quite difficult to apply the traditional research achievement success in the realistic world.

The author provides a new research object: virtual artificial intelligence simulation in virtual environment trying to look for a breakthrough between the traditional research realm and the realistic world. One the one hand, it should abandon the complicatedly and the unpredictable facts of realistic world that is the very indetermination and incompletion preventing the traditional research results from applying to realistic world. On the other hand it should abandon the too simple and idealized static state of traditional research realm that restricts the application of artificial intelligence technique. It should be taken place by virtual digital environment where all object can be controlled. In this environment, from the simple problem to the complicated problem, the static state environment to the dynamic state environment, the assurance knowledge to indetermination knowledge, complete knowledge to incompletely knowledge etc., which are the opposite characteristic of tradition research realm and realistic world characteristic can be simulated. The virtual artificial intelligence has a process of evolution, it's a gradual evolution process from traditional research realm to realistic world.

In a word, AIPlayer's research realm is the transitional that lies between the tradition and reality. From the AIPlayer's target we can find that it involves several research domains of the artificial intelligence, such as :

(1) The pattern recognition
The main target of this research is: Identifying the state of environment and object from the video appearance; identifying the current environment state from the sound. We have accomplished various technology in pattern recognition, these methods are applied to realistic world in a certain degree, however, pattern recognition usually do correspondingly independent work as a independent system such as identifying fingerprint, identifying human's face, identifying action...etc. We have obtained many useful results in the machine vision research realm. However, As a system of integration, the research in robot with machine sense of vision which can exist in realistic world are still carrying on. Compared with the realistic world, the virtual world have many advantages. In AIPlayer the research of pattern recognition has more advantage because the environment reproduction is very good, image mass is controlled, sample book data acquisition is controlled, the problem scale is controlled and limited.

(2) Computer emotion
AIPlayer must contain the emotion representation in ultimate tartet. In AIPlayer basic goal, the emotion representation may be very basic. This kind of research is the hot domain in the field of artificial intelligence studies at present.

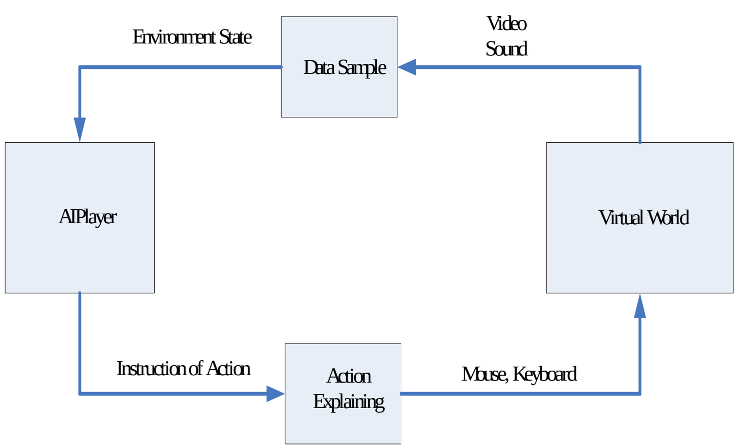


Fig. 1. AIPlayer and Virtual World

(3) Machine learning
In the current artificial intelligence research one of main barrier and development directions is the machine learning. The machine learning, the computer science, the psychology, the cognitive science and so on each discipline all has the closely relative, many theories and the technical questions are still in the progress of research. Having the independent learning ability is the AIPlayer ultimate objective [8]. AIPlayer in the machine learning aspect goal is to Integrate independent learning algorithm. AIPlayer can be regarded as a research platform to each kind of machine learning algorithm, simultaneously, AIPlayer itself also integrated the many kinds of learning algorithm. It is the symbol of succeed that AIPlayer has the independent learning ability.

3 AIPlayer Realization Technology

In this section the author will discuss AIPlayer's realization technology, including architecture of AIPlayer, BSOAA(the core of the AIPlayer), and a prototype system.

3.1 Architecture of AIPlayer

Fig 1 indicates the relation of AIPlayer and virtual worlds. The AIPlayer is constituted by the following components: AIPlayer core the data processing of video, explaining and execution of controlling action. AIPlayer core is a hybrid agent based on intelligent objects, taking charge of the representation and reasoning of knowledge, information obtaining, state analysis decision-making analysis etc. The data processing of video takes charge of obtaining video data, pretreatment, objects identification of virtual worlds. The input is the video data of the virtual worlds, and the output is the objects and their state. The explanation and executing of controlling action, changing the APIlayer core's decision into the action of the mouse keyboard, and obtaining the mouse keyboard's action. The input parts are AIPlayer's decision-making action, the output is the operation sequence of the mouse keyboard. The total architectur of AIPlayer is shown as following fig 2.

Then the AIPlayer core named BSOAA will be introduced emphatically.

3.2 AIPlayer Core: BOSAA

As to increasing the ability of KR and reasoning, a new method named KR method based introspectional object has been introduced into BSOAA. This KR method provided by the author in other paper is KR method based smart object, which is an entity unit that can provide reasoning engine supporting decision. Those entity units can denote some abstract objects such as knowledge units. They can also represent object such as material entity in the world. Inside the smart object, data layer, knowledge layer, data processing layer and the reasoning engine methods are encapsulated. The only way to access smart object from outside is through methods and interfaces in it. In conclusion, smart object is an active object holding multi threads encapsulated lots of KR methods and reasoning engines. Limited by space of this paper, the more details of introspectional object based smart object will not be introduced.

Recently, agent researchers normally agreed to the idea that the behavior of agent was controlled by its mental states. In other words, it means that the mental fact of agent such as beliefs, desires, intentions, capability, etc. are key elements which controlling agent behavior. So, how to figure out the mental states of agent and how these mental stats influent the agent behavior is most important. In the other hand, researchers are debating upon which mental facts are most necessary and basic in depicting agent autonomy. They presented lots of kind of agent model made up of different key mental states. Among those models, the most famous was BDI agent. This model pointed out that the behavior of agent was only controlled by its Beliefs, Desires, and Intentions. So, the most key fact in agent's mental were Belief, Desires and Intentions. But when using BDI agent in Open Environment , many problems there will normally be found . Some of them are as followed:

1) Can't efficiently explain the agent's behavior in dynamic environment: Though BDI agent ever successfully adapted to static environment, the performance of this kind of agent in dynamic environment was not as good as formerly. Because BDI agent didn't involve the effect of the agent's behavior, neither involved the capability of its action, this defect limited its development.
2) Can't resolve the adaptability to the change in Open Environment: As some important fact in mental such as instinct, sensibility, or philosophy didn't involve in BDI agent, the autonomy of this kind of agent was greatly limited in some extreme situation.

In this paper, some key mental facts will be added to BDI agent, include instinct, sensibility or philosophy and all kinds of knowledge relative to agent, especially some knowledge about the effect of the agent's behavior.

So, the definitions of the mental states in BSOAA are as followed:

Definition 1: the mental states of agent is a 8 tuples as M=(SSmtOs,ESmtOs,InstDesires,SenseDesires,Philosophy,Hacts,Lacts,Evts):

1) SSmtOs: Implying Self Smart Object set, including all facts and knowledge about agent itself.
2) ESmtOs: Implying Environment Smart Object set, including all facts and knowledge about environment surrounding agent.
3) InstDesires: Implying agent's instinct desires set.
4) SenseDesires: Implying agent's sense desires set which reasoning from facts and knowledge agent has been known.
5*) Philosophy*: Implying agent's sensibility desires set.
6) Hacts: Implying agent's high layer actions set.
7) Lacts: Implying agent's atomic actions set.
8) Evts: Implying event set, including inside events and outside events.

Based smart object KR method, BSOAA is defined as followed:

Definition 2: BSOAA is 8 tule as (M,SOManager,Upd _InstDisire,Upd_SenseDisire,Upd_Philosophy,Desire_SelHAction_SelLAction_Sel):

1) M: Implying agent's mental states.
2) SOManager: Implying Smart Object Manager.

3) Upd_InstDisire: Implying the function which update instinct desires.
4) Upd_SenseDisire: Implying the function updates rational desires.
5) Upd_Philosophy: Implying the function updates agent's philosophy.
6) Desire_Sel: Implying the function that selects which desire is active.
7) HAction_Sel: Implying the function that selects high layer action.
8) LAction_Sel: Implying the function that selects atomic action.

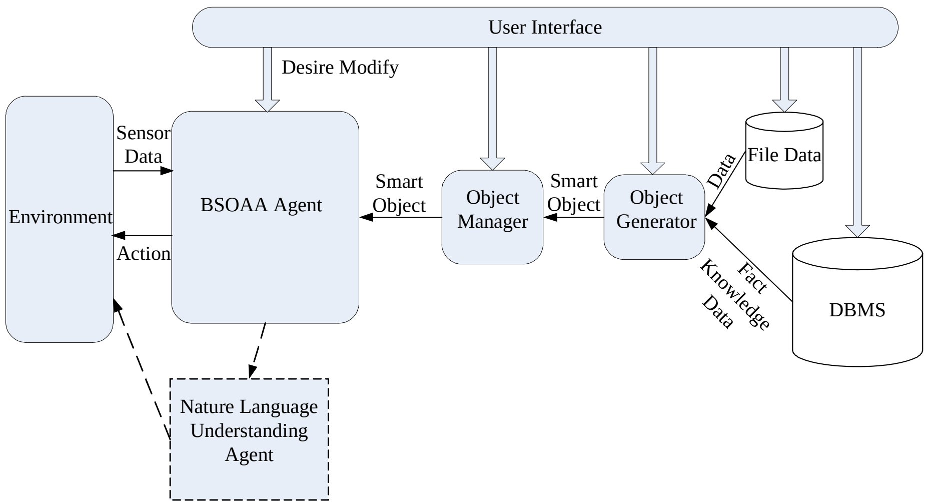


Fig. 2. Architecture of AIPlayer

3.3 Overview of AIPlayer

The author of this paper has structured an application based BSOAA aimed at simulation of virtual human intelligence named AIPlayer that has been shown in figure 2. The purpose of AIPlayer is to simulate the intelligence of virtual human, that means it should have more higher autonomy, more higher adaptability in Open Environment than normal agent. In fact, AIPlaer is living in a digital virtual world. The outstanding characteristic is that AIPlayer will completely ignore geometry character, even the simulation of movement when comparing with other type of virtual human. AIPlayer will focus on simulating intelligence behavior when environment states changing, include interacting with the other agents in the virtual world, collaborating with them. Especially AIPlayer has philosophy, that means its behavior will entirely controlled by itself, live for itself. The architecture of AIPlayer has been shown in figure 2.

The core component is BAOAA agent . The technology of object-oriented will be applied in AIPlayer. The rule of virtual world, knowledge and facts about environment will be saved in database or files according to their properties. When AIPlayer is starting up, Object Generator will be in charge of drawing knowledge or data from the knowledge source or data source. Object Generator will generate all smart objects needed by AIPlayer. Once generated in Object Generator, smart object will be sent to Object Manager. The states of smart objects and the behavior of those will be effect by Object Manager through message. AIPlayer will receive picture taken by video camera every a certain time, it will reason the situation, then take action like a real player. Through user interface the mental states of AIPlayer could be modified, thus instinct desire and philosophy of AIPlayer may be changed. At the same time, data, facts and knowledge in files or in database also may be updated. Even the function of the Object Manager and Object Generator may be modified by script. In figure 2, the dashed rectangle implies future work: Nature Language Understanding Agent (NLU Agent) who can talk with other players in nature language. That will be a new challenge.

4 Conclusion and Future Work

Under the digital, virtual environment, this paper provides a research platform named AIPlayer that can make the study of virtual human's intelligence simulation more convenient. The BSOAA which is core of AIPlayer has also been introduced. In future we will gradually consummate each part of AIPlayer, expands BSOAA `s function, and make it to support natural language understand in a certain context environment.

References

1. Turing. A.: Computing Machinery and Intelligence, Mind 49, 433-460 (1950). Published under the title Can a Machine Think. In: Newman, J.R. (ed.): The World of Mathematics, vol. 4, pp. 2099–2123, Simon & Schuster (1956)
2. Whitby, B.: The Turing Test: AI's Biggest Blind Alley? In: Millican, P., Clark, A. (eds.) Machines and Thought: The Legacy of Alan Turing, vol. 1, pp. 53–62. Clarendon Press, Oxford (1996)
3. Russell, S.J., Norving, P.: Artificial Intelligence: A Modern approach, 2nd edn. Prentice-Hall, Englewood Cliffs (December 20, 2002)
4. http://www.planetunreal.com/gamebots/index.html
5. http://www.unrealtournament.com
6. Wavish, P., Graham, M.: A situated action approach to implementing characters in computer games. Applied AI Journal (1995)
7. http://www.planetquake.com/emporium/
8. Tom, M., Mitchell, T.M.: Machine Learning. McGraw-Hill, New York (March 1, 1997)
9. Clark, T.: The Turing Test as a Novel Form of Hermeneutics. International Studies in Philosophy 24(1), 17–31 (1992)

Understanding RUTH: Creating Believable Behaviors for a Virtual Human Under Uncertainty

Insuk Oh and Matthew Stone

Department of Computer Science
Rutgers, The State University of New Jersey
110 Frelinghuysen Road, Piscataway NJ 08854-8019

Abstract. In pursuing the ultimate goal of enabling intelligent conversation with a virtual human, two key challenges are selecting nonverbal behaviors to implement and realizing those behaviors practically and reliably. In this paper, we explore the signals interlocutors use to display uncertainty face to face. Peoples' signals were identified and annotated through systematic coding and then implemented onto our ECA (Embodied Conversational Agent), RUTH. We investigated whether RUTH animations were as effective as videos of talking people in conveying an agent's level of uncertainty to human viewers. Our results show that people could pick up on different levels of uncertainty not only with another conversational partner, but also with the simulations on RUTH. In addition, we used animations containing different subsets of facial signals to understand in more detail how nonverbal behavior conveys uncertainty. The findings illustrate the promise of our methodology for creating specific inventories of fine-grained conversational behaviors from knowledge and observations of spontaneous human conversation.

Keywords: Uncertainty expression, uncertainty recognition, embodied conversational agent, talking head, RUTH.

1 Introduction

People can achieve natural and effective interaction with each other partly by expressing and recognizing how much they understand or do not understand each other [1]. People have a rich repertoire of nonverbal cues – on their faces, heads and bodies, and in their voices – that they can use to show how much uncertainty they have about the state of the conversation; see for example [2, 3]. Our prior work underscores that viewers can attend to these signals [3]. When asked to rate the uncertainty conveyed in videos of human speakers, judges' ratings were highly correlated with the level of uncertainty originally reported by the speaker.

Preliminary work suggests that conversational agents will also achieve more natural interaction with users when those agents signal what they have understood and signal what they remain uncertain about [4]. In fact, uncertainty communication may well prove particularly important for conversational agents, because we can expect agents to have less reliable perceptual information than people in conversation and

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 443–452, 2007.

less complete knowledge of their partners and their contexts. Conversational agents will therefore have to take risks to optimize their behavior in the presence of uncertainty rather than to find one absolute best way that is guaranteed to work in all situations; and, to collaborate effectively with their partners, they will also need to signal that uncertainty.

As a stepping stone toward designing interactive agents that achieve this ultimate goal, we focus here on what nonverbal behaviors we should add into computer agents to allow them to express their uncertainty in face-to-face interaction and how we can implement these behaviors practically and reliably. We have identified and annotated a set of uncertainty signals through a systematic coding methodology and then implemented the simulated behaviors onto our ECA (Embodied Conversational Agent), RUTH [5]. The implementation exploits the fact that ECAs resemble the appearance of real humans and are capable of carrying out many common behaviors we saw in people, so we can create animations in RUTH by analyzing and specifying the time course of behaviors as observed in natural human-human communication. As a follow-up to our communication and animation research [3, 6], here we aim at evaluating how well people can recognize uncertainty expressions as animated with RUTH, and at using RUTH to discover the effects of different layers of nonverbal signals that we have found in uncertainty communication.

2 Believable Interaction with a Believable Agent

Making an agent believable involves addressing users' expectations for both the appearance and the behavior of the agent [7]. And the bar keeps getting higher – as users gain increasingly diverse experiences with digital entertainment and computer-mediated communication in everyday life, they may develop correspondingly heightened expectations of ECAs. System designers can't realize agents with all the naturalness users might call for. As Nass et al., [8] point out, as designers, we must identify where we can get by with endowing our agents with a transparently synthetic character, and where we must make them more closely resemble humans – on what dimensions and on what level of detail. The evaluation we report here is a check on how well we have managed these design trade-offs.

2.1 RUTH

Our research starts from our embodied conversational agent, RUTH (Rutgers University Talking Head) [5]. RUTH is a freely available cross-platform real-time facial animation system developed by the VILLAGE Laboratory at Rutgers University. It is available at http://www.cs.rutgers.edu/~village/ruth/. A key goal of developing this animation platform has been to offer a methodological tool for developing and testing theories of functions and behaviors that occur during natural face-to-face conversation. Our work on the communication of uncertainty built on the initial design of RUTH and helped shape a number of extensions to the underlying capabilities of the system [6].

Fig. 1. RUTH, our talking head agent

RUTH displays a talking head rendered in three dimensions – the animation shows no hands or torso, just a head and neck. Figure 1 is a snapshot of RUTH smiling. RUTH's appearance is intentionally designed to be somewhat ambiguous as to gender, age, and race. RUTH is most often described as male although it has been seen as a female figure. With its large, stylized eyes, RUTH can be reasonably believed to be a range of ages, from elementary school to young adult. RUTH has the characteristics of a number of different ethnicities. The idea was to appeal to a wide range of users. However, our experience also taught us that it was also possible for certain users to find RUTH a bit discomfiting because RUTH seems to lack a specific identity.

2.2 Encoding Uncertainty Facial Signals onto RUTH

RUTH takes input for behaviors on the face and head, including the possibility of most of the facial actions inventoried in the standard Facial Action Coding System (FACS) [9, 6]. RUTH can work in conjunction with the Festival speech synthesis system to derive a sound file together with animation instructions for corresponding lip, face, and head movement. These specifications then guide a synchronized realization of the animated speech in real time. The work reported here describes silent

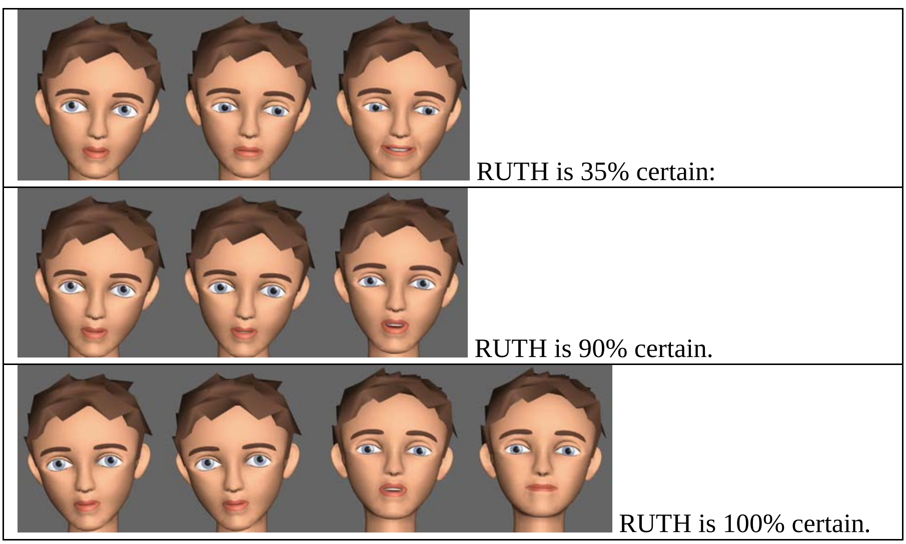


Fig. 2. RUTH's simulation of real humans' uncertainty expressions

animations that were specified directly from the time course of human behaviors; there was no speech synthesis. This illustrates a general feature of RUTH's architecture – the specification for RUTH can be done both in simple and abstract ways, depending on human analyst's coding scheme [5].

In this study, we used animations that were constructed as follows. A FACS coder analyzed a recording of a person's natural utterance, and specified the type of facial *action unit* (AU) that occurred at each moment (AU), as well as its intensity level (on a scale of A-E, from minimum to maximum), and four timing values: onset, apex-start, apex-end, and offset. The intensity levels were converted to numerical parameters for applying deformations in the animation engine so as to mirror the appearance changes seen in the original video. The timing of onset, apex and offset was used to adjust RUTH's generic temporal profile for animating behaviors to match speakers' particular motions. Figure 2 contains snapshot images of three utterances from our data set, as animated with RUTH. More details about the animation methodology and its implications for the technology of conversational animation can be found in [6].

3 Recognition Experiment: Understanding RUTH

In using RUTH to explore signals of uncertainty, we developed a set of expressive resources that allow RUTH to closely mirror an original human performance in terms of the available behaviors and the temporal coordination among them. Nevertheless, the details of RUTH's appearance and motions are highly simplified and stylized, leaving substantial differences with a video of a person talking. A key question is therefore whether RUTH's architecture and the animations we generate with it can succeed in conveying different levels of certainty to a conversational partner. To investigate this question, we carried out a recognition experiment that corresponded as closely as possible the video only condition from our earlier judgment study [3]. The major difference was that, instead of presenting recordings of real human subjects, the animations by RUTH served as the test items.

3.1 Stimuli

We started with ten videos of human subjects giving short answers in conversation. These videos were collected in a conversation experiment designed to elicit utterances with varying levels of certainty; after giving their responses the original subjects were asked how certain they were, giving a self-reported certainty level for each clip. The questions to which subjects responded were "Do you know in what year the University was chartered?" and "Who is the current Rutgers University president?", and arose in a dialogue covering questions a student might ask about the University [3]. The videos were then analyzed and recoded to provide input to RUTH to synthesize a corresponding delivery [6].

We used the resulting specifications to create three sets of RUTH animations as stimuli for the present experiment. The first set, the *full facial behavior condition,* included all the observable nonverbal behaviors in our analysis, including the face, head and eye movements. The second set, the *facial gyration only condition,* included just the deformations that play out on the face, but did not include eye or head movements. The third set, *the head/eye movement only condition,* included movements of the head and eyes, as well as the movements of the lower face driven by the production of visual speech (mostly movements of the lips, tongue and jaw), but did not include the other, expressive movements of the face.

3.2 Participants

All participants were current Rutgers University students. The majority of students were recruited from the Psychology department subject pool system with the exception of a few volunteers. There were total of sixty participants. Fifty five subjects were undergraduate students while five others were graduate students. Each participant was randomly assigned to one of the three conditions. There were nine males and eleven females for full facial behavior condition, fifteen males and five females each for both the facial gyration only condition and the head/eye movement only condition.

3.3 Procedure

Subjects performed the experiment using a web based interface which allowed them to view stimulus videos and to record their judgments about what they saw. Before making each judgment, they could view the video up to three times. For each video, subjects were asked to rate how uncertain the speaker looked on a scale of 0 (completely uncertain) to 100 (completely certain). Subjects were also asked to specify their own level of confidence in making the judgment, as either very confident, somewhat confident, or not confident at all. Subjects saw two practice items and, then, rated ten test items associated with the condition to which they had been assigned. Items were presented in random order to avoid order effects. Upon each rating, a response was automatically stored into a preconfigured MySQL database through a web interface with other information such as subject number, condition, test item, and experiment date. At the debriefing and an exit interview, participants were informally asked about their overall impression on the experiment and the difficulty in accomplishing the tasks.

4 Overall Results

The two primary goals of the recognition experiment were: first, to investigate how well people could recognize the uncertainty expressions that were coded onto ECA and, second, to learn the effects of different layers of the nonverbal signals. Table 1 lists the mean values of each certainty level by the subjects' original self reports and the judges' ratings of the three different conditions of RUTH animation. In addition, Spearman's rho correlation coefficients were computed for these five variables and are presented in Table 2.

Table 1. The mean values of the judges' uncertainty ratings across conditions, ordered by the self-report of the original speaker from whom the animation was modeled

Self-report: certainty level	All: face and head	Face only	Head/eyes only
100	83.2	74.5	73.3
90	75.8	65.5	85.8
80	62.3	64.5	59.1
75	62.1	64.5	62.1
65	61.8	49.9	63.5
50	27.7	74.8	23.5
40	29.4	49.9	25.8
35	25.05	56.9	28.2
10	31.75	50.0	38.9
0	23.7	26.5	36.1

Table 2. Correlation coefficients relating judges' mean ratings with other factors: the original self reports, or judges' ratings of completely expressive animations or ratings of face only

Spearman's rho		All: face/head	Face Only	Head/eyes only
Self report	Correlation Coefficient	**0.915	*0.683	*0.709
	Sig. (2-tailed)	0.000	0.030	0.022
	N	10	10	10
All: face and head	Correlation Coefficient	1	0.524	**0.830
	Sig. (2-tailed)		0.120	0.003
	N		10	10
Face only	Correlation Coefficient		1	0.207
	Sig. (2-tailed)			0.565
	N			10

The most important result here is that judges' ratings of RUTH animations containing all of the analyzed behaviors correlate very highly with the self-reports of the original speakers (r = .915, p < .001). Indeed, when we compared the correlations between judges' ratings of the original videos of human speakers [3] and judges' ratings of the RUTH animations, we also get a very high correlation (Spearman r = .903, p < .001, 2-tailed, N=10). Moreover, judges' scores track uncertainty levels fairly accurately across all conditions. When we made comparisons with the self-reported certainty levels from the conversation experiment, the ratings in the signal identification experiment and in all three conditions from the recognition experiment showed significant associations. The relationship was lower when we removed head and eye movements (r = .683) and pure facial gyration signals (r = .709). Concretely, this means that about 46% of the variance ($.683^2$) in judges' responses to pure facial gyration was accounted for by its linear relationship with speakers' self-reported certainty.

Likewise, about 50% of the variance ($.709^2$) in judges' responses to head and eye movements was accounted for by its linear relationship with speakers' self-reported certainty.

However, the different nonverbal channels seem to be providing judges with very different kinds of information. Note that we found no evidence that judges' responses to an animation with head and eye movements removed covaried with judges' responses to the same animation with facial gyration removed (r = 0.207, p = 0.565 NS). This suggests that the different modalities present complementary information that is reconciled in viewers' overall understanding of a talking speaker, as in the integrated message model of nonverbal communication [10]. Indeed, it may be that facial expressions get their precise meanings in part by how they co-occur temporally with movements of the head and eyes – since movements of the head and eyes seem to signal what an interlocutor is doing to contribute to the conversation (e.g., listening, planning an utterance, presenting information, questioning or revising previous contributions) while other displays seem to serve to appraise how well that ongoing activity is proceeding. It is striking in this connection that the correlation of judges' responses to an animation with the head and eye movements removed had such a weak correlation with judges' responses to the full animation (r = 0.524, p = .12 NS). We hope to investigate this further in future work.

4.1 Judgments for Individual Viewers

So far, we reported the effective reliability of the mean judgments. This shows that there is no systematic bias or information loss in viewers' judgments of the RUTH videos. This does not show, however, that individual viewers recognize the uncertainty of a specific video reliably or agree with one another's judgments of what they see. To explore this question, we computed correlations using all 200 individual data points for each condition. The overall tendencies remained the same as those shown in Table 2. The Spearman's rho correlation coefficient between the self-reported and the rated certainty level was the strongest when all available visual cues were presented to them (r = .635, p = 0.01) and the weakest when head/eye movements cues were all removed (r = .405, p = 0.01). The correlation was .541 when we removed all observable facial gyration signals.

4.2 Judgments for Individual Items

We can also analyze responses by item, and explore the extent to which different animations are equally effective at conveying a particular level of uncertainty. To present these findings, we tabulate the *certainty recognition difference*, which is computed by subtracting the self-reported uncertainty level from a judge's ratings, for subject's ratings of each of the test items in the sample.

Figure 3 graphically displays the variance on each test item, presenting the median, interquartile range, outliers, and extreme cases within the *certainty recognition difference* variable for ten certainty levels: 0%, 10%, 35%, 40%, 50%, 65%, 75%, 80%, 90%, and 100%. In most cases, the median values were close to 0%, meaning that there was no systematic bias in how judges rated it. Judges most accurately rated the

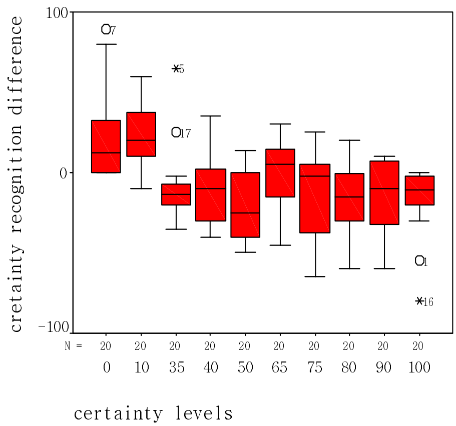


Fig. 3. Difference between the self-reported uncertainty levels and the ratings at the recognition experiment when all signals were presented

100% certainty test item. The variance was the smallest for the 35% certainty. For 0% certainty, there were three individuals who rated 0% as either 80% or 85% certainty. These judges seemed to be confused or forgetful about our operationalized definition of uncertainty – if the agent appeared not to know the answer at all, judges were instructed to give a rating of 0% certainty. During the debriefing, judges sometimes reported that they rated this test item as high certainty because they thought the agent looked "certain" of not knowing the answer. By contrast, the subject who reported 50% certainty showed lots of hesitation and confusion during the conversation experiment. To several judges, this test item offered a more intuitive illustration of high uncertainty than a speaker who quickly indicates that they cannot provide an answer.

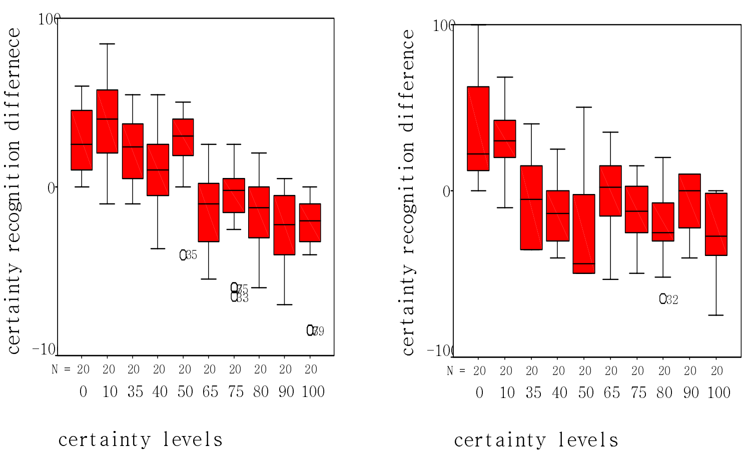


Fig. 4. Certainty recognition differences: pure facial gyration without head/eye movements (left) and head/eye movements only (right)

Figure 4 shows the box plots for the difference distributions for the simplified animations. These visualizations help to suggest the different ways viewers might combine information from facial expressions and head movements to understand an animation. In cases such as the 50% test item, it seems obvious that the head and eye movements – an extended interval of looking away and shaking the head while speaking – had more to do with the raters' very low judgments of certainty than the facial gyration – in this case, a sustained and perhaps apologetic smile. Finally, there were some cases where the signals from the two different channels seemed to be weighted together in subjects' ratings of the overall behavior. Thus, compared with the original values in the case of 35% and 40%, people gave relatively higher certainty ratings with the head/eye movement information and relatively lower ratings with the pure facial gyration signals. If this was a case where the two cues were redundant, that may explain why the full rendering of the 35% test time had the clearest signals among all test items (See Fig. 2).

5 Conclusion

Our results showed that judges rate certainty for animations with RUTH in ways that correlate closely with the self-reported uncertainty of the speaker on whom the animation was based, and with judges' ratings of the original video of the speaker. These relationships strongly suggest that people could recognize different levels of uncertainty not only with another human conversational partner, but also with an embodied conversational agent that used similar cues. Our work thus promises to inform the design of agents that signal their uncertainty as effectively in conversation as people do – and, more generally, to allow us to exploit knowledge and observations of specific human behaviors to enrich human-agent interaction.

Our research also points to how viewers make the judgments they do. The association between the self-assessed uncertainty and the rated uncertainty in all three conditions (*all observable, pure facial gyration, and head/eye movements only*) proved to be highly significant. Head and eye movements were better indicators of uncertainty than pure facial gyrations. But raters made the most reliable judgments when both kinds of information were presented to them in a coordinated manner. In future work, we hope to develop a more detailed model of the way addressees combine these different signals into an integrated understanding of speakers.

The kinds of human computer interaction that ECAs can bring to us are unique and varied. What is new with ECA is that we must now enable an agent to use its embodiment in support of effective interaction. Research still has a long way to go to achieve this. However, it is increasingly practical and reliable to increase agents' believability by crafting communicative expressions like our own for them.

Acknowledgments. We are grateful for the assistance of Mark Frank and Doug DeCarlo at various stages of this research. Our work was supported in part by NSF HLC 0308121 and by the second author's Leverhulme Trust Visiting Fellowship at the University of Edinburgh 2005-2006. We are grateful to Alias for the use of Maya modeling and animation software.

References

1. Clark, H.H.: Using Language. Cambridge University Press, Cambridge (1996)
2. Swerts, M., Krahmer, E.: Audiovisual prosody and feeling of knowing. Journal of Memory and Language 53(1), 81–94 (2005)
3. Oh, I., Frank, M., Stone, M.: Face-to-face communication of uncertainty: expression and recognition of uncertainty signals by different levels across modalities. In: ICA International Communication Association (2007)
4. Nakano, Y.I., Reinstein, G., Stocky, T., Cassell, J.: Towards a model of face-to-face grounding. In: Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL 2003), pp. 553–561 (2003)
5. DeCarlo, D., Revilla, C., Stone, M., Venditti, J.: Specifying and animating facial signals for discourse in embodied conversational agents. Computer Animation and Virtual Worlds 15(1), 27–38 (2004)
6. Stone, M., Oh, I.: Interdisciplinary challenges of generating animated utterances. In: Wachsmuth, I., Knoblich, G. (eds.) Modeling Communication with Robots and Virtual Humans, Springer, Heidelberg (2007)
7. Bailenson, J.N., Blascovich, J.: Avatars. In: Bainbridge's, W.S. (ed.) Encyclopedia of Human-Computer Interaction, pp. 64–68. Berkshire Publishing Group (2004)
8. Nass, C., Robles, E., Wang, Q.: 'User as Assessor' Approach to Embodied Conversational Agents. In: Ruttkay, Z., Pelachaud, C. (eds.) From Brows to Trust: Evaluating Embodied Conversational Agents (Human-Computer Interaction Series), Kluwer Academic Publisher, Netherlands (2004)
9. Ekman, P., Friesen, W.V., Hager, J.C.: Facial Action Coding System (FACS): Manual and Investigator's Guide. A Human Face, Salt Lake City, UT (2002)
10. Bavelas, J.B., Chovil, N.: Visible acts of meaning: an integrated message model of language in face-to-face dialogue. Journal of Language and Social Psychology 19(2), 163–194 (2000)

Automatic, Body Measurements Based Generation of Individual Avatars Using Highly Adjustable Linear Transformation

Andreas Volz, Rainer Blum, Sascha Häberling, and Karim Khakzar

Fulda University of Applied Sciences, Department of Applied Computer Sciences,
Marquardstr. 35, 36039 Fulda, Germany
Andreas.Volz@tux-style.de,
{Rainer.Blum,Sascha.Haeberling,
Karim.Khakzar}@informatik.hs-fulda.de

Abstract. In this paper part of the work of the research project 'IntExMa' is presented, that aims at the development of an interactive, in-store sales support system focusing on the virtual 'try-on' of clothing. Customers are provided with highly realistic, individualized, three-dimensional virtual counterparts, derived from a small number of body measurements. The paper describes in detail the approach followed for the required automatic generation of avatars. It is based on a sophistication of the well-known concept of parametric deformation. The solution presented here is characterized by a range of means that provide highly flexible, fine-grained adjustments of the linear transformation process leading to precisely controllable results.

Keywords: body measurements, parametric deformation, individualized avatars.

1 Introduction

1.1 The IntExMa Project

The goal of the IntExMa project (Interactive Expert System for Made-to-measure Clothing) is to provide an in-store installation for small and medium sized businesses, which provides a virtual shopping experience to the customers. The special context addressed by the project at the moment are made-to-measure shirt shops. Customers of these shops cannot see the final product in real life, as this is specifically assembled and tailored according to their specifications. Although examples of fabrics and other components are available in the store, the final assembly is not visible to the customers before delivery.

This is where the IntExMa system offers added value by providing a "virtual try-on" solution. It consists of a virtual representation of the customer (avatar) on a large display screen wearing the not yet produced shirt. The latter can be change according to the product range offered in the shop. The setup's intention is to create the impression of a "virtual mirror". This way, customers shall get a better understanding of the

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 453–459, 2007.

look and fit of the final product, leading to well-founded buying decisions. In order to reach this goal, an advanced cloth simulation is used to render the virtual shirts close to reality.

1.2 Objectives

User reactions have shown, that the overall acceptance of the IntExMa system by customers does not only depend on the quality of the displayed clothes but also on the appearance of the avatar wearing the virtual shirt. Giving the users the impression to look at their virtual counterparts seems to be an important aspect. To offer a limited selection of avatar shapes has proven not to be sufficient. Avatars are required that are an individual representation of the customers, at least to the required extent for personal identification.

A crucial point is the rather rare data that is available about the clients. The basic concept of IntExMa is to provide an affordable system solution which can seamlessly be integrated into existing point of sale activities and environments. A high-end approach for customer digitization with the help of body scanners or the like is not feasible in this context of use. Therefore, it was decided to design, implement and evaluate an avatar generation system that works with the limited set of body measurements anyway taken from the customer by the sales person by hand. In case of a shirt these are typically less than ten different measures.

This paper presents a method to solve this problem. It describes an approach to create a three-dimensional avatar, based on a given, limited set of measures taken from a real-life person. This method is based on a sophistication of the concept of “parametric deformation” or “keyframe morphing” [1, 2], which is, for example, implemented in the well-known 3D figure design application Poser®.

Keyframe morphing is a method to change the shape of a given three-dimensional object by re-locating its vertices. Vectors for those vertices are defined to describe a linear deformation path. A group of those vectors is called a morph target, which usually covers only a subset of vertices. A scalar defines the impact of this vector and, thus, the dimension of this re-location.

To fulfill the described requirements of the IntExMa system, additional functionality has to be provided: Avatars that match the entered measures closely shall be produced automatically. No other currently available system known to the authors has this required characteristic.

Releasing the results of this work as open source makes them available to the public in order to be used on products with similar demands. A project that could already profit is the MakeHuman Project [3], which is specialized on creating highly realistic three dimensional humanoids (cf. Section 3.4).

2 Deformation Process

In order to transform the base avatar into a model that complies with the desired body measurements, a morphing algorithm is introduced. It uses a measuring algorithm to obtain information about pre-defined measuring sections of the body. These are used to evaluate the correctness of the result, by comparing these values to the reference values. Those may include circumferential as well as range measures, to be able to set the girth of certain body parts as well as their length.

Morph targets are used to define vectors for the deformation of a defined set of vertices from the complete model. At least one morph target is assigned to every measuring section in order to be able to achieve the final deformation value independently from the rest of the measuring sections.

The general process consists of discrete alternating steps of measuring and deformation and ends by reaching the pre-defined values with the given tolerance.

2.1 Deformation Calculation

The deformation procedure shown in this paper is an expert system that makes use of linear transformation of vertices. As a base for this transformation a three-dimensional human model is used. The algorithm itself is independent of the amount of vertices in the model, but it is crucial that the topology of vertices in the base model corresponds to the one of the morph target - although the morph target can and usually does cover a subset of the base model's set of vertices. This way many overlapping and non-overlapping targets can be defined for a base model and applied concurrently, either fully or partly.

The deformation vector is computed by subtracting the vertice in the base model from the associated vertice within the morph target. A scalar m is influencing the application of a target on the deformation. Thus, only the length of the deformation vector is manipulated. A value of $m=1$ doesn't change the length at all. For $0 < m < 1$ the impact of the vector is decreased, resulting in a partial deformation. Negative values of m lead to a reversion of the deformation vector. Usually, applying a reversed deformation leads to unsatisfying results and can only be used in rare cases, recommending only the interval (0,1) for predictable results.

2.2 Measurement

Since multiple targets allow for independent deformation of the body parts, a way to measure these deformation results at various positions within the model independently had to be found. Therefore, we introduce a technique that measures the avatar with so-called hotspots which can be distributed around the body.

This hotspot technique is similar to the well-known concept of feature points, which, for example, is defined by the H-Anim Standard [4] for humanoids, but offers some necessary extensions. A hotspot contains several feature points in a specified order, thus defining a complex measuring section like the girth of an arm. Figure 2 shows an avatar with some measuring sections highlighted.

The hotspot concept is an easy and efficient way, which allows the calculation of the distances between all the ordered feature points hs_i within the hotspot, adding them to the final measure value l, as shown in the following equation:

$$l = \sum_{i=0}^{n-1} \left| \overrightarrow{hs_i} - \overrightarrow{hs_{i+1}} \right| \ . \tag{1}$$

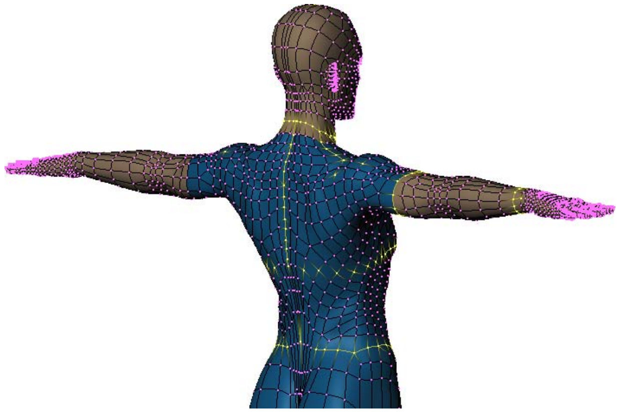


Fig. 1. Avatar with highlighted measuring sections/hotspots

The actual value of the result depends on the scale of the model. Thus, in order to obtain real-world units, the scale of the base model has to be considered.

2.3 Automatic Deformation

Of course, deformations could be done directly by setting the scalars of each morph target. But with the given input data at hand, i.e. base model, morph targets and body measures, an automatic process can perform this deformation iteratively to end up with a model that fits the desired input measures.

The principle of deformation described here is independent from material masses and units. By the definition of a suitable scale the associated measure can be computed to each measuring section. A suitable target, at least within the defined range, must have a linear course. This characteristic is used for the determination of the necessary deformations.

As an initializing step, a linear gradient for a defined range is determined using test measurements. Then a first estimation of the necessary morph value can be computed. The deformation is then applied to the avatar and a new test measurement of the appropriate measuring section is taken. If the linear process is sufficiently precise and the delta is within a defined tolerance this iterative deformation process ends. In many cases the linear value gets very close to the final value, but often it's not within the given tolerance. In this case a loop starts which repeats this process of re-measuring and deformation in order to refine the result until the delta is sufficiently small. The computing time rises proportionally with each additional deformation step until the desired accuracy level is reached.

To further control the deformation process, additional configuration input parameters can be provided. As mentioned before, several targets may be applied to the model and overlapping is possible. Furthermore, two overlapping morph targets may cover a number of vertices that are of relevance for one particular measure. For the iterative process to work properly under these circumstances, an order of application must be provided for the targets.

2.4 Range Parameterization

This deformation method is not limited to changing proportions of a base avatar in order to control its body measurements. Its flexibility allows an application for other purposes, like controlling age or gender, as well. In order to achieve this, range parameterization is applied. This technique maps a real world measure onto a scalar for a given target. This way it is possible to define the application of a target like "young" within the range of 6 to 18 "years", resulting in a scalar of $m=1$ for age 6 and $m=0$ for age 18 (regarding a base model representing age 18). Likewise, additional targets could be added to cover other ranges as well, allowing for a fuzzy-like curve.

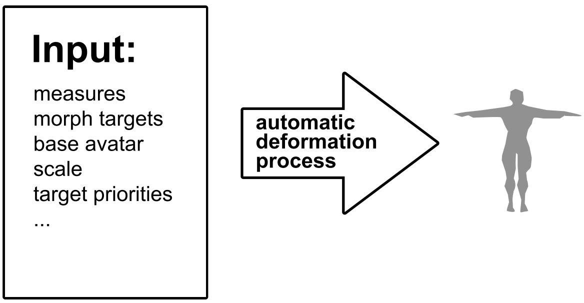


Fig. 2. Principle of the automatic deformation

Fig. 2 depicts the overall process summarizing necessary input elements.

3 System Design and Implementation

3.1 Library Design

The prototypic implementation of the system is object oriented. The classes were divided into two separated libraries, that can be used independently. The first library was named libanimorph. It implements all low-level tasks like transformation and deformation of the models.

For the automatic deformation process a second library, libavatar, was created on top of libanimorph. It contains the routines that handle the measures in order to perform the automatic deformation process.

Both libraries are released under the GNU Lesser General Public License (LGPL).

3.2 Content Generation Tools

In order to support the generation process of the necessary input data, some specialized tools were developed. MakeTarget, a plug-in for the open source 3D modeling software Blender [5], helps to generate morph targets by comparing the base model to a modified version and saving only the modified vertices.

As the deformation system libanimorph uses a special format to store the three-dimensional model data, the tool PyExportMH provides the export of a model loaded in Blender to this format, including vertices, faces, vertex groups, materials and texture coordinates.

Another helper tool uses libanimorph and libavatar directly to calculate scales between real world and virtual construction measures.

In order to be able to analyze a range of results of a given morph target by applying different scalars, plottarget outputs a two-dimensional graph, containing the scalar value on the x- and the result of a measured hotspot on the y-axis.

3.3 Import and Export Capabilities

The format to import data into the system is well-defined, but not the data itself. The import format is similar to Wavefront OBJ, but for various reasons split into several files. In principle, it is possible to use any three-dimensional geometric model as input data, independently of its topology. Using the present principle it is possible, for example, to build an automatic deformation system for a car model or a fantasy figure.

Currently supported export formats are Wavefront OBJ, Renderman RIB, and a proprietary format for the IntExMa project based on OBJ.

3.4 Integration into Existing Software

MakeHuman [3] is a multi-platform and open-source (GPL) software program whose purpose is to provide a versatile, professional and specialized application for parametrical modeling of three-dimensional humanoid characters. The MakeHuman application uses libanimorph as its rendering engine. The automatic deformation library libavatar is currently not used in MakeHuman. Currently, MakeHuman aims at the creation of nice-looking human avatars, not at the automatic reproduction of an individual real human.

In order to provide the users of the IntExMa system with a virtual counterpart, that looks as close to the real person as possible, the avatar generation system presented in this paper was integrated. As soon as new measurements are input by the salesperson the corresponding avatar is generated. Subsequently, a physically-based simulation clothes the virtual human with digitally represented garments. For more details about the IntExMa system please refer to [6].

4 Results and Conclusions

The quality of the results highly depends on the quality and quantity of the input data. The more detailed the base model is and the more morph targets respectively measuring sections are provided, the more precise the reproduction of the measures of the real person will be.

Fig. 3 shows two avatars resulting from the deformation process and dressed with a physically-based simulated shirt. The base model lies somewhere in between both figures. The avatars result from different morph targets that were specified for decreasing and for increasing measures (in relation to the base model). It can be stated, that the automatically generated avatars are visually appealing and suitable for the “virtual try-on” simulation.

Fig. 3. Two avatars resulting from the same base model

At the same time, it must be considered that conventional length and girth body measures of a human do not fully specify the targeted outer characteristics of the respective body. The same set of measures may result from several real human bodies differing significantly in their appearance (muscle mass, posture etc.).

Nevertheless, it is not clear, how realistic the reproduction of the individual customer must be for the purpose at hand, the support of the buying decision and, accompanying, the general acceptance of the system concept.

Integrated into the overall IntExMa system, the efficiency of the presented approach for the “virtual try-on” of clothing and sales support will be evaluated within the scope of usability and acceptance tests with real customers and sales staff in a real made-to-measure tailor shop.

Acknowledgments. This research was carried out during the project IntExMa and financially supported by the German Federal Ministry of Education and Research.

References

1. Magnenat-Thalmann, N., Thalmann, D.: Handbook of Virtual Humans. John Wiley and Sons Ltd, New York (2004)
2. Lever, N.: Real-time 3D Character Animation with Visual C++, Focal Press (2002)
3. MakeHuman: (last retrieved 16/02/07) http://www.makehuman.org/
4. Humanoid animation (H-Anim): (last retrieved 16/02/07) http://www.h-anim.org/
5. Blender: (last retrieved 16/02/07) http://www.blender.org/
6. Khakzar, K., Blum, R., Kohlhammer, J., Fuhrmann, A., Maier, A., Maier, A.: Interactive Product Visualization for an In-store Sales Support System for the Clothing Retail. In: Proceedings of HCII 2007, 12th International Conference on Human-Computer Interaction 2007 (to appear)

A Composite Measure for the Evaluation of Mental Workload

Lee-ming Wang[1], Vincent G. Duffy[1,2], and Yingzi Du[3]

[1] School of Industrial engineering, Purdue University, West Lafayette, IN 47906, USA
{lwang,duffy}@purdue.edu
[2] Regenstrief Center for Healthcare Engineering, and School of Agricultural and Biological Engineering, Purdue University, West Lafayette, IN 47906, USA
[3] Department of Electrical and Computer Engineering, Indiana university-Purdue University Indianapolis, Indianapolis, IN 46202, USA
yidu@iupui.edu

Abstract. Physiological measures have found reliable sensitivity to the variation of mental efforts to tasks of different difficulty levels. The sensitivity needs to be enhanced for further application. This paper proposed a composite measure consisting of three physiological measures, facial skin temperature, eye blinks and pupil dilation. The facial skin temperature will be measured by an infrared camera. One dimensional iris image will be used for the measurement of eye activities. All measurement will be done in real-time and unobtrusively. A preliminary experiment will be conducted for each measure to demonstrate their sensitivity. The combination then will be accomplished by factor analysis and regression analysis. Last, the analysis will focus on the improvement in sensitivity from the combination of individual measures.

Keywords: mental workload, facial skin temperature, pupil dilation, thermography, 1D iris identification.

1 Introduction

Multitasking is common in human daily activity. For example, driving consists of at least two tasks, tracking – maintain vehicles on the designated route, searching direction, and memorizing – keep in mind the direction to destination. Hence, it is reasonable to consider the 'multimodal information' embedding in the human behaviors with a composite evaluation index.

Past research has supported this idea. Previous research [1-4] showed facial skin temperature reveals sensitivity to the variation of mental workload. However, the researchers also noticed that the sensitivity may not be strong enough for practical application in complex tasks [4, 5]. The combination of various measures was suggested to enhance the explanatory power. Some related research on this concern has been done. Ryu et al. [6] found that in a dual-task experiment, none of the selected measures, excluding the subjective evaluation, was able to detect the variation of the mental workload for both tasks. Comparatively, when these measures

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 460–466, 2007.

were combined, the significant difference of mental efforts for both tasks was successfully identified. The authors demonstrated, because the demands for each mental processing resource may be determined by different physiological measures, the combination of physiological measures is a good way to carry out the maximum explanatory power. Lin et al. [7] had a similar conclusion. They applied two composite measures, different combination of eye movement and hand motion, to evaluate the mental efforts of players in a video game. The combination was found be able to detect all the variation of mental efforts with regard to the games of different difficulty levels, while single measures failed to accurately detect that variation.

In this research, a composite measure, consisting of facial skin temperature, pupil dilation, and eye blink interval, is proposed in order to evaluate mental workload as well as user performance. These three physiological measures have been well studied as evaluation indices of mental workload. We believed these measures can reflect the mental demands of whole or partial different processing resources and therefore it is reasonable to combine them to achieve greater explanatory power for tasks in different domains.

Furthermore, in this research, real-time and non-intrusive implementation for each measure is proposed. Non-intrusive measures are preferred because they can introduce the minimum of disruption to the users and can bring most convenience in practical utilization. Real-time is also an important feature. The continuous surveillance of user status in tasks provides all the data over time, not only at the beginning and the end of the tasks [7]. The future work can be extended, with the higher reliability of composite measures, to analyze the effects of 'learning' and 'individual difference' on the evaluation of mental workload. In addition, adaptive automation provides a promising application [8]. If the measure helps a system with self-adjustment functions according the real-time performance of users, we believe that it can make contribution to many applications such as real-time aid system for operators, design of human computer interface, and quality control for complex systems such as e-service.

2 Measures

2.1 Facial Skin Temperature

Variation of facial skin temperature has received some attention as a physiological measure of mental status. This is because they showed the significant relationship to the change of mental status, and on the other hand, the non-contact measurement is practical with thermal infrared camera since human face emits mid- and far- infrared [2].

When sensations occur, the autonomic nervous system acts to mediate the response to the stimulus, thus resulting in the redistribution of blood flow and then the change of local skin temperature [2]. Genno et al [3] showed the temperature change in nose area when subjects experienced stress and fatigue. In a tracking task, the nose temperature dropped significantly at the beginning of the task and at the moment when an unexpected emergency signal alarmed during the task, while other parts of the face did not have observable patterns for change in temperature. In another task inducing fatigue, the nose temperature also showed significant drop. They quantified this relationship and applied the model to evaluate image preference. The results were in accord with the subjective evaluations for individuals. The temperature drop in nose area with increased mental workload/stress was also observed in other studies. Or et al [4] found that drop in nose temperature of drivers in relation to different driving conditions and environments such as virtual driving test/real car driving, city route/highway route, and with/without mental loading task. Veltman et al [5] detected the temperature drop in nose area of subjects in Continuous memory-Task on two difficulty levels. The research, though they have demonstrated the relationship between nose skin temperature and variation of mental workload, suggested further work in the investigation of measure sensitivity with more difficulty levels and the combination of other physiological measures to enhance the sensitivity and accuracy.

2.2 Pupil Dilation

The correlation between pupil dilation and increased mental efforts has been well known for decades. Beatty [9] investigated task-evoked papillary response through a series of studies. Pupil dilation was determined to be as a reliable indicator of mental efforts or processing load of tasks such as memory, language processing, reasoning, and perception. The results from Murata et al [10] also supported the fluctuation of pupil area to be a meaningful measure of mental efforts with the control of respiration. Iqbal et al [11] showed that papillary response, in interactive tasks, was an effective index of mental workload of computer users when the tasks can be decomposed subtasks in hierarchal structure. In addition to its usability as a objective measure, another attractive feature of pupil dilation is that its measurement can be non-intrusive, which causes minimal interference and enables continuous data acquisition. Therefore, it is expected pupil dilation can contribute explanatory power to the composite measure.

2.3 Eye Blink Interval

Eye movement is another promising indicator of mental efforts due to its sensitivity and the possibility of its non-intrusive and continuous measurement [12]. Blink interval is known for its positive correlation with increased mental workload, while blink duration tends to decrease against more intense workload. These physiological responses allow eyes to have more time to collect visual information for tasks. Van Orden et al. [12] found the blink frequency and duration declined as a function of the density of targets in their visuospatial task, where subjects had to identify targets on a computer screen and opt for an appropriate prosecution on a target that was reaching a specific location based on its identification (friend/enemy). Veltman et al. [13] investigated various physiological measures, including eye blinks, in a flight simulator. In their experiments, blink frequency and duration decreased from rest status when subjects perform both two different flight tasks. Moreover, their results showed that blink frequency increased when subjects met more difficult flight condition with arithmetic tasks. It was probably because subvocal activity of the memory process within arithmetic tasks stimulated the muscles of the eyelid, thus forcing the blinks. Similar results were found by Ryu et al. [6]. Eye blink interval revealed weak sensitivity to the mental workload in tracking tasks, but did not in the arithmetic task in which subjects has to memorize the digits to be manipulated. Previous research suggests the feasibility of eye blinks as a physiological measure; therefore in the proposed research, we adopt it as a component of the composite measure.

3 Experiment

Facial skin temperature, pupil dilation, and eye blink intervals are selected in the proposed research. Implementation of the measurement is real-time and non-intrusive in order to get the intermediate state of users and to avoid unnecessary processing load in addition to task-evoked one. In the following sections, section 3-1 will introduce the measurement in the experiment, section 3-2 describes the design of the experiment, and section 3-3 has the detail about the task.

3.1 Measurement

A thermal camera will be used to capture the facial skin temperature from subjects. The measurement locations on the face will be designated as Region of Interest (ROI). Using thermal camera, the background will be dark and the face will be brighter pixels due to the body temperature. The higher the temperature is, the brighter the pixel will be in the image. The selection of ROIs is determined by their capacity to reflect the variance of temperature due to mental workload. The temperature of a ROI is defined as the mean temperature of pixels (or partial hottest pixels) in it. The computation of mean temperature will be done for every frame in the thermal clip. Other factors that might affect subjects' facial temperature such ambient temperature, but it can be reduced by carefully controlling the environment.

A near infrared (NIR) camera will be used to capture the pupil changes and eye blink. NIR cameras are popular used in iris recognition [14, 15]. In the visible light wavelengths, it is difficult to get the detailed iris patterns from dark color eyes. But under the NIR light, small details from the iris (even dark iris) are visible. Du et al [14, 16] devised a methodology to recognize iris pattern using 1-D iris siganture. In their innovative approach, given an iris image, their method will automatically segment the iris out, which includes detection of pupil, limbic, eyelids, and eyelashes. The segmented iris is then transformed into polar coordinates. In the polar coordinates, all radial information will become linear; for example, the pupil area of our interest becomes rectangle, which can ease the task to calculate pupil size. The size of the same iris taken at different times may be variable in the image as a result of changes in the camera-to-face distance. Due to stimulation by light or for other reasons (such as hippus, the natural continuous movement of the pupil) the pupil may be constricted or dilated. These factors will change the iris resolution, and the actual size of the pupil in the image. However, it is shown that the limbic radius would be stable during pupil dilation or constrain. To solve these problems, we should use the ratio between the pupil radius to the limbic radius to measure the dilation of the pupil.

This methodology is feasible even if only partial iris image is available [17]. The measurement of pupil size and eye blink in this research will be based on Du's 1-D iris image approach. Pupil size will be detected for each frame in the clip. In addition, eye blink interval and duration can be estimated using difference images from consecutive video frames. Using proper setting up the NIR illuminater, the glare or the camera noise from the image acquisition can be reduced. In Ref. [14], it is showed that this method is very robust in pupil detection.

3.2 Experimental Design

The first stage is to conduct the preliminary experiment to demonstrate the selected physiological measures as indicators to mental workload induced by the designed tasks. In the present research, the selected physiological measures include facial skin temperature, pupil dilation, and eye blink intervals. Each of the physiological measures is anticipated to be sensitive to the subject's mental workload in tasks. And, on the other hand, it is also expected that these three measures will reveal sensitivity in a somewhat different way to the variation of workload for different difficulty levels. According to the three criteria, proposed by Kahneman [9], for physiological indicators for processing load, the measures in the presented research will be statistically tested for 'within-task', 'between-task', and 'between-individual' difference. An effective measure should be able to reflect the difference from Kahneman's criteria. The 'Within-task' criteria means the sensitivity a measure has for a task (i.e. difference of measurement from rest to performing task), while the 'between-task' criteria further confirms the intensity of the sensitivity (i.e. pairwise difference of measurement among difficulty levels). In the present research, we will focus on these two criteria and leave the third one as future work. Thus, the three measures, as well as a subjective measure, will be applied to tasks of different difficulty levels. ANOVA will be applied in all the statistical tests for the data analysis. And NASA-TLX will be used to obtain subject report immediately after execution of each task.

The second stage is to combine the three measures together. Ryu et al. [6] suggested two methods, factor analysis and regression analysis to achieve the combination. Although in their methods both led to similar results, the two methods will again be adopted in this research. By doing this we can have comparison for each method, and on the other hand, we can also have a chance to see if any redundancy exists in the selection of measures. The statistical test method will be the same as in the previous stage.

3.3 Task

The design of the task is aimed to provide a very-controlled task so that subjects' workload can be correctly evaluated at different levels of difficulty. A dual task is considered in this experiment since the composite measure is expected to be feasible in the complex environment. Subjects will need to perform a tracking task and mental arithmetic simultaneously as in [6].

In the tracking subtask, subjects have to control a rectangular frame by moving the cursor to track a ball on the screen, which travels on a trajectory unknown to subjects as in [3]. If the ball moves out of the frame, a buzzer will ring to alarm the subjects.

The difficulty of the tracking task is determined by the frame size, ball speed and moving path. For example, in a very difficult task a subject will need to use a small frame to enclose a fast ball which also makes many sudden turns. This subtask is designed with five difficulty levels.

When subjects keep the cursor following the ball, meanwhile they also need to perform mental addition on the numbers displayed on the screen and report the results orally. Subjects need to sum up two numbers and each number contains up to three digits, based on how difficult the mental arithmetic subtask is. And the mental addition subtask contains two difficulty levels: EASY and HARD.

Within-subject factorial design will be implemented in this experiment. Therefore, subjects will need to complete the task in ten conditions. Tasks in different conditions will be executed separately with the insertion of rest periods between each execution. Each execution and rest period will last five minutes. Further, it is expected to recruit 100 subjects for the experiment.

In the task, subjects have to identify the motion of the ball and immediately follow it. It is expected that the eye responses such as eye blinks and pupil dilation can reveal the mental activity of users corresponding to visual stimulus. In the meantime, subjects will have mental addition on the numbers displayed aside the ball-tracking subtask on screen. Mental arithmetic in the task may incur the intensive workload in working memory. The tendency toward subvocal rehearsal activity in the memory tasks can suppress the trigger of some eye related activities. Subvocal rehearsal activity may interfere with eye activities because eyelid, eyeball, and tongue share same area of primary motor cortex [13]. It is reasonable to expect that, even in the case that the measure using eye activities show poor sensitivity to subject's mental workload in the task, the other measure, the detection of facial temperature with thermography, can supplement the explanatory power. Hence, the proposed composite measure has the potential to pertain consistent sensitivity to mental workload during the task.

Previous research [4] indicated that facial skin temperature is sensitive, though not consistently, to mental workload in mental math tasks. Considering this perspective, it is also interesting to observe whether the addition of the measure of eye activities can enhance the overall explanatory power.

4 Conclusion

In this paper a composite physiological measure is proposed by combining measurement of facial skin temperature and eye related activities. Thermal images will be captured from an infrared camera to detect skin temperature, and the measurement of eye activities will be fulfilled by the detection of pupil boundary in one dimensional iris images. All measures will be applied to the subjects who will need to track a moving ball and perform mental arithmetic. Ten task conditions will be provided in order to test the sensitivity of the proposed measure. The analysis will focus on the improvement of sensitivity from the combination of individual measures.

References

1. Puri, C., Olson, L., Pavlidis, I., Levine, J., Starren, J.: StressCam: non-contact measurement of users' emotional states through thermal imaging. In: Proceedings of CHI 2005, pp. 1725–1728 (2005)
2. Pavlidis, I., Levine, j., Baukol, P.: Thermal imaging for anxiety detection. In IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications 2000, pp. 104–109 (2000)
3. Genno, H., Ishikawa, K., Kanbara, O., Kikumoto, M., Fujiwara, Y., Suzuki, R., Osumi, M.: Using facial skin temperature to objectively evaluate sensations. International Journal of Industrial Ergonomics 19, 161–171 (1997)
4. Or, C.K.L., Duffy, V.G.: Development of a facial skin temperature-based methodology for non-intrusive mental workload measurement, Occupational Ergonomics (in press 2007)
5. Veltman, H.J.A., Vos, W.K.: Facial temperature as a measure of operator state. In: Proceedings of the 1st International Conference on Augmented Cognition, 2005, pp. 22–27 (2005)
6. Ryu, K., Myung, R.: Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. International Journal of Industrial Ergonomics 35, 991–1009 (2005)
7. Lin, T., Imamiya, A.: Evaluating usability based on multimodal information: An empirical study. In: Proceedings of ICMI 2006, pp. 364–371 (2006)
8. Schmorrow, D., Stanney, K.M., Wilson, G., Young, P.: Augmented cognition in human-system interaction. In: Salvendy, G. (ed.) Handbook of human factors and ergonomics, John Wiley, New York (2006)
9. Beatty, J.: Task-Evoked Pupillary Responses, Processing Load, and the Structure of Processing Resources. Psychological Bulletin 2, 276–292 (1991)
10. Murata, A., Iwase, H.: Evaluation of mental workload by fluctuation analysis of pupil area. In: Proceedings of the 20th Annual International Conference of the IEEE 1998, vol. 6, pp. 3094–3097 (1998)
11. Iqbal, S.T., Zheng, X.S., Bailey, B.P.: Task-evoked pupillary response to mental Workload in human-computer interaction. In: Proceedings of CHI 2004, pp. 1477–1480 (2004)
12. Van Orden, K.F., Limbert, W., Makeig, S., Jung, T.: Eye activity correlates of workload during visuospatial memory task. Human factors 43(1), 111–121 (2001)
13. Veltman, J.A, Gaillard, A.W.: Physiological workload reactions to increasing levels of task difficulty. Ergonomics 41(5), 656–669 (1998)
14. Du, Y., Ives, R., Etter, D., Welch, T.: Use of one-dimensional iris signatures to rank iris pattern similarities. Optical Engineering 45(3), 37201–37211 (2006)
15. Daugman, J.: How Iris Recognition Works. IEEE Transaction on Circuits and Systems for Video Technology 14(1), 21–30 (2004)
16. Du, Y., Ives, R., Etter, D., Welch, T.: A New Approach to Iris Pattern Recognition. In: Proceedings of SPIE 2004, vol. 5612, pp. 104–116 (2004)
17. Du, Y., Ives, R., Bonney, B., Etter, D.: Analysis of partial iris recognition. In: Proceedings of the SPIE 2005, vol. 5779, pp. 31–40.20 (2005)

Brain-Computer Interfaces Based on Attention and Complex Mental Tasks

Jue Wang[1], Nan Yan[1], Hailong Liu[1], Mingyu Liu[1], and Changfeng Tai[2]

[1] The Key Laboratory of Biomedical Information Engineering of Ministry of Education, Institute of Biomedical Engineering, Xi'an Jiaotong University, Xi'an, P.R. China
[2] School of Pharmacy, University of Pittsburgh, Pittsburgh, PA 15260, USA
juewang@mail.xjtu.edu.cn

Abstract. Brain-Computer Interface (BCI) technology is widely used in rehabilitation field. There are two main applications of BCI systems in assistive technology: regain the movements or communications for people with motor disability and neurofeedback for training the subject to emit a specific brain activity. In this study, we introduce two typical applications of BCI systems in our lab. For the first case, the BCI system based on mental tasks classification for people with motor disability is described. An effective features extraction and classification methods of EEG signals were proposed. For the second case, a neurofeedback (NFB) system was established, which utilized Virtual Reality (VR) to create appropriate feedback information which is more interesting, imaginative and interactive than traditional graphical presentations. Visual & auditory (IVA)-continuous performance test (CPT) results show that it can provide an effective therapy for treating attention deficit hyperactivity disorder (ADHD) children.

Keywords: brain-computer interface, mental tasks, neurofeedback, ADHD rehabilitation.

1 Introduction

Brain-computer interface (BCI) has been identified to function as an extra channel between brain and external environment to transmit information bypassing the spinal and peripheral neuromuscular systems [1, 2] and has also found applications in rehabilitation, neuroscience and cognitive psychology. Existing research in applications of BCI is composed of two main areas. For assistive technology, BCI makes it possible for people with motor disabilities to regain the interactions with external environment in order to improve the quality of their lives [2-5]. The second area aims at training the subject to emit a specific brain activity [6]. In this application, BCI becomes a therapy tool which helps subjects recover their cognitive function by consciously altering some features of their electroencephalographic (EEG) signals in order to stay in certain brain state.

In general, a BCI system based on mental tasks classification will find very useful application in rehabilitations, especially in the areas of motor function restoration and

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 467–473, 2007.

the rehabilitation training for people with full loss of movement capability. It can be used to restore communication such as speech capabilities. A BCI system interfaced with virtual reality (VR) can be used as environmental controls such as thermostats, television and even smart home controls. This application will provide safe environments for “locked-in” patients. It can restore the function of peripheral nerves or sent control command to maneuver wheelchairs so as to make it possible for people with motor disabilities to accomplish simple motion by themselves. Finally a BCI system can be valuable in neuron-rehabilitation of disabled people by reinforcing use of damaged or diseased neural pathways.

The objective of this study is to introduce the two typical applications of BCI in our lab. For the first case, complex mental tasks classification was analyzed. Effective features extraction and visualizing methods of EEG signals were proposed for further offline analysis and online application for restoring the movements and communication. For the second case, a neurofeedback (NFB) system was established, which utilized Virtual Reality (VR) to create appropriate feedback information. Compared to traditional graphical presentations, more interesting, imaginative and interactive treatment for attention deficit hyperactivity disorder (ADHD) was achieved.

2 A BCI System Based on Mental Tasks Classification

Currently, in order to restore the movements and communication for people with motor disabilities, the input signal of BCI systems is composed of evoked potentials (EP), slow cortical potentials (SCPs), event related potentials (ERP) and spontaneous EEG related to motor imagery tasks or different mental tasks [7]. And the features representing EEG patterns can be obtained using either frequency [8, 9] or non-frequency [10, 11] domain information of EEG signals. Frequency domain information is more widely used in BCI systems, e.g., Keirn has performed complex mental tasks by using asymmetry ratios and power values at four frequency bands: delta, theta, alpha, and beta [12]. Current methods extract energy information mainly from the whole EEG segments of specific length. However, while mental tasks were performed, EEG segments of any length started from a specific time point are related to different kinds of brain activity information. A convenient way is to use a phase space of different dimensions to represent the mental tasks. In our study, we used energy information of phase-space as features to perform mental tasks classification with Fisher’s Linear discriminant as the classifier.

Phase-space energy decomposition is a feature extraction method with two steps: reconstruction of the phase space from EEG signals, followed by performing energy decomposition in frequency domain from each vector in the reconstructed phase-space. Fisher’s linear discriminant was used as classifiers for dimension reduction of the input patterns. Since Fisher’s criterion focuses on the separation of different classes, it can produce low-dimensional output patterns appropriate for classification. It is also appropriate for online application.

Seven male subjects (ages from 22 to 26) participated in the experiment. Five mental tasks, including baseline, multiplication, letter composing, geometric figure rotation and visual counting were performed by each subject. An EEG recording of seventy seconds constituted a trial. A session comprised five such trials. The middle sixty EEG signals of each trial are used for analyses. Subjects performed one session in one day for each task, and different sessions were recorded on separate weeks.

The EEG signals were recorded from eight channels (F3, F4, C3, C4, P3, P4, O1 and O2) which are referenced to electrically-linked mastoids according to the 10-20 electrode system. A separate channel was used to detect eye blinks using two electrodes placed below and above the subjects left eye. The signals were band pass faltered at 0.1-100Hz and sampled at 1000Hz. There are 140 mental tasks in all for analysis.

Task-pair classification was performed using eighty features that were extracted from frequency bands of 0 to 100 Hz and 8 EEG channels. The average classification accuracy over 140 task-pairs is up to 99.4%. An information transfer rate of 56.8 bits/minutes was achieved. These results imply the good potential use of mental tasks classification in online BCI system.

3 A BCI-NFB-VR System for ADHD Treatment

Currently NFB technology has been mainly used in behavioral medicine as an adjunct to psychotherapy [13-16]. VR is a human-computer interface in which users can move their viewpoint freely in real time [17, 18]. Its purpose is to be constructed a virtual environment (VE), bring about a natural interactivity and make a live-sensation from multimodality.

In this study, a system based on NBF and VR technology for ADHD treatment is established. The NBF utilizes operative conditioning principles selectively enhancing or suppressing EEG waves within preset frequency bands, hence subjects can maintain their brain state in a specific state so as to gradually recover their cognitive function [15, 19]. It is a painless, non-invasive treatment and allows the individual to gain information about his or her EEG wave activity, to use computerized biofeedback system to change their EEG activities.

This ADHD training system utilizes VR to create appropriate feedback, which is more interesting, imaginative and interactive than traditional graphical presentations. During the whole process, the feedback modalities continuously represent brain activity with a minimum constant delay.

The structure of the proposed BCI-NFB-VR system is presented in Fig. 1. This system composes of three modules. The first module is a multi-channels, high input impedance EEG electrodes set. Three electrodes (Cz, Fp1 and Fp2) which are located on the scalp according to the standard “10-20 system” placement are used to collect EEG signals. The electrodes are connected to an embedded system for EEG signals collection and pre-processing. The embedded system has three functions: 1) Convert the analog EEG signals into digital EEG data by sampled 1000 times per second with 12 bits of accuracy; 2) Extract different EEG components and remove EEG artifacts (such as EOG, ECG, etc.); 3) Transfer EEG components to the main controller by USB port.

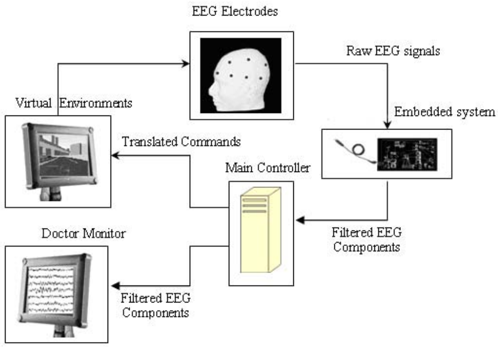


Fig. 1. Structure model of the system. First, raw EEG signals are sampled; then the EEG data are translated into "commands" by signal processing module; lastly virtual environments feedback is presented to subject.

The second module is the main controller, a personal computer, which transfers the filtered EEG components to “commands” signal. The process consists of three parts: Firstly, the EEG signal is decomposed into different frequency bands (Theta: 4-7Hz, SMR: 12-15 Hz, Beta: 15-18Hz) based on wavelet packet analysis method. Secondly, the corresponding power of each sub-band is calculated. The parameters such as θ/β can be calculated form that power information. Finally, the “commands” can be obtained by using the simple fusion rule. If both θ/β and SMR power go beyond a predefined threshold, the “command” is an inspiriting signal; otherwise the “command” is a stop signal.

The “command” signal obtained from the main controller is fed to the third module, the visual environment (VE) games and becomes a parameter affecting the behaviors of the VE games. Values of the “command” correspond to the changes in position, appearance, and size of VE objects or in the subject’s viewpoint of the VE scene. In the process of learning to “control” their brain waves, the subjects will gradually recover their cognitive function. An example of VE feedback game is showed in Fig.2. Three airships are flying in the space. Two of these are controlled by computer. One (the middle one) is controlled by the subject’s “brain wave”. If the “command” signal is an inspiriting signal, the airship controlled by the subject will speed up. Otherwise it will keep unchanged. In the training scenarios, the monitor is placed a few feet away directly in front of the subject. It provides the virtual environments feedback to the subject. The judge rate was over 10/s, which is responsive enough to give the subject a sense of direct control.

Two ADHD subjects (an 8 years old male and an 11 years old female) are recruited for this experiment. During the training, all the subjects were evaluated by an Integrated Visual & Auditory (IVA) - Continuous Performance Test (CPT) [21]. Two major quotients, response control quotient (RCQ) and attention control quotient (ACQ) of the IVA-CPT are obtained for both visual and auditory performance. The RCQ measures impulsivity, while the ACQ measures inattentiveness. These global composite quotient scales allow efficient summary and evaluation of the test results. Table 1 presents subjects’ RCQ and ACQ from pretest to posttest. The RCQ yielded

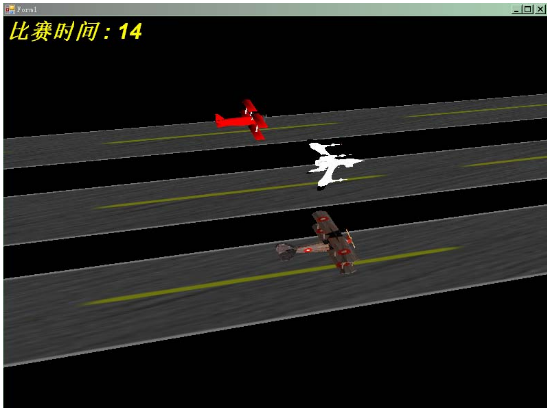


Fig. 2. The interface of the VR environment

Table 1. Subjects' RCQ and ACQ from pretest to posttest by IVA-CPT

Subject	Response Control Quotient (RCQ)		Attention Control Quotient (ACQ)	
	Pre-treatment	Post-treatment	Pre-treatment	Post-treatment
1	51	86	44	81
2	67	110	81	113

significant pre-treatment to post-treatment differences. It shows significant gains in the RCQ. Increases in ACQ are also found to be significant. After one month of use, the treatment proved to be effective. It can be concluded that this BCI-NFB-VR system can be used to improve the attention of children and adolescents suffering from ADHD.

In this study, the NFB-VR system was also regarded as a NFB framework, because all signal processing methods and Visual Environment feedback game are encapsulated in the dynamic link library (DLL). This system can be used to assist the preparation of the application for different rehabilitation training if we select different EEG components and correlative Visual Environment feedback games. So this system will be utilized as an adjunct to other psychotherapy such as anxiety, Depressive disorder and stroke.

4 Conclusion

The BCI method of phase-space energy decomposition was found to reflect more accurately the variation of energy of EEG signal. Using this method, higher classification accuracy can be achieved. This implies that mental tasks classification has good potential to be used in online BCI system for restoring the movements and communications for people with motor disabilities. We developed a NBF system which provides virtual environments to treat ADHD children. Two subjects tested clearly benefited from the treatment. The system is non-invasive and relatively easy for the therapist and the children to use. It can provide an effective therapy tool for treating ADHD children.

Acknowledgments. This work is supported by National Science Foundation, China, Grant No. 30670660 and National Science Foundation, Shaanxi, China, Grant No. 2006k15-G5. The authors would also like to acknowledge the technical support provided by the Biomedical Engineering Institute of Xi'an Jiaotong University.

References

1. Vaughan, T.M., Wolpaw, J.R., Donchin, E.: EEG based communications: Prospects and Problems. IEEE Transactions on Rehabilitation Engineering 4(4) (1996)
2. Wolpaw, J.R., Birbaumer, N., Heetderks, W.J., McFarland, D.J., Peckham, P.H., Schalk, G., Donchin, E., Quatrano, L.A., Robinson, C.J., Vaughan, T.M.: Brain-Computer Interface Technology: A Review of the First International Meeting. IEEE Transactions on Rehabilitation Engineering 8(2), 164–173 (2000)
3. Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.: Brain-computer interfaces for communication and control. Clinical europhysiology 113(6), 767–791 (2002)
4. Birbaumer, N.P.: Brain-computer communication: Self-regulation of slow cortical potentials for verbal communication. Arch. Phys. Med. Rehabil. 82, 1533–1539 (2001)
5. Müller-Putz, G.R., Scherer, R., Pfurtscheller, G., Rupp, R.: EEG based neuroprosthesis control: A step towards clinical practice. Neurosc. Lett 382, 169–174 (2005)
6. Bashashati, A., Ward, R.K., Birch, G.E.: Fractal Dimension-Based EEG Biofeedback System. In: Proceedings of the 25 Annual Intemational Conference of the IEEE EMBS 2003 (2003)
7. Wolpaw, J.R., Loeb, G.E., Allison, B.Z.: BCI Meeting 2005—Workshop on Signals and Recording Methods. IEEE Trans on neural system and rehabilitation engineering 14(2), 138–141 (2006)
8. Wolpaw, R., McFarland, D.J., Vaughan, T.M.: Brain-computer interface research at the wadsworth center. IEEE Trans. Rehab. Eng. 8(2), 222–226 (2000)
9. Fabiani, G.E., McFarland, D.J., Wolpaw, J.R., Pfurtscheller, G.: Conversion of EEG activity into cursor movement by a Brain-Computer Interface (BCI). IEEE Trans. Neural Syst. Rehab. Eng. 12(3), 331–338 (2004)
10. Pfurtscheller, G., Neuper, C., Schlogl, A., Lugger, K.: Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters. IEEE Trans. Rehab. Eng. 6(3), 316–325 (1998)
11. Pfurtscheller, G., Neuper, C., Guger, C., Harkam, W., Ramoser, H., Schlogl, A., Obermaier, B., Pregenzer, M.: Current trends in graz Brain-Computer Interface (BCI) research. IEEE Trans. Rehab. Eng. 8(2), 216–219 (2000)
12. Keirn, Z.A., Aunon, J.I.: A new mode of communication between man and his surroundings. IEEE Trans. Biomed. Eng. 37(12), 1209–1214 (1990)
13. McFarland, D.J., McCane, L.M.: EEG-based communication and control: short-term role of feedback. IEEE Transactions on Rehabilitation Engineering 6, 7–11 (1998)
14. Anon: Controlling pain by watching your brain. New Scientist, 182, 9 ((2004)

15. Lubar, J.F., Swartwood, M.O., Swartwook, J.N., O'Donnell, P.H.: Evaluation of the effectiveness of EEG neurofeedback training for ADHD in a clinical setting as measured by changes in T.O.V.A. scores, behavioral ratings, and WISC-R performance. Biofeedback and Self-Regulation, 20(1), 83–99 (1995)
16. Allanson, J., Mariani, J.: Mind over virtual matter: using virtual environments for neurofeedback training. In: Virtual Reality Annual International Symposium Proceedings of the 1999 IEEE Virtual Reality 1999, pp. 270–273 (1999)
17. Boman, D.K.: International survey: virtual-environment research. Computer 28, 57–65 (1995)
18. Goebel, M., Hirose, M.: Today's VR. IEEE Computer Graphics and Applications 21, 22–24 (2001)
19. Anon: Controlling pain by watching your brain. New Scientist, 182, 9 (2004)
20. Seckler, P., Burns, W., Sandford, J.: A reliability study of IVA: Intermediate Visual and Auditory Continuous Performance Test. In: Presented at the 1995 Annual Convention of CH.A.D.D 1995 (1995)

Fractal Modeling of Human Psychomotor Skills Acquisition Process

James Wanliss[1], Dahai Liu[2], Vadim Uritsky[3], and Michael Wuergler[2]

[1] Department of Physical Science, Embry-Riddle Aeronautical University
Daytona Beach, Florida, 32114, USA
James.Wanliss@erau.edu
[2] Department of Human Factors and Systems, Embry-Riddle Aeronautical University
Daytona Beach, Florida, 32114, USA
{dahai.liu,michael.wuergler}@erau.edu
[3] Department of Computer Science, University of Calgary
Canada
vuritsky@phas.ucalgary.ca

Abstract. Existing research on human skills acquisition studies has shown that learning follows a non-linear pattern, but the exact form remains unknown due to the limitation of traditional experimental methods and lack of systematic modeling of tasks. We applied a non-linear fractal analysis on the time series data produced by human subjects on target-tracking motor learning tasks. Tracking of a non-fractal sinusoid-cosinusoid signal was used as the platform. Our preliminary results suggest that fractal models may prove effective in investigating details of the human learning process.

Keywords: Fractal analysis, human learning, skill acquisition.

1 Introduction

Human psychomotor learning refers to physical or manual skills development, such as skilled performance of a musical instrument, driving a car or flying an airplane; this domain includes physical movement, coordination and use of motor-skill areas. Empirical research on human psychomotor learning can be traced back to the early 19th century when people tried to understand and examine different body movements related to their jobs [18]. Post World War II, with more studies from different disciplines on human motor learning, the development of machine complexity and the emergence of various theories of learning, psychomotor learning studies have become more systematic and achieved a more solid theoretical basis. Many learning theories such as Fitt's law [8] [9], memory drum theory [11], close-loop theory [1] and schema theory [17] [18] have been developed to try to explain the human learning mechanism of motor tasks, and continue to be developed [18] [19]. Many models and theories were developed to attempt to study human motor skills, in terms of its capabilities and limitations [1] [8] [9] [11] [17] [18].

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 474–482, 2007.

Human motor skills can be measured through performance outcome and via subjective feedback or self-report. Outcome performance measures (e.g. movement time, response time, movement accuracy) were used most of time [18]. Of all the performance measures, one particular interest is how to measure the learning process of human psychomotor skills. Traditional studies in this domain applied an empirical approach, for example, researchers attempted to show the effect of discrete levels of some independent variables on the learning. These studies were mostly non-analytic [7] and cannot reveal the continuity effect of this variable, thus the overall learning process is unknown. Most studies collect learning data in discrete stages, i.e., learning data for different trials, thus the data collected did not show too much of the continuity in learning time series. Furthermore, many studies did not take prior learning amount or experience into account – learning was administered for a fixed number of trials for all subjects and data collected tend to be from immediately after a limited learning exercise. Another shortcoming is that they only provide "crude" measure by giving one global mean value [6] [14].

Another approach is to use a performance curve fitting technique [18] to provide a more comprehensive measure of learning and a clearer picture of the learning process. It has been found that human learning follows a nonlinear pattern. That is, the improvement from practice is rapid in the beginning but decreases as the learner becomes more skilled. Newell and Rosenbloom's influential paper [15], titled "Mechanisms of Skill Acquisition and the Law of Practice", studied the relationship between response time and practice trials. They investigated diverse ranges of tasks, and found that the fitting is more of a power function rather than an exponential form [10]. The performance curve was claimed to be a major improvement for learning process estimates [6]. Using this method, the curve equation provides estimates for rate of initial level of performance, rate of improvement, and the asymptotic level of performance for a particular group (e.g., type of learning).

However, there are several drawbacks with curve fitting and traditional regression models. Just as Schmidt and Lee [18] pointed out, since the curve is built on the mean value of group performance data, it is "insensitive to the differences among individuals that arise as a function of practice". That is to say, the performance curve cancels out some of the difference among people for a given trial; this difference may be important to show the true learning improvement trend with practice. The second limitation is that the performance curve also obscures the within-subject variability by averaging the performance over a large group of people. So the generalization and applicability of such a curve is limited, since the behavior of group learning hides many important factors. Thus we may predict that a learning process has achieved its steady state by looking at its group mean asymptote, but in fact, the variability for this time point is still large. Learning curves for other types of learning have found similar limitations [22].

From the literature in psychomotor skill learning studies, it can be concluded that the learning mechanism is still poorly understood due to lack of understanding of the nature of task and the dynamic nature of the learning process thus making it difficult to systematically model the learning process. Existing methodologies in psychomotor skill research imposes limitations on modeling capabilities. No theoretical models have been established from these studies to enable prediction of learning efficiency and transfer. In order to find the true mechanism behind human learning and effects of tasks on human learning of psychomotor skills, we need to go beyond these limits. Traditional research methods are of little use for investigation of this process. It is therefore essential to include well developed mathematical techniques and models into the picture. In our study, we adopted a spatiotemporal fractal analysis to investigate the nature of the human learning process.

2 Fractal Analysis of Non-stationary Time Series Data

2.1 Fractal Nature of Human Learning

The mathematical theory of fractals brings to us a new set of ideas and thus fresh ways of looking at nature, human interactions with the nature, and the learning process. With this viewpoint come new tools to analyze experimental data and interpret the results. The essential characteristic of fractals is that as finer details are revealed at higher magnifications the form of the details is similar to the whole: there is self-similarity. A simple example of a fractal object is a fern, shown in Figure 1. The number of branches increases in a **power law** fashion with each subsequent generation from the first mother branch. If one were to zoom in on a small part, the zoomed in object would resemble the whole. Such scaling and fractal behavior is ubiquitous in the universe. In human physiology, fractal structure has been studied in numerous contexts including the arterial branching system in the kidney, the alimentary tract, the bile duct system, in nerves and muscles such as the heart, and finally the convoluted surface of the brain [2].

Fig. 1. Fractal fern

Since fractal behavior is so universal in nature, it is perhaps not surprising that it is found as well in studies of memory and the process of learning. In an associative memory, the accuracy of identifying an input with the stored memory traces undergoes a transition from very good to very poor, as the number of stored memories increases. In other words, the process of memory also seems to follow behavior consistent with the fractal picture. Memory is obviously a crucial aspect of the learning process [18], so it would appear that fractal ideas are also helpful in understanding the way in which human beings learn. This leads us to consider that human skill acquisition processes influenced by the external task variations will show fractal patterns as well. No literature can found in applying fractal analysis on human learning process studies.

2.2 Detrended Fluctuation Analysis

Novel ideas from statistical physics led to the development of the detrended fluctuation analysis (DFA) [16]. The method is a modified root mean squared analysis of a random walk designed specifically to be able to deal with nonstationarities in nonlinear data, and is among the most robust of statistical techniques designed to detect long-range correlations in time series [3] [4] [20]. DFA has been shown to be robust to the presence of trends [12] and nonstationary time series [5] [13] [21].

The methodology begins by removing the mean, $\overline{B}$, from the time series, *B(t)*, and then integrating

$$y(k) = \sum_{t=1}^{k} [B(t) - \overline{B}] \tag{1}$$

The new time-series is then divided into boxes of equal length, *n*. The trend, represented by a least-squares fit to the data, is removed from each box; the trend is usually a linear, quadratic, or cubic function. Box *n* has its abscissa denoted by $y_n(k)$. Next the trend is removed from the integrated time series, *y(k)*, by subtracting the local trend, $y_n(k)$, in each box.

For a given box size *n*, the characteristic size of the fluctuations, denoted by *F(n)*, is then calculated as the root mean squared deviation between *y(k)* and its trend in each box

$$F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[y(k) - y_n(k)]^2} \tag{2}$$

This calculation is performed over all time scales (box sizes). A power-law scaling between *F(n)* and *n* indicates the presence of scaling

$$F(n) \propto n^{\alpha} \tag{3}$$

where the parameter α is a scaling exponent.

We will apply DFA to our psychomotor experimental data, whose acquisition is described below, in order to explore the possible presence of fractal behaviors.

3 Experimental Method

3.1 Apparatus

The study used a Wacom Graphire4® USB Tablet and pen as the tracking devices. The tracking software was developed by using MATLAB 7. The following Figure 2 illustrates the hardware used and interface of the software.

In this testbed, a square box (target) moves across the screen at a predetermined constant speed following a sinusoid-cosinusoid signal path, $Y_m(r,t)$ The subject is required to use the pen on the Tablet to drive a cross (cursor) to follow this box, and to stay within the box as much as possible. We thus have a spatiotemporal tracking signal, *Y(r,t)*, from the subject. Performance data (the location of the target, and

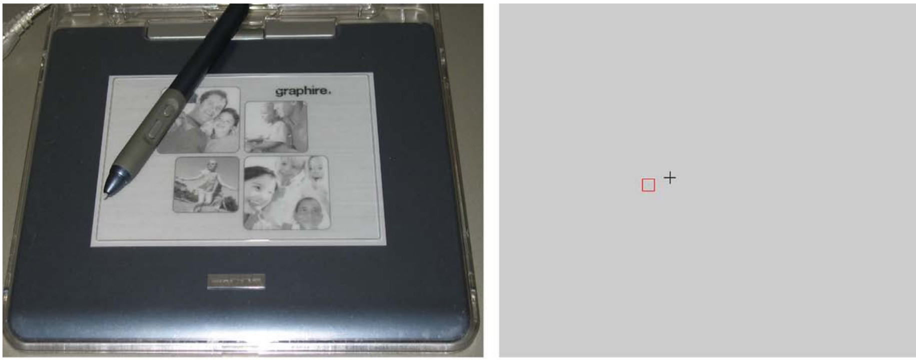


Fig. 2. Experiment testbed: Wacom Tablet and tracking pen (left), tracking software screenshot (right)

location of the cursor) was recorded five times every one second. The tracking performance was assessed by calculating the fluctuation from the actual target location and subject cursor location, $D=Y(t)-Y_m(t)$. Following this we extracted the temporal fractal scaling exponent from the vector fluctuation function. We have no clear *a priori* reason to expect that such a fractal scaling even exists. However, the fractal nature of some human physiology and neurophysiology leads us to suspect this possibility.

3.2 Procedure

We recruited college students to participate in this study. In this initial study we present results from just one subject in order to demonstrate our experimental and analysis techniques. To eliminate the possibility that efficiency will be influenced by memory, the human subject was not notified the path goal *a priori*. Upon arrival to the test location, subjects are briefed on the study objectives and procedures, and sign an informed consent form. They then are instructed to use the stylus and tablet to track the target and allowed access to the stylus and tablet a short time before the experimental task to become familiar with the experiment architecture.

Each session of the task takes 20 minutes, every two minutes of tracking followed by a 1-2 minute break to avoid subject fatigue. Each subject performs two sessions, one using primary hand, and one using the secondary hand (i.e., if he/she is a right handed person, then the right hand is primary, the other hand is the secondary hand). The sequence of primary and secondary hand were randomized to counterbalance the learning transfer effect between two hands.

4 Results

The detailed analysis of tracking data for one subject is now considered. The results for left hand tracking are shown. Figure 3 (top left) shows the fluctuation data, computed by subtracting the tracking from the target sinusoidal values. The subject generally performed tracking quite well. This is clear by the majority of data points clustering around the origin, which corresponds to zero fluctuation error (perfect tracking). There are a few significant deviations that occurred at least once in the positive Y-direction (up direction) and in both the positive and negative X-directions. The actually tracking data in XY are shown in the top right plot, and the fluctuations are shown plotted as a function of time in the lower left plot. Red corresponds to the X-fluctuations, and blue to the Y-fluctuations. It is clear from this plot that the amplitude of the fluctuations is much smaller in the Y-direction, demonstrating more accurate tracking in this direction; this is presumably due to the use of different muscle groups with finer psychomotor control for up/down motion than left/right motion.

Figure 4 shows the probability distribution functions for the X and Y motion. To test whether the data are normal we used a standard Lilliefors' composite goodness-of-fit test. The Lilliefors goodness-of-fit test considers whether the data in the vector D came from an unspecified normal distribution, and returns the result of the test in H. H=0 indicates that the hypothesis that data are normally distributed cannot be rejected at the 5% significance level. For our subject we found H=1; this indicates that the null hypothesis can be rejected at the 5% level. We found that the t location-scale distribution was an excellent fit to the empirical distributions. The t location-scale distribution is useful for modeling data whose distributions have heavier tails than the normal distributions.

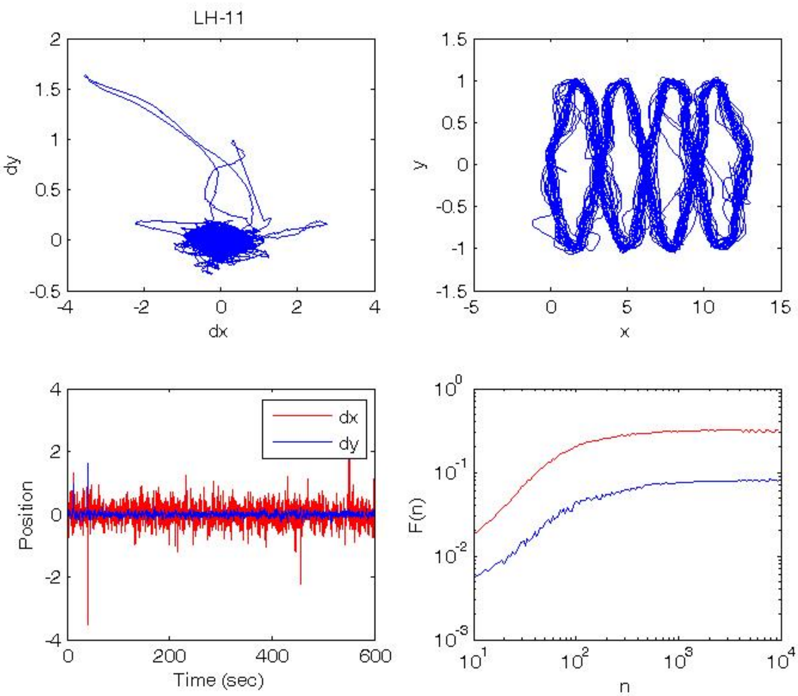


Fig. 3. Left hand human tracking data

The t location-scale distribution has the density function

$$\frac{\Gamma(\frac{\nu+1}{2})}{\sigma\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}\left[\frac{\nu+(\frac{x-\mu}{\sigma})^2}{\nu}\right]^{-(\frac{\nu+1}{2})} \tag{4}$$

with location parameter μ, scale parameter σ, and shape parameter ν. for the fluctuations in the X-direction, we found $\mu = 0.0011 \pm 0.0001; \sigma = 0.253 \pm 0.001; \nu = 5.7 \pm 0.2$ and for the Y-direction $\mu = -0.0062 \pm 0.0003; \sigma = 0.0511 \pm 0.003; \nu = 4.6 \pm 0.1$. The large scale function for the X-fluctuations indicates the much broader range of possible fluctuations in this direction.

Finally we turn to fractal analysis of the perturbations shown in the bottom left plot of Figure 3, which shows the results from DFA. The fractal fluctuation function *F(n)* as a function of the temporal scale size is shown in the lower right plot of Figure 3. A power law trend is apparent for the small scales for both fluctuations in the X- and Y-directions. We fit these between the scales of n=10 to 100. The fractal scaling exponent for the X-direction was $\alpha = 1.22 \pm 0.01$ and for the Y-direction $\alpha = 0.96 \pm 0.03$. The linear correlation coefficients of the raw data are r_x=0.999 and r_y=0.987.

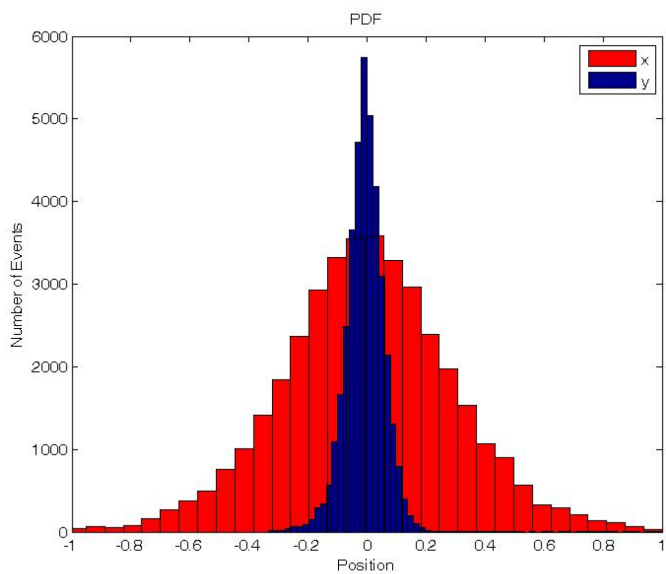


Fig. 4. Probability distribution functions for the X and Y motion fluctuations

5 Discussions and Conclusions

This study applied an innovative mathematical approach in human psychomotor skills learning studies. It is an advance in that a continuum of 'activity' levels can be modeled via fractal mathematical techniques; this advance can be seen both in model used and in terms of task manipulation itself.

In this study, we used DFA to examine the temporal variations of the nonlinear complexity of the psychomotor learning task previously described. We probed the existence of correlated behavior over the entire experimental time series. From our results, DFA analysis on all data demonstrated that the human psychomotor task learning, at time scales up to about two seconds (100 data points), is fractal – strong evidence is the existence of clear power laws for the fluctuations F(n).

This study helped to advance knowledge in the field of human psychomotor skills learning in terms of introduction of new ideas regarding the possible role of fractal processes in learning process efficiency and learning transfer. At this time, understanding of the development of human psychomotor skills is still extremely limited. This is largely due to the dynamic nature and variety of the task, lack of in-depth understanding and controllability of the task itself and the difficulty of measure of stochastic learning process by traditional methodologies. By using a specified task motion model (such as the sinusoid-cosinusoid signal model we applied here), this study provides a well defined structure for researchers to systematically quantify and control the task noise levels thus making the prediction more meaningful. The fractal learning model developed from this study will provide an innovative approach to integrating external causes of the human behavior into the picture, enabling researchers to obtain a more comprehensive picture of the problem. The model will provide detailed information regarding the nature of development of psychomotor skills.

Our study also applied a novel approach in data capturing in the behavior science domain. As stated in previous sections, traditional learning studies utilized discrete experiment data (i.e., different trials performance); this type of data cannot tell too much about the continuity of the learning process and more importantly, it limits the scope of analytical models that can be applied towards these data. The time series data generated from this study will enable many new statistical models to be applied on these data, so a more in-depth analysis can be performed.

To summarize, we have presented a new paradigm for the roots of psychomotor learning. In this paper we have presented only very preliminary experimental evidence in support of this view. We found a fractal scaling behavior was obtained for small temporal scales. These intriguing results will be further elucidated by further experiments on statistically significant samples, such as including a larger sample of human subjects and incorporation of different motion noises into the task such as Brownian motion. We believe this type of research will be of particular value to the field of human motor learning and training, to open a new window for studying human learning behavior within a specific human motor task context, and can provide great insights to those who train people on different psychomotor skills, and provide a valid prediction model on how to design a variable task to maximize the learning efficiency.

References

1. Adams, J.A.: A closed-loop theory of motor learning. Journal of Motor Behaviors 3, 111–150 (1971)
2. Bassingthwaihte, J.B., Liebovitch, L.B., West, B.J.: Fractal Physiology. Oxford University Press, New York (1994)
3. Blok, H.J.: On the nature of the stock market: Simulations and experiments, PhD thesis, University of British Columbia, Canada (2000)

4. Cannon, M.J., Percival, D.B., Caccia, D.C., Raymond, G.M., Bassingthwaighte, J.B.: Evaluating the Scaled Windowed Variance Methods for Estimating the Hurst Coefficient of Time Series. Physica A 241, 606 (1997)
5. Chen, Z., et al.: Effect of Nonstationarities on Detrended Fluctuation Analysis. Phys. Rev. E 65, 041107 (2002)
6. Damos, D.L.: Examining transfer of training using curve fitting: A second look. The International Journal of Aviation Psychology 1(1), 73–85 (1991)
7. Ellis, H.C.: The transfer of learning. The Macmillan Company, New York (1965)
8. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology 47(6), 381–391 (1954)
9. Fitts, P.M.: Perceptual-motor skills learning. In: Melton, A.W. (ed.) Categories of human learning, pp. 243–285. Academic Press, New York (1964)
10. Heathcote, A., Mewhort, D.J.K.: The evidence for a power law of practice. In: Proceedings of the 4th Conference of the Australasian Cognitive Science Society, The University of Newcastle, Australia 2000 (2000)
11. Henry, F.M., Rogers, D.E.: Increased response latency for complicated movements and a memory drum theory of neuromotor reaction. Research Quarterly 31, 448–458 (1960)
12. Hu, K., et al.: Effect of trends on detrended fluctuation analysis. Phys. Rev. E 64, 011114 (2001)
13. Kantelhardt, J.W., Zschiegner, S.A., Koscielny-Bunde, E., Havlin, S., Bunde, A., Stanley, H.E.: Multifractal detrended fluctuation analysis of nonstationary time series. Physica A 316, 87 (2002)
14. Liu, D., Blickensderfer, E., Vincenzi, D., Macchiarella, N.: Transfer of training. In: Vincenzi, D., Wise, J. (eds.): Human factors and simulation. In: review (2006)
15. Newell, A., Rosenbloom, P.: Mechanism of skill acquisition and the law of practice. In: Anderson, J.R. (ed.) Cognitive skills and their acquisition, pp. 1–56. Erlbaum, Hillsdale, NJ (1981)
16. Peng, C.-K., et al.: Quantification of Scaling Exponents and Crossover Phenomena in Nonstationary Heartbeat Timeseries. Chaos 5(1), 82–87 (1995)
17. Schmidt, R.A.: A schema theory of discrete motor skill learning. Psychological Review 82(4), 225–260 (1975)
18. Schmidt, R.A., Lee, T.D.: Motor control and learning: a behavior emphasis. 4th edn. Human Kinetics (2005)
19. Sherwood, D.E., Lee, T.D.: Schema Theory: Critical review and implication for the role of cognition in a new theory of motor learning. Research Quarterly for Exercise and Sport 74(4), 376–382 (2003)
20. Taqqu, M.S., Teverovsky, V., Willinger, W.: Estimators for long-range dependence: An empirical study. Fractals 3, 185 (1996)
21. Wanliss, J.A., Dobias, P.: Space Storm as a Phase Transition. J. Atmos. Sol. Terr. Physics (in press, 2007)
22. Yelle, L.E.: The Learning Curve: Historical Review and Comprehensive Survey. Decision Sciences 10, 302–328 (1979)

Effect of Noise-Enhanced on the Balance Control Ability in Older Adults

Fangfang Wu[1], Rencheng Wang[1], Dewen Jin[1], Xiao Hu[1],
Yiyong Yang[2], JiChuang Zhang[1], and Noboru Youshimura[3]

[1] Division of Intelligent and Biomechanical System
State Key Laboratory of Tribology, Tsinghua University Beijing, 100084, China
[2] School of Engineering and Technology, China University of Geosciences (Beijing), Beijing 100083, China
[3] Faculty of Engineering and Resource Science, Akita University, Akita 010-8502, Japan
wuff@mail.tsinghua.edu.cn

Abstract. Somatosensory function declines with age, and such changes have been associated with diminished motor performance. Noise input can enhance sensory and motor function. We tested the effects of the vibrating stimulation applied at the feet on balance control of 6 healthy elderly and 8 young volunteers. Balance performance was characterized using a FASTRAK System. We calculated four traditional measures of postural sway parameters and five stabilogram-diffusion analysis variables. Among the 14 participants, application of noise resulted in a reduction of seven of nine sway parameters in young participants and eight of nine sway variables in elderly participants. The results suggested that the imperceptible noise, when applied to the feet, could enhance the balance performance of healthy older adults. Therefore, using the noise-based devices may be an effective way to improve the balance control of elderly people.

Keywords: Older adults, Balance control, Balance ability, Vibration.

1 Introduction

Balance is an important functional skill that greatly influences human's ability to perform activities in daily living. Diminished balance ability, often seen in older adults, poses a serious health risk due to the increased likelihood of falling. Approximately one-third of community dwelling older adults aged 65 or over, fall at least once each year, and 10–15% of these falls result in serious injury or fracture [1–3]. Somatosensory deterioration, particularly in proprioception sense, has been identified as a risk factor contributing to falls in the elderly [4-5]. Input noise can enhance sensory [6] and motor [7] function, via a mechanism known as stochastic resonance. Reductions in proprioception at the feet have been shown to adversely affect posture control in older adults. Therefore, a therapeutic intervention designed to enhance feet sense may prove to be beneficial to balance control and the prevention of falls in the elderly. In this study, we extended the investigation of noise-based sensory enhancement to behavioral improvements in human balance control, especifically, for older adults.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 483–489, 2007.

2 Subjects and Methods

Six healthy ambulatory elderly adults (four males and two females, age 61-78, mean age 68±6.9, body mass index (BMI) 18.42-22.98 kg/m2, BMI mean 20.84±1.66 kg/m2) and young adults (six males and two females, age 20-22, body mass index (BMI) 18.59-22.39 kg/m2, BMI mean 20.19±1.36 kg/m2) volunteered to participate in this study. Elderly Subjects were randomly selected and contacted from old people's longevity association for Aged. A self-reported medical history screened potential participants for orthopedic or neurological conditions such as Parkinson's disease, diabetes, peripheral neuropathy, stroke, disabling arthritis, uncorrected visual problems, dizziness or vertigo.

During the experiments, subjects stood comfortably on a platform with their eyes closed and hands at their side (Fig.1). The experimental setup consisted of noise generator and power supply and 3space and data collection computer. Platform was flapped soon by 6s to inset on a piece of flat panel respectively(Fig.1①), according to the stand position of the feet, sole of each foot insets three vibrating elements, the amplitude of which was set independently using potentiometers for each foot. The noise generator consisted of a flat coreless vibration motor (Fig.1②), the noise generator had a white noise signal, vibration frequency is 100Hz or 200Hz (Fig.2).

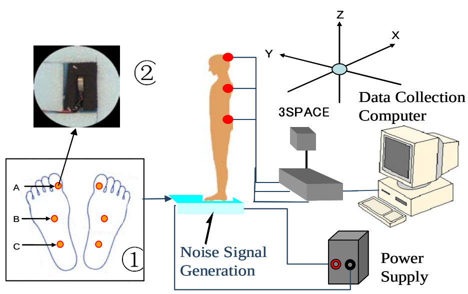


Fig. 1. A diagram of the experimental setup

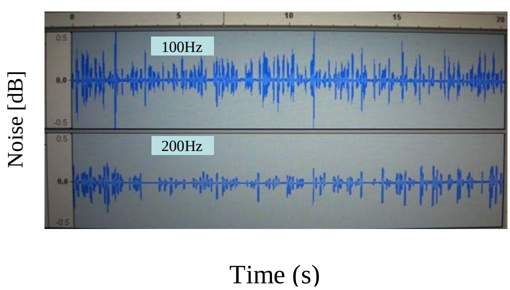


Fig. 2. Vibratory noise signal

To assess whole-body postural sway, displacement of the head-arm-waist segment was measured with the electromagnetic sensor to the left shoulder of each participant.

We used 3space FASTRAK motion track system (made by Polhemus company) to record the displacement during each 60s stance trial. Each subject stood barefoot on the platform, with their arms relaxed at their body sides and their eyes closed. Young and Elderly participants did 6 trials: three with noise presented to the foot and three without noise, the presentation sequence, noise or without noise. The data were sampled at a frequency of 8 Hz. Rest periods of 2min were permitted between each trial.

To characterize balance during quiet standing, we used both traditional stabilogram analyses and stabilogram-diffusion analysis [8-9]. Four traditional sway variables were computed for each trial: the Maximum excursions range of the AP, the Maximum excursions range of the ML, path length and swept area. stabilogram-diffusion analysis(SDA) was performed on the COP trajectories. SDA, as described by Collins and DeLuca, provides a number of physiologically meaningful parameters [8-9]. This analysis revealed that over short-term time intervals during quiet stance, the COP tends to drift away from a relative equilibrium point, while over long-term time intervals, the COP tends to return to a relative equilibrium point. This finding was interpreted as an indication that the postural control system utilizes open-loop control mechanisms over the short-term and closed-loop control mechanisms over the long-term to maintain upright stance. An open-loop control system is one which operates without sensory feedback, and in the case of the human postural control system may correspond to descending commands which set the steady-state activity levels of the postural muscles [8-9], A closed-loop control system, in contrast, operates with sensory feedback, and in the case of the human postural control system, it corresponds to the visual, vestibular, and somatosensory systems[8-9].

In this study, we considered five SDA variables: the critical mean square displacement $<\triangle r^2>_c$, the critical time interval $\triangle t_{rc}$, the effective short-term diffusion coefficient D_{sl}, the effective long-term diffusion coefficient D_{rl}, and the long-term scaling exponent H_{rl}. Within the above modeling framework, the critical mean square displacement characterizes the threshold at which feedback mechanisms are called into play by the postural control system, while the long-term diffusion coefficient and scaling exponent characterize the stochastic activity and antidrift-like dynamics, respectively, of these feedback mechanisms [9]. We focused on long-term (closed-loop feedback) variables because sensory enhancement from noise would affect predominantly closed-loop balance control. Specifically, we hypothesized that the addition of mechanical noise to the feet would lead to a reduction in the feedback threshold (as indicated by a decrease in the critical mean square displacement) and a more tightly regulated control system (as indicated by decreases in the long-term diffusion coefficient and scaling exponent).

Two-way repeated-measures analyses of variance (ANOVA) were used to determine if significant interactions between the main effects of stimulation (without noise versus noise) on postural sway in elderly and young subject (p =0.05).

3 Results

We using arm data to analyzed postural sway, Raw-data resultant planar stabilogram-diffusion plots for representative young and elderly subjects are given in Fig.3. In

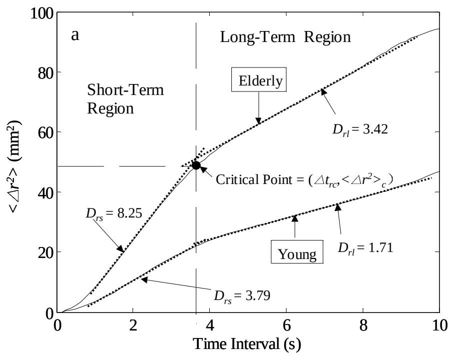


Fig. 3(a). Ling-ling plot of <$\triangle r^2$>$_c$ versus $\triangle t$ for a representative subject

Fig.3(a), Experimental anteroposterior linear stabilogram-diffusion plots (solid line) and fitted regressions (dotted lines) for one representative young subject (age 21 years) and one representative elderly subject (age 78 years) (solid lines). The computed short-term (D_{rs}) and long-term (D_{rl}) diffusion coefficients are shown for each subject. In fig.3(b). Experimental anteroposterior log stabilogram-diffusion plots (solid line) and fitted regressions (dotted lines) for one representative young subject (age 21 years) and one representative elderly subject (age 78 years) (solid lines). The computed long-term (H_{rl}) scaling exponent are shown for each subject.

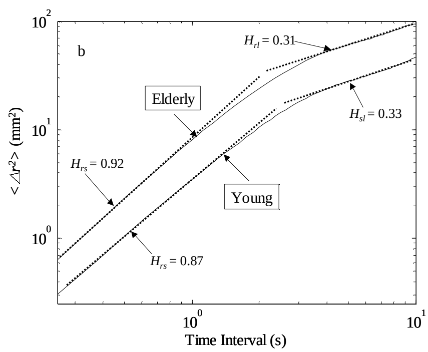


Fig. 3(b). Log-log plot of <$\triangle r^2$>$_c$ versus $\triangle t$ for a representative subject

In Fig.3, several comparative points about these plots should be noted. Firstly, from the linear-linear plots of mean square 3 space displacement versus time interval (Fig. 3(a)), it can be seen that the slope of the short-term region for the elderly subject was substantially larger than that for the young subject, whereas the slopes of the long-term region for the two subjects were relatively similar. These result were reflected in the computed values for the effective diffusion coefficients the short-term effective diffusion coefficient (D_{rs}) for the aged subject was substantially greater than that for the young subject, whereas the long-term effective diffusion coefficients (D_{rl})

for the two subjects were relatively similar (Fig.3(a)). Secondly, it should be noted from Fig.3(a). that the transition region between the short-term and long-term regions for the representative elderly subject occurred over larger mean square displacements and larger time intervals than that for the representative young subject. These differences were reflected in the calculated values for the critical point coordinates the critical mean square displacement ($<\triangle r^2>_{c}$,) and critical time interval ($\triangle t_{rc}$,) for the elderly subject were both larger than those for the young subject. Finally, from the log-log plots of mean square displacement versus time interval (Fig. 3(b)), it can be seen that the slopes of the short-term and long-term regions for the elderly subject were respectively greater and less than those for the young subject. These differences were reflected in the computed values for the scaling exponents the elderly subject's short-term and long-term scaling exponents (H_{rs} and H_{rl}, respectively) were respectively larger and smaller than other those for the young subject (Fig.3(b)). these representative findings, in general, capture the significant differences between the two populations.

Table 1. Comparison of the Postural sway measures with noise and no noise in elderly and young adults

Parameters	Yong			Elderly		
	No Noise	Noise	*p*	No Noise	Nois	*p*
Locus length	194.97±18.9	176.12±21.9	0.007	253.15±23.7	234.37±24.7	0.04
Range MLmax	25.12±6.50	17.51±2.5	0.008	32.28±6.2	26.32±2.3	0.006
Range APmax	13.97±5.5	11.13±2.9	0.065	15.34±5.5	13.26±7.9	0.264
Swept area	155.46±72.8	101.39±55.4	0.01	167.12±42.8	130.73±10.4	0.041
D_{rs}	9.14±3.6	6.49±4.4	0.209	18.48±8.1	9.39±4.5	0.036
D_{rl}	5.16±3.7	1.79±1.7	0.038	5.03±2.4	1.70±1.3	0.014
H_{rl}	0.33±0.2	0.36±0.1	0.68	0.33±0.1	0.385±0.1	0.387
$<\triangle r^2>_{c}$	45.84±13.1	24.11±9.0	0.005	100.22±60.0	37.21±7.9	0.048
$\triangle t_{rc}$	4.12±1.36	2.86±0.52	0.043	3.67±0.72	2.85±0.18	0.038

Eight of the nine sway parameters decreased with noise (Table 1), indicating an overall improvement in balance performance in the vibration condition, in elderly, seven of the nine sway parameters improved significantly (p <0.05). The application of noise resulted in a 18.5% reduction in Range MLmax (p =0.006) and a 5.4% decrease in the maximum AP excursion (p =0.264), a 7.4% reduction in path length (p =0.04), and a 21.8% decrease in swept area (p =0.041) with elderly adults. In the SDA variables, the application of noise resulted in a 48.2% significant reduction in the D_{rs} (p =0.036) and a 66.2% significant decrease in the D_{rl} (p =0.014), a 62.9% significant reduction in $<\triangle r^2>_{c}$ (p =0.048), a 22.3% significant reduction in $\triangle t_{rc}$ (P =0.038) with elderly adults.

4 Discussion

It was found that three of four the traditional sway variables in the elderly with noise stimulation, especially rang ML , rang AP and swept area, trended toward those of young subjects without noise condition. The findings indicate that the traditional sway variables in the elderly subjects decreased to levels of the young with the introduction of noise. Besides, the SDA variables were the same results in D_{rs}, D_{rl},$<\triangle r^2>_c$, $\triangle t_{rc}$ variables. As can be seen in Table 1, that elderly subjects, compared to young subjects, have higher feedback thresholds, approximately the same stochastic activity in their feedback mechanisms, and a greater likelihood after excursions to return to a relative equilibrium position.

The application of noise resulted in significant decreases in six of the nine sway parameters in young participants and seven of the nine sway variables in elderly participants (table1). Subsensory mechanical noise applied to the feet of quietly standing individuals with vibrating to enhanced feedback and reduced postural sway. Different effects noted between young and elderly indicate that elderly people gain more in motor control performance than young people with the application of noise to the feet. This may be due to that the young participants mayt have almost optimum sensory feedback and balance control compared with elderly. Noise such as randomly vibrating, might be effective in enhancement of performance of dynamic balance activities, and could enable older adults to overcome postural instability caused by age related sensory loss.

5 Conclusion

This study shows that subsensory mechanical noise applied to the feet of quietly standing subjects leads to enhanced feedback and reduced postural sway. Therefore, that imperceptible noise, when applied to the feet, could enhance the balance performance of healthy older adults. The findings suggest that noise-based devices may be effective to improve balance control of elderly people.

Acknowledgments. This study was supported by China Postdoctoral Science Foundation (No20060390480) and The Project Sponsored by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry.

References

1. Lavizzo-mourey, R.J., Forciea, M.A.: Geriatric secrets [J], pp. 110–115. Hanley & Belfus, Philadelphia (1996)
2. Prudham, D., Evans, J.G.: Factors associated with falls in the elderly: A community study [J]. Age and Aging 10, 141–146 (1981)
3. Perry, B.C.: Falls among the aged living in a high-rise apartment [J]. Journal of family practice 14, 1069–1073 (1982)
4. Robbins, A.S., Rubenstein, L.Z., Josephson, K.R., et al.: Arch Intern Med, vol. 149, pp. 1628–1633 (1989)

5. Lipsitz, L.A., Jonsson, P.V., Kelley, M.M., et al.: J Gerontol 46, M, 114–121 (1991)
6. Liu, W., Lipsitz, L.A., Montero-Odasso, M., Bean, J., Kerrigan, D.C., Collins, J.J: Noise-enhanced vibrotactile sensitivity in older adults, patients with stroke, and patients with diabetic neuropathy. Arch. Phys. Med. Rehabi 83, 171–176 (2002)
7. Priplata, A., Niemi, J., Salen, M., Harry, J., Lipsitz, L.A., Collins, J.J.: Noiseenhanced human balance control. Phys. Rev. Lett. 89, 238101 (2002)
8. Collins, J.J., De Luca, C.J.: Open-loop and closed-loop control of posture: a random walk analysis of center-of-pressure trajectories [J]. Expermental Brain Research 95(2), 308–318 (1993)
9. Collins, J.J., De Luca, C.J., Burrows, A., Lipsitz, L.A.: Age-related changes in open-loop and closed-loop postural control mechanisms. Exp. Brain Res. 104, 480–492 (1995)

Development of the Virtual-Human Santos™

Karim Abdel-Malek, Jingzhou Yang, Joo H. Kim, Timothy Marler, Steve Beck, Colby Swan, Laura Frey-Law, Anith Mathai, Chris Murphy, Salam Rahmatallah, and Jasbir Arora

Virtual Soldier Research (VSR) Program
Center for Computer-Aided Design, The University of Iowa, Iowa City, IA 52242
jyang@engineering.uiowa.edu

Abstract. This paper presents the background and history of the virtual human Santos™ developed by the Virtual Soldier Research (VSR) Program at The University of Iowa. The early virtual human environment was called Mira™. This 15-degree-of-freedom (DOF) upper-body model with posture and motion prediction was funded by John Deere Inc. and US Army TACOM Automotive Research Center. In 2003 US Army TACOM began funding VSR to develop a new generation of virtual humans called Santos (109 DOFs), which was to be another generation of Mira. Later on, Caterpillar Inc., Honda R&D North Americas, Natick Soldier System Center, and USCAR (GM, Ford, and Chrysler) joined the VSR partnership. The objective is to develop a *new generation* of digital humans comprising realistic human models including anatomy, biomechanics, physiology, and intelligence in *real time*, and to test digital mockups of products and systems before they are built, thus reducing the significant costs and time associated with making prototypes. The philosophy is based on a novel optimization-based approach for empowering these digital humans to perform, un-aided, in a physics-based world. The research thrusts include the following areas: (1) predictive dynamics, (2) modeling of cloth, (3) hand model, (4) intuitive interface, (5) motion capture, (6) muscle and physiology modeling, (7) posture and motion prediction, (8) spine modeling, and (9) real-time simulation and virtual reality (VR). Currently, the capabilities of Santos include whole-body posture prediction, advanced inverse kinematics, reach envelope analysis, workspace zone differentiation, muscle force and stress analysis, muscle fatigue prediction, simulation of walking and running, dynamic motion prediction, physiologic assessment, a user-friendly interface, a hand model and grasping capability, clothing modeling, thermo discomfort assessment, muscle wrapping and sliding, whole-body vibration analysis, and collision avoidance.

Keywords: Virtual humans, predictive dynamics, muscle wrapping.

1 Introduction

Research on digital human modeling and simulation has gained significant momentum in recent years even though several commercial software programs are available, including Jack®, RAMSIS®, and Safework®. Traditionally, digital human

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 490–499, 2007.

modeling and simulation is based on data-driven algorithms. Most models are developed from experiment data and statistic regression. However, computational approaches have been the focus of increasing attention in the human factors and ergonomics community. One reason is that while data-driven approaches can give accurate results for a finite number of tasks, computational approaches can give accurate results for an infinite number of tasks. For example, in posture prediction, we can perform an experiment to obtain data for a finite number of target points and, if the user would like to predict a target point that was not part of the experiment data, the data-driven algorithm would use interpolation to obtain the results. However, computation-based approaches embed the target point in their models.

In this paper, we review the history of the Digital Human Laboratory at The University of Iowa and its research activities. We also introduce the Mira environment from early research results and describe the Santos model and its capabilities.

2 History of Digital Human Lab

The Robotics Lab was launched in 1994 by Dr. Karim Abdel-Malek, and the basic research activities were robotics-related topics such as robotic kinematic and dynamics study, robot control systems, and robotic mechanism design. In 2001 research efforts were extended to human-related studies; for example, studying workstation design for optimal seating position and orientation or studying the difference between how humans and robots move objects. The Digital Human Lab at The University of Iowa was created and included five graduate students and one professor. We had several Unix workstations and PC desktops. We also had Jack software, EON Reality, and a 4-wall virtual reality environment. We developed a digital human environment called Mira, which was financially supported by TACOM Automotive Research Center and John Deere Inc. Research topics included human kinematic model, workspace, posture and motion prediction, biomechanics analysis, placement of manipulators and digital humans for the design, and virtual reality environment. In 2003 the Digital Human Lab was extended to the Virtual Soldier Research (VSR) Program. The research team has 36 members, including faculty, staff, and students. The scope of the research is much more broad, including all topics related to digital human modeling and simulation. The original 4-wall VR system was replaced with a 6-wall Portal system.

3 Mira Environment

The first generation of digital humans developed by the Digital Human Lab is called Mira, and the living environment for this digital human is EON Reality. Figs. 1 and 2 show snapshots of this digital human. This digital human has 15 degrees of freedom (DOFs) from the waist to the right hand, as shown in Fig. 3. It has the following capabilities: (1) posture prediction (Mi et al., 2002; Abdel-Malek et al., 2001), (2) motion prediction (Abdel-Malek et al., 2006), (3) reach envelope analysis (Abdel-Malek et al., 2004; Yang et al., 2005; Yang et al., 2006), (4) placement (Abdel-Malek et al., 2004), and (5) layout design (Abdel-Malek et al., 2005). This approach

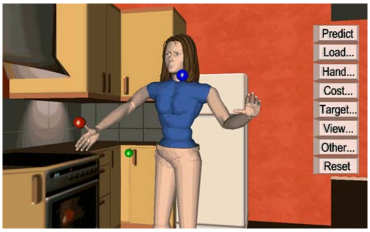


Fig. 1. MiraTM environment

Fig. 2. Testing using Mira^{TM}

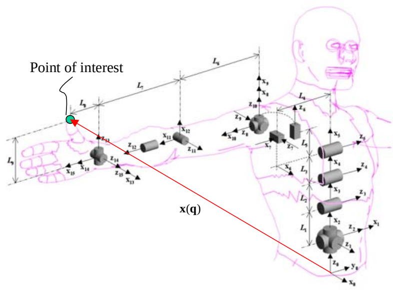


Fig. 3. Model for torso and right upper extremity

formulates the design process as an optimization problem. Fig. 4 is a snapshot of the cross section of the workspace for the restrained driver. Figs 5-8 show results for this model where the posture prediction results are compared with IKAN (Tolani et al., 2000).

4 Virtual Human Santos

In 2003, the VSR team was formed and several large projects were initiated. Sponsors include US Army TACOM TARDEC, Natick Soldier Research Center, Caterpillar Inc., USCAR, Rockwell Collins, and HONDA R&D North Americas. There are fourteen thrust areas for the new digital human modeling and simulation research.

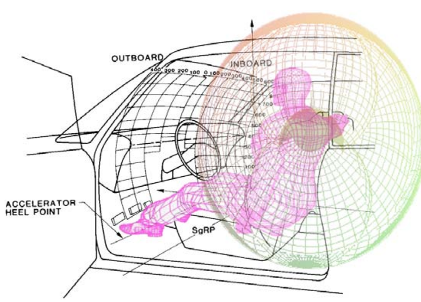


Fig. 4. Cross section of the workspace for the restrained driver

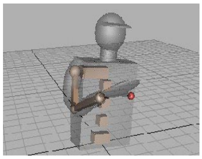


Fig. 5a. Posture prediction example 1 Mira result

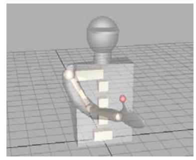


Fig. 5b. Posture prediction example 1 IKAN result

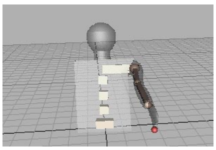


Fig. 6a. Posture prediction example 2 Mira result

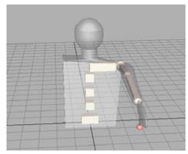


Fig. 6b. Posture prediction example 2 IKAN result

Santos is realistic digital human model that includes anatomy, biomechanics, physiology, and intelligence in real-time. It has 109 DOFs in the kinematic model, thousands of muscle lines in the musculoskeletal model, and a real-human appearance. The kinematic model is shown in Fig. 9, and the musculoskeletal model is shown in Fig. 10. The work to build muscle lines is ongoing, and it is an interactive real-time muscle wrapping model.

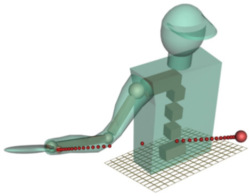
Major projects and capabilities are listed as follows:

(1) Real-time VR

A primary focus at VSR is virtual reality (VR), modeling the world on computers with superior realism. However, we continually strive to take VR one step further—to model the world in real time, thus providing instantaneous feedback during human simulations. This allows for additional realism as well as the opportunity for users to actually interact with the virtual world. A key component in our development of premiere virtual reality environments is a state-of-the-art facility, a six-wall, 3-dimensional virtual reality room called the Portal. The Portal is the only PC-based six-wall virtual reality environment in the world. The interior of the Portal is a large 3m x 3m x 3m room. It works with a high-speed native Super XGA DLP projection system with active stereo. Each of the six screens has a resolution of 1024 x 1024 (SuperXGA). The six-screen cube provides a viewing range of 360 degrees horizontally and 360 degrees vertically.

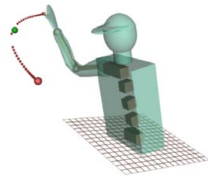
(2) Optimization-Based Human Prediction

A key element in our efforts to model humans is a novel optimization-based approach to posture and motion prediction. In fact, much of what we do uses optimization as a foundation. This foundation stems from a solid history of expertise in the field of optimization, not only at the Center for Computer Aided Design (CCAD), but also at The University of Iowa in general. With this approach, joint angles essentially provide design variables, constrained by joint limits based on anthropometric data as



(a)

Fig. 7a. Upper body motion prediction point-to-point



(b)

Fig. 7b. Upper body motion prediction via point

Fig. 8. Layout design

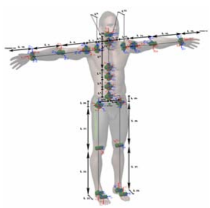


Fig. 9. Santos kinematic model

Fig. 10. Musculoskeletal model

well as other physical constraints. Human performance measures that represent physically significant quantities, such as energy or discomfort, provide the objective functions. We contend that human posture and motion is *task-based*, meaning it is governed by different performance measures, depending on which task is being completed. Using this approach affords the digital human a substantial amount of autonomy, enabling an avatar to react to infinite scenarios rather than having to draw on prerecorded motions. With the optimization approach, incorporating additional capabilities or features is usually a matter of simply introducing new constraints and/or objective functions. In addition, our optimization-based approach allows us to operate in real time. Current digital human modeling methods and commercial codes do not have the capability to actually predict realistic motion in real time.

(3) Rigorous Human Performance Measures

Research concerning the human performance measures addresses fundamental questions about the prediction of human motion, resulting in realistic postures, path trajectories, and locomotion. The primary objective of this work is to mathematically model physically significant quantities. With metrics like energy, the process is relatively simple. However, with more subjective quantities, such as discomfort, the process involves determining which factors have been shown experimentally to contribute to discomfort and then how to represent each factor mathematically. Thus far, we have developed a comprehensive model for energy, discomfort, effort, maximum joint torque, two models for vision (and how the tendency to actually view objects affects posture), and joint displacement. In addition, we are working toward performance measures for muscle fatigue and time to exhaustion. The results of this research are the cornerstone of real-time optimization and real-time kinematics/dynamics models.

(4) Multi-objective Optimization

We recognize that in various situations, posture and/or motion may not be dictated by just one performance measure; it may be necessary to combine various measures. Consequently, we have developed substantial expertise in the field of multi-objective optimization and currently conduct research on how to effectively combine multiple performance measures. With multi-objective optimization problems, there is no single solution point. Rather, there is a set of infinitely many solutions, called the Pareto optimal set, with each point representing a different posture. We are able to depict this set of points and use its structure to glean insight as to how functions should be combined. We have studied more than 60 different methods and have developed new methods as well.

(5) Kinematic Posture and Motion Prediction

Although optimization is used extensively with VSR capabilities, the primary application is with human posture and motion prediction. Currently, we are accelerating toward having the most versatile, extensive, realistic, and computationally efficient posture prediction software package available. In addition to modeling the above-mentioned performance measures, we have developed a new method for dual-arm coordination (predicting posture to multiple kinematic chains that depend on common degrees of freedom, such as the spine). We can constrain any number of end-effectors to any point, line, or plane. We can stipulate the orientation of any local coordinate system. In addition, we incorporate self-avoidance; Santos is essentially aware of himself. All of the features can be applied to the whole body. Currently, we are conducting research on methods for incorporating whole-body movement (as opposed to limb movement with the hip fixed), as when Santos kneels down or leans his whole body forward. Our success with posture prediction is being extended to motion prediction, where optimization is used to predict time-histories of joint angles rather than just a single set of static angles. With notion prediction, time-histories are represented by B-splines, and the parameters for these splines provide the design variables for the optimization problem.

(6) Zone Differentiation

As part of the work with posture prediction, we have developed a novel approach to depicting and analyzing reach envelopes, which not only indicates which points in space are accessible, but also provides contours of values for the various performance measures. Depicting a reach envelope (or workspace) essentially provides a tool for product design. However, the workspace indicates whether or not a point is reachable, but does not necessarily indicate which points are most desirable (most comfortable, require the least effort, etc.). Thus, we have developed an optimization-based approach for workspace zone differentiation. In addition, we are conducting research on voxel processing methods for effectively visualizing 3-dimensional clouds of data, such as those that are provided with our zone differentiation capabilities.

(7) Real-Time Dynamics

This research effort focuses on adding significant capabilities to the motion subtended by digital humans. To date, there is no single digital human modeling program that enables its avatars to act and react to dynamic forces. Tasks such as pulling a lever, feeling the weight of an object, pushing a load, or simply throwing an object of known mass and inertia are not possible. Fundamentally, the optimization formulation for kinematic motion prediction can be augmented by including the equations of motion as constraints. In this way, we avoid cumbersome and time-consuming integration. This optimization-based approach avoids having to use typical forward dynamics or recursive dynamics to solve equations of motion, can easily accommodate highly complex human models, incorporates externally applied loads, and can potentially function in real time. Dynamic motion prediction becomes especially complex for activities that involve gait and balance, such as walking and running. Again, we take advantage of optimization, in conjunction with the Zero Moment Point (ZMP) approach.

(8) Modeling Physiology

In addition to predicting human motion, we are able to model physiological indices, such as heart rate, breathing rate, oxygen consumption, energy consumption, body temperature, and physiological strain index. Oxygen consumption is a common measure of energy expenditure during sustained tasks in humans. Aerobic mechanisms for the regeneration of the required chemical sources of energy can be maintained for extended periods of time at submaximal levels of exercise. If the limit of aerobic energy metabolism is reached, anaerobic metabolism is required. While this metabolic process is efficient, most individuals cannot maintain it for more than several minutes. Further, time-to-exhaustion is inversely proportional to oxygen uptake. Therefore, the estimation of oxygen consumption provides a method to estimate how long an activity could be maintained and how well it would be tolerated for a variety of task scenarios.

(9) Task-Segmentation

In order to model complex multi-faceted tasks, we are developing a new methodology that involves segmenting motion into smaller subtasks that are simulated using the above-mentioned techniques and then stitched together. The general process is described as follows. First, using motion capture data of a soldier conducting a specific task(s), the motion will be segmented into a series of subtasks. For each subtask, the initial and final configurations for the digital human (a virtual soldier) will be determined and modeled with various constraints. Which constraints are used depends on the nature of the subtask. The digital human is then used to predict the motion between the initial and final configurations within each subtask while providing feedback for various indices. The protocol will be repeated for each subtask so that a user of the final environment will be able to request that the virtual soldier conduct a task under varying conditions of loading and mobility (e.g., different armor configurations and different external loading conditions).

As an example, consider the task outlined in Section 2. Motion capture will provide an understanding of the kinematics performed by a typical soldier for executing this task. The second step will be to segment the task into smaller subtasks as follows: (i) Run for 30 meters from point A to point B, (ii) Change postures from running to a diving position (B to C), (iii) Assume prone position (from diving at C to prone at D). For each of these subtasks, the initial and final conditions are modeled with constraints in the optimization formulation, and which constraints are most appropriate depends on the subtask being modeled.

(10) Mathematical Modeling of Clothing

In many circumstances, clothing can and does impact human performance while also providing protection from exposure to a wide spectrum of external hazards. Our research objective is to describe clothing with computational models and to then exercise these models to realistically predict how different types of clothing (of varying fabrics, cut, and fit) impact human performance when engaging in physical tasks such as running, climbing, and walking. With a knowledge of how clothing impacts human performance, the clothing can be redesigned to improve performance. This research combines many long-standing issues in solid mechanics, such as the following: (i) Modeling flexible systems (clothing) as shells and membranes at arbitrarily large deformations, (ii) Contact modeling between clothing and the human, and self-contact between portions of the same garment, (iii) Developing suitable constitutive models that account for the fibers comprising the fabric, and their evolving woven or knitted structure in the garment.

(11) Hand Modeling

This research addresses the complex biomechanics of the human hand and wrist. The long-term objective of this effort is to develop a three-dimensional model of the human hand and wrist for the purpose of kinematic and dynamic modeling, with the ability to predict grasps autonomously for any type of object. The research will couple digital image segmentation routines, realistic motions, and advanced modeling techniques to determine the joint torques, reaction forces, etc. The ability to incorporate realistic forces and force distributions via a force-feedback system holds the potential to immerse the subject in the virtual task, thereby improving the overall perception of task performance and the overall level of fidelity. A 25-DOF hand model with which users can interact has been developed. Artificial intelligence (AI) is used as Santos's brain to guide him to grasp objects or interact with his environment autonomously.

(12) Skeletal Muscle Modeling and Simulation
Skeletal muscles are important components of a virtual human. It is muscle contraction that causes the movement of the bones and consequently the motion of the body. The stress level in the muscle directly relates to a human's feeling of comfort. The objective of this research is to develop a computational framework that allows for the consideration of skeletal muscle in human modeling. While most commercial digital human environments do not have skeletal muscle modeling capabilities, we believe that this study will be crucial to the development of new capabilities for analysis and simulation.

In order to incorporate the limitations of muscles on the complete body, we have devised a relatively simple but computationally efficient muscle-modeling approach based on readily available data for torque-velocity relationships at various joints. In this way, we simplify the modeling problem by working in the joint space rather than the more complex muscle space.

Muscles are capable of generating more force at lower shortening velocities, and minimal force production at high shortening velocities. A similar trend applies to the relationship between torque and velocity. Validating simulations in joint space is easier than validating the performance of individual muscles, since the determination of joint torque velocity data can be obtained in vivo, whereas force velocity data for individual muscles is not feasible in humans. During the dynamic motion predictions for each subtask, the required joint torques needed to accomplish the predicted tasks are calculated. By plotting the first derivative of the joint angle versus time predictions (angular velocity versus time) against the joint torque versus time predictions we can determine the predicted joint torque versus angular velocity profiles for each subtask. These predicted curves are compared to the standard values for specific joint torque-velocity (T-V) profiles, readily available in the literature, in order to estimate the perceived level of difficulty of a subtask.

5 Summary

This paper reviews the research activities at VSR and the history of Santos, a physical-based digital human environment. Kinematics and dynamics are married with optimization techniques and are the core of this digital human. This is groundbreaking work in the human factors and ergonomics community because traditionally everything comes from experiments. However, we are investigating different human performance measures that govern human motion and formulating the human motion as an optimization problem. We also use experiments to validate our work.

Acknowledgements. This research work was supported by John Deere Inc., US Army TACOM, US Army Soldier System Center at Natick, Caterpillar Inc., Honda R&D North Americas, and USCAR.

References

1. Abdel-Malek, K., Yu, W., Jaber, M.: Realistic Posture Prediction. In: Proceedings of SAE Digital Human Modeling for Design and Engineering, Arlington, VA, USA (2001)
2. Abdel-Malek, K., Yu, W., Yang, J., Nebel, K.: A Mathematical Method for Ergonomic Based Design: Placement. International Journal of Industrial Ergonomics 34(5), 375–394 (2004)
3. Abdel-Malek, K., Yang, J., Brand, R., Tanbour, E.: Towards Understanding the Workspace of Human Limbs. Ergonomics 47(13), 1386–1405 (2004)
4. Abdel-Malek, K., Mi, Z., Yang, J., Nebel, K.: Optimization-Based Layout Design. Journal of Applied Bionics and Biomechanics 2(3-4), 187–196 (2005)
5. Abdel-Malek, K., Mi, Z., Yang, J., Nebel, K.: Optimization-Based Trajectory Planning of Human Upper Body. Robotica 24(6), 683–696 (2006)
6. Abdel-Malek, K., Yang, J., Marler, T., Beck, S., Mathai, A., Patrick, A., Arora, J.: Towards a New Generation of Virtual Humans. International Journal of Human Factors Modelling and Simulation 1(1), 2–39 (2006)
7. Kim, J.H., Abdel-Malek, K., Yang, J., Marler, T.: Prediction and Analysis of Human Motion Dynamics Performing Various Tasks. International Journal of Human Factors Modelling and Simulation 1(1), 69–94 (2006)
8. Mi, Z., Yang, J., Abdel-Malek, K., Mun, J.H., Nebel, K.: Real-Time Inverse Kinematics for Humans. In: Proceedings of 2002 ASME Design Engineering Technical Conferences, Montreal, Canada, September 29-October 2, 2002 (2002)
9. Tolani, D., Goswami, A., Badler, N.: Real-Time Inverse Kinematics Techniques for Anthropomorphic Limbs. Graphical Models 62(5), 353–388 (2000)
10. Yang, J., Abdel-Malek, K., Nebel, K.: On the Determination of Driver Reach and Barriers. International Journal of Vehicle Design 37(4), 253–273 (2005)
11. Yang, J., Marler, R.T., Beck, S., Abdel-Malek, K., Kim, J.: Real-Time Optimal Reach-Posture Prediction in a New Interactive Virtual Environment. Journal of Computer Science and Technology 21(2), 189–198 (2006)
12. Yang, J., Kim, J.H., Abdel-Malek, K., Marler, T., Beck, S., Kopp, G.R.: A New Digital Human Environment and Assessment of Vehicle Interior Design. Computer-Aided Design (in press)

Validation of Predicted Posture for the Virtual Human Santos™

Jingzhou Yang[1], Salam Rahmatalla[1], Tim Marler[1], Karim Abdel-Malek[1], and Chad Harrison[2]

[1] Virtual Soldier Research Program, Center for Computer-Aided Design
The University of Iowa, Iowa City, IA 52242
[2] Honda R&D Americas, Raymond, OH
jyang@engineering.uiowa.edu

Abstract. Digital human modeling and simulation plays an important role in product design, prototyping, and manufacturing: it reduces the number of design iterations and increases the safety and design quality of products. Posture prediction is one of the key capabilities. It is especially useful in the design of vehicle interiors for checking the reachability of buttons and determining comfort levels. This paper presents the validation of predicted posture for the virtual human Santos. The predicted posture is a physics-based model and is formulated as a multi-objective optimization (MOO) problem. The hypothesis is that human performance measures (cost functions) govern how humans move. We chose 12 subjects from four different percentiles, all Americans (female 5%, female 50%, male 50%, and male 95%). Four realistic in-vehicle tasks requiring both simple and complex functionality of the human simulations were chosen: reaching a point at the top of the A-pillar, the radio tuner button, the glove box handle, and a point on the driver's B-pillar seatbelt adjuster. The subjects were asked to reach the four target points, and the joint centers for wrist, elbow, and shoulder and the joint angle of elbow were recorded using a motion capture system. We used these data to validate our model. The validation criteria comprise R-square and confidence intervals. The results show that the predicted postures match well with the experiment results, and are realistic postures.

Keywords: Predicted posture, MOO, human performance measures, validation, virtual humans.

1 Introduction

Posture prediction is an important component within the human modeling and simulation package. There are three major types of posture prediction. The first is an experiment-based method (empirical-statistical approach) in which the posture comes from experiments and statistic regression; this approach does not need to be validated. The second is an inverse kinematic method and the third is a direct optimization-based method, and other approaches. The approaches that are not directly from experiment do need to be validated.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 500–510, 2007.

The objective of this study is to validate in-vehicle optimization-based posture prediction using experiments. In this paper we present the details of the experimental protocol, criteria for validation, and a three-domain validation method.

In the empirical-statistical approach, data are collected either from thousands of experiments with human subjects or from simulations with three-dimensional computer-aided human-modeling software (Porter et al., 1990). The data are then analyzed statistically in order to form predictive posture models. These models have been implemented in the simulation software along with various methods for selecting the most probable posture given a specific scenario (Beck and Chaffin, 1992; Zhang and Chaffin, 1996; Faraway et al., 1999). This approach is based on actual human data and thus does not need to be verified in terms of realism.

The inverse kinematics approach to posture prediction, which uses biomechanics and kinematics as predictive tools, has received substantial attention. With this approach, the position of a limb is modeled mathematically with the goal of formulating a set of equations that provide the joint variables (Jung et al., 1995; Kee et al., 1994; Jung and Choe, 1996; Wang, 1999; Tolani et al., 2000). Zhang and Chaffin (2000) introduce an optimization-based differential inverse kinematics approach for modeling 3-D seated dynamic postures and validation. Griffin (2001) gives a review of the validation of biodynamic models. Park et al. (2004) present a memory-based human motion simulation. It uses an optimization method as the motion modification algorithm to fit the specific motion scenario. It also uses experiments to validate the model. Wang et al. (2005) demonstrate the validation of the model-based motion in the REALMAN project. However, all of these approaches are restricted to relatively simple models.

Yang et al. (2004) introduce a direct MOO-based posture prediction for a high-degree-of-freedom human model in real-time. This paper presents a three-domain method for validating the predicted posture. The validation scenario consists of the in-vehicle seated reaching tasks.

This paper is organized as follows. Section 2 briefly reviews Santos's kinematic model and the multi-objective optimization-based posture prediction model. Section 3 presents validation tasks, subject selection, experimental protocol, data collection and analysis, validation criteria selection, and the detailed validation process. Section 4 presents the conclusion and discussion.

2 MOO-Based Posture Prediction

In this section we present the kinematic model for the virtual human Santos, determine the physics factors that affect human postures, derive the mathematical models (human performance measures) of these factors, and formulate the redundant inverse kinematics problem as a MOO problem.

2.1 Santos's Kinematic Model

The human body is a complex system that includes bones, muscles, tendons, and nerves. The human skeletal model can be simplified as a mechanism with rigid links connected by kinematic joints. One anatomy joint could have one or more kinematic

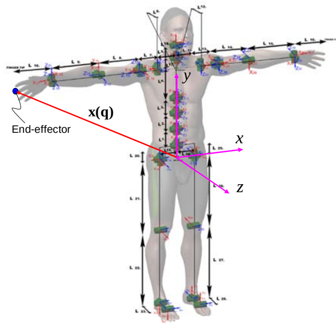


Fig. 1. Santos's kinematic model

joints. For example, the shoulder joint has three revolute kinematic joints, while the elbow joint entails only one revolute joint. Therefore, we can model the human as a system with high degrees of freedom (DOF). Fig. 1 shows a 109-DOF Santos kinematic model (Yang et al., 2005, 2006). It has five open loops starting from the root at the hip point to the right hand, left hand, head, left leg, and right leg.

The end-effector, or point of interest, is the fingertip, toe, or a point on the head. The position vector of the end-effector with respect to the global frame attached to the hip point $\mathbf{x(q)}$ is defined by the following:

$$\begin{bmatrix} \mathbf{x(q)} \\ 1 \end{bmatrix} = {}^{0}\mathbf{T}_{1}{}^{1}\mathbf{T}_{2} \cdots {}^{n-1}\mathbf{T}_{n} \begin{bmatrix} \mathbf{x}_{n} \\ 1 \end{bmatrix} \tag{1}$$

where ${}^{i-1}\mathbf{T}_i$ is a (4×4) transformation matrix from frame i to frame $i-1$ determined by the Denavit-Hartenberg method (Denavit and Hartenberg, 1955). $\mathbf{q}=\begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix}^T$ is the joint angles from the hip to the end-effector. $\mathbf{x}_n$ is the end-effector vector with respect to frame n.

2.2 Posture Prediction

Posture prediction is defined as follows. We try to find the configuration (joint angles) of a human when he or she reaches for a target point using the fingertip or other end-effector (point of interest on body). Because the human model has a high number of DOFs, this problem has multiple solutions, or is a redundant problem. We have proposed a MOO-based method (Yang et al., 2004) in which physics factors govern how humans move. This approach ensures autonomous movement regardless of the scenario and can be implemented in real time.

There are many physics factors that govern human postures. The first is that human posture gravitates to a "neutral posture," and there are different task-related neutral postures. For example, when humans stand to achieve a task, the neutral posture is one in which the arms are straight down at the side, and the neck, torso, and legs remain straight in the frontal plane. When humans sit, the neutral position is one in which the torso leans on the seat back, knees bend, feet rest on the floor, and arms are on the arm rests of the seat. The second factor that governs human posture is potential energy. Humans use one posture instead another to save energy. For example, humans use the arm before the torso or clavicle because the mass of the arm is much smaller than the mass of the torso. The third factor is that initial posture affects the predicted posture, and the fourth is that joints with tendons try to avoid stretching those tendons. The fifth factor is vision; humans try to see a target as they reach for it.

Based on these physics factors, we developed several mathematical cost functions (human performance measures). They are joint displacement, effort, change of potential energy, visual displacement, and discomfort (Yang et al., 2004; Marler, 2005, Marler et al., 2006). The MOO-based approach is different from traditional inverse kinematic methods or the experiment-based approach. It enforces human performance measures to drive the posture. Therefore, this approach is more generic and can be used in all different scenarios. This MOO problem is formulated as follows:

To find: $\mathbf{q}$

to minimize: Human performance measures

subject to: Constraints

where constraints include the requirement that the end-effector should touch the target point and the requirement that joint angles should be within their limits.

3 Validation of Postures

In this section, we will describe the validation tasks, subject selection, data collection using a motion capture system, experimental protocols, experiment data variation analysis, validation criteria, and the three-domain approach to the validation.

3.1 Tasks to Validate

This study focuses on the in-car environment. We selected realistic in-vehicle tasks that test both the simple and the complex functionality of the human simulations. Fig. 2 shows the four tasks that were chosen for the experiment. Task 1 requires reaching the point at the top of the A-pillar, a simple reach task. Task 2 requires reaching the radio tuner button, a slightly difficult reach task. Task 3 requires reaching the glove box handle, a difficult reach task. Task 4 requires reaching a point on the driver’s B-pillar seatbelt adjuster. This is a complex task that requires reaching across the body and turning the head to see the target. The general procedure for achieving a task is as follows: the subject holds the steering wheel using both hands for the initial posture, then maintains the left handhold and uses the right-hand index finger to touch the target point.

3.2 Subject Selection

To cover a larger driver population, auto designers choose a range of percentiles from 5% female to 95 % male. Therefore, in our experiment, we chose four different populations, all Americans: 5% female, 50% female, 50% male, and 95% male. Also within percentiles, three subjects are selected.

3.3 Data Collection and Experimental Protocol

Optical systems have many applications in biomechanical studies (Hagio et al., 2004; Robert et al., 2005; Rahmatalla et al., 2006). In the motion capture process, a number of reflective markers are attached over bony landmarks on the participant's body, such as the elbow, the clavicle, or the vertebral spinous processes. As the participant walks or carries out a given physical task or function, the position history of each marker is captured using an array of infrared cameras. A potential problem with passive markers, though, is *occlusion*, where the markers do not appear in enough of the camera shots due to blockage of the line of sight between the marker and the cameras by objects in the scene or by other parts of the subject's body. In this work, redundant markers (more than the minimum required) were used to compensate for occluded markers. The time history of the location of the reflective markers was collected using a Vicon motion capture system with eight cameras at a rate of 200 frames per second.

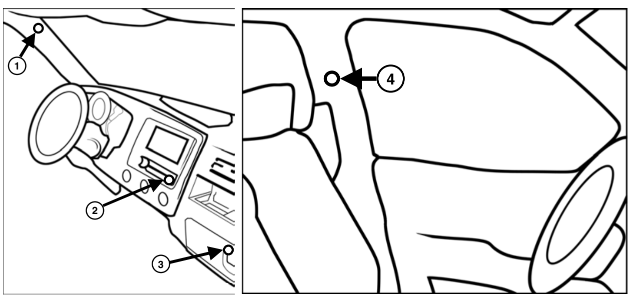


Fig. 2. Four in-vehicle tasks

Several marker placement protocols have been introduced in the literature for studying various types of motion. Among these protocols, plug-in gait is a typical protocol that has been adopted by systems such as Vicon and LifeMOD. In the plug-in gait protocol, markers are attached to bony landmarks on the subject's body to establish local coordinate systems on various segments of the body. Joint centers and joint profiles can then be obtained using these coordinate systems. Methodologies for calculating joint center locations and link lengths of humans are available and have been somewhat successful (Halvorsen et al., 1999). However, because of the occlusion problem, it is very hard to accomplish this goal with a high degree of accuracy for people in seated positions with the existence of obstacles in the environment. In this work, due to the complexity of the capturing environment for a seated person inside a car and due to the limited number of cameras available at the time of the experiments (eight), redundant markers were attached to the upper part of the subject's body to estimate joint center locations and to compensate for the missing markers, as shown in Fig.3. The marker placement protocol was designed to facilitate the process of obtaining the location of the joint centers of the upper part of the subject's body (right wrist, right elbow, right shoulder, and hip joint) during the experiments. In the marker placement protocol, one marker was attached to the end effecter (the end of the middle finger), three markers were attached to the wrist joint, and three markers were attached to the elbow joint. The shoulder joint is very complicated, so four markers were used to estimate the location of this joint center, and two markers were used to estimate the location of the clavicle (one on the clavicle and one on the T4). In all cases, the joint center location was estimated by the intersection of the lines connecting the markers on each joint.

3.4 Trial-to-Trial and Subject-to-Subject Data Variation

As described in above section, we obtained the joint centers for the right wrist, elbow, and shoulder and the right elbow angles using a motion capture system. The objective was to correlate the relationship between the simulation results and experiment results. However, we had to transform all data to the same coordinate system. In this study, we used Santos Spine1 above the hip joint as the common coordinate system.

When we did the experiment, each subject reached one target point for five trials. For trial-to-trial variation, we use subject S95M3 to illustrate the results. For subject-to-subject variation, we demonstrate variation within one percentile (the three S95M subjects). The results are shown in Fig. 4. The vertical axis denotes the variation from the mean value of the distance from the target point to the origin, and the horizontal axis represents the four tasks. As shown in Fig. 4, a single subject's trial-to-trial variation is small and within ± 10mm and ± 3degrees . Subject-to-subject variation within percentiles is only slightly larger than trial-to-trial variation. The variation values are within ± 25 mm and ± 6 degrees. Therefore, it is appropriate to average the data from all five trials and reasonable to choose a representative subject within the percentiles.

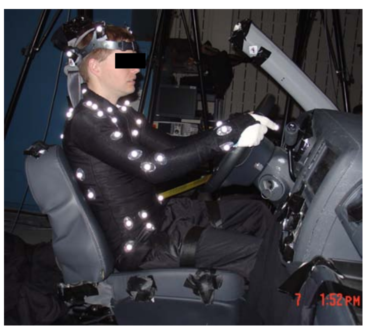


Fig. 3. Experiment markers

3.5 Validation Criteria

There are several criteria for validation, though they are not independent. In this section, we summarize the criteria for general model validation, which are R^2 , confidence intervals of mean, regression, and other statistic values.

The coefficient of determination R^2 is *the relative predictive power of a model*. Generally, *the closer it is to one, the better the model is*. However, R^2 is not an absolute measure for a good model. There are different factors that affect the values of R^2, such as the range of X values, different patterns of X values, average values of X, and randomness. Therefore, when R^2 is close to 1, it does not necessarily mean that the model is good; it only indicates that the model can represent the experiment data very well. Likewise, when R^2 is small, it does not necessarily mean that the model is not good. We have to carefully study the data themselves and also check whether the slope is close to 1. The 45 degree line with respect to the X axis denotes that the experiment and simulation values are the same. This is another important parameter for validating the model.

A confidence interval gives an estimated range of statistic values that is likely to include an unknown population ρ, the estimated range being calculated from a given set of sample data. Confidence intervals are calculated at a confidence level (a certain percentage), usually 95% (α =0.05), but we can also produce 90%, 99%, 99.9%, and other confidence intervals for the unknown parameter ρ. Confidence intervals are more informative than the simple results of hypothesis tests because they provide a range of plausible values for the unknown parameter. Confidence intervals of mean, regression, and slope of the linear regression are used for validation of the model.

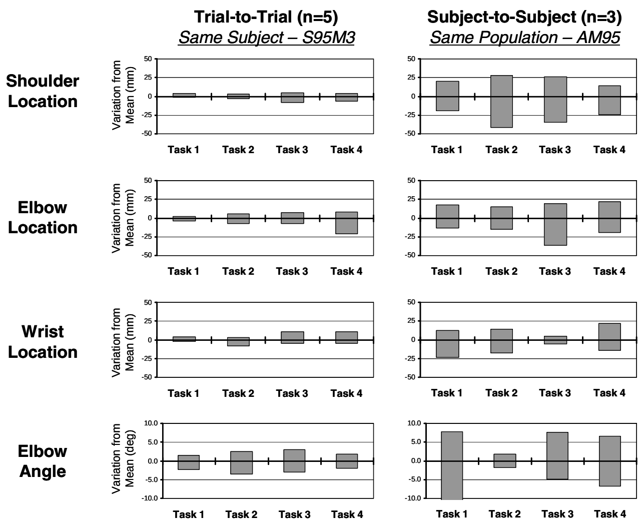


Fig. 4. The variation in the mean values for the distance between the target and the origin

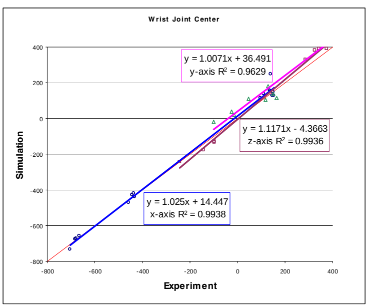
3.6 Three-Domain Approach

In this study we have developed a three-domain approach to validate the predicted posture. This approach validates the predicted posture in three domains: across percentiles, within percentiles, and with respect to Task 2.

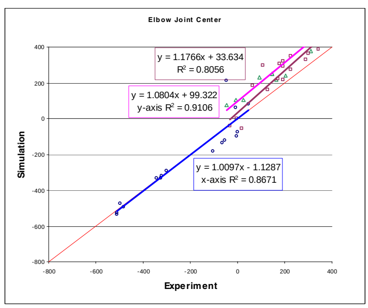
We have selected one representative subject for each percentile. That means there is one subject from each of these percentiles: 5% female, 50% female, 50% male, and 95% male. Therefore, across percentiles we consider only four subjects and consider all tasks together. We first illustrate the coefficient of determination R^2. Then, we demonstrate only the confidence intervals for the elbow joint angles. Figs. 5 (a)-(c) are regression plots of Cartesian coordinates of x, y, and z for the joint centers of the right wrist, elbow, and shoulder, respectively. Fig. 5(d) is the right elbow joint angle regression plot. The red straight line is at 45 degrees with respect to the X axis. All R^2 values satisfy $R^2 \geq 0.7$, and the slopes of all regressions are close to 1.

For confidence intervals, we demonstrate only the elbow joint angle in this paper; however, the procedure to determine the confidence intervals is the same as for other joint centers. The confidence interval of regression for the elbow joint angle with a 95% confidence level is $0.1115 \leq \rho^2 \leq 0.7967$. The confidence interval for the slope of the regression of the elbow joint angle with a 95% confidence level is within $(0.7033, 1.3038)$. The confidence interval for mean with 95% confidence level is shown in Fig. 6 (a).

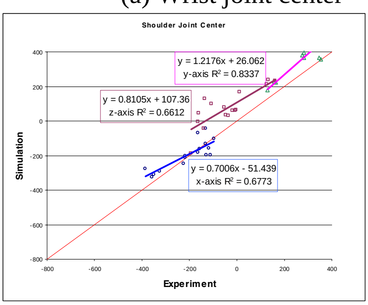
Within 50% male category, we analyze the regression similar to the cases across percentiles. All R^2 values satisfy $R^2 \geq 0.7$, except in the case of Shoulder x. The confidence intervals are straightforward, as in the examples above. In this case, the confidence interval of regression for the elbow joint angle (95% confidence level) is



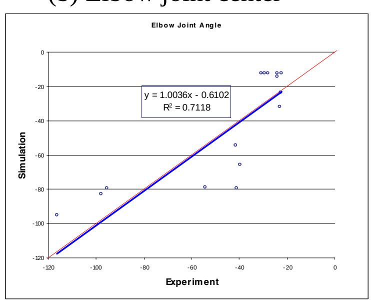
(a) Wrist joint center



(b) Elbow joint center

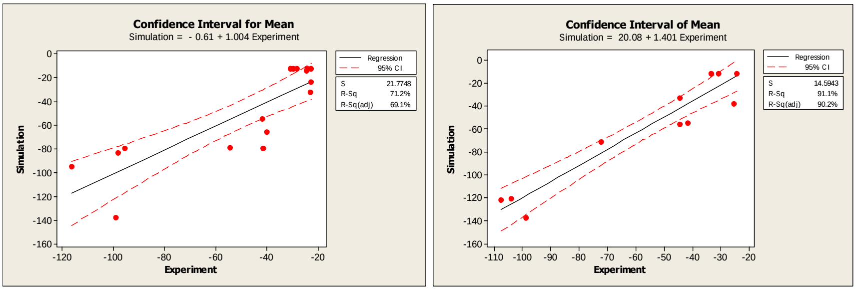


(c) Shoulder joint center



(d) Elbow joint angle

Fig. 5. Regression plots across percentiles



(a) Cross percentile (b) Within a percentile

Fig. 6. Confidence interval for mean

$0.4999 \le \rho^2 \le 0.9509$. The confidence interval for the slope of the regression of the elbow joint angle (95% confidence level) is within $(1.075, 1.727)$. It is obvious that one is not within this range, which means the accuracy is not as good as it is across percentiles. The confidence interval for mean with 95% confidence level is shown in Fig. 6(b).

Regarding to Task 2, most R^2 values are very small. Does this tell us that the model is not good? The answer is NO. In these plots, most data points are clustered together. This shows the property of "pure randomness." That means it is not appropriate to validate our model in this domain. However, this also tells us that our posture prediction model does not capture the properties of gender because we can use one set of data from male 50% to replace the data from female 50% because there is no significant difference.

4 Conclusions and Discussions

This study presents a systematical validation procedure of the predicted posture. In general, although there were errors from all different sources, such as the motion capture system, the method that transforms markers on the skin to joint centers or joint angles, the human model, and the posture prediction model, the validation process was successful and the predicted postures were within the required error limit. Human postures can be validated using a few selected key markers, and every degree of freedom does not need to be tracked. The results show that it is best to track key joint angles and the location of joint centers, not the markers on the skin. We used a three-domain approach to validate the predicted postures, and all plots contain a wealth of information. Generally, R^2 is a metric for the degree of precision of the model. The slope of the regression is another criterion to indicate the accuracy of the model. However, because we used a limited number of samples, it was necessary to investigate confidence intervals to predict the range of values that was likely to indicate the unknown population.

The results show that regression with respect to a task is not appropriate for validating the posture because it is a pure randomness problem. They also show that there are some areas in which our model could be improved. The posture prediction model could be improved in the following areas: (1) an advanced shoulder model considering coupled degrees of freedom and coupled joint limits; (2) gender within the model; (3) hip movement; (4) neck and head model that is connected to the spine; (5) cognitive modeling aspects. In the meantime, we should do further experiments with an increased population.

Acknowledgements. This work is supported by Honda R&D Americas, Inc. The authors would like to thank all other research members at VSR for their contribution to the Santos environment, especially Steve Beck, Anith Mathai, Chris Murphy, and Amos Patrick. The authors also thank Dr. Ying Zhang from the Department of Biostatistics at The University of Iowa for help on statistics aspects for this research work.

References

1. Denavit, J., Hartenberg, R.S.: A kinematic notation for lower-pair mechanisms based on matrices. Journal of Applied Mechanics 77, 215–221 (1955)
2. Faraway, J.J., Zhang, X.D., Chaffin, D.B.: Rectifying postures reconstructed from joint angles to meet constraints. Journal of Biomechanics 32, 733–736 (1999)
3. Griffin, M.: The validation of biodynamic models. Clinical Biomechanics 16(1), S81–S92 (2001)
4. Hagio, K., Sugano, N., Nishii, T., Miki, H., Otake, Y., Hattori, A., Suzuki, N., Yonenobu, K., Yoshikawa, H., Ochi, T.: A novel system of four-dimensional motion analysis after total hip athroplasty. Journal of Orthopaedic Research 22(3), 665–670 (2004)
5. Halvorsen, K., Lesser, M., Lindberg, A.: A new Method for Estimating the Axis of Rotation and the Center of Rotation. Journal of Biomechanics 32, 1221–1227 (1999)
6. Jung, E.S., Kee, D., Chung, M.K.: Upper body reach posture prediction for ergonomic evaluation models. International Journal of Industrial Ergonomics 16, 95–107 (1995)
7. Kee, D., Jung, E.S., Chang, S.: A man-machine interface model for ergonomic design. Computers & Industrial Engineering 27, 365–368 (1994)
8. Marler, R.T.: A Study of Multi-objective Optimization Methods for Engineering Applications, Ph.D. Dissertation, University of Iowa, IA (2005)
9. Park, W., Chaffin, D., Martin, B.: Toward memory-based human motion simulation: development and validation of a motion modification algorithm. IEEE Transactions on Systems, Man, and Cybernetics-Part. A: Systems and Humans 34(3), 376–386 (2004)
10. Porter, J.M., Case, K., Bonney, M.C.: Computer workspace modeling. In: Wilson, J.R., Corlett, E.N. (eds.) Evaluation of Human Work, pp. 472–499. Taylor & Francis, London, UK (1990)
11. Robert, J.J., Michele, O., Gordon, L.H.: Validation of the Vicon 460 Motion Capture SystemTM for Whole-Body Vibration Acceleration Determination, ISB XXth Congress-ASB 29th Annual Meeting, Cleveland, Ohio, July 31 - August 5, 2005 (2005)
12. Rahmatalla, S., Xia, T., Contratto, M., Wilder, D., Frey-Law, L., Kopp, G., Grosland, N.: 3D Displacement, Velocity, and Acceleration of Seated Operators in Whole-Body Vibration Environment using Optical Motion Capture Systems. In: The Ninth International Symposium on the 3-D Analysis of Human Movement, Valenciennes, France, June 28-30, 2006 (2006)
13. Tolani, D., Goswami, A., Badler, N.: Real-Time Inverse Kinematics Techniques for Anthropomorphic Limbs. Graphical Models 62(5), 353–388 (2000)

14. Wang, X.: A behavior-based inverse kinematics algorithm to predict arm prehension postures for computer-aided ergonomic evaluation. Journal of Biomechanics 32, 453–460 (1999)
15. Wang, X., Chevalet, N., Monnier, G., Ausejo, S., Suescun, A., Celigueta, J.: Validation of a model-based motion reconstruction method developed in the REALMAN project, 2005-01-2743. In: Proceedings of SAE Digital Human Modeling for Design and Engineering Symposium, Iowa City, Iowa, June 14-16, 2005 (2005)
16. Yang, J., Marler, R.T., Kim, H., Arora, J., Abdel-Malek, K.: Multi-Objective Optimization for Upper Body Posture Prediction. In: 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Albany, New York, USA, August 30-September 1, 2004 (2004)
17. Yang, J., Marler, T., Kim, H.J., Farrell, K., Mathai, A., Beck, S., Abdel-Malek, K., Arora, J., Nebel, K.: Santos^{TM}: A New Generation of Virtual Humans. In: SAE 2005 World Congress, Cobo Center, Detroit, Michigan, April 11-14, 2005 (2005)
18. Yang, J., Marler, R.T., Beck, S., Abdel-Malek, K., Kim, J.: Real-Time Optimal Reach-Posture Prediction in a New Interactive Virtual Environment. Journal of Computer Science and Technology 21(2), 189–198 (2006)
19. Zhang, X., Chaffin, D.B.: Task effects on three-dimensional dynamic postures during seated reaching movements: an analysis method and illustration. In: Proceedings of the 1996 40th Annual Meeting of the Human Factors and Ergonomics Society, Philadelphia, PA, Part 1, vol. 1, pp. 594–598 (1996)
20. Zhang, X., Chaffin, D.: A Three-dimensional Dynamic Posture Prediction Model for Simulating In-vehicle Seated Reaching Movements: Development and Validation. Ergonomics 43(9), 1314–1330 (2000)

Modeling of Multi-organization Performance for Emergency Response

Yufei Shu[1] and Kazuo Furuta[2]

[1] RISTEX, Japan Science and Technology Agency,
[2] Graduate School of Engineering, Tokyo University,
100-0004 Tokyo, Japan
shuyufei@ristex.jst.go.jp

Abstract. To response severe disaster, such as hurricane, tsunami, or terrorists attack, multi-organization, such as government, companies, media, and public are involved. How to make these inter- or intra dependent organizations cooperate timely and accurately when they face uncertainty and stressful situation is the primary concerns. In this paper, we proposed multi-organization performance model, which introduced organizational factors, individual factors and focused on communication and collaboration among different organizations. Based on it we developed multi-agent emergency response system. This IT solution has capability to integrate drills into simulated response activities and reproduce the drills. To demonstrate how the system can be used as a powerful tool for supporting and evaluating the performance of multi-organization in disaster management and response, the test simulation is carried out using a scenario extracted from real-life emergency drills.

Keywords: Multi-organization performance, Multi-agent system, Emergency response, Disaster management.

1 Introduction

Statistics show globally, more than 2.5 billion people were affected by floods, earthquakes, hurricanes and other natural disasters between 1994 and 2003. A 60 percent increase over the previous two 10-year periods, U.N. officials reported. Not only protect from natural or man-made disasters, but also mitigate the effect when disaster occurs is recognized as most important way to save the society from losing life and property. Response to disasters in a timely and effective manner can reduce deaths and injuries, contain or prevent secondary disasters, and reduce the resulting economic losses and social disruption. Based on these considerations, emergency response systems are prepared before hand. Responding to large-scale disasters, such as sever hurricane, earth-quake, or terrorists attack it is a usual case that multi-organization involving many layers of government, commercial entities, media, and the public are involved. These entities work together to save lives, preserve infrastructure and community resource, and reestablish normality within the community. How to make these inter- or intra- dependent organizations cooperate

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 511–520, 2007.

timely and accurately in making critical decision when they confront uncertainties and stressful situations during crisis is the primary concerns.

As Klein (1997) claimed that human beings recognize the situation and make decision well under stressful situation if they have experienced the similar case. However, the experience of severe natural or man-made disasters is very less on hand. Because these kind accidents are rare happened and characteristics of risk are hardly to know. In addition, it is hardly to forecast how the multi-organization cooperatively response to severe disaster, because the multiformity of collaboration and diversity of communication among the organizations. The problem that emergency responders faced is that from those a few disaster experiences, it is hardly to develop proper emergency response system or assess the effectiveness of it that had already been developed. To enhance the capability of situation awareness, decision-making under crisis situation, emergency responders are trained to go through similar scene. One way is to conduct real-life drills. However, real-life drills are limited in many aspects. It is expensive in terms of time and resources. It is also hard to organize. Furthermore, detail planning is required to instrument drills. This makes it is almost impossible to test out response to the scenarios by “if-then” analysis. The other way is IT based solution, the computer simulation, which can help addressing many of the challenges brought forth by the need for emergency response preparedness. It is much more flexible and can test various responses by setting different simulation scenarios. As US Department of Defense (DoD) found that use of simulations instead of live exercise for training can reduce the training cost to one-tenth (Jain and McLean, 2003). It is recognized as a good alternative and complementary method to training responders and assessing performance of emergency response system.

A numbers of emergency response simulation system have been developed (Cohen, etc., 1989; Avesani, etc., 2000; NSSL, 2003; LANL, 2003; A-TEAM, 2005). By reviewing the state of the art, we noticed that major parts of the emergency response studies focus on the model disaster itself, for instance, the fire dynamics inside buildings. What is not studied in as much detail is modeling response activity, especially modeling response activity of inter, intra-dependent organizations. For the emergency response, not only we need to know how the disaster propagate, how to evacuate, but also we should find out how can effectively cooperate between different organizations because it plays heavy role to mitigate the effect of disaster. When the severe disaster occurs, multi-organization cooperate and coordinate to deal with emergency situations. It becomes very necessary to model and assess the performance of cooperated multi-organization. Although it is important, how to model multi-organization performance, how to assess that if the emergency response system is well designed is still not established well.

Our study will explore the design of organizational models for emergency response to increase the efficiency of entities organizations under the conditions of uncertainty and rapid change. Therefore, the aim of this study is firstly to understanding the underlying mechanism of multi-organization cooperation and to propose a multi-organization performance model, which will focus on communication and collaboration between different organizations. Since the real socio-technological system is so complex, the models need to be designed to abstract the critical points of interaction within each organization, as well as the critical points of interaction among the organizations. Secondly, based on proposed team model, we build a parallel and distributed multi-agent system to simulate emergency response activity of multi-organization, that agents play different roles to drive simulation. In addition, various modules are embedded in the system to capture information, make decisions, and execute action.

2 Multi-organization Performance Model

2.1 Multi-organization Cognitive Process

There are numbers of research efforts have been targeted at modeling small team behavior in the context of specific application domain, such as industrial process control, aviation, or ship navigation, etc. For example, Furuta and Shimata (1999) analyzed team cognitive process in open and interactive style. Milici (1996) used fixed coordination pattern to study team performance. Shu etc. (2002) modeled team behavior in a dynamic scenario to simulate operator team response to accident in nuclear power plant. Shu etc. (2005) studied team situation awareness and developed criteria to assess it. These works have been proved successfully reflect the essential characteristics of team cognitive process in engineering application domain. However, since man-made disaster risk has increased due to a possibility of terrorist attacks in the world wide. Also, statistics shows that natural disasters also have increasing intention. The analyzing of team performance has been shifted to the context of social-technology system, especially, multi-organization participating system for emergency response. Analysis sophisticated multi-organization performance, the multiformity of collaboration and diversity of communication within organizations should be toke into account.

Considering the fact, that members of large team can be divided into smaller groups or arranged in a hierarchical structure, we will analysis multi-organization performance based on our previous research on small team behavior. We have proposed TBNM model using a serial sub-model to describe context of team performance (Shu, etc. 2002). The model defined task structure, event propagation, team factor, and individual factor, which determine the basis of communication. In this study, we adopt this TBNM model to describe multi-organization performance. We analyze multi-organization cognitive process within a hierarchical organizational structure, because it is a usual case within sophisticated social-technical system. In the multi-organization, since assigned tasks are different among the organizations, emergency response task is carried out via communication and collaboration. Multi-organization's emergency response process consists of recognition of information, decision-making, planning and execution as shown in Fig.1. The collaboration or communication scheme is not static due to changing of environmental situations. It is vitally affected by social interaction. Organization or member judge, decide, plan, and act based on the knowledge they have and the information they obtain from the environment and other members. Collaboration and information exchange among organizations or members are accomplished through communication and embedded in the cognitive process. It is triggered by task structure, organization structure, individual factor (e.g. distribution of knowledge, authority among members, etc.). This team-oriented cognitive process emphasized that problem solving is

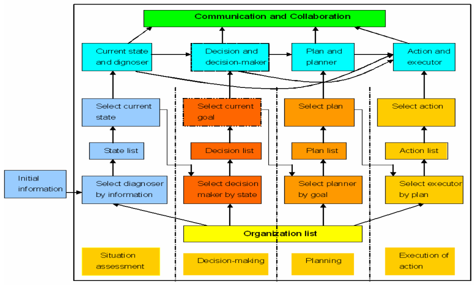


Fig. 1. Multi-organization cognitive process

accomplished through communication and coordination between members. It is domain independent, which can describe the multi-organization behavior from small operator teams to local government offices.

2.2 Role-Based Agent Model

In the hierarchical multi-organization, each organization or member play different role to carry out the emergency response task. The organization or member is modeled as a rule-based human agent embedded in the team cognitive process with input-recognition-decision-plan-action cycle. Depending on the concrete simulation requirement, agent can be designed according to the scale of participants and granularity of organizations by configuring knowledge base for each organization.

An agent obtains information or resources from the environment and others agents, recognizes situation, decides, plans referring to the knowledge base, and then acts, as shown in Fig.2. We predefine the profile of each agent, including name of agent, position in the organization, resource they equipped, contact point with other agent, individual factors such as knowledge, duty, and authority distribution among members over the organizations, and emergency responding procedures. Individual factors contribute to determining path of communication. The distribution of knowledge defines knowledge about emergency response task that each agent has. The distribution of authority defines authority to make a decision that each agent has. Although some organizations know how to execute the action well, they are not allowed to make decision of execution of certain action, if they do not have authority to do so. Distribution of authority is used to find an appropriate decision-maker among the cooperative organizations. Distribution of duty means that each agent takes the different charge of execution. It used to find an executor of action through the entity of organizations.

Disaster information and resource as initial information are input to the agent model. Resources that each agent equipped are defined initially. According to the different disaster drill setting, we can design different contents and allocation of resources that each agent holds. These resources, such as workforces, telephones, fax, are required for executing certain actions. When agent uses resource to execute actions, the number of resources decreases. When agent obtains a new resource message, the number of resource increases.

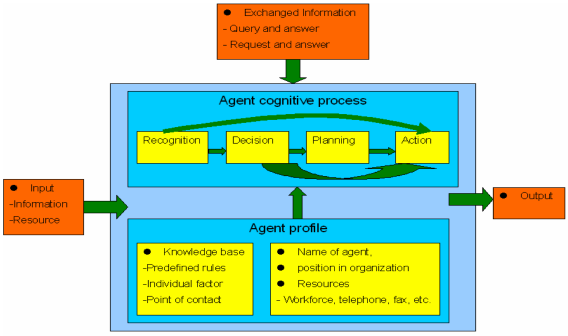


Fig. 2. Agent model

When the new message is received, based on the contents of message, agent recognizes the current situation and appropriate decision are determined in this step. Since emergency response is shaped in light of individual factors, such as distribution of authority, duty and knowledge, primitive response can be determined by referring knowledge base. However, some responses are depending on the situation. These responses are determined by experience. Rules and protocols used in this study were extracted from the emergency drills and result of interviews emergency responders.

After decision made, a typical and important response plans or actions should be decided, which were constructed by goals-means task analysis (GMTA). They are prescribed in knowledge base extracted from emergency response procedures or manuals, such as resident evacuation, dispatch fire-fighter forces or star emergency transportation. Procedure is series of actions, it consists preconditions and detailed actions. Each action requires certain preconditions, resources, time period, and effects. When preconditions satisfied, the action will be initially automatically.

Usually, cooperation and communication among organizations are required to accomplish emergency response task. It is achieved by exchanging message among agents. A message contains various type of basic data used for representing information and resource exchanges in the real situation, such as information sender, receiver, time and contents of communication. We classified the verbal protocol (utterances) using the message type extracted from drills and used them in the simulation system. The utterance includes inform, query, and request. ACL-based (Agent Communication Language) messages transfers are used in this study for exchanging information. The disaster-related information including contents of communication between agents is storied in the messages storage.

3 Multi-agent Emergency Response System

Emergency response system for dealing with nuclear disaster had been developed in our previous work (Kanno and Furuta, 2004). To enhance the ability of handling large spectrum of disaster, such as natural or man-made disaster, the system is implemented under various situations, especially, applied to various social-technological systems. The domain independent proposal is developed to handle different types of crisis and complete different emergency response tasks. Since multi-organization usually use synthetic environments for mission rehearsal exercises, training, or testing the response plans. We are developing multi-agent simulation system based on the above model to reproduce the synthetic environments and represent multi-organization emergency response in it. The simulator mimics the emergency response activities of multi-organization and how each organization interacts with other organizations. Each organization is modeled as an agent that can receive and recognize information obtained from the environment and other agents as well as make decision based on its own knowledge. Each agent is assigned a role for accomplishment of entity emergency response task. Communication and coordination between agents are modeled as information exchange. In order to achieve full potential as a supporter and assessor of emergency response system, the simulator is constructed as a knowledge base system with different functional capabilities. The architecture of it consists of several different technical modules: team cognitive process module, agent cognition module, disaster scenario module, interface module, communication channel, and knowledge base, as shown in Fig.3. The system is developed in program language C++ and Java.

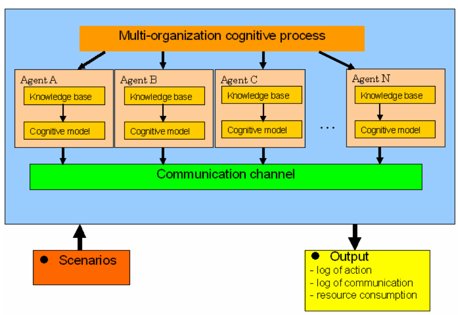


Fig. 3. Architecture of emergency response system

Multi-organization cognitive process module: describes the process of how multi-organization accomplish emergency response task via coordination and communication. It provides a platform to describe disaster propagation, emergency response task, behavior of each agent, and communication paths.

Agent cognition module: this module simulates behavior of organization or member embedded in the disaster response process, such as police, fire fighter, medical staff, or state government agencies. Depends on the scale of simulated disaster, the scale of agents and granularity of organization can be a small group or city or state government offices. The profile of each agent is prescribed, such as name of the agent, predefined responding procedures (e.g. precondition, action), individual factors (e.g. authority, duty), contact point to the other agent (e.g. where or whom the information should sent to). Resources (e.g. phone, fax, and car) of each agent are described in the resource storage. The contact point and responding procedures are extracted from real-life drill and interview with disaster responder. At any given time, agents are associated with a given task in the disaster response synthetic environment.

Communication channel: this module provides platform for information exchange among the agents. In our study, we use CORBA (Common Object Request Broker Architecture) to transfer information. An agent was implemented as a distributed object constructed on the CORBA platform. Message exchange among the agents is carried out via CORBA APIs.

Scenario module: This module describes the disaster-related information, such as event propagates, information exchange among the agents. Scenario stores sets of messages corresponding to the real-life drill scripts that are used by the drill participants. We represent scenario as a snapshot taken at every time unit. This snapshot is reproduced as a set of information which includes information sender, receiver, time, contents of the information, etc. It is used as inputs of the simulation system and then delivered to the proper agent according to the contents of message.

Interface module: The interface allows that users interact with the simulation environment and set input to the simulator. The user can start the simulation, select the response members, chose the granularity of organizations through the interface. The outputs of simulation system are logs of communications, actions, and resource consumptions for each agent. By analysis these simulation logs, the process of multi-organization collaboratively response to the emergency and the performance of the multi-organization can be evaluated and visualized via GUI. The GUI is developed in a programming language Java.

4 Emergency Response Simulation

The most important aim of this IT solution is to support and evaluate emergency response system. To demonstrate how this IT solution can be used as a powerful tool for supporting and evaluating the performance of multi-organization in disaster management and response context, the test simulation is carried out. Firstly, we represent the real-life drill through the computer-based simulation. Secondly, we test the response procedure under different drill settings and assess the emergency response system by changing design of agent profiles, or structure of organizations.

4.1 Simulation Scenario

Since there is high probability of sever earth quake in Kanto area of Japan in the future 50 years, the major Japanese estate company conducts real-life drill every year to train staffs responding to sever earthquake. The drill role assignment and drill scripts that they delivered to the real-drills participants are used in our test simulation.

According to the drill settings and scripts, the simulation scenarios are extracted and profiles of agents are predefined. The emergency response system is divided in to five hierarchical agent groups: headquarter, information section, administration and advertising section, restoration section, and owners. Each group consists of number of agents in a hierarchical structure, equipped with fixed phone, satellite phone, fax, and PC, etc. Agent profiles, such are individual factors (e.g. role, authority, duty, etc.) and resources (e.g. equipment, workforce, etc.) were designed in light on the drills scripts. The knowledge base of each agent is extracted from the earthquake response procedures and manuals. When the disaster message is received, all agents have to work cooperatively to judge the situation, and make corresponding decision. Communication and cooperation paths among the agents are determined according to drills. Real-life drill was conducted in 2 days, totally costing 2 hours and 20 minutes. We will reproduce it on the computer based simulation system within 2 hours and 20 minutes. When we change the contents of agent profile, or organizational structure, we can analyze emergency response system under different settings. In this test application, two groups of simulations setting are prepared by changing agent profiles. One situation is that authority is centralized at headquarter group, and the other is de-centralized among different agents.

4.2 Testing Results and Discussions

Outputs of simulation are log of communication and action of each agent, and resource that each agent used. The response activity of entity organization is shown in Fig.4. This figure is a combined form of the communication flow diagram and the action flow diagram of multi-organization. In the figure, x axis represents time flows from left to right. Communication between agents includes information and resource exchange, e.g. call, fax, etc. is represented as an arrow, while the no-communication actions, such as meeting, setting, etc. corresponds to the colored lines labeled with the action name. Each horizontal line donated to the agent or agent group. When organization model described in agent profile has been changed, as expected, the performance of multi-organization is changed accordingly.

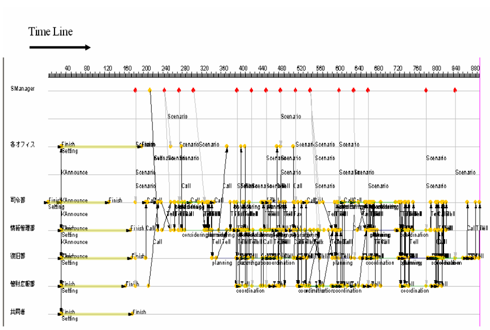


Fig. 4. Task and cognitive action flow

We analyze the outputs of simulation under different settings. Fig.5 and 6 shows the workload at headquarter, administration section, and information section under different simulation settings respectively. Here, workload refers to the number of active and queue actions per head in the group. We can notice the workload is changed accordingly as expected. By distributing the authority of decision-making, the workload on headquarter agent is decreased specifically. We also can see that workload on the advertising group is higher, due to it is required to deal with gather information, report to headquarter, and executing certain tasks.

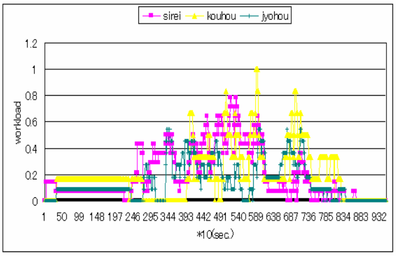


Fig. 5. Authority is centralized at headquarter section

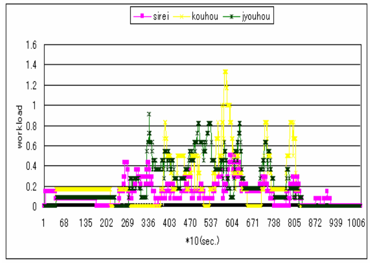


Fig. 6. Authority is decentralized in each section

By represent the real-life drill, we acquire insights for effective emergency response. For example, we can find the bottleneck of coordination in the drill scenario or identify the vulnerability of emergency response system. We also can see the workload at the high level of hierarchical organizations tended to be accumulated because of limitation of information process capacity. It indicates that emergency response system should be designed according to the capacity of processing information.

Through the simulation, we investigated how cognitive process and workload of each agent changes corresponding to the change of emergency response system. Through analyzing the simulation results, we can evaluate the efficiency of the emergency response system and optimize the structure of emergency response system, resource distribution and response procedures, manual.

5 Conclusions

In order to support and assess various emergency response systems, we have proposed rule-based multi-organization performance model, which focused on the cooperation and communication among organizations. Based on the proposed model, we have developed organizational activity simulator by taking earthquake disaster example. The test simulation was carried out and results were compared with real-life drills. Through the test, we confirmed that model reflects the multi-organization behavior very well and the computer based simulator has capability to represent the real-life emergency response drills and to evaluate emergency response system.

Nevertheless, since our model simplified the complex real environment, the model has limitation. The integrated simulation model requires that data transfer from one sub-model to another in the right format. Building the library of objects requires a large effect for simulator implement in the disaster context. This type or simulator, however, in the context of emergency management, is suitable for training emergency responder involved in the disaster by supplying “if-then” scenarios. Also, it can provide various criteria for evaluating performance of existent emergency response system and optimizing organization structure, response procedures or manuals.

Acknowledgments. Work described in this paper was sponsored by Japan Science and Technology Agency.

References

1. Avesani, P., Perini, A., Ricci, F.: Interactive case-based planning for forest fire management. Applied Intelligence 13, 41–57 (2000)
2. A-TEAM (2005) http://www.ess.co.at/A-TEAM/
3. Cohen, P.R., Greenberg, M.L., Hart, M., Howe, A.E.: Trial by fire: Understanding the design requirements for agent in complex environments. AI Magazine 10, 32–48 (1989)
4. Furuta, K., Shimada, T., Kondo, S.: Behavioral simulation of a nuclear power plant operator crew for human-machine system design. Nuc. Engne and Desg. 188, 97–109 (1999)
5. Jain, S., Mclean, C.: A framework for modeling and simulation for emergency response. In: Proc. Winter simulation conference 2003, pp.1068–1076 (2003)
6. Kanno, T., Furuta, K.: Multi-agent simulation of emergency response in nuclear disaster. In: Proc. Int'l Conf Probabilistic safety assessment and management (PSAM7), pp. 389–394 (2004)
7. Klein, G.: The Recognition-primed decision (RPD) model: looking back, looking forward, Naturalistic decision making, pp. 285–292. Erlbaum, Nahwah (1997)
8. LANL, 2003. Interdependent infrastructure modeling, simulation, and analysis project (SOFIA) http://www.lanl.gov/orgs/d/d4/infra/sofia.html
9. Milici, A., Wu, S.J., Apostolakis, G.: The use of the dynamic flowgraph methodology in modeling human performance and team effects. In: Proc. Int'l Conf Probabilistic safety assessment and management (PSAM3) pp. 1–6 (1996)
10. NSSL. National severe storms laboratory's warning decision support system (2003) http://www.nssl.noaa.gov/wrd/msd/wdss.html
11. Shu, Y., Furuta, K., Kondo, S.: Team performance modeling for HRA in dynamic situations. Reliab Engne Syst. Safety 2002 78, 111–121 (2002)
12. Shu, Y., Furuta, K.: An inference method of team situation awareness based on mutual awareness. Cogn, Tech. and Work 7, 272–287 (2005)

Estimating Mental Fatigue Based on Multichannel Linear Descriptors Analysis

Chong Zhang, Chongxun Zheng, Xiaomei Pei, and Xiaolin Yu

Key Laboratory of Biomedical Information Engineering of Education Ministry,
Xi'an Jiaotong University, 710049 Xi'an, China
xlyu.czhang@gmail.com, cxzheng@mail.xjtu.edu.cn, pei@zy165.com,
yu__xiao__lin@163.com

Abstract. In this paper, two different mental fatigue experiments, sleep deprivation and prolonged time of work, were designed to investigate their effect on mental fatigue respectively. Three parameters of multichannel linear descriptors i.e. Ω, Φ and Σ are used to measure the level of mental fatigue for the first time. When mental fatigue level increases, the brain neurons are restrained, this results in the increase of synchrony degree between the distributed electroencephalogram (EEG). Therefore, spatial complexity Ω, field strength Σ and field changes Φ of regional brain region would reduce. The experimental results also show that the average values of Ω, Φ and Σ of EEG decrease with the prolonging of sleep deprivation and working time, and then also decrease with the increase of mental fatigue level. The average values of Ω, Φ and Σ of EEG strongly correlate with mental fatigue. The multichannel linear descriptors of EEG are expected to serve as the indexes to evaluate mental fatigue level objectively.

Keywords: mental fatigue, Electroencephalogram (EEG), multichannel linear descriptors.

1 Introduction

Mental fatigue is a term to cover the deterioration of mental performance due to the preceding exercise of mental or physical activities [1]. D. van der Linden et al defined mental fatigue as a change in psychophysiological state due to sustained performance [2]. This change in psychophysiological state has subjective and objective manifestation, which include an increased resistance against further effort, an increased propensity towards less analytic information processing, and changes in mood. Sustained performance, in this definition, does not necessarily involve the same task but can extend over different tasks that require mental effort, such as fatigue because of a day in the office. Fatigue, especially mental fatigue, is inevitable for office workers and in life in general. Working on cognitive demanding tasks for a considerable time often leads to mental fatigue, which can impact behavior of mental task. Mental fatigue is usually related to a loss of efficiency and disinclination to effort. In industry, many incidents and accidents have been related to mental fatigue.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 521–529, 2007.

It is important to manage and cope with mental fatigue so that the workers do not damage their health cases. Therefore, the management of mental fatigue is important from the viewpoint of occupational risk management, productivity, and occupational health.

One of the interesting questions in mental fatigue research is in what way cognitive control of behavior changes under mental fatigue. Lorist et al used behavioural and EEG-data to study the effects of time-on-task (i.e., mental fatigue) on planning and task switching [3]. The EEG-data of their study showed that with increasing time-on-task there was a reduced involvement of those brain areas that are associated with the exertion of executive control (the frontal lobes). It is hypothesized that mental fatigue is indicated by the degradation of the ability to sustain an adequate performance in mental tasks, due to the exertion of mental and physical activities during preceding work times [1]. It was reported that subjects protected their performance by spending more effort in the unfavorable conditions: after several hours of work and after continuous work without short rest breaks [1]. The most important change in performance reported by many authors in relation to fatigue is the deterioration of the organization of behavior [4,5]. But the ways of evaluating mental fatigue now are based on a subjective sensation rather than on an objective assessment of changes in psychophysiological state, and it does not allow for a clear discrimination between different levels of mental fatigue. Moreover, there is a poor association between physiological and more subjective measures of fatigue. Recent years, a variety of psychophysiological parameters have been used as indicators of mental fatigue, with EEG perhaps being the most promising [6]. It can reflect the electrical activity of the central neural system (CNS). Therefore, monitoring EEG may be a promising variable for use in fatigue countermeasure devices.

2 Material and Method

This study is conducted to provide some additional insights into mental fatigue and its underlying processes, tries to use a new approach of multichannel linear descriptors to analyze mental fatigue induced by sleep deprivation and prolonged time of work, and aims to investigate the relationship between different mental fatigue state and multichannel linear descriptors of EEG. Five subjects performed two different experiments. First, three subjects were chosen for 24-hour sleep-deprived mental fatigue experiment respectively to explore the relationship between different mental state and the corresponding average values of Ω, Φ and Σ, which can also be used to test whether the circadian rhythm and sleep affect the average values of Ω, Φ and Σ. Then three subjects participated in two mental fatigue states experiment to explore how the average values of Ω, Φ and Σ in related the prolonged time of work.

2.1 Experiments and Data

Five subjects were chosen for mental fatigue experiments. They were both male and right-handed college students. The work time periods of the college are a.m. 9:00~12:00 and p.m. 1:00~5:00. The electrodes were placed at Fp1, Fp2 and Fz reference to the 10-20 system. The impedances of all electrodes were kept below

5KΩ. The reference electrode was pasted to the skin just above the tuber of the clavicle at the right bottom of the Adam's apple. The experimental environment kept quiet and the temperature was around 20 °C. Subjects were seated in a comfortable chair throughout the experiment. They were asked to simply relax and try to think of nothing in particular. The data were recorded using a PL-EEG Wavepoint system and were sampled at about 6 ms' interval. In order to investigate the relationship between three multichannel linear descriptors and mental fatigue, two mental fatigue experiments were designed as follows:

Experiment 1: Mental Fatigue Experiment under 24-hour Sleep-deprived States. This experiment was carried out by two subjects. The light was on during the experiment. Subjects were sleep deprived from a.m. 11:00 to next day for 24 hours. The first EEG data segment was recorded at a.m. 11:00 while subjects felt wide wake and alert which was denoted state 1-1. The second EEG data segment was recorded at a.m. 11:00 of the next day while subjects felt a little sleepy and fatigue which was denoted state 1-2. The recording duration of each state was about 5 minutes and the subjects' eyes were closed.

Experiment 2: Two Mental Fatigue States Experiment. To explore the relationship between the mental fatigue and three multichannel linear descriptors of EEG and to avoid the physiological accommodation of the circadian rhythm and the sleep influences on mental fatigue, two mental fatigue states experiment was designed. In this experiment, all three subjects were concentrating on reading for a whole afternoon and were required to take notes while reading. The first EEG data segment was recorded at p.m. 4:00 for five minutes while subjects felt a little tired which was denoted state 2-1. The second EEG data segment was recorded at p.m. 5:00 for five minutes while subjects felt very tired which was denoted state 2-2. The light was off and subjects' eyes were closed when EEG data were recorded.

2.2 Multichannel Linear Descriptors

Wackermann proposed a $\sum - \phi - \Omega$ system for describing the comprehensive global brain macrostate [7,8]. Let us consider N EEG samples in the observed time window at K electrodes to construct the voltage vectors $\{u_1, \cdots u_N\}$, where each $u_i (i = 1, \cdots N)$ corresponds to the state vector representing the spatial distribution of EEG voltage over the scalp at the ith sample. The data are assumed to have been already centered to zero mean and transformed to the average reference. Then Ω, Φ and Σ can be calculated as follows [7,8,9]:

$$m_0 = \frac{1}{N} \sum_i \|u_i\|^2 . \tag{1}$$

$$m_1 = \frac{1}{N} \sum_i \left\| \frac{\Delta u_i}{\Delta t} \right\|^2 , where\ \Delta u_i = u_i - u_{i-1} . \tag{2}$$

$$\Sigma = \sqrt{m_0 / K} \tag{3}$$

$$\phi = \frac{1}{2\pi}\sqrt{\frac{m_1}{m_0}} \tag{4}$$

The covariance matrix is constructed as:

$$C = \frac{1}{N}\sum_n u_n u_n^T \ . \tag{5}$$

The eigenvalues $\lambda_1 \cdots \lambda_K$ of matrix C is calculated, then Ω complexity can be obtained:

$$\log\Omega = -\sum_i \lambda_i' \log\lambda_i' \ . \tag{6}$$

where λ_i' is the normalized eigenvalue.

In the $\Sigma-\phi-\Omega$ system, K-dimensional voltage vectors constructed from the simultaneous EEG measurements over K electrodes with time varying are regarded as the trajectories in the K-dimensional state space. By the three linear descriptors, the physical properties of the EEG trajectory and then the brain macrostates are characterized. Φ reflects mean frequency of the corresponding field changes; Ω measures the spatial complexity of the regional brain region, which decomposes the multichannel EEG data into spatial principal components and then quantifies the degree of synchrony between the distributed EEG by the extension along the principal axes;. Larger value of Ω corresponds to the low synchrony. It can be seen that by EEG over the K electrodes the three linear descriptors describe the different brain macorstate features of the interested brain regions.

2.2 Analysis of the Mental Fatigue EEG

In order to get the data of spontaneous EEG signals, a FIR filter with bandpass 0.5-30 Hz was used. And the three-channel EEG data over Fp1, Fp2 and Fz were centered to zero mean value and transformed to the average reference. To eliminate the effect of EEG signal power on Ω complexity, the data were normalized by mean power of each channel before calculating Ω complexity. Then Ω, Φ and Σ were calculated separately for electrode arrays (Fp1, Fz) and (Fp2, Fz). Ω complexity quantifies the amount of spatial synchrony so that it reflects the degree of synchronization of two spatially distributed brain processes over the contralateral, ipsilateral and mid-central regions respectively; Σ reflects the corresponding regional field power; Φ characterize the speed of regional field changes of the contralateral, ipsilateral and mid-central regions respectively.

One minute EEG data of each trial in different mental fatigue state are selected to be analyzed. To obtain the time courses of Ω, Φ and Σ, the first 1000 sample segment is chosen as basic data segment and 17 samples (0.1s) as step length. By shifting the

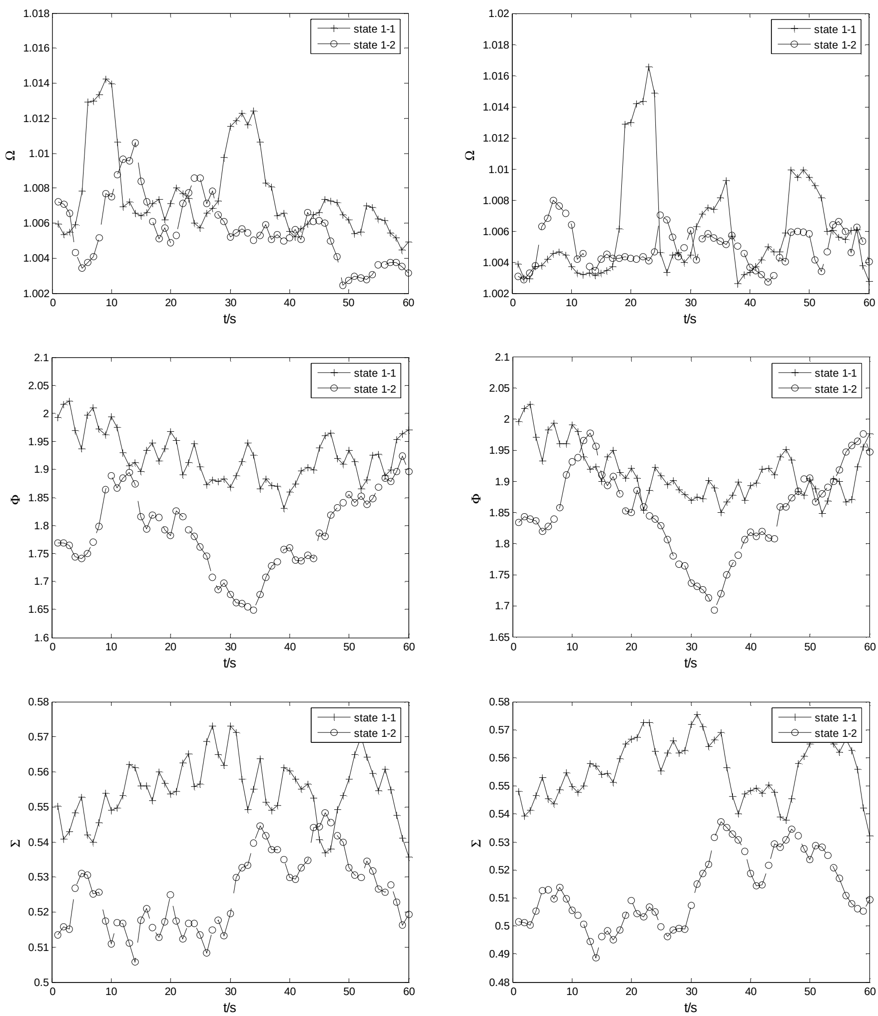


Fig. 1. The time courses of Ω, Φ and Σ between Fp1 and Fz, Fp2 and Fz under 24-hours sleep-deprived states (Experiment 1), the first EEG data segment was recorded at a.m. 11:00 while subjects felt wide wake and alert which was denoted state 1-1, and the second EEG data segment was recorded at a.m. 11:00 of the next day while subjects felt a little sleepy and fatigue which was denoted state 1-2.

basic data segment with step length from the start to the end of the trial, calculating Ω, Φ and Σ across all the trials for each segment. To eliminate the influences of Ω, Φ and Σ's fluctuation, the mean value within a second is calculated. So the time courses of Ω, Φ and Σ are obtained.

Alterations in the person's state of mind are associated with physiological alterations at the neuronal level and sometimes also at the regional brain level [10]. Functionally, central neural system circuitry is complex, consisting of a number of afferent connections, efferent connections and loop ('feedback') connections [11].

Two distinct loops, a motor loop and an association or complex loop, connect basal ganglia with neocortex [12]. The association or complex loop connected caudate with inputs from the cortical association areas and the final output from basal ganglia is projected to the prefrontal cortex. Recent findings have provided considerable evidence that cortex, basal ganglia and thalamus are all linked by the re-entrant circuits [13]. So prefrontal cortex has close connection with mental fatigue, and Ω, Φ and Σ calculated from EEG data of prefrontal cortex in different mental fatigue states can yield information which reveals short time changes due to mental fatigue.

3 Results

For each mental fatigue state of subject 1, the continuous mean values of Ω, Φ and Σ between Fp1 and Fz, Fp2 and Fz for one minute EEG data are calculated. The analysis results of 24-hour sleep-deprived states are given in figure 1. For subject 2, the time courses of Ω, Φ and Σ are similar to those of subject 1.

The independent-samples t test is used to analyze the mean values of Ω, Φ and Σ which achieved in state 1-1 and state 1-2. The results of statistical analysis suggest that the average values of Ω, Φ and Σ for every subject under two different sleep-deprived states have significant differences in general. The results of mean comparison after 24-hour sleep deprivation are shown in table 1.

Table 1. T test of mean comparison under 24-hour sleep-deprived states (mean $\pm$ s.d., $n = 60$)

			Ω	Φ	Σ
subject 1	(Fp1, Fz)	state 1-1	1.0077±.0026	1.9237±.0437	.5547±.0090
		state 1-2	1.0057±.0019**	1.7879±.0709**	.5257±.0109**
	(Fp2, Fz)	state 1-1	1.0061±.0034	1.9167±.0425	.5561±.0106
		state 1-2	1.0049±.0013*	1.8519±.0718**	.01274±.0016**
subject 2	(Fp1, Fz)	state 1-1	1.0319±.0401	1.7501±.0770	.5333±.0362
		state 1-2	1.0060±.0023**	1.4679±.2011**	.4732±.0135**
	(Fp2, Fz)	state 1-1	1.0342±.0503	1.7248±.0934	.5404±.0402
		state 1-2	1.0058±.0018**	1.4213±.2041**	.4764±.0182**

We calculated respectively how significantly the average of differences between the mean values of Ω, Φ and Σ for every subject under two different sleep-deprived states using the t-test. The significance measures are shown in boxes. *P<0.05, **P<0.01, significantly different from the value at state 1-1.

From figure 1 and table 1, it can also be seen that there is a large decrease in the average values of Ω, Φ and Σ under state 1-2 (a.m. 11:00—a.m. 11:05 in the first day) compared to that under state 1-1 (a.m. 11:00—a.m. 11:05 in the next day) not only between Fp1 and Fz but also between Fp2 and Fz. It may means that the influence of sleep loss on mental fatigue is very important and the average values of Ω, Φ and Σ decrease with the prolonging of sleep-deprived time.

For each different mental fatigue state of three subjects under experiment 2, the continuous mean values of Ω, Φ and Σ between Fp1 and Fz, Fp2 and Fz for one

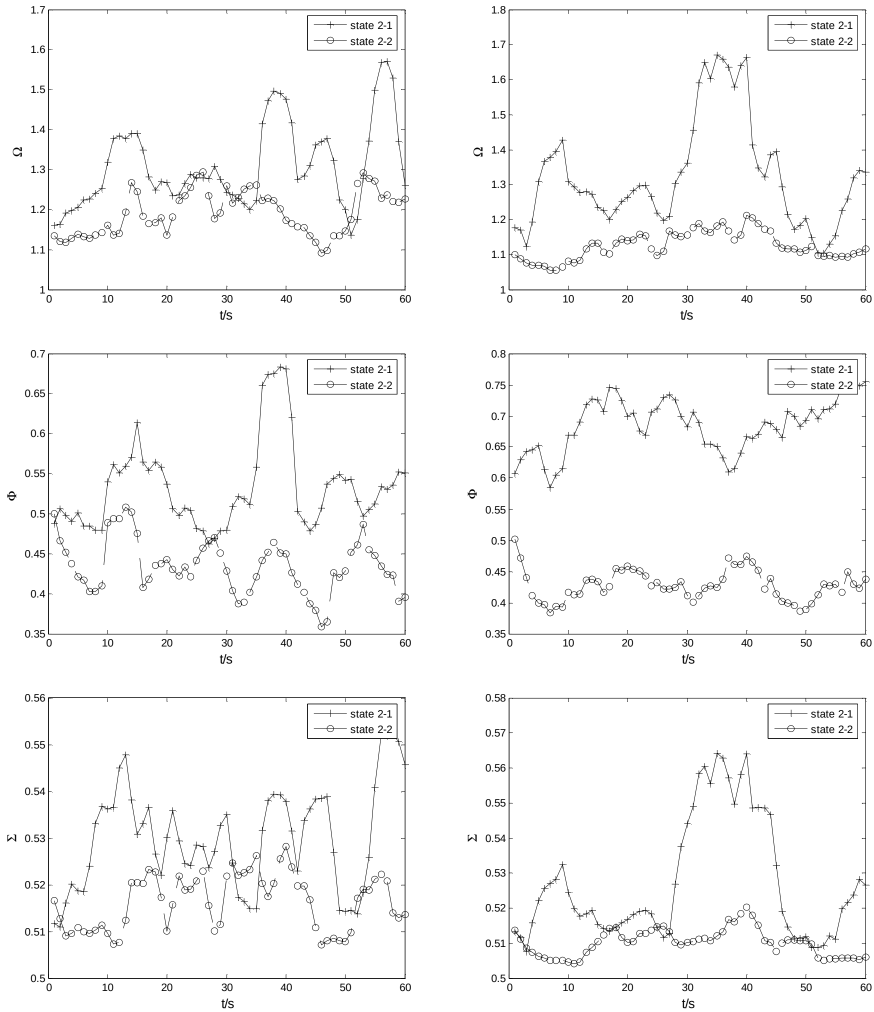


Fig. 2. The time courses of Ω, Φ and Σ between Fp1 and Fz, Fp2 and Fz under two mental fatigue experiment (Experiment 2), the first EEG data segment was recorded at p.m. 4:00 while subjects felt a little tired which was denoted as state 2-1, and the second EEG data segment was recorded at p.m. 5:00 while subjects felt very tired which was denoted as state 2-2.

minute EEG data are calculated. The analysis results are given in figure 2. The time courses of Ω, Φ and Σ for the other two subjects are similar to those of subject 3.

The independent-samples t test is used to analyze the mean values of Ω, Φ and Σ which achieved in state 2-1 and state 2-2. The results of statistical analysis suggest that Ω, Φ and Σ of every subject under two mental fatigue states have significant differences. The results of mean comparison under two mental fatigue states are shown in table 2.

Table 2. T test of mean comparison under two mental fatigue states（mean ± s.d., $n = 60$）

			Ω	Φ	Σ
subject 3	(Fp1, Fz)	state 2-1	1.3090±.1054	.5329±.0548	.5300±.0111
		state 2-2	1.1913±.0554**	.4345±.0342**	.5163±.0059**
	(Fp2, Fz)	state 2-1	1.3221±.1544	.6853±.0431	.5273±.0172
		state 2-2	1.1247±.0399**	.4279±.0248**	.5101±.0039**
subject 4	(Fp1, Fz)	state 2-1	1.2702±.0975	.8993±.0761	.5281±.0106
		state 2-2	1.1975±.1002**	.4307±.0471**	.5144±.0086**
	(Fp2, Fz)	state 2-1	1.2120±.0781	.9201±.0877	.5202±.0096
		state 2-2	1.1029±.0864**	.2915±.0425**	.5160±.0082*
subject 5	(Fp1, Fz)	state 2-1	1.2945±.1095	.9580±.3892	.5379±.0114
		state 2-2	1.2563±.0814*	.4181±.0353**	.5239±.0104**
	(Fp2, Fz)	state 2-1	1.2970±.0948	.8661±.3281	.5436±.0093
		state 2-2	1.1898±.0444**	.3594±.0232**	.5291±.0113**

We calculated respectively how significantly the average of differences between the mean values of Ω, Φ and Σ for every subject under two mental fatigue states using the t-test. The significance measures are shown in boxes. *P<0.05, **P<0.01, significantly different from the value at state 2-1.

From figure 2 and table 2, it can be seen that there are obvious differences for the values of Ω, Φ and Σ between state 2-1(p.m. 4:00—p.m. 4:05) and state 2-2(p.m. 5:00—p.m. 5:05).That is, the values of Ω, Φ and Σ and the mental fatigue state are correlative. Furthermore, the values of Ω, Φ and Σ under state 2-1(p.m. 4:00—p.m. 4:05) are larger than that under state 2-2(p.m. 5:00—p.m. 5:05). It indicates that the values of Ω, Φ and Σ increase with work time prolonging. The average values of Ω, Φ and Σ between Fp1 and Fz, Fp2 and Fz, strongly correlated with the level of mental fatigue.

4 Discussion

Ω complexity measures the spatial complexity of regional brain region and indicates the degree of synchronization between functional processes spatially distributed over different brain regions; Φ reflects mean frequency of the corresponding field changes; Σ describes the field strength of the regional brain region. When mental fatigue level increases, the brain neurons are restrained, this results in the increase of synchrony between the distributed EEG. Meanwhile, the frequency of the corresponding field changes and field strength of the regional brain region reduce.

In all current experiments, the average values of Ω, Φ and Σ between two channels Fp1 and Fz, Fp2 and Fz, are correlated with the mental fatigue state. The experiments show that the average values of Ω, Φ and Σ of EEG decrease with the prolonging of sleep-deprived time or work time. Moreover, the prolonged time of work and sleep deprivation have a greater effect on mental fatigue, and they induce the state of mental fatigue. So the average values of Ω, Φ and Σ between two channels FP1 and Fz, Fp2 and Fz, increase with the intensifying of mental fatigue level in two mental fatigue experiments. This fact may suggest that there are stronger independent, parallel, functional processes active between the two regions, which is just strongly correlative with the level of mental fatigue. This result is also consistent with the findings of Szelenberger et al (Szelenberger, 1996) which found highly significant overall effect, with Ω complexity progressively decreasing in the sequence of sleep stages [8,14].

In sum, mental fatigue is likely to be an integrated phenomenon with complex interaction among central and peripheral factors, physiological and psychological factors, and so on. All factors appear to mutually influence each other. In two mental fatigue experiments it is proved that the multichannel linear descriptors of EEG are strongly correlative with the mental fatigue state. This method may be useful in further research and is able to evaluate mental fatigue level objectively.

Acknowledgments. This work is supported by National Nature foundation of China under the grant No.30670534. The authors would like to thank Xiaoping Li of University of Singapore for providing the data.

References

1. Meijman, T.F.: Mental fatigue and the efficiency of information processing in relation to work times. International Journal of Industrial Ergonomics 20, 31–38 (1997)
2. Van der Linden, D., Frese, M., Meijman, T.F.: Mental fatigue and the control of cognitive processes: effects on perseveration and planning. Acta Psychologica 113, 45–65 (2003)
3. Lorist, M.M., Klein, M., Nieuwenhuis, S., et al.: Mental fatigue and task control: Planning and preparation. Psychophysiology 37(5), 614–625 (2000)
4. Broadbent, D.E.: Is a fatigue test possible now? Ergonomics 22(12), 1277–1290 (1979)
5. Holding, D.H.: Fatigue. In: Hockey, G.R.J. (ed.) Stress and fatigue in human performance, Wiley, New York (1983)
6. Lal, S.K.L., Craig, A.: A critical review of the psychophysiology of driver fatigue. Biological Psychology 55, 173–194 (2001)
7. Wackermann, J.: Beyond mapping: estimating complexity of multichannel EEG recordings. Acta Neurobiol. Exp. 56, 197–208 (1996)
8. Wackermann, J.: Towards a quantitative characterization of functional states of the brain: from the non-linear methodology to the global linear description. Int. J. Psychophysiol 34, 65–80 (1999)
9. Pei, X.M., Zheng, C.X., Zhang, A.H., et al.: Discussion on Towards a quantitative characterization of functional states of the brain: from the non-linear methodology to the global linear description by J. Wackermann. International Journal of Psychophysiology 56, 201–207 (2005)
10. Delgado, P.: Depression: the case for a monoamine deficiency. J. Clin. Psychiatr. 61, 7–11 (2000)
11. Abhijit, C., Peter, O.B.: Fatigue and basal ganglia. Journal of the Neurological Sciences 179, 34–42 (2000)
12. Delong, M., Georgopoulos, A.P.: Motor functions of the basal ganglia. In: Brookhart, J.M., Mountcastle, V.B., Brooks, V.B. (eds.): Handbook of physiology, Bethesda: American Physiological Society, pp. 1017–1061 (1981)
13. Alexander, G.E., Crutcher, M.D.: Functional architechture of basal ganglia circuits: neural substrates of parallel processing. Trends Neurosci. 13, 66–271 (1990)
14. Szelenberger, W., Wackermann, J., Skalski, M., et al.: Analysis of complexity of EEG during sleep. Acta Neurobiol. Exp. 56, 165–169 (1996)

Chinese Visible Human Data Sets and Their Applications

Shao-Xiang Zhang[1], Pheng Ann Heng[2], Zheng-Jin Liu[1], Li-Wen Tan[1], Ming-Guo Qiu[1], Qi-Yu Li[1], Rong-Xia Liao[1], Kai Li[1], Gao-Yu Cui[1], Yan-Li Guo[1], and Yong-Ming Xie[2]

[1] Department of Anatomy, College of Medicine, Third Military Medical University, Chongqing 400038, China
[2] Deprtment of Computer Science and Engineering, Chinese University of Hong Kong, China
zhangsx@mail.tmmu.com.cn

Abstract. The research on digital visible human is of great significance and application value. The US VHP created the first digital image data set of complete human (male and female) in 1995. To promote worldwide application-oriented VHR, more visible human data sets representative of different populations of the world are in demand. The Chinese Visible Human (CVH) male (created in Oct. 2002) and female (created in Feb. 2003) project achieved greater integrity of images, easier blood vessel identification, and were free of organic lesion. The most noteworthy technical advance of CVH Project was construction of a low temperature laboratory, which contributed to prevention of small structures (including teeth, concha nasalis, and articular cartilage) from falling off of the milling surface. Thus, better integrity of images was ensured. So far, we have achieved acquisition of five CVH data sets and their volume visualization on PC. 3D reconstruction of some organs or structures has been finished. The work of segmentation of a complete data set is also under way. Although there is still a distance to go to make visible human meet the application-oriented needs in various fields, we're taking our first step toward future application by acquiring new data sets, performing segmentation and setting up a platform of computer-assisted medicine.

Keywords: Chinese Visible Human (CVH), dataset, 3-D reconstruction, digital anatomy, computer assisted medicine (CAM).

1 Introduction

Digital visible human has become a hot topic in anatomy and computer science as well. Due to demand for digital visible human in various fields related to human morphology, including medicine, sports, bionics, advertisement making, and the aerospace industry, Visible Human Research (VHR) is undoubtedly promising and far-reaching in the digital era. We are just at the beginning, however, to visualize the

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 530–535, 2007.

human body because of its intricacy. Therefore, visible human rendering is at the center of attention of scientists because of its great significance and valuable applications. Take medical education for example. With the Visible Human database, computer simulation of the live human body will provide evolutionary transformation in anatomical education. Digital bodies constitute one necessary medium that can be used to train future generations of students in anatomical practice and computer-assisted surgery [5]. The visible human may also provide the basis for modern medical image interpretation and diagnosis (i.e., beta ultrasound, computed topography [CT], magnetic resonance imaging [MRI]). The possibilities for future application of visible human data sets seem endless, e.g., radiotherapy, interventional diagnosis and interventional therapy, and development of microsurgical techniques.

The creation of the US Visible Human Male and Female marked the start of VHR [1, 4, 5]. To promote worldwide VHR with the purpose of instituting a worldwide application-oriented biomedical images library, more data sets representative of different populations of the world are in demand.

Encouraged by the theoretical importance and vast vistas of applications of VHR, many Chinese scholars represented all medical specialties and relevant areas of computer science have been paying great attention to the research progress and working fruitfully by using the data set from US VHP. They think that the China should contribute to the worldwide VHR through developing a normal human dataset representing an Asian population. The CVH Project (CVHP) calls for collaboration between human anatomists, experts in computer graphics and imaging and experts in medicine.

Our research team at the Third Military Medical University has been involved in study on Human Sectional Anatomy and computerized 3D reconstruction since 1985. We have been focusing on Human Thin Sectional Anatomy since 1990. With grants from the National Science Foundation of China for Distinguished Young Scholars, we began construction of our special low temperature lab and technique research on acquiring the CVH data set in 1999. Data acquisition of the first CVH Male data set was finished at the Third Military Medical University, and 3D visualization of it was achieved at the Chinese University of Hong Kong in Oct. 2002 [6]. The International Workshop on Visible Human Research was convened at the Third Military Medical University in January 2003 for an idea exchange on VHR. The data set acquisition for the CVH Female was completed in February 2003, marking China’s status as the second country in the world to develop a visible human male and female data set [3, 10]. So far, five CVH data sets have been acquired by our team.

2 Acquisition of the Data Set

2.1 Specimen Preparation

Cadavers of each sex were donated by the citizens of Chongqing, China. A key aspect of the CVH project was to use cadavers that were from relatively young adults (20-40 years), and of typical height (1600-1900 mm) and weight (e.g. no evidence of obesity or emaciation). Those meeting these criteria were screened macroscopically to exclude those with any superficial evidence of organic pathology. The remaining cadavers then underwent preliminary CT and MRI examinations to exclude those with internal lesions or pathology. The final cadavers were then transported to the Third Military Medical University Imaging Center to capture CT and MR images, which were kept to compare with later anatomical images, while those cadavers with internal lesions or pathology were further excluded.

2.2 Vascular Perfusion

To facilitate blood vessel identification, we performed vascular perfusion before specimen embedding. The CVH male: femoral artery was perfused with a 20% gelatin solution, which was colored red with food dye. 1400ml of the gelatin solution was perfused into the arterial system. The femoral cut was then sutured layer by layer. The CVH female: the femoral artery was perfused with 20% red gelatin solution. 1150ml of the gelatin solution was perfused into the arterial system.

2.3 Sectioning of the Cadaver

To keep the ice block hard enough to keep the cutting surface slick and to avoid the ejection of small segments of tissue from the block, we constructed a low temperature laboratory (5.0m long ×5.0m wide × 2.2m high) that could be maintained at or below –25°C. The milling machine was placed in the laboratory, while its electronic control system was kept outside. For sectioning, the ice block was transferred to the aboratory and mounted to the cryomacrotome table of the milling machine.

Sectioning of the intact block was done using an improved TK-6350 numerical control milling machine with a milling accuracy of 0.001mm (The numerical control system was made in Japan and the mill made in France; engineers from our team and the Hanzhong Machine Tool Factory, China, designed and implemented necessary modifications).

Slices of each body were then milled layer by layer, from head to toes, at –25°C in the low temperature aboratory. The serial cross-sections were photographed with a Canon high-resolution digital camera and scanned into an animation computer.

2.4 The Data Sets Acquired

There are 2518 anatomical cross-sectional images in total for the first CVH male. The data set also includes CT, MR, and radiographic images. Axial and coronal MR images of the head and neck and axial sections through the rest of the body were obtained at 3.0-mm intervals and in matrices of 512 ×512 pixels (256 grey levels, approximate pixel size 440μm). CT data consisted of axial scans through the body at 1.0 mm intervals (0.1mm at the skull base). CT images are 512 × 512 pixels, where each 256 bits pixel value is related to electron density of the specimen at that point. The MR and CT axial images have been aligned with the anatomical cross-sections. All digital images of milled surfaces had 6,291,456 (3072×2048) pixels. The data file of each section occupies 36 MB. The complete data files occupy 90.65 GB.

3640 anatomical cross-sectional images were acquired for the first CVH female. The visible female data set has the same characteristics as the male with the following exceptions: The serial sections were sampled at 0.25- mm intervals for head and 0.50-mm intervals for other regions. CT data consisted of axial scans through the entire body at 1.0-mm intervals. The data file of each section occupies 36 MB. The complete data files occupy 131.04 MB.

The man whose body was used as the third CVH (male) was 21 years old at the time of death. He was 1820 mm tall and weighed 66kg. CT data consisted of axial scans through the head at1.0 mm intervals, and that through the rest of the body at 2.0 mm intervals. Axial and coronal MR images of the head and neck were obtained at1.5 mm intervals, and axial sections through the rest of the body were at 3.0 mm intervals. The total 18398 anatomical cross-sectional images for the third CVH were obtained at 0.1 mm intervals. Acquisition of all slices had been sampled at a resolution of 10,989,056（4064×2704）pixels. The data file of each section occupies 62.9 MB. The complete data files occupy 1157.23 GB. This data set was finished in April 2003.

The fourth and the fifth CVH data sets were acquired in July and October 2003, respectively. The women whose body were used as the fourth and the fifth CVH were both 25 years old at the time of death.

The woman for the fourth CVH data set was 1620 mm tall and weighed 57.5kg.Acquisition of all slices of both data sets had been sampled at a resolution of 10,989,056（4064×2704）pixels, and the data file of each section occupies 62.9 MB. The serial sections were sampled at0.25- mm intervals for head, 0.50-mm intervals for trunk, 1.0- mm intervals for other regions. 3020 cross-sectional images were acquired in total. The complete data files occupy 189.9GB.

The woman for the fifth CVH data set was 1700mm tall and weighed 59kg.It had been sampled at the same resolution with the fourth CVH. But the serial sections were sampled at0.2- mm intervals through the body. 8510 cross-sectional images were obtained in total. The complete data files occupy 535GB.

3 Applications of the Data Sets

3.1 Virtual Anatomy

To date, five gigantic digital human data sets of Chinese origin have been collected by our CVH group. The highest resolution CVH dataset is about 1143 GB in size, it consists of 18,200 cross-section color digital images (4064×2704 resolution, 24 bits per pixel) at 0.1 mm intervals. We have processed these digital human datasets and completed the 3D reconstruction of five Chinese Visible Human. Based on our developed imaging and rendering algorithms, real-time stereoscopic volume visualization of multi-gigabyte digital human datasets becomes possible on PC platform. At least 10 –time speed-up of rendering performance has been achieved as compared to previous standard volume rendering methods. A powerful virtual anatomy software package has been developed to interactively dissect CVH data. It provides users with highly flexible navigation controls to examine the human anatomical structures with extremely fine details and photo-realistic rendering.

The 3-D reconstruction of CVH (by volume rendering reconstruction) slices can be stereoscopically viewed on a tiltable projection table. In our visualization toolkit, we have integrated 3-D input devices in order to allow user to choose different cross-sectional and volumetric views interactively. By pressing a set of 3-D menu buttons using a stylus-like input device (virtual scalpel), orthogonal-triplanar or monoplanar cross-sectional views can be switched interactively. These tools improve the navigation and study of different anatomical structures. Though we can only display 3-D cross-sectional images, we're positive that it is the rudiment of real time digital anatomy.

3.2 Virtual Ultrasound

Ultrasound images are cross-sectional images of the body detected by a probe. Being held against the skin and moved over its surface, the probe detects echo signal and sends it down to the ultrasound machine. Thereby, images of the internal organs and structures are real-time displayed on the monitor. Based on 3D reconstructed CVH model set up by volume rendering reconstruction, real-time display of human cross-sectional images was performed in any direction. When we move the mouse of the computer, just like moving the probe of an ultrasound machine, corresponding cross-sectional image of the body can be real-time displayed. Suppose we connect the probe to our human visualization system, as well as to the ultrasound machine, the ultrasonographer will refer to image from normal human body when observing images of the patient. That is to say, during the examination, the video monitor displays cross-sectional images of internal organs and structures of the patient. At the same time, the other computer displays cross-sectional image of Visible Human model corresponding to ultrasound cross-section. As the images of Visible Human model are clearer and more meticulous, and come from normal human body, this system will provide the ultrasonographer with normal handy control so as to contribute to correct Ultrasound diagnosis.

3.3 Platform of Computer Assisted Medicine (CAM)

VHR is of great application value in various fields related to human morphology, including medicine, sports, automobile and mechanism production, even aerospace industry and military affairs. We will focus on setting up the platform of Computer Assisted Medicine. That is, based on the acquired datasets of human anatomical structure, image processing, segmentation, classification, registration and reconstruction by rendering are preformed to achieve 3D reconstruction of each system, organ and region. Next, the complex structures of human body are precisely demonstrated on computer by image synthesis. Then, digital model for seperating organs is built. And then, a computer-based basic platform system is planned to be set up, which has image dataset of normal human anatomical structures and can be further developed to serve clinical diagnosis and treatment including clinical image diagnosis, interventional diagnosis, surgical simulation, surgery navigation, radiotherapy localization.

By connecting with Ultrasound machine, CT machine, MR machine and radiotherapy localization machine, this basic platform system can real-time display the image of the normal body corresponding to the image being displayed by the diagnosis machine so as to improve accuracy of clinical diagnosis and treatment. This can be better manifested in the following aspects:

1. Regarding sectional image diagnosis such as Ultrasound, CT, MRI and PET, corresponding sectional image of the same location and the same plane can be

extracted from the basic platform system and displayed simultaneously with the sectional image of the patient in examination.As the images displayed by the basic platform come from the photos of milled human specimen taken with digital camera, they are far clearer than that of CT, MRI and PET. Thus provide the doctors with realistic normal control and facilitate correct judgment on the image of the patient.
2. With interventional duct, corresponding human ductwork can be extracted from the basic platform system and simulation of the pathway of the ducts will be achieved, which will help the surgeons become well-informed of the corresponding anatomical structures and foresee possible problems in surgery.
3. With regard to endoscopy, corresponding 3D images of human ducts (e.g. trachea, bronchia, alimentary tract, etc.) and lacuna (e.g. subarachnoid space, joint cavity, peritoneal cavity, etc.) can be extracted from the basic platform system so as to simulate process of endoscopy.
4. As for surgery, 3D images data of surgical locus can be extracted from the basic platform system so as to plan for and simulate the process of surgery.
5. And for tumor radiotherapy localization, data of 3D structures of radiotherapy locus can be extracted from the basic platform system. So irradiation position can be simulated. Together with refering to biological irradiation doses for specific tumors, optimum irradiation position can be determined, and proper irradiation dose can also be worked out.

Acknowledgments. The work described in this paper was supported by National Natural Science Fund of China (NSFC) for Distinguished Young Scholars (Grant number: 39925022), National Natural Science Fund of China (NSFC) (Grant number: 60473128), and Research Grants Council of the Hong Kong Special Administrative Region of China (Grant number: CUHK 4223/04E).

References

1. Ackerman, M.J.: The visible human project: a resource for education J. J. Acad. Med. 74, 667–670 (1999)
2. Chung, M.S., Kim, S.Y.: Three-dimensional image and virtual dissection program of the brain made of Korean cadaver. Yonsei Med. J 41, 299–303 (2000)
3. Chung, M.S., Park, H.S.: Another trial for making serially sectioned images (Visible Korean Human). In: International Workshop on Visible Human, pp. 2–9, Chongqing, China (2003)
4. Spitzer, V.M., Ackerman, M.J., Scherzinger, A.L., Whitlock, D.: The visible human male: a technical report. J. Am. Med. Inform. Assoc. 3, 118–130 (1996)
5. Spitzer, V.M., Whitlock, D.G.: The Visible Human dataset: The anatomical platform for human simulation. Anat. Rec. 253, 49–57 (1998)
6. Zhang, S.X., Liu, Z.J., Tan, L.W., et al.: Number one of Chinese digitized visible human completed. Acta Academiae Medicinae Militaris Tertiae 24, 1231–1232 (2002)
7. Zhang, S.X., Heng, P.A., Liu, Z.J., et al.: The Chinese Visible Human (CVH) Datasets incorporate technical and imaging advances on earlier digital humans. J Anat 204, 165–173 (2004)
8. Zhang, S.X., Heng, P.A., Liu, Z.J., et al.: Creation of the Chinese Visible Human Data Set. Anatomical Record (Part B: New Anat.) 275B, 190–195 (2003)

Visualizing User Experience Through "Perceptual Maps": Concurrent Assessment of Perceived Usability and Subjective Appearance in Car Infotainment Systems

Xianjun Sam Zheng, James J.W. Lin, Salome Zapf, and Claus Knapheide

Siemens Corporate Research
Princeton, NJ 08540, USA
{sam.zheng,jameslin,salome.zapf.ext,aus.knapheide}@siemens.com

Abstract. Users' perceptions of the appearance and the usability of an interactive system are two integral parts that contribute to the users' experience of the system. "Actual usability" represents a system value that is revealed either during usability testing and related methods by experts or during use by the target users. Perceived usability is an assumption about a systems' usability that has been made prior to, or independent of, its use. The appearance of a product can inadvertently affect its perceived usability; however, their relationship has not been systematically explored. We describe an approach that uses "perceptual maps" to visualize the relationship between perceived usability and subjective appearance. A group of professional designers rated representative car infotainment systems for their subjective appearance; a group of usability experts rated the same models for their perceived usability. We applied multidimensional scaling (MDS) to project the ratings into the same Euclidean space. The results show certain overlap between the perceptions of product appearance and usability. The implications of this approach for designing interactive systems are discussed.

Keywords: User experience, usability, perceived usability, appearance, design, interactive systems, multidimensional scaling, visualization.

1 Introduction

Users' perceptions of an interactive system and its usability are two integral parts that influence the users' purchasing decision, and they also contribute to the emerging user experience of the system. The general appearance of a product can inadvertently affect its perceived usability. For instance, a clean or neat design may lead to the assumption that the system is also easy to use; whereas a cluttered or busy design may lead to the impression of poor usability. The relationship between subjective appearance and perceived usability, however, has not been systematically explored.

Both subjective appearance and perceived usability are the results of complex cognitive processes, involving multiple assessment criteria. The assessment of appearance may involve asking whether a design looks professional or casual, exciting or boring; the perception of usability may use criteria like whether the

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 536–545, 2007.

interface looks self-explanatory or provides easy navigation. It is important to understand which appearance attributes are more tightly linked to the perception of usability, for two reasons. On one hand, the findings can help understand how these two integral parts interact to shape the users' experience of a product, as both are happening prior to and determining a purchase decision or pre-shape a user's attitude towards the system. On the other hand, the understanding can also guide a designer to use elements of overall appearance to generate a target perceived usability.

In this paper, we use car infotainment systems as a case study and describe an approach that uses "perceptual maps" to visualize the relationship between perceived usability and subjective appearance. We also discuss design implications.

1.1 Car Infotainment System Design

In recent years, the market for car infotainment systems has been booming, projected to increase its worth to $56 billion globally in 2011 from $27 billion in 2004 (SA, 2005). Consumers demand multifunctional systems including navigation, communication, information, and entertainment. For car radios alone, drivers desire to connect their iPods or other mp3 players, navigating and playing their favorite music while driving. The soaring need is partially driven by the trend that people want to extend their life style from their homes and offices into the mobile car environment, maintaining a consistent user experience as they know it from the usage of iPods and other devices.

In responding to this huge market demand, many companies, including car manufactures, OEMs, and aftermarket product vendors, are rushing to put out their latest designs and products. Established standards, i.e., the well-known row of 'radio buttons' at the lower margin of the car radio, are being given up by most vendors, and a wide variety of designs can be seen in the market place: touch screens, hard keys, soft keys, small or large displays, and designs with many or very few hard buttons.

Successful HMI design not only needs to incorporate the ever-growing wealth of information and functionality, but also should take into account the requirements of driving safety and human factors. The driving environment is dynamic and can become risky or even dangerous, and the HMI design of a car information system should minimize the interference effect with the primary task: driving. A design that looks attractive and is also perceived as easy-to-use might be key for a positive user acceptance of the system. Note that the actual usability of the product is critical because of the safety concern as the system is being used; perceived usability, however, will influence users whether to purchase the product or not. Ideally a truly safe and usable product will be perceived as such and be attractive at the same time and for that same reason. We believe that the reconciliation of the need to consider both: attractive appearance and perceived usability should be based on a deep understanding of the relationship between them. In the following, we review the methods that allow studying the users' perception of design.

1.2 Methods to Study Perceptions of Design

Several methods have been used to study the users' perceptions of appearance, including constructing perceptual maps (e.g., Chuang et al., 2001, Petiot & Yannou, 2004, Hsiao & Wang, 1998) and Kansei Engineering (Nagamachi, 1995). The method of perceptual mapping is the focus of our paper, and the basic idea is to build a multi-attribute perceptual space in which each product is represented by a point.

This approach typically involves several steps. First is to generate the evaluation attributes of the products (e.g Osgood et al., 1975). The attributes of subjective appearance can be solicited by asking people to vocalize their feelings or impressions towards a certain family of actual products or the images of these products. The final list of attributes is often composed of multiple pairs of antonymous adjectives. Then the next step is to collect people's subjective evaluation of the products based on the defined attributes. Lastly, multidimensional scaling (MDS) is applied to translate the evaluation data of products into a geometrical representation of a perceptual space. Principle component analysis or factor analysis are used to reduce the dimensionality of the space, typically yielding a low dimensional visualization that can be interpreted by researchers.

In the past, the approach of constructing perceptual maps has been mainly used to study the users' subjective appearance of design. Here, we used the car infotainment system as an example to apply the similar approach to study the users' perceptions of appearance and usability as well as their relationship. The method session below describes the approach in more details.

2 Method

2.1 Selecting Car Infotainment Systems

Twenty representative models of car infotainment system were initially chosen from the Consumer Electronic Show (2007), the Convergence Conference (2006), local stores (e.g., Best Buy, Circuit City), or the Internet (vendors' websites or product catalogs). The good quality, color image of each model was printed out on 1: 1 ratio according to its physical size.

A pilot study was conducted, which involved seven professional user interface (UI) designers to evaluate the degree of dissimilarity between each pair of the car infotainment systems on a scale of 1-5 (1: minimum dissimilarity; 5: maximum dissimilarity). We ran the MDS on the averaged dissimilarity matrix, and a Euclidean perceptual space was generated with each model represented as a point and the distance between any two objects depicting the level of their perceptual dissimilarity.

The goal was to find the most representative models of car infotainment system. Based on the results of MDS, 15 car infotainment systems with the least perceptual similarities were selected, and their images are shown below (Figure 1), arranged in an alphabetic order by their labels.

Subjective appearance attributes. In the pilot study, we also asked the user interface designers to vocalize their first impressions or feelings of each car infotainment system. A list of the adjectives that they used to describe the design appearance was recorded. These adjectives were then classified and clustered: the words with similar meanings were combined, and those with opposite meanings were paired. In the end, 14 pairs of antonymous words mentioned most frequently were selected as the attributes of subjective appearance (see Table 1).

Fig. 1. Fifteen most representative models of car infotainment system

Usability attributes. The usability attributes were derived from driving human factors research (e.g., Wickens, et al, 2003) as well as the established human factors design guidelines for car infotainment system (e.g., Stevens et al., 2002). The major considerations are that the driving environment is dynamic and risky, and the HMI design of the car infotainment system should minimize the interference effects with the primary driving task. Because maneuvering a car drivers heavily rely on visual monitoring, a good HMI design should support non-visual input (i.e., blind controllability) and the task operations should allow interruptions (i.e., interruptability). Some other GUI usability evaluation criteria (e.g., Nielsen's usability heuristics (1994)), if applicable to the HMI design of the car infotainment system, were also incorporated into the list.

Note that because we only consider *perceived* usability, we did not include those usability attributes that seem to depend more on interactions with the systems, e.g., feedback, error tolerance etc. Four human factors researchers were engaged in selecting and discussing the attributes of usability; and a list of seven attributes was finalized (see Table 2).

Table 1. Fourteen pairs of antonymous adjectives of subjective appearance attribute for car infotainment systems

Sample scale:	*active*	*1 2 3 4 5 6 7*	*passive*
aggressive – calm	masculine – feminine	organized – cluttered	innovative – conventional
dynamic – static	expensive – cheap	professional – casual	stylish – inelegant
exciting - boring	robust – fragile	square – organic	

Table 2. Seven perceived usability attributes for car infotainment systems

Scale:	*Very poorly supported 1 2 3 4 5 6 7 Very well supported*
Blind controllability	The ability to perform certain actions that require no or minimum visual input.
Limited interruptibility	The level of interference with the primary driving task when operating the system.
Minimize cognitive load	Systems should minimize the cognitive load, and the design should consider human cognitive limitations (e.g., showing a limited number items per page).
Consistency	The HMI design should be consistent in all aspects.
Ease of navigation	Systems should be easy to navigate. Commonly used functions are easily accessed.
Self-explanatory	The HMI is intuitive, clearly indicating current status and providing sufficient cues for performing actions.
Ease of Use	The system is easy to use.

2.2 Collecting Evaluation Data

The method that we used to collect the evaluation data of subjective appearance and perceived usability is similar to the semantic differential method (Osgood et al., 1957; Chuang et al., 2001).

Subjective appearance. Four experienced professional UI designers (2 females and 2 males) were asked to rate the 15 car infotainment systems for their subjective appearance (Figure 1) based on each of the attributes defined above (Table 1). The rating scale is 1-7, with 1 indicating the strongest feeling towards an adjective, and 7 towards its antonym.

Perceived usability. Four experienced usability experts (1 female and 3 males) rated the same models for their perceived usability based on each of the perceived usability attributes defined above (Table 2). The rating scale is also 1-7, with 1 meaning very poorly supported, and 7 very well supported.

2.3 Running MDS Analysis and Visualizing the Perceptual Maps

After collecting both sets of the rating data, the average matrixes of both perceptions of appearance and usability were calculated (Table 3 & 4). Multidimensional Analysis of Preference technique (MDPREF) was applied to translate the average evaluation data into a geometrical representation of a perceptual space. Principle component analysis was used to reduce the dimensionality of the space, yielding interpretable low dimensional visualizations. (More detailed information for this method can be found at http://marketing.byu.edu/htmlpages/books/pcmds/mdpref.html).

Table 3. The average matrix of perceptions of 15 car infotainment systems based on 14 subjective appearance attributes. Values range from 1 (left attribute) to 7 (right attribute).

Product Code Appearance	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
aggressive – calm	3.0	2.5	2.3	5.8	4.8	3.0	4.8	4.0	2.8	3.5	3.5	2.3	4.0	2.0	4.0
dynamic – static	2.5	3.0	4.0	5.8	3.5	3.0	4.3	4.3	2.3	5.3	4.3	2.8	3.3	4.0	3.5
exciting - boring	3.3	2.0	5.3	6.8	2.5	1.8	3.5	4.0	2.8	4.8	4.5	2.5	3.5	3.8	3.3
expensive – cheap	4.0	3.0	5.3	6.5	2.0	2.0	2.8	1.8	2.8	2.3	3.0	1.8	2.3	2.5	2.3
innovative – conventional	4.5	2.3	5.8	6.5	2.5	1.5	3.3	2.8	3.5	3.5	3.5	2.0	3.3	1.8	3.5
masculine – feminine	4.5	1.8	4.0	4.3	4.5	3.5	3.0	3.0	3.3	2.0	3.0	1.8	3.8	2.3	3.3
organized – cluttered	5.3	2.5	5.5	3.5	3.3	3.3	2.3	2.5	2.5	2.3	3.5	6.0	3.0	4.3	2.5
professional – casual	5.5	3.0	5.0	3.5	2.8	2.0	2.3	3.3	4.0	2.0	3.5	5.5	2.5	2.3	3.0
robust – fragile	4.5	1.5	4.3	3.5	3.3	2.3	3.5	3.3	3.0	2.3	2.3	3.3	3.8	2.8	2.8
square – organic	5.8	5.0	2.5	2.5	4.3	5.3	4.3	3.0	5.5	1.3	1.8	3.5	2.3	1.8	2.8
stylish – inelegant	3.5	2.0	5.8	6.8	2.0	1.3	2.5	2.0	2.8	5.8	5.0	3.8	4.3	4.8	4.3

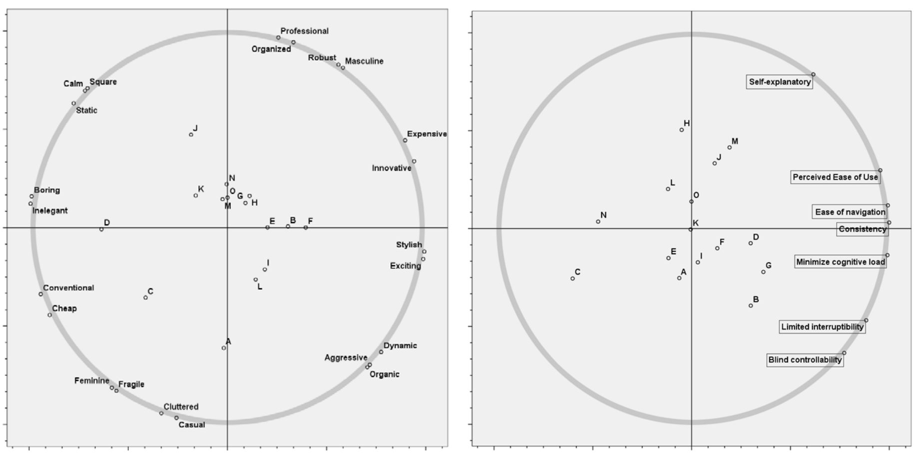
Table 4. The average matrix of perceptions of 15 car infotainment systems based on 7 usability attributes. Values range from 1 (very poorly supported) to 7 (very well supported).

Product Code Usability	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Blind controllability	3.5	5.0	2.0	4.0	3.0	3.3	4.0	1.8	3.3	2.5	3.0	2.3	3.0	2.0	3.0
Limited interruptibility	3.5	4.8	2.3	4.8	3.3	3.5	4.5	2.5	3.5	3.5	4.0	3.0	3.0	2.0	3.0
Minimize cognitive load	3.5	4.3	1.8	4.0	3.3	5.0	4.5	3.8	4.0	3.0	3.0	3.0	3.8	1.5	3.5
Consistency	4.3	5.0	3.0	4.8	4.0	4.0	5.8	4.5	5.0	5.0	4.3	4.5	4.5	3.0	4.3
Ease of navigation	3.3	4.5	2.3	4.8	3.5	4.3	5.3	3.8	3.8	4.5	3.8	3.5	4.5	3.5	3.8
Self-explanatory	3.0	3.8	1.8	4.5	3.0	3.5	3.5	5.3	3.3	4.8	4.0	4.3	5.3	2.8	4.3
Perceived Ease of Use	3.5	4.0	1.8	4.5	3.5	4.3	4.8	3.5	3.5	4.8	3.8	3.3	5.3	2.5	4.0

3 Visualization Results

3.1 Perceptual Space of Subjective Appearance

Figure 2a depicts the perceptual space for a set of 15 car infotainment systems (letters A to O) and the vector model of the subjective appearance attributes, obtained from the MDPREF analysis. The Euclidian distance between any two objects represents the level of their perceptual dissimilarity. For instance, the far separations between models J and A or D and F suggest that the level of their perceptual dissimilarity of subjective appearance are high; where as the proximity of models B and F suggests that they are perceived similar. Also taking into account of the vector model of the subjective appearance attributes, we can see that the car infotainment system J is assumed to have high values on the attributes of Calm, Square, and Static (the calculation requires some imagination; it requires to project the point J to the attribute vectors, and then to measure the distance to the origin). Whereas the car infotainment system A has high values on the opposite subjective appearance attributes: Dynamic, Organic, and Aggressive. Similarly, model F is perceived to be Innovative, Stylish, and Exciting; yet the model D is perceived to be Conventional, Inelegant, and Boring. Note that the above observations are consistent with the average ratings of subjective appearance as shown in the Table 3. The advantage of using the perceptual map is to visualize the relationship of many objects in an intuitive manner, and the visualization can also help discover result patterns.



(a) (b)

Fig. 2. Two different 2-dimensional perceptual spaces for a set of 15 car infotainment systems (as denoted by alphabetical letters A to O), the vector model of the subjective appearance attributes (a) (denoted by the points on the big grey circle with attribute names), and the vector model of the perceived usability attributes (b) (labels with rectangular)

3.2 Perceptual Space of Perceived Usability

Figure 2b shows the perceptual space for the same set of car infotainment systems but with the different vector model of the perceived usability attributes. Because we did not use a pair of antonymous adjectives to describe each perceived usability attribute, the car infotainment models with positive distance (to the origin) after the projection to some vector is perceived to be supportive. For instance, the model H and M are perceived to support Self-explanatory, whereas model C is perceived to be highly

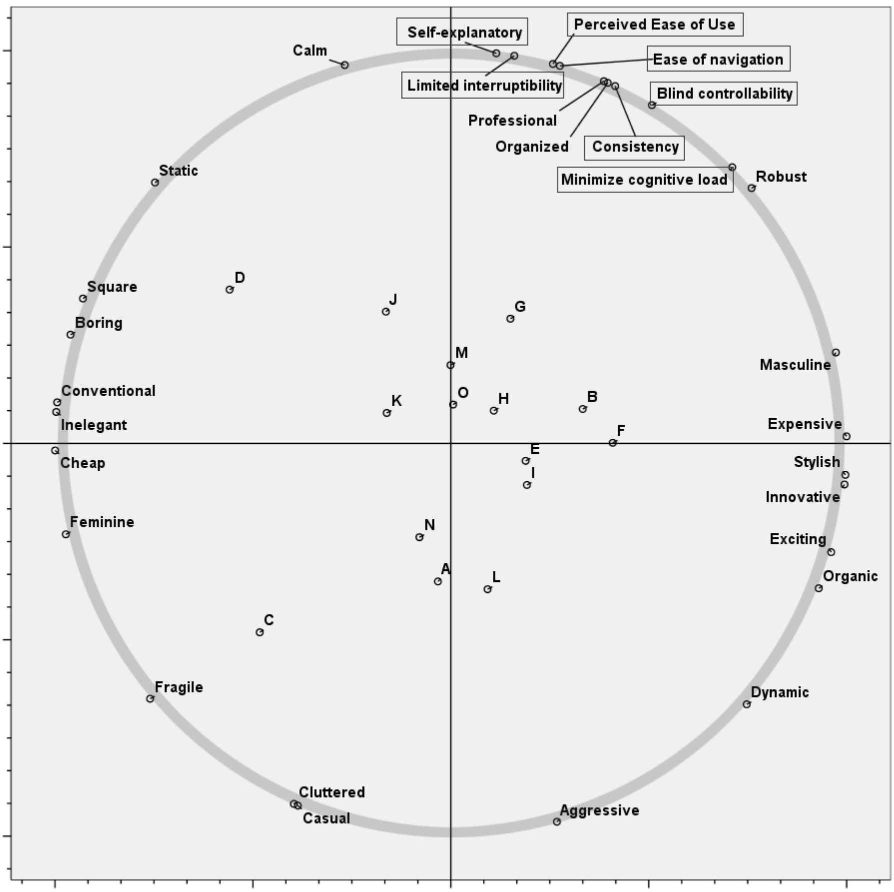


Fig. 3. The perceptual space for a set of 15 car infotainment systems (letters A to O) and the vector models of both the subjective appearance attributes (labels without rectangular) and the perceived usability attributes (labels with rectangular)

unsupportive of the attribute of Self-explanatory. In general, the model G, D, and B are perceived to support Perceived Ease of Use, Ease of Navigation, Consistent and Minimize Cognitive Load, and the model B and G also are perceived to support Blind Controllability and Limited Interruptability.

3.3 Combining the Two Spaces

To explore the relationship between the subjective appearance and the perceived usability, we projected both ratings into the same Euclidean space. Figure 3 shows the results of the perceptual space of the 15 car infotainment systems with both sets of vectors of the subjective appearance and the perceived usability attributes. The figure shows the proximity between some subjective appearance attributes and some usability attributes suggesting certain overlap between the perceptions of product appearance and usability. For instance, the Professional and Organized subjective appearance attributes are perceived to be closely related to the perceived usability attributes of Consistent, Ease of Navigation, and Perceived Ease of Use; the Robust is positively related to the Minimize Cognitive Load. The findings suggest that a design with Professional and Organized appearance will also be likely perceived to be Ease of Use; whereas a design the Casual and Cluttered appearance will be perceived to be unsupportive to Ease of Use, i.e., difficult to use. Figure 3 also shows that several subjective appearance attributes are perceived to be orthogonal to the perceived usability attributes, e.g., Innovative and Stylish vs. Self-explanatory. The finding indicates that whether the design style of a product is Innovative or Conventional, Stylish or Inelegant, may have little influence on users' perception of Self-explanatory. The implications for designing interactive systems are discussed below.

4 Discussion and Conclusion

The subjective appearance and the perceived usability of an interactive system are two integral parts that contribute to the user experience of the system. In this paper, we use the approach of "perceptual maps" to explore the relationship between the perceived usability and subjective appearance of the car infotainment systems. The findings confirm that the appearance of a product will affect its perceived usability. In the context of car infotainment system, professional or organized HMI appearance will also be perceived as ease-of-use and ease of navigation. However, not all appearance attributes are tightly linked to the perceived usability: whether the design is stylish or not may have little influence on the perceived ease-of use.

These findings have important implications for designing interactive systems. Appearance of a product is a combination of design style, use of colors, shapes, layouts, and other design elements. Designers are often focusing on making product to look innovative, stylish, or exciting, yet they may ignore another important factor that also influences users' purchase decision---perceived usability. By also considering the appearance attributes that are closely linked to the perception of usability, the designer can achieve the goal of both attractiveness and perceived ease-of-use. Our approach is generic, and it can be applied to various product designs. The "perceptual maps" provide an intuitive way to explore the relationship between the two.

One caveat of this study is that we only examined the perceived usability; the actual usability is also critical, especially for car infotainment systems due to the safety concern. Whether the actual usability is consistent with the perceived usability and how they may interact will be a topic for future studies.

Acknowledgments. We thank Robert Rauschenberger for his valuable inputs and comments. We also thank Dirk Dudenbostel for his support of this project.

References

1. Chuang, M.C., Chang, C.C., Hsu, S.H.: Perceptual factors underlying user preferences toward product form of mobile phones. International Journal of Industrial Ergonomics 27, 247–258 (2001)
2. Hsiao, S.W., Wang, H.P.: Applying the semantic transformation method to product form design. Design Studies 19, 309–330 (1998)
3. Nagamachi, M.: Kansei engineering: a new ergonomic consumer-oriented technology for product development. International Journal of Industrial Ergonomics 15, 3–11 (1995)
4. Nielsen, J.: Heuristic evaluation. In: Nielsen, J., Mack, R.L. (eds.) Usability InspectionMethods, John Wiley & Sons, New York, NY (1994)
5. Osgood, C.E., Suci, G.J., Tannenbaum, P.H.: The Measurement of Meaning. Illinois press, Urbana, IL (1957)
6. Stevens, A., Quimby, A., Board, A., Kersloot, T., Burns, P.: Design Guidelines for Safety of In-Vehicle Information Systems. Transport Research Laboratory (2002)
7. Strategy Analysis: Market Report of In-vehicle Entertainment Systems 2004-2011 (2004)
8. Wickens, C.D, Lee, J., Liu, Y., Gordon-Becker, S.: Introduction to Human Factors Engineering. Prentice-Hall, Inc, Upper Saddle River, NJ (2003)

This page intentionally blank

Part III

Medical and Rehabilitation Applications

Digital Human Modelling: A Global Vision and a European Perspective

Gordon J. Clapworthy[1], Peter Kohl[2], Hans Gregerson[3], S.R. Thomas[4], Marco Viceconti[5], D.R. Hose[6], D. Pinney[6], John Fenner[6,*] , K. McCormack[6], P. Lawford[6], S. Van Sint Jan[7], S. Waters[8], and P. Coveney[9]

[1] Univ. of Bedfordshire, UK
[2] University of Oxford, UK
[3] Aalborg Hospital, Denmark
[4] Centre National de la Recherche Scientifique, France
[5] Istituti Ortopedici Rizzoli, Italy
[6] Univ. of Sheffield, UK
j.w.fenner@sheffield.ac.uk
[7] Univ Libre Bruxelles, Belgium
[8] University of Nottingham, UK
[9] University College London, UK
Dept of Medical Physics, University of Sheffield,
Floor I, Royal Hallamshire Hospital, Glossop Road, Sheffield, S10 2JF, UK

Abstract. The Physiome is an umbrella term that refers to human modelling with mathematics and computational methods, accommodating cross-disciplinary science (chemistry, biology, physics) and a breadth of dimensional and temporal scale (sub-cellular to organs, sub-microsecond to tens-of-years). The Virtual Physiological Human is a European initiative that is intended to provide a unifying architecture for the integration and cooperation of multi-scale physiome models, thereby creating a predictive platform for the description of human beings *in silico*. Unlike the Genome, the challenge of the Physiome may be almost unbounded, as the desire for increased detail imposes a continuing pressure for ever-finer data granularity, and the necessary Information Technology (IT) infrastructure spawns innovations that surpass conventional solutions. It is foreseen that maturing physiome activities will increasingly influence medicine, biomedical research and commercial developments, and the central role of IT highlights the need for a specific and robust IT infrastructure.

The European Commission has experience of supporting challenging technical endeavours through its Framework Programmes of research, and in the forthcoming 7th Framework Programme, it will invite researchers from within and outside Europe to unite in seeking solutions to key issues of the Physiome. The Virtual Physiological Human (VPH) investment programme will establish the necessary infrastructure and address the grand technical challenges identified by experts. This paper examines the background to the strategy and the ways in which the programme's structure has been determined.

Keywords: Physiome, Virtual Physiological Human, infrastructure, modelling.

* Corresponding author.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 549–558, 2007.

1 Introduction

The Virtual Physiological Human (VPH) is a methodological and technological framework that once established, will enable collaborative investigation of the human body as a single, complex system. It will be a tool to facilitate the sharing of observations, an environment for the investigation of hypotheses, supporting the development and validation of theories, integrating all observations and hypotheses into a dynamic understanding of human physiology and pathology. The VPH refers to the supporting infrastructure that makes this possible but it is not the model itself.

The need for such an architecture has emerged from the European scientific community, which has eagerly embraced the concept of the Physiome, as described by Hunter et al [1]. The current partitioning of health science along traditional lines (scientific discipline, anatomy, physiology etc) is considered an artificial and inefficient framework for such an integrated and dynamic description of human biology. In contrast, the VPH is tendered as a comprehensive solution to this disparate challenge, and the first steps of its development are the subject of this paper.

2 The First STEPs to the VPH

Development of the VPH requires scientific focus, funding and forethought, and consequently the European Commission (EC) has instigated a think-tank that operates as a European Coordination Action, funded by the e-health unit of the EC as part of the 6^{th} Framework Programme. It has the acronym STEP (**S**trategy for **T**he **E**uro**P**hysiome) and is tasked with considering and recommending strategies that promote the development of the VPH, providing focus for VPH activities within the 7^{th} Framework Programme (FP7) between 2007 and 2013. STEP will publish its report [2] in spring 2007 ("*Seeding the EuroPhysiome: A Roadmap to the Virtual Physiological Human*"). It is intended to inform the decisions of the EC with respect to the funding of VPH and physiome research, and will emphasise the need for efforts that are descriptive, integrative and predictive across scientific disciplines and physical measurement scales. Contributions to this roadmap have been critiqued through a consensus process that has involved the broadest of groupings, from the academic community, industry, clinical experts and a wide range of professional associations.

By reference to the emerging STEP roadmap, this paper outlines the EC commitment to digital human modelling and simulation, highlighting its interest in developing an infrastructure capable of supporting global physiome activity. Critically, the EC recognises that the influence of the VPH extends beyond the scientific domain, and inevitably, will have an impact on society as a whole.

3 The VPH and Its Motivation

Dissatisfaction with current modelling approaches from clinical, industrial and academic quarters offers strong justification for the development of the VPH and emphasises the belief that a cross-discipline approach is the only credible way forward in the 21st century [3]. Anticipated societal benefits also add significant support to the endeavour.

Clinical: A feature of current medical practice is specialization, and bridges between clinical disciplines are weak. Consequently, western medicine struggles with the management of multi-system or multi-organ pathologies and the VPH offers a means of breaking down these barriers by providing an inherently cross-disciplinary tool that can alter the way in which we approach physiology and pathology. Ideally, close integration with the healthcare system will encourage its adoption as an effective clinical decision support tool.

Industry: The VPH will be a resource to aid product design and development, reducing risk and cost. Innovative resources will accelerate time-to-market, with the potential to strengthen the medical related industries. Approval of the VPH by regulatory authorities will help to reduce (but not totally remove) the need for clinical trials and animal testing.

Academia: The narrow expertise of individual researchers is characteristic of scientific focus, which all too often, produces modellers, engineers and experimentalists who do not understand each other; frequently none of them appreciates the clinical impact of their joint work. This is an impediment to the development and understanding of science and healthcare, and again the VPH offers a means of crossing these barriers, expanding horizons and facilitating communication between those participating in biomedical science.

Society: Societal needs are a conglomeration of personal needs and expectations, which includes ambitions for quality of life and an anticipation that taxes are used to improve the lives of citizens. This is consistent with a vision of the VPH as a tool that can have far-reaching effects on the breadth of human experience, from improved healthcare to enhanced quality of life.

In essence, motivation for VPH development is justified by the pervasive nature of the technology and its benefits, reinforced by the international perspective on the VPH as presented below.

4 International Context and Common Objectives

Numerous examples can be cited that illustrate the relevance of the VPH to a diversity of current projects. These extend beyond Europe, and include the USA and the Asia-Pacific region. Activities that have resonance with the VPH (funded under EC FP6) include AneurIST, ImmunoGrid and the Living Human Digital Library (LHDL). Alongside EuroPhysiome activities, there are a number of other Physiome projects around the world, the major ones being the National Simulation Resource (NSR) Physiome project from the USA and the IUPS Physiome, coordinated by the University of Auckland, NZ. Numerous successful initiatives are present across the globe [2] (e.g. LYMFASIM, the Virtual Soldier Research programme, Heart Physiome etc.), and to reduce the risk of fragmentation, the authors of the STEP roadmap recommend the establishment of a World Physiome initiative – a lightweight coordinated action among all physiome-related projects around the globe.

It is important to recognize that there are many fundamental activities common to these projects that need not be reinvented with each project but could usefully be developed and administered through a single infrastructure platform or resource such as the VPH. Overlap with VPH objectives is particularly apparent in the health computing and health-grid arena (e.g. the SHARE initiative for Health-Grid computing). There is a growing belief that the VPH can only be successfully implemented by using the resource-sharing mechanisms provided by a Grid computing infrastructure, but barriers within current Grid technologies need to be overcome before full-scale VPH implementation becomes possible.

5 Research Challenges

The vision for the VPH is the result of a complex consultation and consensus-building exercise involving a large number of internationally recognised experts. This cross-discipline complementarity comes to the fore when the challenges for VPH implementation are considered. The two principal research challenges are identified as:

1. Clarifying the nature of the scientific problems and principles pertinent to the VPH and how they might be solved, and
2. Identifying the ICT tools that should be developed to address these problems.

It is widely accepted that the true, grand research challenge lies in understanding biological function. This is the focus of the first point, and requires a multi-faceted approach to biomedical endeavours, since the complexity of a human being allows for almost limitless inter-connections and interactions between systems that can be open, responsive and adaptive. An abundance of data provides a wealth of descriptive information, but to understand it fully and determine physiological function, models are needed that closely link descriptive data with underlying biology.

The pivotal role of information technology as a solution to this challenge is self evident, and leads to consideration of the second research challenge (point 2 above). This is predominantly a challenge of integration: not only the integration of processes spanning large temporal and spatial scales and the integration of descriptive data with predictive models, but also integration across disciplines. A concerted effort is required [4], [5] to ensure that adequately supportive and appropriate infrastructures, frameworks and technologies are developed. To address the multi-scale aspect of the VPH, a collection of large databases, accommodating models and data at the various scales, is required. A variety of software tools is also needed to author, run and visualise these models, along with a defined common language to address the disparate components.

6 Scientific Challenges

The VPH research challenges discussed above can be further dissected into Challenges in Prediction, Description, Integration and Information, each examined below:

6.1 Challenges in Prediction

Model interaction and complexity, and issues of inhomogeneity and inter-subject variation, fall into this area, as does the difficulty of problem identification and boundary setting. The identification of gaps in our knowledge of human biology and pathology is also important, along with the clinically essential area of model and data validation.

6.2 Challenges in Description

It is prudent to recognise that the presence of data should not be taken to mean it is accurate or complete. Instruments, models and systems that produce data introduce artefacts and set limits that translate to uncertainties in measurement. The introduction of data collection protocols could limit these effects, as detailed records of data collection methods would avoid misinterpretation. A solution is the inclusion of appropriate metadata, with consequences for all those involved in data collection. However, this rigour has advantages, for if high quality data is traceable to source, there may be justification for its use in physiological and clinically applicable models.

6.3 Challenges in Integration

Integration in this context refers to the establishing of appropriate connections at the interfaces between differing temporal and spatial scales, for both models and data, and to the integration between academic and clinical specialties. The interface between a model and an appropriate data set is an important link for validation, and demonstrates the need for comprehensive, flexible and extensible mark-up languages such as CellML[6] and FieldML etc.

6.4 Challenges in Information and Communications Technology (ICT)

The primary challenge for ICT is to develop a coherent suite of tools capable of addressing the scientific challenges described above. One of the core requirements is for complex databases that collect and classify models and data and act as guides to modelling efforts in Europe and beyond. This requirement illustrates the need for a standard framework that can support the coupling of different models. Since the VPH will involve very large amounts of data (petabytes, at 2007 standards) smart strategies will be needed for cataloguing, storage and sharing.

A dedicated and distributed VPH computing resource is an attractive solution and would expedite the integration of grid computing technologies and middleware (amongst other things) with biomedical research. However, since the VPH by its very nature operates at the forefront of technology, many solutions do not yet exist, but there are prototypes – such as the Application Hosting Environment (RealityGrid [7]) – that can already act as an interface with federated grids. The template provided by the BioSimGrid is another example, since this is designed to act as a repository of simulation data specifically for biomolecular dynamics. BioSPICE also has a similar, open-source, modular infrastructure.

7 Impact Analysis

The roadmap speculates about the future impact of the VPH. Effects will be not only on biomedical research and clinical practice but also in the associated industry sectors and society at large.

Impact on biomedical research will be evident through the provision of an unprecedented collaborative architecture on an international scale. The main benefit for the clinical arena will be the facilitating of patient-specific treatment, along with better cooperation between the various clinical specialisations. Indirectly this will also have an impact on the cataloguing of human physiology and pathology through the collation of studies and data gathered under the VPH umbrella.

The most immediate impact to industry will come in the medical devices sector where the VPH will offer improved technical excellence, a reduction in the development time of new products and amelioration of the risks associated with their introduction.

Societal impacts would take longer to manifest themselves, reliant on the postulated outcomes described above. Reduced public expense and healthier citizens will be the primary anticipated effects, but other less-obvious consequences (workforce size, educational diversity, vehicle safety, ergonomics etc) will follow.

8 Ethical, Legal and Gender Issues

Since the impact of the VPH will spread beyond the realms of normal scientific influence, it is perhaps more important than usual that these wider aspects are addressed comprehensively. The VPH poses numerous ethical [8] and legal dilemmas that need explicit attention. For instance, guidance in the event of an adverse clinical outcome resulting from the VPH is a very important legal area to consider alongside the balance between freedom of information and the security and privacy of patient data. Copyright and the sharing of data (IPR) is a topic in need of urgent and specific legal clarification.

Social and political dilemmas are also a feature of the VPH since it must not compromise ethical ideals valued by society. Indeed, the VPH could influence perennial issues such as social equality because it may identify race and gender characteristics that challenge our preconceptions. Typically, a gender-neutral approach to patient care is currently used in modern medicine since there are comparatively few gender specific diseases such as osteoporosis. However the VPH has the potential to identify gender-sensitive behaviours that are much more subtle, creating the prospect of gender-specific treatments and improved outcomes.

9 Dissemination, Exploitation and Sustainability

Some of the consequences of the VPH could require a shift in cultural behaviour, for example in the working practices of clinicians; the VPH may therefore be viewed by some as a threat. As with any other major change, the VPH requires sensitive and comprehensive education in its use and benefits. Dissemination and profile-raising campaigns need careful planning so that developers and end-users alike may be engaged. Ultimately the long-term sustainability of the VPH will be dependent upon its ability to influence the health and well being of all citizens.

10 VPH Recommendations That Issue from the Roadmap

Following consideration of the many issues raised, the STEP roadmap offers a number of recommendations for the concrete implementation [2] of the VPH in each of the following areas: Infrastructure, Data, Models, Validation, Dissemination and Sustainability. Firstly, it is suggested that continued funding for a network of experts is confirmed so that the available momentum generated by the STEP action can be sustained and translated into VPH success. A suitable instrument might be an FP7 'Network of Excellence' (NoE) that would track all physiome-related projects and facilitate dialogue to encourage cross-fertilisation of ideas.

10.1 Infrastructure

A scalable but resilient infrastructure should be designed as a priority. The VPH will harness very large quantities of information, and its sustainability will be dependent upon its ability to accommodate this data. An adaptive infrastructure that can deal with the challenges of transfer, organisation, indexing and storage at the rate of terabytes per day is required. Success is critically dependent upon the development of a suitably robust IT infrastructure. More widely, a 'VPH community' should be formalised, willing to help develop and enforce the standards to which the VPH should be built.

10.2 Data

Standards for data-collection should be established as a priority. Projects for the collection of data for cell-cell and cell-tissue interaction could be favoured since this will enhance the understanding of the underlying biology. Ideally, data should be collected in a manner consistent with written protocols. This would provide data that is traceable so that the quality of the information can be assured enabling separation of high quality data from that obtained as a by-product of some other activity, unlikely to carry the same levels of assurance.

Data curation (storage, protocols and standards) is critical to the success of this venture and requires specific attention and clarification of IPR issues for data sharing if the VPH is to be successful. The brief of the VPH includes dealing with patient data gained from a clinical setting. The utmost care must be given to the creation of relevant security protocols. Rigorous mechanisms for quality assurance are a necessity when considering integration of clinical and non-clinical data.

10.3 Models

Standards for the interconnection of models should be established as a priority. Development of simulation tools and models is to be encouraged, however the output of these activities needs to be validated to ensure their suitability. Sharing and coupling of models also requires support, with dedicated effort directed at support for physiome mark-up languages.

10.4 Validation

Standards for model validation are also an early necessity. Since it is envisaged that the VPH will be used in a clinical setting, this imposes a strict requirement for validation. Methods are needed to ensure that models are consistent with their claims, accompanied by explicit statements concerning assumptions and limitations along with care in the acquisition of validation data and acknowledgement of errors and limitations.

10.5 Dissemination

A number of strategies are required for the effective dissemination of the VPH concept. There is a recognised need for:

- Continued 'evangelisation' for the VPH through effective channels
- Representation at conferences to cement the VPH community
- Representation of the VPH in education
- Standardisation
- Coordinated action to endorse and disseminate the VPH vision
- European and global harmonisation of the Physiome effort

10.6 Sustainability

Undoubtedly interim effort (and finance) demonstrating the viability of VPH concepts in the short term will be needed before the VPH gains sufficient credibility to become sustainable in the long-term. Suitable EC vehicles include the aforementioned Network of Excellence (NoE), but also Integrated Projects (industrial involvement with short-term impact) and STRePs (Specific Targeted Research Project – smaller consortia with focused outcomes). A mechanism for accommodating non-government finance and sponsorship (e.g. industry) is also a necessity.

11 Discussion

The Physiome is a long-term vision, one in which eventually, an integrated multi-scale model describes all aspects of human biology from genes and molecules through cells to organs and the whole organism, all acting as one. Functional elements of this mosaic are beginning to appear, such as the genome (genes), heart physiome, etc. However, these currently operate as ‘stand-alone’ models and do not interact with one another. If they were to interact, a much better picture of the entirety of human biology could be gained to enhance understanding of diseases and their treatment. The Physiome will therefore comprise a kit of parts, together with assembly tools, that will allow the investigator to construct models matched to the tasks in hand, and to output data in forms appropriate to the interventions, investigations or designs required. This has implications for the clinical decision-making process along with much wider benefits to health, industry and society. This paper postulates that a suitably robust and adaptive infrastructure is a critical part of the endeavour and needs to be available to all users worldwide if it is to be a success.

The grandest challenge of the Physiome lies therefore not in the construction of a single model of enormous complexity, but in the need to establish a vast, expandable IT framework. Within this, models of any complexity can be constructed from available quality-assured sub-components and data, such that reliable results can be obtained with the minimum of effort. A secondary goal is to begin to populate this framework with the raw content (tools and data) that users will require. The final element will be the inclusion of proven, documented and validated models of physiological subsystems that can offer immediate benefit. The vision of the European Community (EC) is that over the next ten years a suitable infrastructure is built to accommodate the framework of technology and methods supporting Physiome efforts worldwide. The technological component is primarily focused on information and communications technology and the management of knowledge, with innovations that improve access, exploration and understanding of the content accumulated within the VPH. The methodological component relates to a unified representation of concepts, data and models, development of new methods of processing and modelling and the integration of models across all scales.

The roadmap to the EuroPhysiome[2] was crafted by those working daily in the field, assisted by information from a wide range of potential industrial and clinical users. Many components with which the Physiome will be populated are already available from specialists: Cardiome, Epitheliome, GIOME etc. The challenge is now to establish the framework that will allow the mutual interaction of these components, and to extract their true potential to improve the health and well being of the world's citizens. The principal European mechanism by which VPH developments can occur is through the funding instruments available within FP7, arranged around four 'work-programmes': Cooperation, Ideas, People and Capacities, each with a number of themes. EC funding for VPH activities exceeds 70 million euros, and can complement external sources of funding (e.g. grants that are available from national agencies and industry). It is expected that the health and ICT areas will account for approximately 50% of all EC research spending in the FP7 Cooperation programme with ICT being the largest priority theme for FP7. An estimated 500million euros will be required to fund the development and deployment of the VPH. These are substantial sums and indicate the magnitude of the challenges to be addressed.

The creation of the VPH poses a number of challenges across multiple disciplines, from mathematics and ethics to genetics and clinical integration. Unquestionably, the VPH is a grand challenge for the European research system as a whole. However, this challenge is larger than any human subset, whether European, American or Asian; it is a truly global endeavour, and the EC is inviting global participation as it seeks to add European resources to the quest to understand humankind.

12 Conclusion

The Physiome concept is simultaneously intriguing and challenging, and it has profound implications for the future of human modelling and biomedical science. The European response presented in this paper, is to invest in the VPH as an underpinning IT infrastructure designed to unify disparate projects and expertise, creating a platform that offers the functionality of a multi-component Physiome. The challenge is technological and methodological, and Europe recognises the necessity for international cooperation, seeking collaboration within and beyond its borders as a means of consolidating the vision that is the Virtual Physiological Human.

Acknowledgements. STEP (Strategy for the EuroPhysiome) was funded as a Coordination Action under the European Commission Framework 6 programme (FP6-2004-IST-4).

References

1. Hunter, P., Robbins, P., Noble, D.: The IUPS human Physiome Project. European Journal of Physiology 445(1), 1–9 (2002)
2. STEP_Consortium, Seeding the EuroPhysiome: A Roadmap to the Virtual Physiological Human (2007) http://www.europhysiome.org/roadmap
3. McCallin, A.M.: Interdisciplinary researching: exploring the opportunities and risks of working together. Nursing & Health Sciences 8(2), 88–94 (2006)
4. Kohl, P., et al.: Computational modelling of biological systems: tools and visions. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 358, 579–610 (2000)
5. Gavaghan, D., et al.: Mathematical models in physiology. Philosophical Transactions of the Royal Society London, Series a, 2006 vol. 364(1842), pp. 1099–1106 (2006)
6. Lloyd, C.M., Halstead, M.D., Nielsen, P.F.: CellML: its future, present and past. Progress in Biophysics & Molecular Biology 85(2-3), 433–450 (2004)
7. Cohen, J., et al.: RealityGrid: an integrated approach to middleware through ICENI. Philosophical Transactions of the Royal Society London, Series a, 2005 vol. 363(1833), pp. 1817–1827 (2005)
8. Bassingthwaighte, J.B.: The macro-ethics of genomics to health: the physiome project. Comptes Rendus Biologies 326(10-11), 1105–1110 (2003)

ICT Methodologies to Model and Simulate Parts of Human Body for Prosthesis Design

Giorgio Colombo[1], Stefano Filippi[2], Paolo Rissone[3], and Caterina Rizzi[4]

[1] Politecnico di Milano, Dipartimento di Meccanica
Via G. La Masa, 34 – 20153 Milano, Italy
[2] Università di Udine, DIEGM
Via delle Scienze, 208 – 33100 Udine, Italy
[3] Università di Firenze, Dipartimento di Meccanica e Tecnologie Ind.li
Via S. Marta, 3 – 50139 Firenze, Italy
[4] Università di Bergamo, Dipartimento di Ingegneria Industriale
Viale G. Marconi, 5 – 24044 Dalmine (BG)-Italy
giorgio.colombo@polimi.it, filippi@uniud.it,
paolo.rissone@unifi.it, caterina.rizzi@unibg.it

Abstract. The work presented in this paper refers to the implementation of a product development process based on the use of virtual model of the human body to design specific custom-fit product, such as a prosthesis socket (interface between the residual limb and the mechanical part of the prosthesis). It considers the integration of advanced ICT tools coming from the reverse engineering, the physics-based modelling and simulation, and the rapid prototyping fields. The paper describes problems related to the implementation of each step within a real socket development process.

Keywords: Product customization, Prosthesis design, Physics-based simulation, Human body modeling.

1 Introduction

Custom-fit products, especially those with an interface with the human body, cannot be completely designed using methods and tools developed for other industrial fields (i.e., mechanical field) for mass production or for modular products, but it is necessary to realize “ad hoc” methodologies.

The work presented in this paper refers to this context with the aim of verifying the implementation of a design paradigm for a specific custom-fit product, a prosthesis component, through the integration of advanced ICT tools, allowing the evolution from a hand-made production to a computer assisted realisation of highly customised products. We consider the case of the socket (interface between the residual limb and the mechanical part of the prosthesis) both for trans-femoral or trans-tibial amputees since it requires a high level of customisation. The socket plays the fundamental role in the comfort and functionality of the prosthesis. It has to be manufactured in the correct way both from anatomical and biomechanical point of view because each movement between stump and socket can hugely reduce the prosthesis control causing unsafe and pain during walking.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 559–568, 2007.

The research has been developed in the framework of an Italian PRIN (Research Project of National Interest) Project named DESPRO (Integration of Innovative Methodologies to DESign and Develop Custom-fit Products: Application and validation for a Socket of a Lower Limb PROsthesis) funded by Italian Research Ministry. The consortium comprises four universities, University of Bergamo, Florence, Udine, and Polytechnic of Milan collaborating with an Italian prosthesis manufacturer named Centro Protesi INAIL, Budrio (BO).

2 Traditional Process and Knowledge Acquisition

Before proceeding with the implementation of the new design paradigm, we studied the traditional process followed by orthopedic technicians to design non-standard components of lower limb prosthesis. The product and process knowledge has been acquired analyzing scientific literature and through interviews with domain experts (medical, paramedical and technical staff) and with the patients that will wear the prosthesis. This activity has been particularly important since it permitted to define key characteristics of the socket and consequently possible improvements both for the product and for the design process.

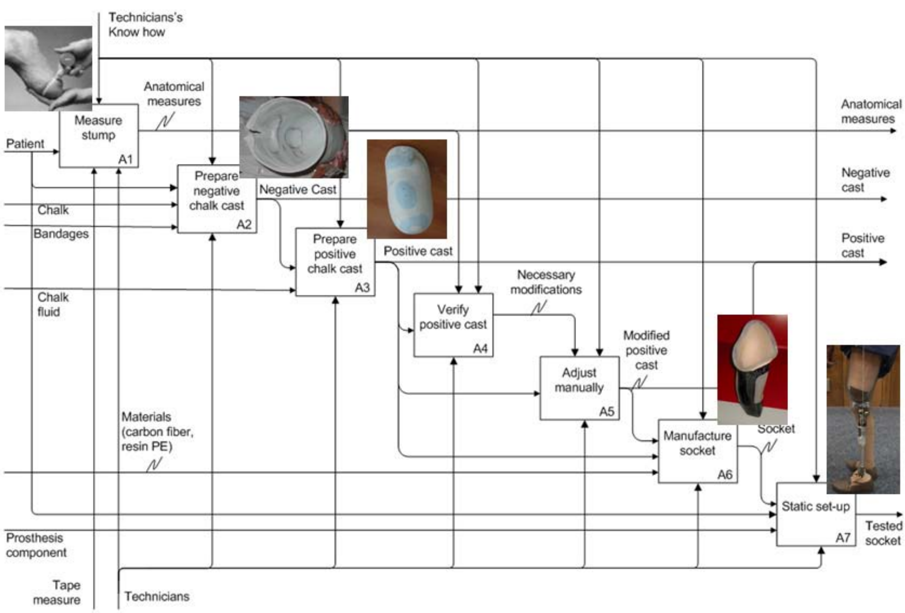


Fig. 1. Design and manufacturing process of a socket

First, the design and manufacturing of high-quality sockets must fulfill the following principles: accurate measurement of the stump geometry, perfect close fitting of the prosthesis to the stump, good response to forces and mechanical stress, safety, and each single area of the socket must have a tight connection to the stump anatomy without affecting blood circulation.

Both trans-tibial and trans-femoral sockets are manufactured starting from a positive chalk cast. The cast can be manufactured following a fully hand-made procedure or partially based on digital tools, such as CAD/CAM systems. In the last case, the procedure for trans-tibial and trans-femoral is slightly different but the output is always the 3D model of positive cast for the CAM system and, thus, to produce the physical prototype of the positive cast. However, in Italy, most of laboratories producing prostheses are SMEs and the design and manufacture of a socket is almost a hand made activity carried out by skilled technicians. Figure 1 portrays the main phases of the As-Is production process represented with and IDEF0 diagram [1].

In literature, we can find various researches on ICT tools to support some single steps of the process [2-11]; however, there is not a validation of a complete process based on digital data and tools. This highlights the need of new design methodologies that take into account materials used for the socket, the stump structure (i.e. bones, muscles and skin) and manage their mutual interactions.

3 The New Development Process

We propose a new design paradigm based on the use of virtual models of human body to design specific custom-fit product, such as the prosthesis socket. In such a context, different issues related to the human body should be considered, e.g., the acquisition of stump morphology, generation of a complete virtual model that includes both the external shape (skin) and the geometry of internal parts (muscles and bones) and mechanical characterisation of the stump to be able to simulate correctly the socket-stump interaction.

Figure 2 shows the new design paradigm completely based on computer-aided tools and on the modelling and simulation of the two interfacing parts (human body and socket). It has been thought to implement best practices used by orthopaedic technicians and finalised to ensure high-level products independently of the competencies of the domain expert that manufactures the socket. It consists of five main steps, where following tools are integrated (Figure 3):

- reverse engineering tools for the automatic (or semi-automatic) acquisition of patient's morphology and bony-muscular structure (in our case the residual limb);
- a physics-based modeler allowing the designer to represent both the human body's parts and the socket as composed by different materials;
- an environment for physics-based simulation to reproduce the real behaviour of socket-stump system and to verify the product functionalities;
- Rapid prototyping tools for the realization of physical prototypes to test and validate the virtual product and to identify adjustments.

The first activity is devoted to the acquisition of the stump geometry for the reconstruction of its digital model and the corresponding socket model. Since ten years ago, several non-contact reverse engineering techniques and medical imaging

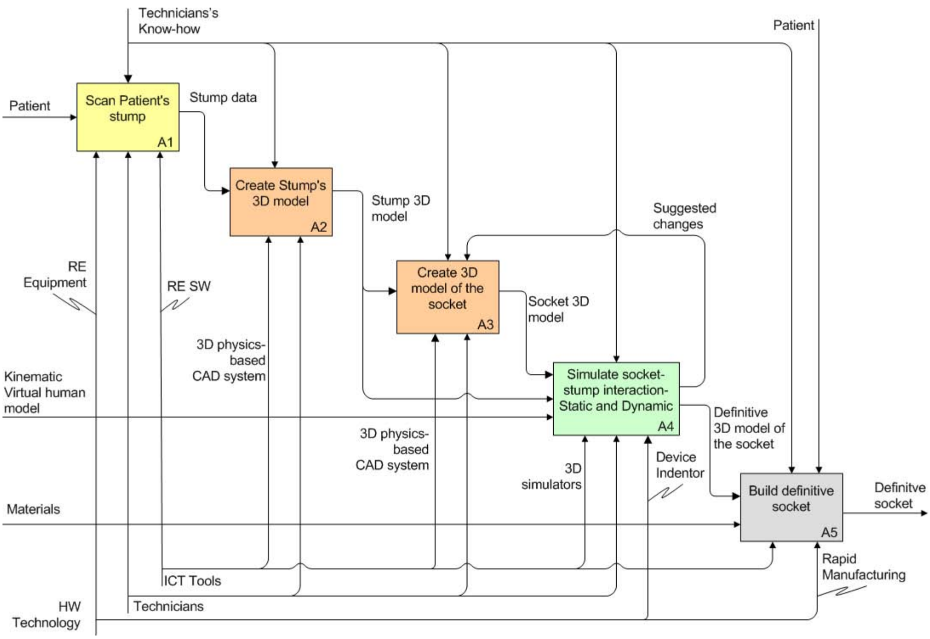


Fig. 2. The new product development process

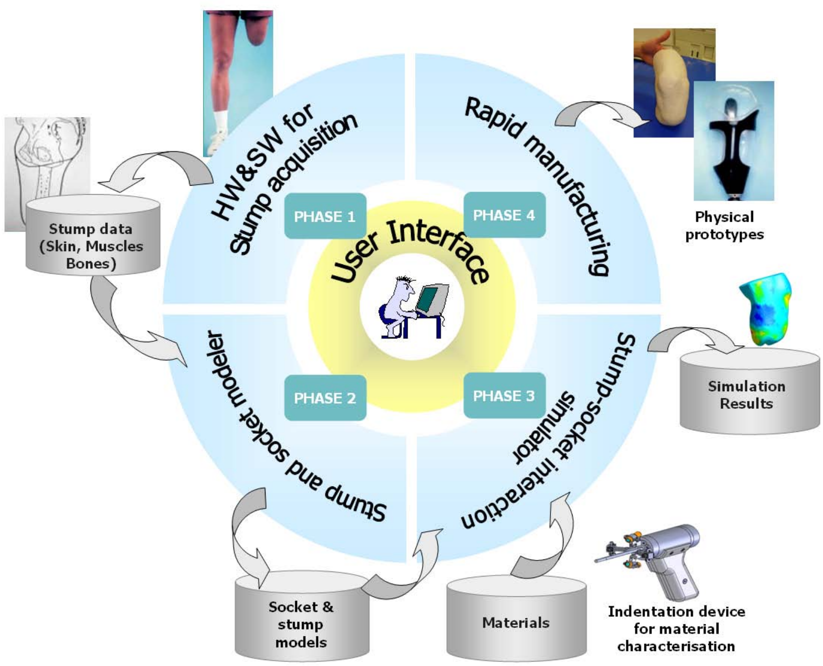


Fig. 3. The design process phases and related tools

have been tested. First techniques for digitalizing the 3D geometry of the residual limb were based on the acquisition of bi-dimensional silhouettes [2], but more recently, MRI (Magnetic Resonance) and CT (tomography) have been considered in order to build a model as a set of cross-sectional images, by distinguishing bones, soft tissue and skin of the residual limb [3], [4], [5]. We adopted both reverse engineering based on laser technology and medical imaging techniques to obtain a model that includes both stump external shape and its inner parts (muscles and bones).

Concerning stump and socket modelling, CAD/CAM systems available on the market for prosthesis design (e.g., TracerCAD - www.tracerCAD.com, Ossur CAD - www.ossur.com) permit to model the positive cast model and are not able to model objects made by different materials (e.g., for the stump: skin, tissue, bones...).

Therefore, the choice of the product model representation is one of the most important aspects of the whole project. In fact, the representation used to codify both the geometric information and the physical-mechanical features of the product can heavily influence the following simulation and manufacturing phases. Several modelling techniques and representation schemes have been considered; the choice has been the physics-based modelling (both the Particle-based modelling and Finite Element Method) [6], [8], [13] for what concerns the simulation activities and the parametric solid modelling for the rapid prototyping activities.

The aim of the simulation phase is to optimize the shape of the prosthesis socket according to fitting and wearability criteria. The simulation of the stump-socket biomechanical interaction is a very difficult task since the tissues of the residual limb have a strong non-linear mechanical behaviour and exhibit very large deformations under the loads that the stump experiences during the gait cycle. Particle-based [6] and the Finite Element [12] approaches have been investigated. While particle-based approach has been widely used only for real-time simulation purposes, the Finite Element Method has been adopted to simulate the interaction between the prosthesis socket and the stump according to the tasks of this work. However, the use of this method is still strongly limited by the material characterisation adopted to simulate the behaviour of the soft tissues. The optimization task requires material models that take into account the non-linearity of the human soft tissues and their distribution around the bones of the stump [9], [14], and [15]. According to the purposes of the road – map developed in this work, the authors are developing a device to measure the mechanical characteristics of the stump to characterise the material model for the simulation task.

Finally, rapid Prototyping (RP) can play an important role in this process reducing the time to deliver the definitive socket to the patient. Countless are the applications of RP in the medical field, specifically regarding prosthesis development, and last ones have been mainly considered for the production of the stump positive plaster cast and for socket manufacturing [10], [11]. In the context of this project, the effort is directed towards two main strategic directions: the application of the RP technologies and the re-engineering/optimization of the product oriented to the RP technology, aiming at the generation of the physical prototypes of the anatomical parts and the artefacts.

4 Design Process Experimentation

Main efforts have been spent to demonstrate the practical implementation of the proposed approach and highlight the aspects to be improved. In the following, we will describe preliminary results of experimentation carried out until now with reference to each step of the new design paradigm. Both trans-tibial and trans-fermoral amputees have been considered; however, the complete process has been mainly validated for trans-tibial prosthesis.

4.1 Geometry Acquisition

A non-contact laser scanner (Minolta Vivid VI-9i^{TM}) has been adopted to acquire the external shape of the stump. We also acquired the geometry of the related positive plaster cast to compare the stump external shape with the socket internal surface and determine critical zones. Computer Tomography (CT) and Magnetic Resonance Imaging (MRI) have been adopted for the internal structure, respectively for bones and soft tissues and muscles. During acquisition, patient's posture has been determined using a support device that permitted to maintain lower limb position adopted for manual measurements. In addition, markers have been used to identify anthropometric standard points. For further details, see [16].

Figure 4 shows the geometry acquired for a trans-tibial amputee: (A) point clouds, (B) STL from CT, and (C) a MRI image with markers.

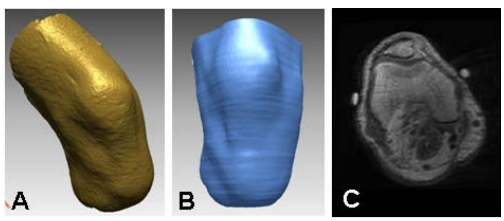


Fig. 4. Acquisition of stump external and internal shape

4.2 Stump and Socket Modelling

The stump geometric model has been generated integrating the data acquired through the three mentioned technologies [16]. Two types of geometric model have been reconstructed: one tasseled (STL) and another based on NURB surfaces.

The geometric model of the skin, derived from laser points clouds, ensures a high quality in morphological details necessary for the simulation of stump-socket interactions. For the internal parts both CT and MRI data have been used. CT and MRI could be considered quite similar from the shape reconstruction point of view. They are based on different technologies (X-ray and magnetic resonance of particles respectively) but both have as output a data volume consisting in a set of images, strictly ordered, and representing anatomical sections. Starting from this data, the 3D reconstruction of the anatomical structures is quite straightforward.

The three different models have been assembled aligning them into the same reference global system taking into account markers used during acquisition phase (Figure 5a).

Regarding the socket model, we initially reconstructed it from the points of clouds acquired for the positive plaster cast (Figure 5b). Twice are the objectives: first compare the stump and socket model in order to identify main differences and, then, provide reference data to evaluate simulation results.

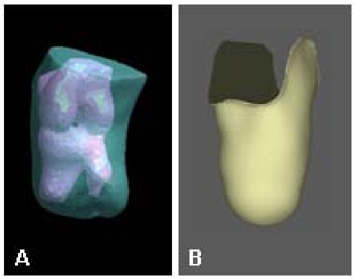


Fig. 5. Geometric models: a) stump model, b) socket model

4.3 Stump-Socket Simulation

The authors consider this task as a fundamental step to find the optimal morphology of the socket. We analysed two different aspects: the wearability (donning simulation) and patient's motion (gait simulation). By the donning simulation, the designer can verify if the shape of the socket allows the wearibility. In this way the presence of dangerous undercuts, which can produce stresses and large deformations of the soft tissue, can be easily identified. The gait simulation allows the analysis of the biomechanical behaviour of the socket-limb interface during the patient's walking. We first focused the attention on donning simulation adopting FE approach.

2D and 3D models, implicit and explicit codes and different types of material models are under investigation. A hybrid method based on indentation tests and Finite Element analysis has been developed to characterize the material models of the stump soft tissue. This allows defining ad hoc material models according to the anatomy of the residual limb of each patient so that more accurate results in the pressure distribution can be expected. Two different types of simulation are under investigation. The first, based on 2D elements, analyses meaningful sections of the stump and the socket. The second uses 3D elements and studies the complete interactions between the two mentioned parts. When the 2D simulations furnish good results, the second more complex simulation based on 3D models permits to verify the fitting and the wearibility of the whole 3D socket. If these verifications are satisfied, final model of the socket is reached. Otherwise, other modifications to the 3D geometrical model of the socket are planned and a new cycle of 2D and 3D simulations is performed.

In the approach based only on 3D model, the fitting and the wearibility simulations are performed without any preliminary 2D simulation. The task aims at verifying if the designer and the orthopaedic technician are able to optimize the socket shape dealing directly with 3D more complex geometrical models. In order to speed up the optimization of the socket shape the integration of optimization tools in the simulation task is, also, under investigation.

Figure 6 shows an example of a preliminary simulation of wearibility realized by using a 3D mesh and the explicit solver LS-DYNA; in this case, the bucket sort-searching algorithm is used as contact model between the socket and the limb.

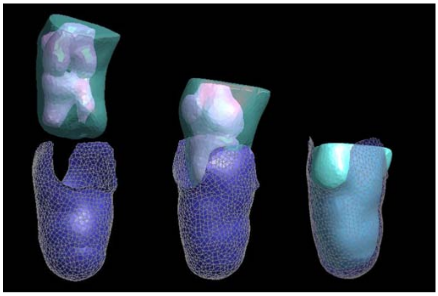


Fig. 6. Wearibility simulation

Figure 7 shows the pressure distribution obtained by a preliminary 3D simulation of the fitting between the socket and the stump under the body weight loads.

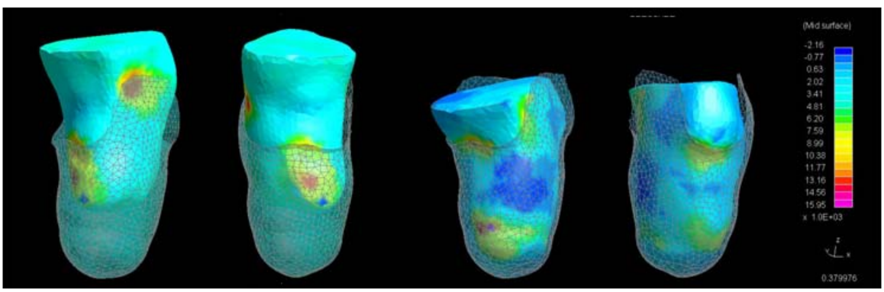


Fig. 7. Pressure distribution

For the simulations shown in Figures 6 and 7, only linear models have been used; for the bones, one with Young's modulus equal to 10 GPa, for the socket, a second linear elastic material with Young's modulus equal to 1.5 GPa, and, for the soft tissues, a third linear model with Young's modulus equal to 0.2 GPa according to [17].

4.4 Rapid Prototyping

For what concerns the rapid prototyping, the activities now are focused on the definition of the product model, using the data coming day by day from the simulation phase. In a short time, the model of the first socket designed inside the framework described in this paper will be translated into the STL format and the best rapid prototyping technology will be selected and used to generate the physical representation. We think to generate a SLA model of the socket and to replicate it using vacuum casting. As soon as these models are available, the testing phase will take place to validate the design and engineering choices and to compare the results with those ones coming from the simulations. Based on a recursive algorithm, the simulation packages will be fine-tuned, to get the result of avoid the rapid prototyping phase and generate a correct result already at the simulation stage.

5 Conclusions

This paper describes a new design process based on the use of virtual humans and specifically conceived for a prosthesis component, the socket. Different ICT technologies, from reverse engineering to physics-based simulation and rapid prototyping, must be integrated to evolve from a hand-made approach to a new one computer assisted. The proposed design paradigm has been experimented both for trans-tibial and trans-femoral amputees and main problems related to the implementation of each step have been generally discussed. Problems related to stump measurement and modeling have been completely identified, and results from acquisition and modelling of the stump and of positive plaster cast provided useful guidelines for socket modelling and physics-based simulation. However, even if results obtained encourage us to proceed, some issues are still open and require further efforts, especially for the socket of trans-femoral amputee. Future research activities will be concentrated in particular on 3D socket modelling, the evaluation of more proper nonlinear models of material for highly deformable tissues, (e.g., hyperelastic and viscoelastic models). Finally, the feasibility and usefulness of simulations using "virtual human" will be considered and to estimate the effects of prostheses modifications

Aknowledgements. The authors would like to thank Massimiliano Bertetti, Daniele Bonacini and Grazia Magrassi from Politecnico di Milano for their invaluable research work on stump modeling, Barbara Motyl from Università di Udine and, Federico Rotini and Francesco Saverio Frillici from Università di Firenze.

References

[1] http://www.idef.com
[2] Schreiner, R.E., Sanders, J.E.: A Silhouetting Shape Sensor for the Residual Limb of a Below-Knee Amputee. IEEE Transactions On Rehabilitation Engineering 3(3) (1995)
[3] Santosh, G.Z., Sanders, J.E., Turkiyyah, G.M.: Automated Hexahedral Mesh Generation from Biomedical Image Data: Applications in Limb Prosthetics. IEEE Transactions On Rehabilitation Engineering 4(2) (1996)
[4] Ming, Z., Mak, A.F.T., Chung, A.I.K., Chung, K.H.: MRI investigation of musculoskeletal action of transfemoral residual limb inside a prosthetic socket. In: Proceeding 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Part 5/6, October 29-November 1, 1998, Hong Kong, China, pp. 2741–2743 (1998)

[5] Shuxian, Z., Wanhua, Z., Bingheng, L.: 3D reconstruction of the structure of a residual limb for customising the design of a prosthetic socket. Medical Engineering & Physics (27), 67–74 (2005)
[6] Nedel, N.P., Thalmann, D.: Real Time Muscle Deformations using Mass-Spring Systems. In: Computer Graphics International (1998)
[7] Lee, W.C.C., Zhang, M., Jia, X.: Load transfer mechanics between trans-tibial prosthetic socket and residual limb dynamic effects. Journal of Biomechanics 37, 1371–1377 (2004)
[8] Lee, W.C.C., Zhang, M., Jia, X., Cheung, J.T.M.: Finite element modelling of the contact interface between trans-tibial residual limb and prosthetic socket. Medical Engineering & Physics (26), 655–662 (2004)
[9] Tönuk, E., Silver-Thorn, M.B.: Nonlinear viscoelastic material estimation of lower extremity residual limb tissues. Journal of Biomechanical Engineering 126(2), 289–300 (2004)
[10] Cheng, T.K., Sin, P.L.V., Fye, T.K., Lin, L.S.: Automation of prosthetic socket design and fabrication using computer-aided-design/computer-aided-engineering and rapid prototyping techniques. In: Proceeding First National Symposium on Prosthetics and Orthotics. Singapore (1998)
[11] Freeman, D., Wontorcik, L.: Shererolithiography and prosthetic test socket manufacture: a cost/benefit analysis. Journal of Prothestics and Orthotics 10(1) (1998)
[12] Mollemans, W., Schutyser, F., Cleynenbreugel, J.V., Suetens, P.: Tetrahedral mass spring model for fast soft tissue deformation. LNCS, vol. 2673(1), pp. 145–154. Springer, Heidelberg (2003)
[13] Faustini, M.C., Neptune, R.R., Crawford, R.H.: The quasi-static response of compliant prosthetic sockets for transtibial amputees using finite elements methods. Medical Engineering & Physics N. 28, 114–121 (2006)
[14] Tönuk, E., Silver-Thorn, M.B.: Nonlinear elastic material property estimation of lower extremity residual limb tissues. IEEE Transaction On Neural Systems And Rehabilitation Engineering 11(1), 43–53 (2003)
[15] Zheng, Y.P., et al.: State – of – the – art methods for geometric and biomechanical assessments of residual limbs: a review. Journal of Rehabilitation Research and Development. 38(5) (2001)
[16] Colombo, G., Bertetti, M., Bonacini, D., Magrassi, G.: Reverse Engineering and rapid prototyping techniques to innovate prosthesis socket design. In: Proceeding SPIE-IS&T Electronic Imaging, San Diego, CA, vol. 6056, 60560P (2006)
[17] Lee, W.C.C., Zhang, M., Jia, X., Cheung, J.T.M.: Finite element modelling of the contact interface between trans-tibial residual limb and prosthetic socket. In: Medical Engineering & Physics N. 26, 655–662 (2004)

Simulating Cancer Radiotherapy on a Multi-level Basis: Biology, Oncology and Image Processing

Dimitra D. Dionysiou[1], Georgios S. Stamatakos[1], and Kostas Marias[2]

[1] In Silico Oncology Group, Laboratory of Microwaves and Fiber Optics, Institute of Communication and Computer Systems, School of Electrical and Computer Engineering, National Technical University of Athens, Iroon Polytechniou 9, GR-157 80 Zografos, Greece
dimdio@esd.ece.ntua.gr, gestam@central.ntua.gr
[2] Biomedical Informatics Laboratory, Institute of Computer Science (ICS), Foundation for Research and Technology - Hellas (FORTH), Vassilika Vouton, P.O. Box 1385, 71110 Heraklion, Crete, Greece
kmarias@ics.forth.gr

Abstract. Tumour growth and response to radiotherapeutic schemes is a markedly multiscale process which by no means can be reduced to only molecular or cellular events. Within this framework a new scientific area, i.e. *in silico* oncology has been proposed in order to address the previously mentioned hypercomplex process at essentially all levels of biocomplexity. This paper focuses on the case of imageable glioblastoma mulriforme response to radiotherapy and presents the basics of an essentially top-down modelling approach, aiming at an improved undestanding of cancer and at a patient-specific optimization of treatment.

Keywords: Radiotherapy, Modelling, Glioblastoma, *In silico* oncology.

1 Introduction

Tumour growth and response to radiotherapy is a markedly multiscale process of obvious clinical importance, which spans from the atomic level, where the primary interaction of radiation with matter takes place, to the organism level of which the survival constitutes one of the most important goals of radiotherapy. Within this framework a new scientific area emerges, *in silico* oncology, a complex and multiscale combination of sciences and technologies in order to simulate malignant tumour growth and tumour and normal tissue response to therapeutic modalities at all levels of biocomplexity.

The *In Silico* Oncology Group, National Technical University of Athens, has adopted an essentially "top-down" modeling approach and developed a number of *hybrid* discrete Monte Carlo / cellular automata and continuous differential equation simulation models of tumour growth and response to therapeutic modalities. The models range from tumour growth and radiotherapy response *in vitro* to the clinical tumour response to radiotherapeutic and chemotherapeutic schemes *in vivo* [1]-[8].

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 569–575, 2007.

The aim is to better understand cancer and related phenomena and to optimize therapeutic interventions by performing *in silico* (on the computer) experiments based on the individual data of the patient. In this paper the case of imageable glioblastoma multiforme tumour response to radiotherapy will serve as a paradigm in order to present the basics of this modelling philosophy.

Obviously cancer is far from being a purely deterministic phenomenon. Instead it seems to behave like a mixture of deterministic (e.g. sequence of cell cycle phases) and stochastic (e.g. radiation cell kill probability) processes. Subsequently, stochasticity aspects should always be taken into account. Nevertheless, the more critical knowledge becomes available, the more deterministic the cancer phenomenon appears to become. An illustrative example supporting this hypothesis is that more detailed knowledge of the genetic status of a tumour may lead to a better prediction of its response to therapeutic interventions, thus to an apparently more deterministic tumour behaviour.

2 Outline of the Basic Features of the Simulation Model

In the following paragraphs a brief outline of the simulation model is presented. For a detailed description refer to -. The model is based on the clinical, imaging, histopathologic, and molecular data of the patient. The clinician delineates the tumour and its metabolic subregions on the available imaging data by using a dedicated computer tool (Fig.1).

A prototype system of quantizing cell clusters included within each geometrical cell (GC) of a discretizing mesh covering the anatomic area of interest lies at the heart of the proposed simulation approach. Cell cycle phase durations and imaging based metabolism distribution define i.a. the quantization equivalence classes considered.

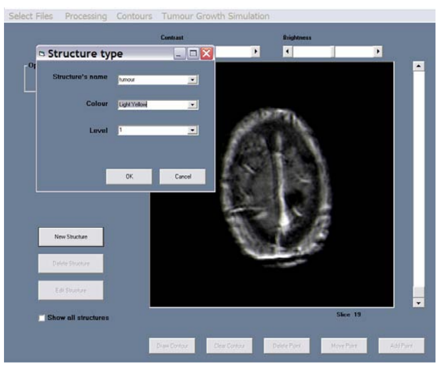


Fig. 1. A dedicated computer tool for the delineation of structures of interest upon the imaging data

In each time step of the simulation the geometrical mesh is scanned and the updated state of a given GC is determined on the basis of a number of algorithms, simulating various macroscopic mechanisms such as tumour expansion/shrinkage and mechanical boundary conditions as well as the effect of irradiation on the tumour under consideration, which can be briefly summarized as follows:

Each GC of the mesh belonging to the tumour initially accommodates a Number of Biological Cells (NBC). Each equivalence class in each GC corresponds to the phase in which its cells are found, within or out of the cell cycle: G1, S, G2, M, G0, Necrosis, Apoptosis (for details on the adopted cytokinetic model refer to [2]).

Cell killing by irradiation is described by the Linear Quadratic or LQ Model:

$$S(D) = \exp[-(\alpha D + \beta D^2)] \tag{1}$$

where S(D) is the surviving fraction after a (uniform) dose D (Gy) of radiation to a population of cells. The parameters α (Gy^{-1}) and β (Gy^{-2}) are called the radiosensitivity parameters of the LQ model. Processed molecular data can be used in order to *perturb* the radiobiological cell kill parameters about their population based mean values. The LQ model parameters constitute one possible way to incorporate the influence of genetic determinants, such as the p53 gene status, into the simulation model. In general, the model offers the advantage of readily adapting parameters related to the status of genetic determinants such as the p53 gene (LQ Model parameters, possibility of apoptosis, G_1 phase duration etc.). Once the prognostic value of specific genetic data becomes well established, the status of the relevant genetic indicators can be easily incorporated into the simulation model, leading to its substantial refinement.

The simulation of tumour expansion or shrinkage is based on the following rules:

In case that the actual number of alive and dead (but still existing) tumour cells contained within a given GC is reduced to less than NBC/10, then a procedure which attempts to “unload” the remaining biological cells in the neighboring GCs takes place. In case that at the end of the unloading procedure the given GC becomes empty, it is assumed to disappear from the tumour. An appropriate shift of a chain of GCs, intended to fill the “vacuum”, leads to tumour shrinkage.

On the other hand, if the number of alive and dead cells within a given GC exceeds NBC+NBC/10, then a similar procedure attempting to unload the excess cells in the surrounding GCs takes place. In case that the unloading procedure fails to reduce the number of cells to less than NBC+NBC/10, then a new GC emerges. Its position relative to the “mother” GC is determined using a random number generator. An appropriate shifting of a chain of adjacent GCs leads to the expansion of the tumour.

3 The Image Analysis Perspective

The most crucial step from the image analysis perspective is the extraction of relevant information, i.e. the use of some kind of process (e.g. segmentation) to identify important structures and features in the images (e.g. tumours can be segmented using a pharmacokinetic model of gadolinium uptake with Contrast-Enhanced MRI, CE-MRI). In order to produce a 4D prediction of tumour response to therapy it is essential to identify the region of interest (i.e. segment the tumour from the background tissue) and if possible, estimate different properties or tissue sub-categories within the tumour (e.g. differentiate necrotic from proliferating tissue, dense/highly vascularised tissue from fat, etc.) on the images prior and post treatment. For contrast enhanced data (e.g. CE MRI), the fact that malignant tumours exhibit an increased vascularity, since they begin to grow their own blood supply network, can be exploited. For this reason when the contrast agent is distributed, malignant masses enhance faster.

Recently, a methology has been proposed that can be applied along with the standard clinical practice associated with dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) [9], with the aim of optimizing the imaging technique so as to provide more accurate and precise tumour information, which will permit a higher level confidence in the validation of the *in silico* model.

Another point of importance concerns the so-called geometrical normalization, which refers to the establishment of correspondences in temporal imaging data in order to better assess changes [10]. Geometrical changes occur in soft tissue imaging but also in newer applications such as molecular imaging and microarray imaging. Non-rigid alignment or registration is required to compensate for such differences. This problem does not only pertain to the multi-modal scenarios, as a temporal acquisition of the same modality will still likely involve registration in order to facilitate comparison.

Dedicated 2D, 3D and 4D visualization techniques can be used to obtain detailed and intuitive representations of the tumours under study. The visualization of the region of interest can provide an enhanced 3D/4D picture of the biological problem. Virtual cuts of the tumour and the adjacent anatomic area can reveal the inner structure of the tumour and normal tissue under consideration as well as their cytokinetic activity distribution.

4 Clinical Validation, Indicative Simulations and Results

The simulation model of imageable glioblastoma response to radiotherapy has already been clinically validated to a substantial degree [8], by performing a series of simulations corresponding to various arms of the RTOG study 83-02 [11]. The observed agreement of the simulation results with the results of the RTOG 83-02 study implies that the simulation model has successfully captured and integrated the critical biological aspects that determine the clinical outcome. Furthermore, it strengthens the proposition that advanced cancer integrative radiobiology might explain and predict the therapy outcome despite the complexity of the clinical setting.

Fig.2 presents characteristic comparative simulation results, for a typical glioblastoma with mutant p53 gene, irradiated according to three different radiotherapy schedules: a standard fractionation scheme (2 Gy once a day, 5 days a week, 60 Gy in total), an accelerated fractionation scheme (2Gy twice a day, 5 days per week, 60 Gy in total) and a hypofractionation scheme (6Gy once a week, 60 Gy in total). In these simulations, crucial parameters such as the α and β parameters of the LQ model, the cell cycle duration, the clonogenic cell density and the cell loss factor have been adjusted according to the pertinent literature for glioblastomas [12]-[17].

These simulation results are biologically reasonable and in agreement with clinical experience. The glioblastoma tumour with mutant p53 gene is so radioresistant that all

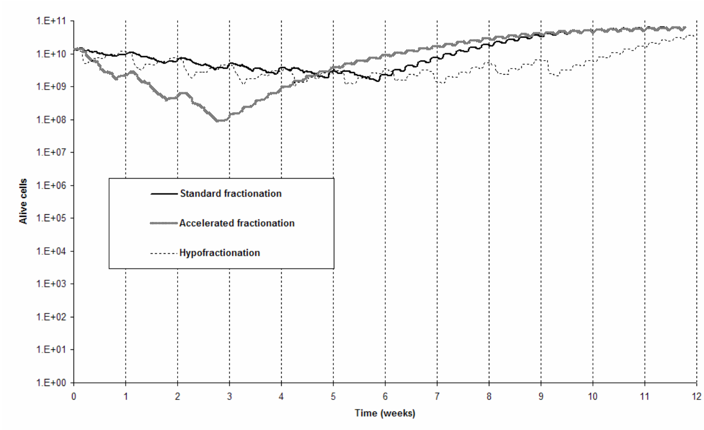


Fig. 2. Indicative radiotherapy simulation results. The number of surviving tumour cells as a function of time from the start of radiotherapy, for a typical glioblastoma with mutant p53 gene, irradiated according to three different radiotherapy schedules: a standard fractionation scheme (2 Gy once a day, 5 days a week, 60 Gy in total), an accelerated fractionation scheme (2Gy twice a day, 5 days per week, 60 Gy in total) and a hypofractionation scheme (6Gy once a week, 60 Gy in total).

radiotherapy schedules fail to hinder clonogenic cells from rapidly proliferating during therapy. This type of comparison of different fractionation schemes for specific model parameters can be quite meaningful. The accelerated fractionation scheme seems to be particularly efficient in terms of tumour cell kill. It achieves maximum cell kill at a specific instant compared to the other schedules. Nevertheless, its duration is smaller and as a result, if it fails in eradicating “all” tumour cells, as is the case for the particular tumour considered, tumour repopulation begins earlier. At the other extreme, hypofractionation is advantageous in terms of the duration of tumour control, but achieves less tumour cell kill. Of course, in clinical practice the choice of the appropriate radiotherapy schedule depends both on the expected tumour cell kill and the expected normal tissue complications.

Obviously, experimental and clinical feedback should always be used in order to improve the reliability of the model. Long term clinical testing and adaptation procedures are in process, by comparing the model “predictions” with clinical data before, during and after a radiotherapy course. In parallel, all the involved phenomena are constantly being studied, in order to keep pace with the ever-accumulating scientific knowledge.

Currently, a substantial extension of the *In Silico* Oncology Group simulation models to the nephroblastoma (Wilm's tumor) and breast cancer cases is being performed within the frame of the European Commission funded project ACGT (Advancing Clinico-Genomic Trials on cancer), in collaboration with several clinical and technological institutions across Europe.

References

1. Zacharaki, E.I., Stamatakos, G.S., Nikita, K.S., Uzunoglu, N.K.: Simulating growth dynamics and radiation response of avascular tumour spheroid model validation in the case of an EMT6/Ro multicellular spheroid. Computer Methods and Programs in Biomedicine 76, 193–206 (2004)
2. Dionysiou, D.D., Stamatakos, G.S., Uzunoglu, N.K., Nikita, K.S.: A computer simulation of in vivo tumor growth and response to radiotherapy: new algorithms and parametric results. Computers in Biology and Medicine 36, 448–464 (2006)
3. Dionysiou, D.D., Stamatakos, G.S.: Applying a 4D multiscale in vivo tumor growth model to the exploration of radiotherapy scheduling: the effects of weekend treatment gaps and p53 gene status on the response of fast growing solid tumors. Cancer Informatics 2, 113–121 (2006)
4. Dionysiou, D.D., Stamatakos, G.S., Uzunoglu, N.K., Nikita, K.S., Marioli, A.: A Four Dimensional In Vivo Model of Tumour Response to Radiotherapy: Parametric Validation Considering Radiosensitivity, Genetic Profile and Fractionation. J. theor. Biol. 230, 1–20 (2004)
5. Stamatakos, G.S., Dionysiou, D.D., Zacharaki, E.I., Mouravliansky, N.A., Nikita, K.S., Uzunoglu, N.K.: In silico radiation oncology: combining novel simulation algorithms with current visualization techniques. In: IEEE Proceedings: Special Issue on "Bioinformatics: Advances and Chalenges vol. 90, pp. 1764–1777 (2002)
6. Stamatakos, G.S., Antipas, V.P., Uzunoglu, N.K.: Simulating chemotherapeutic schemes in the individualized treatment context: The paradigm of glioblastoma multiforme treated by temozolomide in vivo. Comput Biol Med. 36, 1216–1234 (2006)
7. Stamatakos, G.S., Antipas, V.P., Uzunoglu, N.K.: A spatiotemporal, patient individualized simulation model of solid tumor response to chemotherapy in vivo: the paradigm of glioblastoma multiforme treated by temozolomide. IEEE Transactions on Biomedical Engineering 53(8), 1467–1477 (2006)
8. Stamatakos, G.S., Antipas, V.P., Uzunoglu, N.K., Dale, R.G.: A four dimensional computer simulation model of the in vivo response to radiotherapy of glioblastoma multiforme: studies on the effect of clonogenic cell density. British Journal of Radiology 79, 389–400 (2006)
9. Zacharopoulou, F., Marias, K., Georgiadi, E., Tollis, G., Maris, T.G.: Optimized MR Imaging methology for tumour characterization. In: Proc. 2nd International Advanced Research Workshop on In Silico Oncology, Kolympari, Chania, Greece, September 25-26, 2006, pp. 46–47 (2006)
10. Marias, K., Margaritis, Th., Zacharopoulou, F., Georgiadi, E., Maris, T.G., Tollis, G., Behrenbruch, C.P.: Multi-level analysis and information extraction considerations for validating 4D models of human function. In: Proc. 2nd International Advanced Research Workshop on In Silico Oncology, Kolympari, Chania, Greece, September 25-26, 2006, pp. 48–50 (2006)
11. Werner-Wasik, M., et al.: Final report of a phase I/II trial of hyperfractionated and accelerated hyperfractionated radiation therapy with carmustine for adults with supratentorial malignant gliomas. Cancer 77, 1535–1543 (1996)
12. Haas-Kogan, D.A., Yount, G., Haas, M., Levi, D., Kogan, S.S., Hu, L., Vidair, C., Deen, D.F., Dewey, W.C., Israel, M.A.: p53-dependent G1 arrest and p53 independent apoptosis influence the radiobiologic response of glioblastoma. Int. J. Radiat. Oncol. Biol. Phys. 36, 95–103 (1996)

13. Kocher, M., Treuer, H., Voges, J., Hoevels, M., Sturm, V., Mueller, R.P.: Computer simulation of cytotoxic and vascular effects of radiosurgery in solid and necrotic brain metastases. Radiother. Oncol. 54, 149–156 (2000)
14. Jones, B., Dale, R.G.: Mathematical models of tumour and normal tissue response. Acta Oncol. 38, 883–893 (1999)
15. Hegedues, B., Czirok, A., Fazekas, I., Babel, T., Madarasz, E., Viscsek, T.: Locomotion and proliferation of glioblastoma cells in vitro: statistical evaluation of videomicroscopic observations. J. Neurosurgery 92, 428–434 (2000)
16. Nakajima, M., Nakasu, S., Morikawa, S., Inubushi, T.: Estimation of volume doubling time and cell loss in an experimental rat glioma model in vivo. Acta Neurochir 140, 607–613 (1998)
17. Tribius, S., Pidel, A., Casper, D.: ATM protein expression correlates with radioresistance in primary glioblastoma cells in culture. Int. J. Radiat. Oncol. Biol. Phys. 50, 511–523 (2001)

Applied User Performance Modeling in Industry – A Case Study from Medical Imaging

Marcela Esteves, Tobias Komischke, Salome Zapf, and Antje Weiss

Siemens Corporate Research User Experience
Princeton, NJ 08540, USA
{marcela.esteves,tobias.komischke,salome.zapf.ext}@siemens.com
Siemens Medical Solutions
Information & Knowledge Management, Customer Relationship Management
91052 Erlangen, Germany
antje.peters@siemens.com

Abstract. To determine ways of improving a user interface – so that a person's task produces fewer errors or takes less time – is a major goal of user interface designers. Usability testing is an established and proven method for revealing these user performance metrics, but is rather time-consuming, resource intensive and requires at least functional prototypes. Therefore, it may not always be the optimal choice during the development of very complex, expensive and context-specific applications like those in medical imaging. A promising alternative is to simulate user performance with computational models based on psychological models. In this paper the *Performance Modeling Inventory (PMI)*, which was developed based on Card, Moran and Newell's (1980) Keystroke Level Model for estimating user performance with medical imaging systems, is introduced for the first time.

1 Introduction

Compared to consumer products, investment goods are different in one aspect that impacts the role of usability engineering: buyer and user are not the same person. The product is acquired by an organization and made available to internal or external users. So, whereas in the case of a consumer product like an MP3 player, the user has the decision to buy the product (and upon dissatisfaction also to return it), an individual user of an investment good typically has a rather limited voice in the buying decision.

For both product categories, the assumed and real user experience influences the buying decision, and the overall usability target criteria defined in ISO 9241-11 (1997), namely effectiveness, efficiency and satisfaction, apply. However, owing to the high costs involved (such as investment and maintenance costs), an organization like a hospital that wants to buy a Computed Tomography scanner may greatly benefit from usability metrics as inputs for their buying decision, as those usability metrics can be converted into business metrics like payback or return on investment (Bias & Mayhew, 2005).

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 576–585, 2007.

For example, if, through the improved usability of a medical imaging system, the average processing time for one patient can be reduced by one minute (these are example numbers), the hospital can process more patients and calculate the monetary benefit right away. Thus, usability engineering can be a powerful sales argument in the area of investment goods. The challenge is how to generate meaningful usability metrics.

Usability metrics such as user task completion time[1] or problem solving performance can be assessed by different means. Once a functional prototype exists, user testing can reveal performance times and error rates as well as qualitative aspects. As precise measurement of quantitative data can be difficult; commercial software like, for instance, Morae© by Techsmith can help capture mouse movements, task times and errors in a quantitative way (for a review of Morae, see Shrestha, 2007). However, in early stages of developing user interface concepts or when comparing different products, a functional prototype may not exist.

For such situations, the use of computational models, which are based on psychological principles, show promise. The first modeling approaches Keystroke Level Model ("KLM;" Card, Moran & Newell, 1980) and GOMS ("Goals, Operators, Methods, Selection Rules;" Card, Moran & Newell, 1983), introduced the "Model Human Processor" with the recognize-act cycle of the cognitive processor as the basic principle of operation. Whereas KLM and GOMS enable rapid predictions of quantitative parameters for simple tasks, later models like Soar (Newell, 1990), EPIC (Kieras & Meyer, 1997) and ACT-R (Anderson & Lebiere, 1998) account for more complexity and include aspects like 'for instance' rule-based actions, multi-tasking and learning processes.

All of the aforementioned cognitive models can be applied by using software tools, which are designed in order to assist researchers and user interface designers to estimate usability metrics. These applications are usually freeware and can be downloaded from the respective researcher's web page (e.g., Carnegie Mellon University's CogTool, 2006). Nonetheless, the use of these tools requires a certain amount of design experience, as, for example, the storyboard for analysis might have to be implemented by using HTML or may need other advanced design formats.

The options described above for analyzing user performance data differ in time and effort needed to carry them out. User testing is a rather elaborate way of collecting data collection method, as not only a functional prototype has to be designed, but the process of testing also takes time. Computer based modeling tools require only a mock-up and a pre-defined standard workflow; but even though their software typically claims to be easy to use, it requires installation and training before being applicable.

In industry, the development of new applications is often characterized by a tremendous pressure on development times. The competitor's products may have to be analyzed within the same time frame as the new prototype is developed. The analysis and development cycles overlap and the complexity of many systems makes it hard to design early prototypes solely for user testing. What is needed is a more

---

[1] We define "task" as the aim of achieving a specified goal and "workflow" as actions performed during the task.

efficient way of using the modeling approach, so that user interface designers can quickly fill in the needed parameters for calculating cognitive model algorithms of the system.

2 Medical Imaging and Its Specific Requirements for the Quantification of User Performance

Medical imaging is often understood as the production and interpretation of internal images of the human body. Examples of medical images include X-ray, Computer Tomography, and Magnetic Resonance images of specific parts of the body. Radiologists interpret those images and document the inferred diagnosis in a report, later distributed to the requesting physician and to the patient.

Medical images are typically accessed by radiologists through a RIS/PACS (Radiology Information System/Picture Archive and Communications System) system, which replaces hard copy formats (i.e., films) of reviewing and managing images. The RIS component allows users to retrieve patient data and create reports, whereas the PACS component allows users to navigate and interact with images. A typical radiologist's working environment comprises one landscape-oriented RIS monitor, 15 inches or larger, and at least two portrait-oriented PACS monitors, 21 inches or larger. With this equipment, they usually read between 30 and 60 cases of varying complexity a day.

User performance while working with RIS/PACS systems is critical, as the quality of patient healthcare is greatly improved by the report turnover time. In addition, because reports are reimbursed by health insurance companies, the financial return of the hospital or private practice is directly connected to the number of reports created, and to how fast they are delivered.

Improving the overall usability of RIS/PACS systems allows for both improved patient care through faster responses to life threatening questions and major financial gains and return on investment for the healthcare institutions. Therefore, a major goal for usability engineering in the area of medical imaging is to gather information about user task performance times for the most common workflows.

3 Recent Methodological Advances in Medical IT

Previous attempts to use metrics to measure user task performance for medical IT systems include the UTAPAM (User Task Performance Assessment Method) method (Komischke, 2005). It assumes that the performance of a user carrying out a task can be described quantitatively by five easy-to-assess parameters: mouse clicks, scrolling, ASCII input, screen transitions and reading load. Unlike many methods that are based on observations, UTAPAM neither requires an authentic context of use nor a running software or real users carrying out tasks. A user interface specification, along with a task collection, an appropriate service catalogue definition, and adequate knowledge about the users' mental model, is sufficient.

UTAPAM was applied during the development of computerized physician order entry systems (Komischke, 2005) and a radiology information system (Freymann, in press) in order to quantitatively analyze their respective efficiency, compared to other products or concepts.

The results obtained from UTAPAM cannot be as precise as those from more sophisticated quantitative methods that use an authentic context of use or that are model-based (like KLM and GOMS); yet as an approximation, UTAPAM was proven to be very helpful in comparing different software – including in very early stages of the development process.

The experience with using UTAPAM in industry shows that this method is easy to manage and apply. Other than subjective ratings or vague qualitative findings, UTAPAM yields quantitative data that can be presented in an easy-to-understand way. The results (e.g., radar diagrams) also proved valuable for customer presentations to benchmark the products against other versions or competitors.

One limitation of UTAPAM is its lack of consideration of execution times. The method neither takes into account the time it takes a user to perform her mental and motor actions, nor the response time of the system.

In order to address this shortcoming, we applied parameters from the KMS/GOMS (Card et al., 1980) and the Model Human Processor (Card et al., 1983) to create the Performance Modeling Inventory (PMI) – a set of guidelines and templates for calculating performance metrics. The objective was to provide user interface designers with an easy way to carry out user performance estimates without special software requirements. The core of the PMI is a simple Excel template with pre-defined formulas calculating task times for different user actions. The PMI was first designed specifically for the task of analyzing user performance with medical applications. It differentiates between three types of actions: navigating, typing and reading. Figure 1 shows the formulas we adapted from Card et al.'s (1980) human processor model to calculate execution times for the three different action types.

All actions require mental preparation time, consisting of an effort for the initial eye movement and for perceptual and cognitive processing. Navigating and typing

Action Type	Formula for Task Time (in seconds)										
Navigation	Tnav	=	TM	+	TH	+	TP*NM	+	TK*NC	+	TR
		=	0.4	+	0.4	+	1.1*NM	+	0.2*NC	+	0.5
Keystrokes	Tkey	=	TM	+	TH	+	TK*NC	+	TR		
		=	0.4	+	0.4	+	0.2*NC	+	0.5		
Reading	Trea	=	TM	+	TP*NM						
		=	0.4	+	1.1*NM						

Fig. 1. PMI Task Time Calculation Formulas. With: Tnav, Tkey, Trea (task time for different action types), TM (constant for effort of mental preparation), TH (constant for effort of homing hands on keyboard/mouse), TP (constant for executing mouse/eye movement), NM (number of mouse movements), TK (constant for effort of one mouse click/keystroke), NC (number of mouse clicks/keystrokes), TR (constant for system response time), NE (number of eye movements). All constants derived from Card et al. (1980, 1986). System response time estimated.

additionally include time for homing the hands on the keyboard or mouse, and are also assigned constants for clicking or typing effort as well as for system response time (adapted for medical imaging software from Williams and Buehler, 1997). Reviewing actions only require mental preparation and an initial eye movement, whose effort is estimated in accordance with mouse movement efforts, to take into account reading efforts.

To derive a task time estimate, user interface designers only need a workflow description with the right sequence of steps for the task to be analyzed, and an image of the corresponding user interface (screenshot or mock-up). They then need to enter the number of mouse/eye movements and keyboard/mouse entries in the Excel template in order to calculate the formulas described above. To support this step, we provide an optional PowerPoint template, where mouse/eye movements can be counted as the number of lines of a grid to be passed between two successive steps of a workflow. Figure 2 shows a generic PMI PowerPoint template, which uses a Microsoft Word document as an example. It is used by pasting an image of the user interface and by marking the different steps of a workflow with numbered tags. As software in medical imaging runs on multiple monitors, a 4x4 grid is used for each screen, although other grid options are provided for single monitor software with a higher density of screen elements.

This step can be skipped for simple user interfaces or short workflows, as its single purpose is to support the visualization of the various workflow steps. The core of the

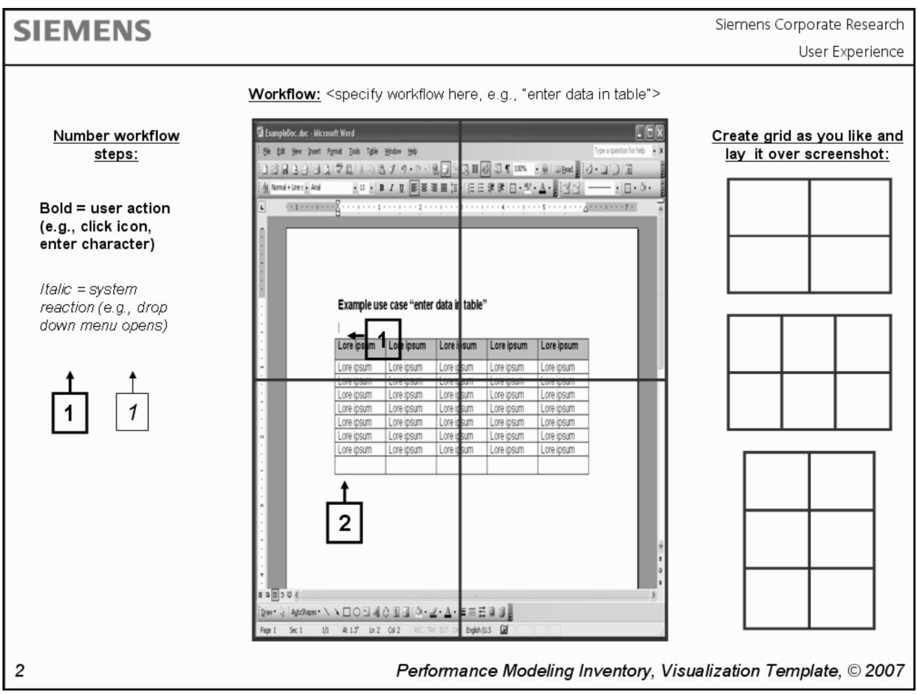


Fig. 2. PMI Visualization Template

PMI_Template_Calculation.xls

	A	B	C	D	E	F	G	H
1	**Enter task here, e.g. "edit document"**	**Steps**	**Name of system**	**mouse movement**	**eye movement**	**# of Mouse clicks/ key strokes**	**time per task in seconds**	**Notes** (e.g. features improving/jeopardizing workflow)
2	**Enter workflow here, "e.g. "open file from Explorer"**	**<could insert number of steps here manually>**		**0**	**0**	**2**	**1.7**	**all time per task fields in grey are sums of the time spent on substeps!**
3		number steps, e.g. "1.", "2." etc.	define steps, e.g "double click file"	0	0	2	1.7	Example calculation for navigation action (mouse movements and click)
4	**Enter workflow here, e.g. "enter data in table"**	**<could insert number of steps here manually>**		**0**	**0**	**7**	**2.7**	
5		number steps, e.g. "1.", "2." etc.	define steps, e.g. "enter *'loreley'* in free cell"	0	0	7	2.7	Example calculation for typing action S(keyboard entry)
6	**Enter workflow here, e.g. "review table"**	**<could insert number of steps here manually>**		**0**	**1**	**0**	**1.5**	
7		number steps, e.g. "1.", "2." etc.	define steps, e.g. "review information"	0	1	0	1.5	Example calculation for reading action (only eye movement)
8								

General_Calculation / Competitor_Calculation / Comparison_Calulation

Fig. 3. PMI Calculation Template

PMI is the above mentioned Excel template, where the steps of the workflow(s) to be analyzed have to be entered. Figure 3 shows the template with the generic Microsoft Word example.

The template provides the user interface designer with sample fields and formulas for the three action types, navigating, typing and reviewing. Row 3 contains the formulas for navigation actions, row 5 those for typing actions and row 7 those for reading actions. To start a new analysis, the correct row for the workflow's steps has to be copied and pasted. For the analysis, the name and steps of each workflow have to be specified for the pasted template rows in columns A and B. The number of mouse and/or eye movements has to be entered in columns C and D as the number of lines passed on the PMI Grid between the end point of the previous step and the starting point of the current step. After filling in the number of mouse clicks or key strokes, the task time estimate is automatically calculated and can be summed up over different steps of one workflow. Every important cell in the template has an explanatory comment of how to use it, so that the documentation doesn't need to be consulted while working with a case. The templates can be adapted for comparisons of two systems without needing to know anything more than how to operate Excel. An application example of how PMI was used in medical imaging is described in the following section.

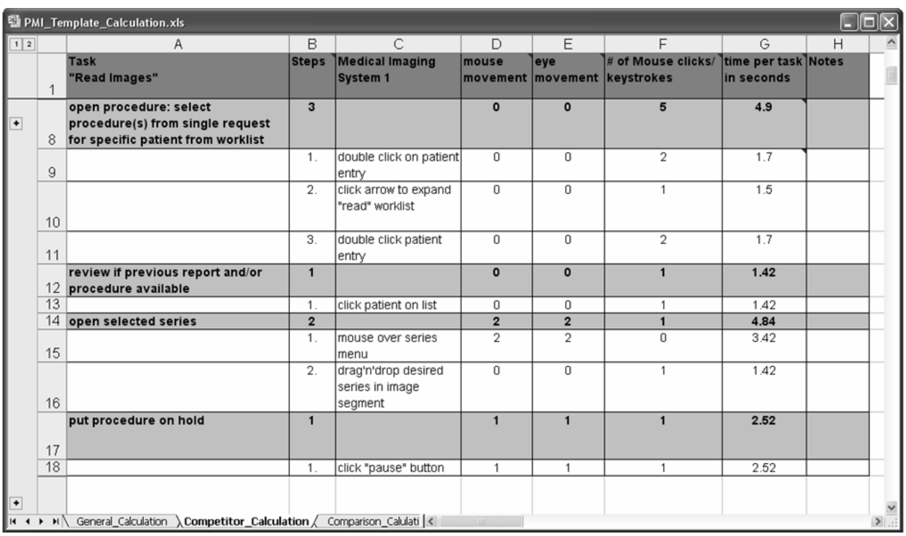
4 Case Study: Competitor Analysis in Medical Imaging

In the course of improving the user interface for a new release of a medical imaging system for radiologists, a competitor analysis, which also required quantifications of user performance, was carried out. Various analysis steps prior to using the PMI revealed the existence of 52 workflows for one of the most relevant tasks of the system. For these 52 workflows "real life" mock-ups of the competitor system were created and tagged with a workflow description for each task. The same documentation was carried out for the competing Siemens system, so that, in the end, each workflow was visualized by numbered tags on the various screenshots and mock-ups for both systems. At this stage of the process, the PMI visualization template was used as a way not only to prepare for the user performance modeling analysis, but to provide workflow visualizations concurrently for development, sales and management.

To carry out the main performance analysis, the resulting visualization files were used to fill the PMI calculation template with content. This was easily achieved by adding rows for each step in one workflow and by counting the number of mouse clicks as well as the number of grid lines crossed by the mouse or eye. Given that, in medical imaging, some workflows require a transition between RIS and PACS monitors, we adapted the PMI calculation template accordingly by simply adding a column for the additional system. Mouse movements between the two systems were then quantified by adding one more movement to the number of all grid lines to be crossed on both monitors. We filled one Excel template for each system and then created a new comparison sheet by duplicating all columns of the PMI calculation template and referencing to the different systems' calculations. All in all, the completion of the analyses took only one day for 104 workflows (52 for each of the two systems analyzed), with each workflow taking only about 3 minutes to quantify.

Results

Using the PMI during the competitive analysis prooved to be very useful for user interface design decisions. Not only did the process of the analysis take no more than one day, but its results also revealed several patterns. Figure 4 shows an example of one of the PMI templates used to estimate user performance metrics for one competitor. For presentation purposes, the majority of the analyzed workflows are collapsed in this view, so that only four example workflows are visible. Also, the above mentioned second column for transitions between the RIS and PACS systems is collapsed, so that only examples from navigating within one monitor are shown. What can be seen here is how task times are composed of mouse movements and clicks. Some workflows are time consuming mainly to the number of clicks required; other workflows take long, because they necessitate extensive mouse movement.



PMI_Template_Calculation.xls

	A	B	C	D	E	F	G	H
1	**Task "Read Images"**	**Steps**	**Medical Imaging System 1**	**mouse movement**	**eye movement**	**# of Mouse clicks/ keystrokes**	**time per task in seconds**	**Notes**
8	**open procedure: select procedure(s) from single request for specific patient from worklist**	**3**		**0**	**0**	**5**	**4.9**	
9		1.	double click on patient entry	0	0	2	1.7	
10		2.	click arrow to expand "read" worklist	0	0	1	1.5	
11		3.	double click patient entry	0	0	2	1.7	
12	**review if previous report and/or procedure available**	**1**		**0**	**0**	**1**	**1.42**	
13		1.	click patient on list	0	0	1	1.42	
14	**open selected series**	**2**		**2**	**2**	**1**	**4.84**	
15		1.	mouse over series menu	2	2	0	3.42	
16		2.	drag'n'drop desired series in image segment	0	0	1	1.42	
17	**put procedure on hold**	**1**		**1**	**1**	**1**	**2.52**	
18		1.	click "pause" button	1	1	1	2.52	

General_Calculation | Competitor_Calculation | Comparison_Calulati

Fig. 4. PMI calculation template filled in for Medical Imaging System

This pattern is even more obvious when compared between different competitors. In our overall analysis, the two systems considered showed different task times in only 20 out of 52 workflows. This number alone facilitated the decision about prioritization of front end design issues. The decision regarding which workflows needed to be addressed how urgently was further facilitated by the observation that one competitor required generally both more clicks and more mouse movements for the workflows, whereas the task times of the other competitor mainly comprised only large numbers of mouse clicks. Therefore, one requirement derived for optimizing the user interface was to decrease the required mouse-travel. Furthermore, the close association between workflows and usability metrics revealed qualitative patterns. For these patterns, workflows with a shared characteristic (e.g., allowing the user to deviate from the default system behavior) were associated with very similar metric result (e.g., high time per task). These associations were of great relevance in understanding the user interface concept strategies of the systems. The qualitative patterns revealed by the metrics also made evident how specific user interface components (e.g., access to tools) impact the overall system usability, with special regards to time per task.

5 Conclusions

In spite of the effort involved in visualizing the workflows, making use of the PMI for user performance estimates in medical imaging systems had several advantages. As only Microsoft Office software was needed in order to work with the PMI, neither additional time for installation nor effort for getting to know how to use a new tool were required. The individuals using the templates felt that these were easy to use, and carrying out the necessary calculations did not take much time. The calculation template could furthermore be adapted to compare performance of two different systems without any special programming knowledge. Another advantage of using this technique is the format of the analysis results. As stakeholders are familiar with Microsoft Excel, the tables created, as well as the charts, are easy to read and not only show averaged task times but give visual clues as to how the different times are constituted. This permits an inference regarding the effect of possible user interface design trade-offs.

One constraint of the PMI, when used in domains other than medical imaging, is the simplification of constants used in the formulas. Although derived from validated models (Card et al, 1980, 1986), the PMI applies average values as constants instead of calculating exact values for each step. This is especially relevant for estimating mouse travel times, because most other software tools use Fitt’s law (Fitts, 1954) to estimate travel times based on exact measurements between starting and ending point of the movement (e.g., John & Salvucci, 2005). As, in medical imaging, users work with multiple, portrait-oriented 21-inch monitors, the estimate of 1.1 seconds as assumed by Card et. al (1980) for moving the mouse from one part of a 4x4 grid to another seems appropriate. However, for other applications with only one screen, this constant should be adapted.

To address this and other shortcomings, some modifications of the PMI could be beneficial. The calculated estimates could be compared with time measurements from the real context of use (hospital setting, real radiologists working with the real systems). This would yield a measure of validity of the model. Furthermore, domain-specific templates for other industries, like, for example, supervisory control or telecommunication systems, need to be designed, and their applicability has to be evaluated. For adjusting the formulas to different domains, other constants (e.g., system response time) as well as different types of input devices (e.g., stylus, special control panels) must be taken into account. Finally, different levels of user expertise could be considered, as the current method assumes an optimal expertise level, which may be justified for users of medical imaging systems, but might not be applicable for domains with a higher diversity among users. For competitive analyses, the assumption of an optimal user yields more accurate results because it evens out the different experience levels of real users.

The successful application of the PMI reviewed in this paper suggests that tailoring templates to different domains is a worthwhile effort toward providing user interface designers with an analysis tool to support well-founded user interface design decisions.

References

Anderson, J.R., Lebiere, C.: The atomic components of thought. Erlbaum, Hillsdale, NJ (1998)

Card, S.K., Moran, T.P., Newell, A.: The keystroke-level model for user performance time with interactive systems. Communications of the ACM 23(7), 396–410 (1980)

Card, S.K., Moran, T.P., Newell, A.: The psychology of human-computer interaction. Erlbaum, Hillsdale, NJ (1983)

Card, S., Moran, T.: User technology—from pointing to pondering. In: White, J.R., Anderson, K. (eds.) Proceedings of the ACM Conference on the History of Personal Workstations, Palo Alto, California, January 09–10, 1986, pp. 183–198. ACM Press, New York (1986)

Bias, R.G., Mayhew, D.J. (eds.): Cost-justifying usability. Elsevier, San Francisco (2005)

Carnegie Mellon University. The CogTool Project. Overview (2006) (Retrieved January 25, 2007) http://www.cs.cmu.edu/ bej/cogtool/index.html

Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology 47, 381–391 (1954)

Freymann, M.: Validierungsmethode UTAPAM: Einsatz und Nutzen im User-Centered Design Prozess (Validation method UTAPAM: Application and Benefit in the User-Centered Design Process). i-com, Zeitschrift für interaktive und kooperative Medien. Munich: Oldenbourg (in press)

ISO 9241. International Standards Organisation DIS 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs). Part 11: Guidance on usability. Beuth: Berlin (1997)

John, B.E., Salvucci, D.D.: Multi-purpose prototypes for assessing user interfaces in pervasive computing systems. IEEE Pervasive Computing 4, 27–34 (2005)

Kieras, D.E., Meyer, D.E.: An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. Human-Computer Interaction 12, 391–438 (1997)

Komischke, T.: Usability engineering in the area of computerized physician order entry. In: Proceedings of HCI International 2005, vol. 6, CD-ROM. Lawrence Erlbaum Associates, London (2005)

Newell, A.: Unified theories of cognition. Harvard University Press, Cambridge, MA (1990)
Shrestha, S.: A Review of Morae© for Usability Testing. Usability News, 9 (2007) (Retrieved January 29, 2007) http://psychology.wichita.edu/surl/usabilitynews/91/MoraeAnalysis.html
Williams, M.G., Buehler, J.N.: A study of program entry time predictions for application-specific visual and textual languages. In: Wiedenbeck, S., Scholtz, J. (eds.) Papers Presented At the Seventh Workshop on Empirical Studies of Programmers, Alexandria, Virginia, pp. 209–223. ACM Press, New York (1997)

Computer Aided Lumbar Support Design and Application

Susanne Frohriep and Jan Petzel

Leggett & Platt Automotive Europe, Frankenstrasse 150A, 90461 Nuremberg, Germany
Susanne.frohriep@lpautomotive.eu, Jan.petzel@lpautomotive.eu

Abstract. Support properties in the lumbar area have been identified as a major factor for the comfort impression on automotive seats. However, the relationship of human body characteristics to specific properties of seats is not clear. Ongoing research at L&P Automotive Group analyzes seat dimensions and human body characteristics in order to evaluate the man-machine interface of the seated human. Computer tools employed are 3D measuring systems, pressure mapping and subjective comfort assessment. Specific measurement routines have been developed to allow for correlation of human dimensions and seat characteristics. Using computer supported methods for measuring, data collection and data processing in combination with engineering knowledge and knowledge about other soft facts, comfort assessment processes can be developed and effectively used in order to obtain better, meaning more comfortable products.

Keywords: lumbar support, comfort assessment, automotive seating, man-machine interface.

1 Introduction

In today's automotive industry, the market is becoming tighter, and developing differentiations has become increasingly difficult between competitors. In the vehicle interior, many components define the vehicle's character, one of the most important for the comfort impression being the seat [1], which may function as a differentiator.

Seats provide a range of functions, some of which contradict each other. A seat has to support the load of the passenger, meaning loads between 20kg representing a child without safety seat and 150kg representing a heavy adult. At the same time, the seat should feel soft and comfortable for the entire range of passengers. With almost all other features of the seat, similar contradictory relations can be found.

Therefore, it is necessary to define reasonable assumptions and restrictions for seat design specifications. One of these criteria is seating comfort. Seating comfort is a multi-dimensional phenomenon. It is yet unclear, which elements compose the totality of the seating comfort impression. Yet, single dimensions can be measured and assessed, and are used in the important process of optimizing single comfort aspects. These can be considered as measurable hard facts. Besides these hard facts, many soft facts influence the perception of comfort and the overall comfort assessment, such as

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 586–594, 2007.

preconceptions or conditioning of the test subjects. In many cases, these soft facts cannot be measured; they have no values or units. An assessment is only possible by using the human being as the measuring gauge. Still, these soft facts influence the human perception of the hard facts.

These complex relations between the various aspects of seating comfort have not yet been resolved by computer solutions, but with single dimensions computer technology offers valuable support.

2 Comfort Parameters and Measurement Methods

Measuring an automotive seat can encompass a range of parameters that all will have some influence on subjective comfort. Which ones constitute the major factors determining the total comfort perception is a matter of ongoing discussion. One proven fact is that the entire comfort impression is rarely better than the worst parameter. This means, e.g. if an otherwise perfect seat had bolsters that were too narrow and therefore uncomfortable, the whole seat would be rated to the level of uncomfortableness of these bolsters. We call this the limiting comfort factor.

Engineers are used to utilizing fixtures for keeping vehicle components in place. There are numerical descriptions and simulation methods for designing and optimizing these connections (e.g. pertaining to loads, vibration etc.). For designing seats, the occupant constitutes the complex interface. Often, if one parameter is improved, another changes for the worse: For instance, if a hard sitting surface is made softer, it might be more comfortable, but the position of the driver might change because of altered deformation parameters, leading to a more difficult operation of the vehicle. Predefined tools seem to have only a limited use in this design process. Instead, grown expert knowledge is put to use in vehicle seat design and optimization, which is often considered proprietary to the applicant. Therefore, failsafe guidelines in this field are difficult to compile.

In seating comfort research, some parameters have been identified as major influences on comfort perception, being pressure distribution on the sitting surface [2], and relations of dimensions and angles of the seated human in the surface geometry of the seat [3]. In the methods section, computer technology employed by L&P Automotive to further lumbar support development and design will be discussed.

2.1 Pressure Distribution Measurements

The body surface of the occupant and the corresponding pressure values can be considered as a numerical description of the human-seat interface. The distribution of pressure values on the contact surface is probably the most important comfort factor of the seat. Pressure distribution in the man-machine interface is measured by pressure mats. There are various manufacturers, and pressure mapping has become a wide-spread tool for seat assessment. The contact surface is covered with a grid of x times y measuring cells delivering values into a matrix with differing resolutions for different mat types. The result is a pressure map showing zones of pressure values in different colors. Such a map does not translate directly into a comfort assessment [4], and therefore, the matrix of values is processed in different ways in order to reduce complexity. One of the standard processes is to define body zones and integrate the values in the specific zones. Other methods use one row of sensors as a two-dimensional cut through the matrix and consider the curve to represent the comfort perception. Variations of the above methods exist, applying additional steps, but none of them can directly measure comfort. Still, the measurements can be used to specify the development target, and that marks their great value for seating technology. In this field, currently there is no reverse engineering, i.e. it is not yet possible to define a desired pressure distribution and deduce a seat design from it.

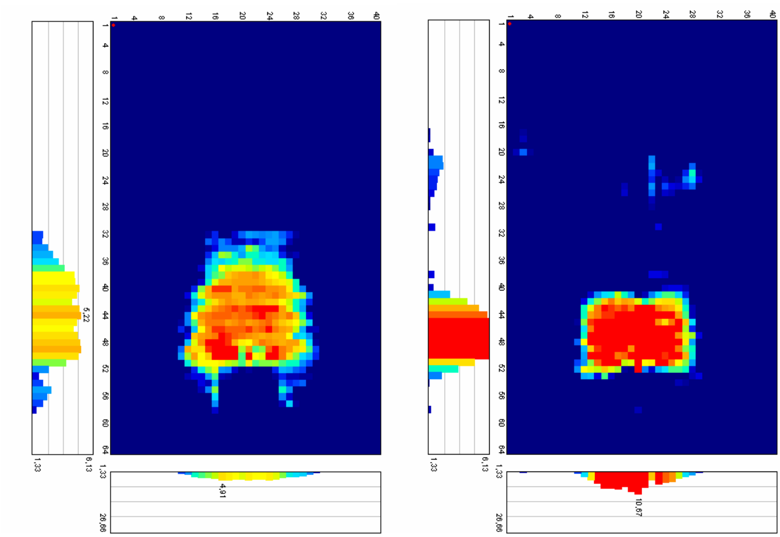
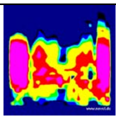
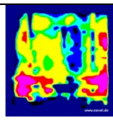


Fig. 1. Output of pressure distribution measurements, cross-section views. Left: Pressure map output of a protruded lumbar basket system showing the "spinal relief" function of the lumbar basket where lower pressure values are generated directly on the spine. Right: Pressure map output of an air bladder lumbar with a simple rectangular bladder exhibiting high load on the spine.

In seat assessment, static pressure mapping in a mock-up or seating buck is mainly employed for rating seats. Dynamic pressure mapping can show the response of the occupant in a driving situation and assess dynamic comfort. In comparing two types of active bolster systems on a test track (July 2006; L&P research), dynamic pressure mapping was able to pinpoint differences (see table 1). An electromechanical system (1) and a pneumatic system (2) were compared for four test subjects in 23 driving situations. While the average pressure was higher for system 1, the peak pressure was higher for system 2, illustrating the fact that the pressure occurred more localized in the pneumatic system and more widespread in the electromechanical system. Furthermore, the standard deviation for all values was higher for the pneumatic system, pointing to a varied response compared to a more constant response of the electromechanical system. Linking to subjective comfort assessment, system 1 was clearly preferred over system 2 due to its rapid actuation and reliable body support.

Table 1. Overview output dynamic pressure mapping, comparison of two active bolster systems for 4 test subjects in 23 driving situations

 1  2	Peak Pressure [kPa]		Average Pressure [kPa]	
	1	2	1	2
Average	17.4	18.2	4.7	3.8
Standard Deviation	2.9	7.2	0.4	1.0

Seat design blunders can be directly derived from the pressure map, for instance when areas of high pressure occur at the rim of the contact area. However, discomfort cannot be deduced from the pressure value alone. A certain amount of pressure in localized areas, for instance in the lumbar region of the back is rated positively by most test subjects [2; 5]. Details of human comfort perception can only be assessed by combining the pressure information with the subjective rating of the pressure distribution in the seat.

No international standards are available that would define a reference body to be used for measuring pressure distributions in seats. Conventionally, test subjects of the 5^{th}, 50^{th} and 95^{th} percentile are used for this purpose, though the population and measure that determine these percentiles are often not defined. For the assessment of seat back pressure distribution, sitting height and body weight should be the measures determining the selected percentiles. Some developers use the SAE-HPM I manikin as a standard body. While this yields comparable data sets for different seats, it does not represent any class of human occupant, because it consists of rigid surfaces that have nothing in common with human occupant pressure maps. Some companies have developed soft dummies to replace the rigid shells of the HPM. This constitutes a more realistic method, but none has been standardized so far.

A side result from the pressure measurement is the contact surface size (with some measurement devices). The surface size can also be used for comfort assessments in connection with micro-climate evaluation.

2.2 Assessment of Seat Dimensions

A vehicle seat needs to fit the driver as well as any piece of clothing s/he would wear. If the dimensions of the seat are not suitable for the individual driver, this seat will ultimately not satisfy any customer, even if all other aspects were ideal.

The geometry of the seat is measured by 3D measurement machines using optical or mechanical sensors. The outcome is three-dimensional objects such as points, curves or even surfaces of the seat contour. Out of the sum of possible measurements, a useful subset needs to be chosen. For comfort features of the seat, the apex height of the lumbar support, its vertical travel and horizontal adjustability as well as special seat features (e.g. adjustable bolsters, see fig. 4) constitute highly relevant measures. With the 3D-data, the adjustment ranges of movable parts of seats can be depicted, and the general fit of the seat can be related to human body measures. Furthermore, this data can be used for quality checks as to the initial design data of the CAD systems.

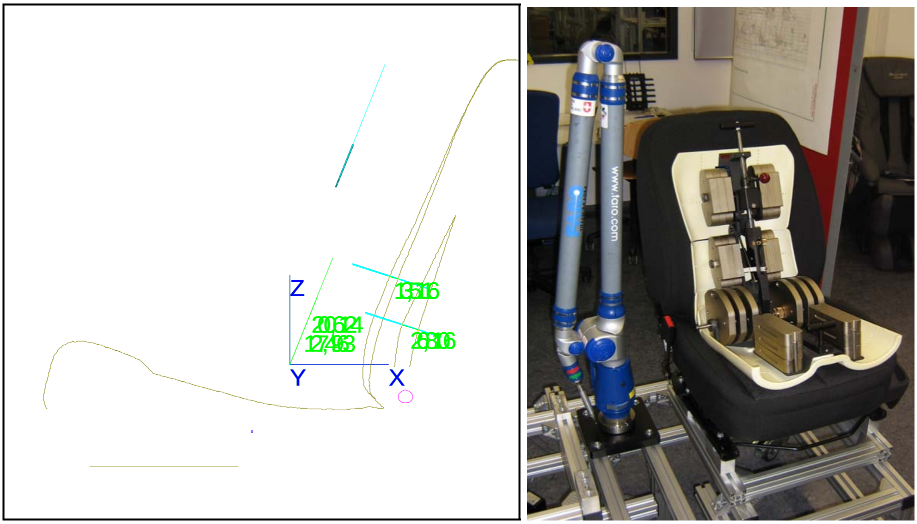


Fig. 2,3. Output of 3D measurement of a seat / Measurement setup on seating buck with 3D measurement device. The standard measurement includes A-surface scans of seat cushion and backrest, in the conditions lumbar retracted and lumbar protruded, measurement of lumbar support contour in retracted and protruded state, as well as HPM II divot points, manikin backline, and seat elements for further CAD alignment such as seat pivot.

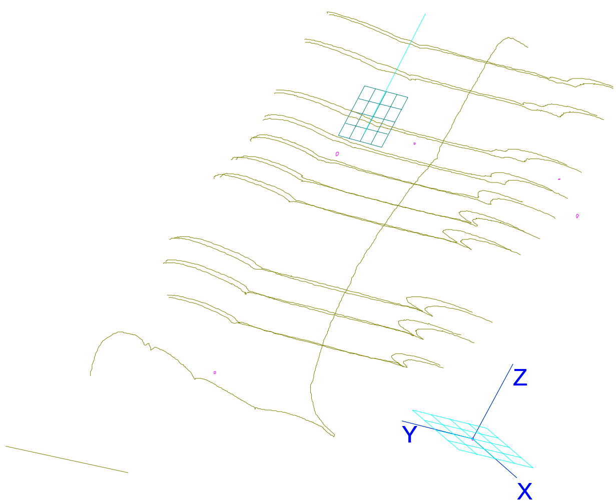


Fig. 4. Output of 3D measurement of a seat, measuring adjustable bolsters on predefined planes

The results of the 3D assessment can be related to subjective comfort ratings in order to build up expert knowledge on the effect of seat design elements. Also, they can be related to specific body data of user populations. Another major function is to specify seat characteristics for the development phase.

2.3 Assessment of Subjective Comfort with the ASDQ

The assessment of subjective comfort is performed through questionnaires. Designing such questionnaires is a complex matter. Single questions need to be relevant, easy to understand and based on the accepted rating culture. Furthermore, the rating system should be neutral and not influence the values rated. For data gathering and evaluation, computer use is a powerful assistance that ensures continuity and safeguards against data loss and transfer errors.

With increasing sample size, the questionnaire method becomes one of the most powerful tools in ergonomic assessment. If surveys are always executed according to an identical procedure, a growing database and growing experience turn the questionnaire and the group of test subjects into extremely useful and economic tools.

At Leggett & Platt, a questionnaire for analyzing seating discomfort was developed, the Automotive Seating Discomfort Questionnaire (ASDQ) [6]. It makes use of visual analogue scales (VAS), which provide accurate quantitative estimates of pain and well-being, factors involved in rating discomfort. The ASDQ consists of 20 questions that are grouped in sub-sets to rate different elements of the seat: trim, cushion, backrest and lumbar. In the testing involved for developing the ASDQ, it was proven that individual ratings of discomfort remained consistent over time and the ASDQ could show them with a significant level of reliability. While the original questionnaire was performed on paper by placing slashes through lines, the current version has been adapted for computer use (see fig. 5). The underlying database allows for comparison of different seats tested with all or sub-groups of test subjects (see fig. 6).

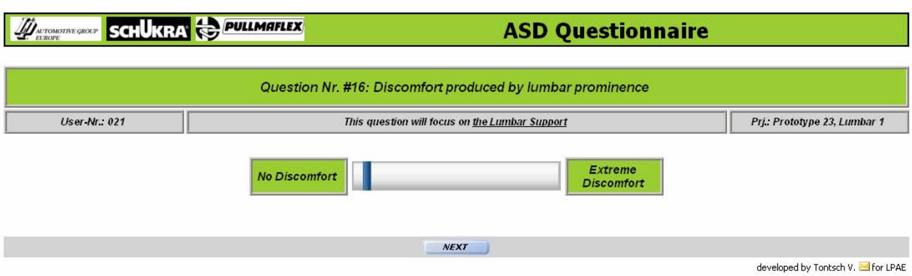


Fig. 5. The ASDQ employs a VAS slider bar for a direct quantitative estimate of seat characteristics. The slider is shown on a computer screen and moved by pulling it with the mouse.

Gathering data in various projects for different seats and configurations allows for a wide-based data comparison between individual discomfort ratings and assessed seat characteristics as dimensions and resulting pressure distribution.

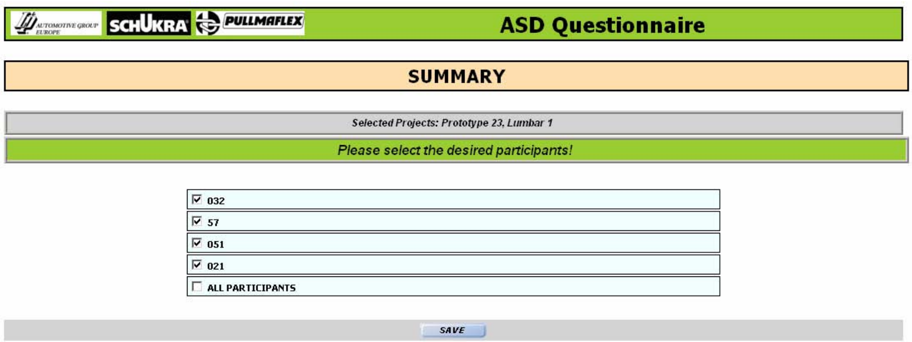


Fig. 6. Menu for choosing participants in the ASDQ analysis

3 Analysis of Occupant Accommodation

The compilation of data gathered for seats needs to be correlated to the seated occupant. Body measures corresponding to specific seat comfort elements are typically lacking in standard body measure data. Furthermore, there is a dispute about advantageous body postures linked to the prevalence of (lower) back pain in the driving population assumed to result from unfavorable postures. A posture that feels comfortable to the occupant does not necessarily represent a posture that is healthy, as is known for extreme kyphotic postures (hyperflexion) in seated positions which lead to ligament creep and loading of structures [7].

3.1 Body Measures and Accommodation

In order to attain body measures of occupants relating to specific comfort elements, Leggett & Platt continues to perform research in-house and in cooperation with external institutions. Measurements of occupants have been performed in a car seat specifically prepared for this purpose to allow for measurement of the human back during sitting. The height of occupants' lumbar vertebrae 1, 3 and 5 were assessed under the condition of lumbar support (a) retracted and (b) protruded, and heights above seating reference point (SRP) as well as lumbar spine angles were analyzed. Data emerging from the lumbar spine study were especially lumbar spine length, heights of vertebrae 1, 3, 5 and lumbar spine apex angle while seated in a car seat.

The resulting range of L3 heights above SRP amounted to 100mm, the average height above SRP being 150mm. Therefore, the working hypothesis was to design in such a way that the apex of the lumbar support would lie at 150mm above SRP. However, it was not clear whether the comfort impression would match the target of L3 height or would differ from it. Numerous follow-up studies in specific seats have shown that the most comfortable position is slightly lower than the average value, probably more on the height of L4. This height not only allows for a support of the lumbar spine, but also of the pelvis, which is essential for maintaining a well-supported back contour.

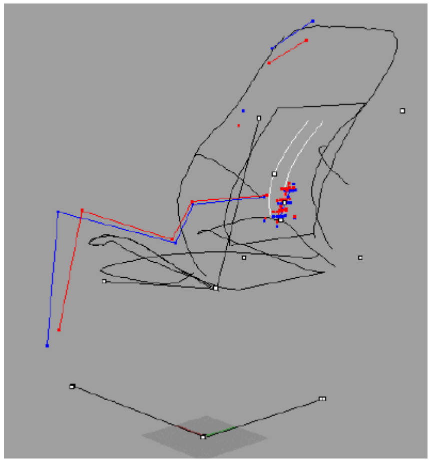


Fig. 7. Assessment of lumbar spine characteristics of occupants in a car seat, ex. Friess, p. 20

3.2 Posture and Accommodation

The prevalence of back pain among the driving population points to the problem of posture in car seats. The adjustment of seats is left to the occupants, who decide according to their individual comfort impression. Yet, comfortable and healthy postures are not always similar. The agreement among researchers is not high; though some facts are undisputed: a) hyper-flexion is the most damaging posture b) sitting posture should be dynamic, not static [7]. Other than that, there is disagreement, even about seemingly basic concepts such as whether sitting or standing is the more strenuous posture for the human spine.

Comparing standing and sitting position, sitting causes flexion of the spine due to pelvis rotation, flexion increasing as backrest angles increase. As it is agreed that excessive flexion is harmful to the spine, methods for safeguarding against it should be employed. Adjustable lumbar supports can prevent excessive flexion without leading to high extension (lordotic posture), as Reed has already demonstrated in 1996 [8]. Leggett & Platt research studies have shown that lumbar spine apex angles in fully protruded lumbar position averaged at 173°, thus a mildly extended posture. In practice, test subjects tend to slide their pelvis forward during prolonged sitting (driving), thus being able to reach a straight or slightly flexed spinal contour while using an adjustable lumbar support.

3.3 Comfort in Vehicle Seats

Ultimately, the posture and movement patterns of human occupants in vehicle seats during their driving life in daily use will determine the long-term comfort of a seating system. In developing seating systems, usage should be modeled to allow for assessing (dis)comfort in the design phase.

In order to arrive at a comfort assessment for a specific seat, available information is pieced together and evaluated, combining specific results and background knowledge. This process is circular, as there is no defined starting point and information feeds into it continuously, with factors influencing each other. Assessing seat dimensions will yield a first notion about accommodation of user groups, which is enhanced by pressure maps and rounded off by questionnaires.

Dynamic seat elements as active bolsters, active lumbar supports or massage functions need to be evaluated under dynamic conditions. For this, computer support is invaluable, advancing assessment from a mere subjective miscellany to quantifiable and demonstrable results.

4 Future Perspectives

In practical seat development, it is most likely impossible to quantify all elements factoring into comfort evaluation. Therefore, it is highly unlikely that one computer tool which can perform a complete comfort evaluation will ever be available. Relating questionnaire results to measured parameters of the seat is the key to repeatable and reliable comfort assessments and even to comfort predictions.

Using all computer supported methods for measuring, data collection and data processing in combination with our engineering knowledge and knowledge about soft facts, comfort assessment processes can be successfully developed and effectively used to obtain better, meaning more comfortable products. This is similar to any other learning process of the brain, except for the fact that if we manage to install this procedure in our organizations, this knowledge is not lost with the passing of time. Ergonomic knowledge will be stored in a company memory system and can be developed further by each new generation of operators.

References

1. Andreoni, G.: A Methodological Approach for the Analysis of the Car Driver's Posture. SAE Technical Paper 1999-01-1927, pp. 1–5 (1999)
2. Mergl, C.: Entwicklung eines Verfahrens zur Optimierung des Sitzkomforts auf Automobilsitzen. Herbert Utz Verlag (2006)
3. Babirat, D., et al.: Körpermaße des Menschen – Komfortbereiche der Gelenkwinkel der Körpergelenke. Forschungsbericht der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin Fb 818, BAUA Dortmund (1998)
4. Gyi, D., Porter, J.: Interface Pressure and the Prediction of Car Seat Discomfort. Applied Ergonomics 30, 99–107 (1999)
5. Dhingra, H., et al.: Discomfort, Pressure Distribution and Safety in Operator's Seat – A Critical Review. Agr. Eng. Int. V, 1–16 (2003)
6. Smith, D.R., Andrews, D.M., Wawrow, P.: Development and Evaluation of the Seating Discomfort Questionnaire (ASDQ). Int. J. Industrial Ergonomics 36/2, 141–149 (2006)
7. Wilke, H.-J. (ed.): Ergomechanics 2 – Second Interdisciplinary Conference on Spinal Column Research. Shaker Verlag, Aachen (2006)
8. Friess, M.: Analysis of the Lumbar Spine for the Design of Lumbar Supports in Car Seats: Anthropometric Parameters, Internal Research Report, Anthrotech (2004)
9. Reed, M.: Lumbar Support in Auto Seats: Conclusions from a Study of Preferred Driving Posture. SAE Technical Paper 960478, pp. 19–28 (1996)

Facilitating Pronunciation Skills for Children with Phonological Disorders Using Human Modelling

Jennifer George[1] and Paul Gnanayutham[2]

[1] SAE Institute, United House, North Road, London, N7 9DP, United Kingdom
[2] Department of Computing, University of Portsmouth, Buckingham Building, Lion Terrace, Portsmouth PO1 3HE, United Kingdom
jennifer.george@sae.edu, paul.gnanayutham@port.ac.uk

Abstract. Children are often forced into mainstream schools while special needs schools are being shut down. Children with phonological disorders in mainstream schools go through fear, shame and frustration of not being understood or misunderstood. The proposed research attempts to address this issue by way of designing and developing a prototype for an assistive device that would help special needs children in mainstream education. This will also help children overcome problems that result from lack of confidence. This work has been very much a pioneering work and has achieved the target it pursued.

Keywords: Phonological disorders, speech disability, pronunciation, interface, sounds and children.

1 Introduction

The seven organs of speech are tongue, teeth, lips, muscles of the lower jaw, hard palette, soft palette and teeth ridge (also known as the alveolar ridge). When the air released from the lungs passes through the mouth, according to the organs of speech used in formation the different sounds and words are formed. Sander [1] created a Speech Sound Development Chart identifying the range of typical sound development in children according to their age. When unusual patterns of this developmental process occurs either in able or disabled children the advice of Speech and Language Therapists (SLT) is sought. These unusual patterns may be diagnosed as either disordered or simply delayed [2]. Disordered speech may occur due to poor auditory discrimination, difficulty in forming sounds, structural difficulties, verbal dyspraxia or hearing difficulties [3]. These disorders on its own can be mild, moderate or severe and on the other hand be connected to other disorders. It may also be as a result of basic communication disorders. The child's chronological age and developmental level may differ according to the mental and biological growth of each child. Buckley, [4] states that the main difficulties experienced by children with Downs Syndrome can be categorised as difficulties in hearing, auditory perception and processing, difficulties in clear speech production and greater difficulty in learning grammar than vocabulary. Children with learning disabilities need more high quality learning opportunities in order for them to learn and remember the forming of sounds, yet they

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 595–605, 2007.

get much less opportunities because of their slower progress. This tool used multimedia to create an application that any support worker or parent will be able to use with minimum training with the children to enhance communication with their peers. Having established the need for a device to enhance pronunciation, the research problem was "*should there be a group of interaction paradigms or one novel interaction paradigm that can be personalised, to enhance the performance of pronunciation skills for children with speech impairments?*" was explored. The investigation was carried out using still images, two-dimensional and three-dimensional animations and evaluated with focus group and individual participants with speech impairments.

2 Assistive Devices

Many multimedia interfaces have been developed as both Augmentative Alternative Communication (AAC) and Assistive Technology (AT) to aid speech disabilities [5]. Children with severe communication impairments use alternative communication modes such as manual signs and visual symbols, which is a substitution for spoken words [6]. However, most of these devices include a speech recognition element. When children have difficulty in pronouncing clearly it is not possible for these devices to assist them. This research will only be dealing with AT as this encourages speech and does not attempt to replace it with alternative methods.

3 Methodology

The design, development and experiments were carried out in two cycles and two phases (Fig. 1.). The two cycles were iterative processes that made the final versions of the group and personalised interfaces possible prior to conducting the experiments. Phase one experiment was conducted on the general interface to explore *"Should there be a group of interaction paradigms to enhance the performance of pronunciation skills for children with speech impairments or…"*and phase two experiments were conducted with the personalised interfaces to explore *"….should there be one novel interaction paradigm that can be personalised to enhance the performance of pronunciation skills for children with speech impairments."*

4 Design and Development

A focus group shown in Table 1, centred around two disabled and eight able was used to evaluate the interfaces during the design and development process. Changes were made to the interface in an iterative process. The interface to be one that can be used on a PC or MAC platform with a standard CD-DVD Rom and speakers (Fig. 2 and 3). The relevant audio clips for the animations were recorded using Pro Tools. They were recorded in both male and female voices and lasted for almost three seconds. The information as to how each sound was produced was obtained from the teaching material used by the Wendy Whatmore Academy of Speech and Drama. Jolly Phonics is a speech sound training system used by many therapists as derived from the various

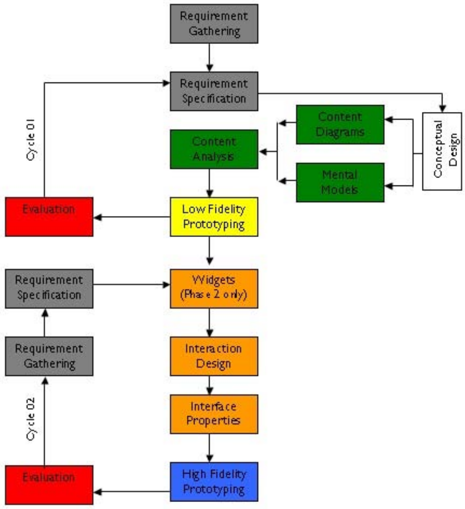


Fig. 1.

Table 1. Focus Group

Capacity	Gender	Age	Institute
Developer of interfaces for brain injured	M	51	University of Portsmouth
Sibling/Carer of adult with Down Syndrome	F	30	Somerset
Speech Therapy Consultant	F	64	Redhill, Speech and Language Therapist
Assistant Speech Therapist	F	29	Radcliffe Infirmary, Oxford
Speech Therapist	F	42	Ormerods, Oxford
Speech Therapist	F	38	Oxford City PCT
Adult with Down Syndrome	F	32	Somerset
Multimedia Developer	M	29	SAE Institute, London
Multimedia Developer	M	31	SAE Institute, London
Child Down Syndrome	M	10	Croydon

interviews. Sounds in order of frequency of use were selected for prototype and the relevant actions for these sounds were obtained from the Jolly Phonics online resources for the 2D animation.

Maxon's Cinema 4D was used due to the simplicity at the same time the ease of bone animation. This would be exported as a .mov file. As evaluation may also be conducted online and the online support available Macromedia Flash was chosen to create the main part of the interface. The interface could be exported to a CD or DVD and used for evaluation.

Facility to start again is provided as this could be used for reinforcement. The user should be able to quit the interface at any time and for this purpose, an exit button has been included. Volume could be controlled using the hardware's volume control. Relevant 2D animation to demonstrate an example of the sound is included to help the child relate to the sound. Each 3D Animation was included with the relevant sound produced to demonstrate how the organs of speech are used in formation to create sound. The 2D and the 3D animations were synchronised with the relevant audio. There was no need for a pause button as the animation would not last for more than 3-5 seconds. The start and stop buttons could be replaced with a replay button. The 3D window had to be significantly larger than the 2D window.

The Jolly Phonics actions for sounds would be used to design and develop the example animations [7].This would be exported as .swf and thereafter imported into the authoring interface. The standard system of learning sounds used by the NHS is Jolly Phonics, which consists of the 42 sounds in English. This method divides the sounds into seven groups. The first group of six sounds have been chosen according to the frequency in three-letter words compared to the second group. Similarly the less used sounds come later on in the list. When these two groups are compared to the International Phonetic Alphabet (IPA) it can be concluded that all organs of speech are being used during the learning of these twelve sounds [8]. This list has a good mixture of both consonants and the high frequency vowels. Thus, the first twelve sound. Sounds from the Jolly Phonics System s, a, t, i, p, n, ck, e, h, r, m and d were chosen. Jolly Phonics also gives actions for each sound to help the children remember letters that represent it. This idea was adopted in order to give the child a demonstration in the form of 2D animation to help the child remember and associate with the sound produced [7].

3D models have been used to demonstrate semi-transparent human models and the formation of different sounds inside the mouth. The head and shoulder of a boy and a girl with well defined parts inside the mouth were modelled. The lips, teeth, teeth ridge (alveolar ridge), tongue, hard palette, lower jaw, soft palette, Velar (back of the tongue and near the soft palette), uvula (extension at the back of soft palette) are visible.

Modelling, texture and lighting, animation and camera movement and rendering of clips using the audio files of the twelve chosen sounds are to be created in C4D. The animations are based on the organs used in formation for each of these sound as defined below. Initially, the mouth can be animated and thereafter the can be exported separately for both the boy and the girl. Each animation may last up to three seconds showing clearly the firm contact and release of the relevant organs of speech in formation.

Gimson [9] has described the production of each sound in the English language. This was used to create animation in order to distinguish, explain the formation and model the creation of each of these sounds. The letters below depict the sound that is produced and not the alphabet. The animations would always start from a relaxed position i.e. Tongue lying flat and upper teeth over lower teeth and following animation regain this position. What happens in between is described below.

Various shapes were considered for the shapes of buttons and the interface objects [10]. Rectangle was chosen having researched into Gestalt Laws and psychology of visual perception [11]. Previous investigation show that users have emotional reactions to colours and fonts, this interface gave the option for making these changes to suit any user. This study also showed that white or yellow text on blue background was more readable [12] this was taken as the default setting for the interface while lilac is to be the background colour for the individual animations. The buttons are to be yellow turning into pink on rollover. The active sound would also be a lighter yellow. The 3D animation colours are natural and do not change. Still images will be used to display the front view of the mouth in 2D form. There will be two moving images, one demonstrating the sound made using side view and the other indicating an example of the sound made. This will illustrate movement of the mouth necessary to produce the sound. Two types of audio were used for the interface. An introductory sentence with instruction to click on buttons and a thank you on exit has been used. For each demonstration of sound, the sound produced is played. The final version of the group interface was taken as the template for the personalised interface. This interface was used for experiments carried out with participants during which time participant observation and action research were undertaken.

In this section, the theory of widgets was used to design the preferences page to give personalisation functionality to the interface. Layered approach was used to design the preferences functionality of the interface. Layout changes for the main page also had to be made occasionally to accommodate the personalisation. Options were given in choosing between male and female voices, switching the 2D and 3D animations off and a choice of window size adhering to usability issues.

To cater to the visually impaired and different types of disability with different levels of distractively options in changing the window size and turning in completely off were given. The instructors voice was also an option to choose from a male of female voice. Although this interface had various interface to cater for different age groups, an option to minimise or exit the window which was distracting the user e.g. An older child may not need the 2D animation for examples.

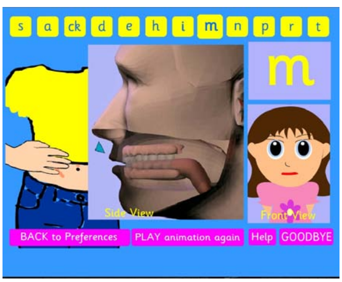


Fig. 2.

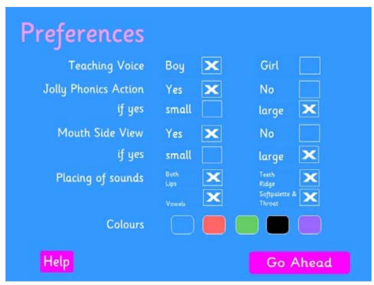


Fig. 3.

Following the development of the final interface of both group and personal, experiments were carried out to test the research question. This interface was used for experiments carried out with participants during which time participant observation and action research were undertaken.

5 Experiments

Phase one of the research lasted 6 weeks and produced the final general interface. Phase two of the research lasted 8 weeks and produced the final personalised interface. This research created the personalised interface effectively by conducting extensive experiments with a focus group of able participants before using the

Table 2. Details of the participants used in experiments

Part. No	Institute	Age	Gender	Clinical Diagnosis
1	Home	7	Male	None
2	Home	4	Male	None
3	Home	5	Female	None
4	Home	6	Female	None
5	Home	7	Male	None
6	Home	9	Male	Autism
7	Home	13	Male	Downs Syndrome
8	Home	6	Female	None
9	Home	5	Male	None
10	Home	36	Female	Downs Syndrome

interface with the disabled participants. Experiments 01 was carried out following the low fidelity prototype development evaluation, experiment 02 was carried out during phase 01 and experiment 03 was carried out during phase 02. Statistical analysis was carried out and conclusions were reached.

Table 2 gives details of the participants who evaluated the group interface. They were between ages 5 years to 36 years, 4 females and 6 males, 7 able and 3 disabled.

Experiment 01

This phase of the study investigated;

1. The number of attempts needed to pronounce each sound
2. The amount of training needed
3. The sounds that cause difficulty

Table 3. Results of experiment without interface

	Number of Attempts per Sound									
Participant	1	2	3	4	5	6	7	8	9	10
s	5	6	5	4	7	5	4	6	7	9
a	3	2	1	2	2	6	3	3	2	1
ck	5	6	7	5	5	10	10	4	8	2
d	3	4	5	4	2	6	7	3	4	1
e	6	4	3	4	7	3	4	7	2	3
h	7	5	4	3	6	10	9	6	4	8
I	2	1	4	2	3	3	2	3	4	6
m	2	1	2	2	2	2	3	2	3	3
n	6	5	6	5	4	9	8	5	6	5
p	3	4	5	4	3	7	5	4	3	5
r	3	8	9	7	3	10	10	3	7	10
t	7	6	5	5	8	9	10	8	7	6

The participants were asked to repeat the 12 sounds after the researcher (Table 3). The results obtained are shown in figure 4. Each sound had to be pronounced and the anatomy explained. A maximum of ten attempts were given for each sound by the participant. Thereafter they were indicated as '10'. Average attempts made by the participants to pronounce each sound correctly were calculated (Table 4). No training was required to carry out this experiment. Difficulties in pronouncing sounds were taken note of. Participant 6 had difficulty in repetitive attempts. Participant 7 was enthusiastic but had difficulty understanding explanations. Overall difficulty was found in pronouncing s, ck, h, n and r. Participant 10 had specific high frequency sounds such as a and m but found others confusing and difficult. Difficulty was also

Table 4. Average attempts without Interface

Sound	Average Attempts Made
s	5.80
a	2.50
ck	6.20
d	3.90
e	4.30
h	6.20
i	3.00
m	2.20
n	5.90
p	4.30
r	7.00
t	7.10

noted in distinguishing t from d and n and r. The next experiment will be conducted using the developed interface to investigate whether further progress can be made in pronunciation for this group of participants.

Experiment 02 - This phase of the study investigated whether;

1. Disabled participants can be grouped together, when developing interfaces.
2. Users can use Group Interface with minimal training.
3. Optimum setting for group interface can be obtained with focus group, which can be used as default when evaluating the interface with disabled users.

Table 5. Results for General interface

	Number of Attempts per Sound									
Participant	1	2	3	4	5	6	7	8	9	10
s	1	3	3	3	2	3	1	1	1	5
a	1	3	3	3	2	3	2	1	1	1
ck	3	2	2	2	1	2	2	1	1	2
d	1	1	3	3	1	1	3	1	1	1
e	1	3	2	2	1	1	3	3	1	2
h	1	2	2	2	1	3	2	1	1	4
I	1	1	2	2	1	1	2	1	1	2
m	1	1	2	2	1	1	2	1	1	3
n	1	1	2	2	1	4	2	3	1	4
p	2	2	3	2	2	1	1	3	2	3
r	1	1	2	2	1	5	4	2	1	9
t	2	2	3	3	2	1	3	3	1	6

Apparatus used was, Computer with CD-DVD driver and speakers, interface program. The functionality of the interface was explained to the participant by way of demonstration. It had to be explained to the participant that the 3D model in the interface was what the inside of the mouth looked like. Then the participant was asked to imitate the 3D animation for each sound. Maximum of ten attempts were given for each sound. No pre-requisites or training was necessary to carry out the experiments (Table 5). Statistical analysis was conducted from the data obtained to choose the optimum initial settings for the interface. Average of attempts made were calculated, in order to compare profiles.

T-test was conducted on the two sets of data (without any interface and with group interface). The t-test shows that the two sets of data are significantly different at

Table 6. Success Rate for General Interface

Sound	Average Attempts Made
s	2.30
a	2.00
ck	1.80
d	1.60
e	1.90
h	1.90
i	1.40
m	1.50
n	2.10
p	2.10
r	2.80
t	2.60

the 0.1% level (0.000013). Hence it can be deduced that using the group interface is 99.9% better than using no interface. In this phase of the research, participants were grouped and experiments were conducted to evaluate the group interface with minimum training. Results were recorded; data analysed and statistical analyses were carried out. Outcomes are to be compared with the results from the personalised interface in the next phase of the research.

Experiment 03

This study investigated whether;

1. Using the Personalised Interface improved the performance of the disabled user in comparison with the group interface
2. Personalised Interface can be reconfigured anytime
3. Users can use Personalised Interface with minimal training.
4. Optimum setting for Personalised Interface can be obtained with able focus group, which can be used as a starting point when evaluating the interface with the participants.

The functionality of the interface was explained to the participant by way of demonstration. Whenever necessary it had to be explained to the participant that this was what the inside of the mouth looked like. The participant was asked to imitate the 3D animation for each sound. Maximum of ten attempts were given for each sound. The participant was given a choice of colours to choose from for the background. Whenever the participant appeared distracted due to the 2D animation, the interface was used without the 2D animation or the size of the window was controlled depending on the priority of need. Statistical analysis was conducted from the data obtained to choose the optimum initial settings for the interface. Similar to experiments 01 and 02, the average attempts made were calculated, in order to compare profiles.

T-test was conducted on the two sets of data (group interface and with group interface). The t-test shows that the two sets of data are significantly different at the1% level (0.00014). Hence it can be deduced that using the personalised interface is 99% better than using the group interface.

Table 7. Results for personalised interface

	Number of Attempts per Sound									
Participant	1	2	3	4	5	6	7	8	9	10
s	2	2	3	2	1	3	2	2	2	3
a	2	3	3	3	2	4	2	2	2	1
ck	3	2	3	3	1	4	1	1	1	2
d	2	2	4	2	2	3	4	2	2	1
e	2	2	3	3	2	3	4	4	1	2
h	2	3	3	3	2	5	3	2	2	3
I	2	2	3	2	1	2	3	2	2	1
m	1	2	2	3	1	1	1	1	2	1
n	2	2	3	2	2	6	3	4	2	3
p	1	3	2	3	3	3	3	1	2	3
r	2	2	3	3	2	7	5	3	2	6
t	3	3	4	4	3	3	4	4	2	6

Table 8. Success Rate of Personalised Interface

Sound	Average Attempts Made
s	2.20
a	2.40
ck	2.10
d	2.40
e	2.60
h	2.80
i	2.00
m	1.50
n	2.90
p	2.40
r	3.50
t	3.60

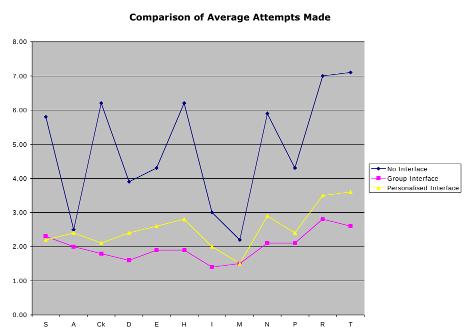


Fig. 4. Comparison of Average Attempts Made

In this phase of the research, participants individually evaluated the interface and experiments were conducted to evaluate the personalised interface with minimum training. Results were recorded; data analysed and statistical analyses were carried out. Outcomes were compared with the results from the group interface in the previous phase of the research. The comparison in figure 4 indicated that the performance of the personalised interface was much better than the group interface.

Figure 4 shows the number of attempts made to pronounce each sound using the various interfaces. The graphs shows the performance of each interface supporting the t-test results calculated earlier. The group and personalised interfaces have a clear advantage over the no interface data. However the general interface goes on to improve the pronunciation even further as indicated in the graphs above. Participants 2 and 9 were also able to pronounce the sounds without the audio indicating that the 3D component of the interface was accurately designed to indicate the actual movements of the mouth.

6 Conclusion

This research carried out a survey of AAC tools and AT devices to find out whether there were any devices that could be used by children with speech impairment to enhance their pronunciation skills so that they can communicate with more confidence in main stream schools. This need arose since the schools for special needs were being closed down and children with special needs were made to join the main stream education. The survey indicated the lack of an assistive device to enhance the pronunciation skills for this group of children with phonological disorders. There were challenges involved with developing and evaluating interfaces for children with phonological disorders, approaches chosen for this research, ethical issues involved and the structure of the investigation. Various methods and methodologies were considered and appropriate methods and methodologies were chosen for the research while others were ‘excluded by including’. Having established the need for a multimedia interface to enhance pronunciation, the research problem on the type of interface needed was answered successfully. Two interfaces were developed to answer the research question. The developed multimedia interface grouped the sounds according to the mechanism in which they are produced in the mouth and also had options in choosing the size and colour of the window and the trainer’s voice. The developed interfaces were evaluated by a focus group of ten participants. The qualitative and quantitative evaluations of the developed interfaces for group and personalised indicated that the participants of the personalised interface performed with a better success rate than the group interface as indicated by the t-test and the graph. Even if solutions are found for mental effects biological factors such as having a larger tongue in proportion to the mouth (e.g. Down’s Syndrome), may not change. Therefore, there will be a continuous need for further research in this area. Some indication of the future research is dealt with in the following paragraphs.

7 Future

Pattern languages may assist user centred design process [13][14]. One way to ensure usability is to treat human abilities as input to design rather than merely evaluating prototypes. Freeman [15] proposed a theory of brain dynamics to construct external representations by observing brain patterns. He stated that minds were nothing but activity of neurons and that thought and ideas were processed according to rules, discovered by psychologists. With an external input device for brain waves this interface can be used for the brain injured during speech therapy. Gnanayutham and his team [16] designed and developed a soft keyboard for non-verbal quadriplegic brain injured persons. Future research in this area could connect brain waves to the interface developed in this research thereby taking this research to the next stage where the brain injured non-verbal community could be integrates into this research.

References

1. Sander, E. K.: (1972) (Accessed 29th August 2006) http://stcloud.k12.mn.us/ecassessment/assets/SpeechSoundDevelopmentChart.pdf#search=%22When%20are%20speech%20sounds%20learned%20-%20Sander%2C%20Eric%20K%22
2. Lemon, K., Bucke, L.: Interviewed (18th July 2006)
3. Bowen, C.: PsychNet-UK (2005) (Accesses on 17th March 2006) http://members.tripod.com/Caroline_Bowen/ phonol-and-artic.htm
4. Buckley, S.: Improving the speech and language skills of children and teenagers with Down syndrome, Down Syndrome news and Update. The Down Syndrome Educational Trust, Great Britain, vol. 1(3), pp. 111–128 (1999)
5. AbilityHub.: (2006) (Accessed on 13th March 2006) http://www.abilityhub.com/aac/aac-devices/htm
6. Sevcik, R.A., Romski, M.A., Adamson, L.B.: Research directions in augmentative and alternative communication for preschool children, Disability and Rehabilitation, vol. 26(21/22), pp. 1323–1329, Taylor & Francis Ltd. (2004)
7. Jolly Learning: (2006) (Accessed 27th July 2006) http://www.jollylearning.co.uk/jp.htm
8. Ager, S.: (2006) Omniglot (Accessed 29th August 2006) http://www.omniglot.com/writing/ipa.htm
9. Gimson, A.C.: An Introduction to the Pronunciation of English, Edward Arnold Publications (1962)
10. Moran, S., Leiser, D.: The limits of shape constancy: point-to-point mapping of perspective projections of flat figures. Behaviour & Information Technology 21(2), 97–104 (2002)
11. Ware, C.: Information Visualization, pp. 203–213. Morgan Kaufman Publishers, USA (2000)
12. Laarni, J.: Effects of color, font type and font style on user preferences. In: Stephanidis, C. (ed.) Adjunct Proceedings of HCI International 2003, June 2003, pp. 31–32. Crete University Press, Crete (2003)
13. Maybury, M.: Intelligent user interfaces: An Introduction, Tutorial presented at MITRE Corporation (2003) http://www.1uiconf.org/03tutorial1.html
14. Mahemoff, M.J., Johnston, L.T.: Principles for a usability-Oriented Pattern Language, Ozchi. In: Australian Computer Human Interaction Conference, 1998, p. 132 (1998)
15. Freeman, W.J.: Brain dynamics in the genesis of trust as the basis for communication by representations. In: AAAI 1996, pp. 1327–1328. MIT Press, Cambridge (1996)
16. Gnanayutham, P., Bloor, C., Cockton, G.: Discrete Acceleration and Personal Titling as Brain-Body Interface Paradigms for Neurorehabilitation, CHI2005, Oregon, USA (April 2005)

FPF-SB: A Scalable Algorithm for Microarray Gene Expression Data Clustering

Filippo Geraci[1,3], Mauro Leoncini[1,2], Manuela Montangero[1,2], Marco Pellegrini[1], and M. Elena Renda[1]

[1] CNR, Istituto di Informatica e Telematica, via Moruzzi 1, 56124, Pisa, (Italy)
[2] Dipartimento di Ingegneria dell'Informazione, Università di Modena e Reggio Emilia, Via Vignolese 905 - 41100 Modena (Italy)
[3] Dipartimento di Ingegneria dell'Informazione, Università di Siena, Via Roma 56 - 53100 Siena, (Italy)
filippo.geraci@iit.cnr.it, leoncini@unimo.it, montangero.manuela@unimo.it, marco.pellegrini@iit.cnr.it, elena.renda@iit.cnr.it

Abstract. Efficient and effective analysis of large datasets from microarray gene expression data is one of the keys to time-critical personalized medicine. The issue we address here is the scalability of the data processing software for clustering gene expression data into groups with homogeneous expression profile. In this paper we propose *FPF-SB*, a novel clustering algorithm based on a combination of the Furthest-Point-First (FPF) heuristic for solving the *k*-center problem and a stability-based method for determining the number of clusters *k*. Our algorithm improves the state of the art: it is scalable to large datasets without sacrificing output quality.

1 Introduction

Personalized Digital Human Modeling is one of the new challenges at the frontier of medical practice, medical research and computer science. The relatively recent family of microarray based technologies holds the promise of personalized diagnosis based on tracing the metabolism of cells of individual patients. However, several obstacles still lay on the path to exploiting the full potential of these technologies [23]. One issue is, for instance, the scalability of the data processing software for clustering gene expression data into groups with homogeneous expression profile. In this paper we tackle this problem by proposing *FPF-SB,* a clustering algorithm based on a combination of the Furthest-Point-First (FPF) heuristic for the *k*-center problem [11] and a stability-based method for determining the number of clusters *k* [22]. The experiments we report here demonstrate that our novel algorithm is scalable to large datasets without sacrificing output quality.

The FPF heuristic is known to attain a result that is within a factor two of the optimum clustering according to the *k*-center criterion. This theoretical guarantee, coupled with a small computational complexity and with a careful implementation, makes this algorithm an ideal candidate for attaining scalability without sacrificing

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 606–615, 2007.

quality. The FPF algorithm constructs the clusters incrementally (that is the k-clustering is obtained by a refinement of the (k-1)-clustering) but it needs to know the number k of clusters in advance. For this reason, we apply a stability-based technique for cluster validation by *prediction strength* [22] that guides the selection of the "best" number of clusters in which the dataset should be partitioned. Our experiments point out that FPF-SB does maintain the original FPF scalability properties.

We apply the FPF algorithm directly to the input real data, overcoming the limitations of a similar approach [7] in which only a restricted set of artificially generated data points is used (meant to uniformly cover the parametric space of possible experimental outcomes). This approach has been shown not to be scalable.

The scalability of our algorithm can find applications in a number of different settings, in terms of both dataset dimension and number of experiments that need to be performed (hence, running time). Concerning the former, the technology of Tiling Arrays is capable of producing a complete profile of the transcript index of an individual genome (up to 50,000 transcript sequences and 60 tissues and cell lines can be mapped in a single chip experiment [17]). Such a technology, adapted towards the needs of personalized medicine, promises to be able to screen a vast range of different pathological conditions. FPF might be used in this context, where the great amount of data to be managed is one of the main issues for the existing techniques.

In some applications there is the need to repeat the experiments many times and to have a prompt result. For example, data taken at different times from the same patient in a healthy state and in a pathological state could be clustered to highlight differences in the metabolism due to the pathology, filtering out the background effects of healthy individual metabolic profile.

State of the art. The papers by Eisen et al. [6], Alon et al. [1] and Wen et al. [24] have shown the rich potential of microarray gene expression data clustering as an exploratory tool for finding relevant biological relationships amidst large gene expression datasets. Since the late nineties a growing body of knowledge has been built up on several algorithmic aspects of the clustering task (see, *e.g.*, a general survey in [15]). Among the most popular approaches we can broadly find those "distance based" such as k-means [21], Self Organized Maps (SOM) [20], Hierarchical Agglomerative Clustering (HAC) [6], and several variants thereof. A second broad class of algorithms is graph-based (CLICK [18], CAST [3] and CLIFF [25] all use weighted min cuts (with variants among them)). Other large families are those models-based [16], fuzzy-logic-based [2], or based on principal component analysis (PCA) [12]. Among the main issues for clustering there are the problem of guessing the optimal number of clusters [22] and that of cluster validation [10][26]. In biological data analysis a further issue is to provide metrics supported by ad hoc external biological knowledge [14].

There is no clear winner in the area of clustering algorithms for gene expression data as any method has strong and weak points: high quality usually results into high computational costs, while high efficiency usually results into poor quality, in both cases affecting the scalability. In our approach we show that the two goals (quality and scalability) are not in contrast: high scalability need not entail lower quality.

Microarray for digital human modeling and personalized medicine. Research on Digital Human Modeling and Simulation[1] is expected to produce models of human biology at all relevant physical scales such as, among the others, proteins, cells, and tissues. Among the computational tools that will be required by digital human models, microarray algorithms will play an important role in order to, *e.g.*, analyze diseased versus normal tissues, profiling tumors and studying gene regulation during development [13]. Also, in case of disease modeling, data extraction and processing will have to be repeated many times (say, before and after a therapy or drug exposure), which is of course another reason for the availability of very efficient algorithms. In medical practice, the adoption of patients' molecular information can give a boost to personalized medicine, and the methods that will make this possible include gene variations and expression level analysis [27]. While cost effectiveness and the availability of the data processing facilities needed by personalized medicine may be a decade ahead from now [28], the coming years will likely see an increasing number of medical protocols based on patients genotype and gene expression level information.

2 Preliminaries

Clustering. Let $N=\{e_1,..,e_n\}$ a set of n vectors of m components, a partition $\tilde{N}=\{N_1,\dots,N_k\}$ of N is a *clustering*, where each N_t, for $t\in\{1,2,\dots,k\}$, is called a *cluster*. Given a clustering $\tilde{N}$, two elements $e_i,e_j\in N$ are mates according to $\tilde{N}$ if they belong to a cluster $N_t\subseteq\tilde{N}$. Thus, mates are "similar" in a sense that may vary depending on the application.

Distance function. Given two vectors $e_i,e_j\in N$ (with components $e_{s,t}$, $s\in\{i,j\}$ and $1\leq t\leq m$), we denote with $d_{i,j}$ their distance and we say that they are similar (resp. different) if $d_{i,j}$ is small (resp. large). We use a distance measure based on the *Pearson Coefficient*, one of the most used in the context of gene expression microarray data clustering, defined as follows:

$$P(e_i,e_j)=\frac{\sum_{t=1}^{m}(e_{i,t}-\mu_i)(e_{j,t}-\mu_j)}{\sqrt{\left(\sum_{t=1}^{m}(e_{i,t}-\mu_i)^2\right)\left(\sum_{t=1}^{m}(e_{j,t}-\mu_j)^2\right)}} \tag{1}$$

where μ_i and μ_j are the means of e_i and e_j, respectively.

The Pearson Coefficient is a measure of similarity but it is not a distance. To come up with a measure suitable for the metric space method, we define $\delta_{i,j}= 1 - P(e_i,e_j)$, $0\leq\delta_{i,j}\leq 2$. Even if $\delta_{i,j}$ is widely accepted as a valid dissimilarity measure in gene expression analysis, it still violates the triangle inequality constraint. However, since $d_{i,j}$, the square root of $\delta_{i,j}$, is proportional to the Euclidean distance between e_i and e_j [5], it can be used with algorithms that need a metric space.

---

[1] http://www.fas.org/dh/publications/white_paper.doc.

The *k*-center problem. We approach the problem of clustering microarray data as the one of finding a solution to the *k*-center problem, defined as follows:

Given a set of points N on a metric space M, a distance function $d(p_1,p_2) \geq 0$ satisfying the triangle inequality, and an integer k, a k-center set *is a subset $C \subseteq N$ such that $|C|=k$. The* k-center problem *is to find a k-center set that minimizes the maximum distance of each point $p \in N$ to its nearest center in C; i.e., minimizes the quantity $\max_{p \in N} \min_{c \in C} d(p,c)$.*

The problem is known to be NP-hard, approximable within a factor 2 by FPF [11], and not approximable within a factor 2-ε for any ε >0 [8].

In our case, *N* is represented as a $n \times m$ matrix, where *n* is the number of gene probes in the dataset and *m* is the number of conditions tested on each probe, the metric space *M* is R^m where $d_{i,j}$ is the distance measure defined above. We apply a variation of the FPF algorithm computing *k* by way of the *Stability-Based* method in [22].

3 Main Algorithm

In this section, we describe FPF-SB by first presenting a variant of the FPF algorithm and then our implementation of the stability-based method for determining *k*.

Clustering algorithm. FPF is based on a greedy approach: it increasingly computes the set of centers $C_1 \subset \ldots \subset C_k$, where C_k is the solution to the problem. The first set C_1 contains only one randomly chosen point $c_1 \in N$. Each iteration *i*, with $1 \leq i \leq k-1$, has the set of centers C_i at its disposal and works as follows:

1. for each point $p \in N \setminus C_i$ compute its closest center c_p, *i.e.*, c_p satisfies:

$$d(p,c_p) = \min_{c \in C_i} d(p,c); \tag{2}$$

2. determine the point $p \in N \setminus C_i$ that is farthest form its closest center c_p and let it be c_{i+1}; *i.e.*, c_{i+1} satisfies:

$$d(c_{i+1} c_{c_{i+1}}) = \max_{p \in N \setminus C_i} d(p,c_p); \tag{3}$$

3. $C_{i+1} = C_i \cup \{c_{i+1}\}$.

At each iteration a new center is added to the set of centers that is being computed. The algorithm stops after *k*-1 iterations giving as result the set C_k. Observe that, at step *i+1*, there is no need to recalculate the distance of p to all centers, but just the distance $d(p,c_i)$, the distance to the unique center c_i added during the previous iteration. Then, just compare this distance with $d(p,c_p)$, the minimum distance to centers of C_i. According to the result of this comparison, c_p can be updated or not. Hence, if for each *p* the value $d(p,c_p)$ is stored, then each iteration can be executed in *O(n)* space and a *k*-center set can be computed in *O(kn)* distance computations.

To actually compute a clustering associated to such a k-center set, N is simply partitioned into k subsets $N_1,\ldots,N_k$, each corresponding to a center in C_k and such that

$$N_j = \{p \in N \mid c_p = c_j\}. \tag{4}$$

In other words, the cluster N_j is composed of all points for which c_j is the closest center, for each j=1,…,k. Here we use the FPF algorithm with the technique improvement described in [22]. Taking advantage of the triangle inequality, the modified algorithm avoids considering points that cannot change their closest center. To this aim, at each iteration i we mantain, for each center $c_j \in C_i$, the set N_j of points for which c_j is the closest center, defined as in (4), for j=1,…,i (*i.e.*, we build the clusters associated to intermediate solutions). We store the points in N_j in order of decreasing distance from c_j. When we scan the set of points to find the closest center, we follow the order given by the N_j's: given $p \in N_j$, with $1 \le j < i$, if $d(p,c_i) \le 0.5\, d(c_j,c_i)$ then we stop scanning N_j, as there cannot be any other point closer to c_i than to c_j. The distances between centers must be stored, requiring additional $O(k^2)$ space. As a consequence, storage consumption is linear in n only provided that $k \le n^{1/2}$.

Stability-based technique. The FPF algorithm must be fed with the number k of clusters into which N has to be partitioned. To appropriately estimate this number, here we use an efficient variant of the prediction strenght method developed by Tibshirani et al. [22]. First we briefly describe the original prediction strength method, and then give details of our implementation. To obtain the estimate, the method proceeds as follows. Given the set N of n elements, randomly choose a sample N_r of cardinality η. Then, for increasing values of t (t=1,2,...) repeat the following steps: (i) using the clustering algorithm, cluster both $N_{ds}=N \setminus N_r$ and N_r into t clusters, obtaining the partitions $C_t(ds)$ and $C_t(r)$, respectively; (ii) measure how well the t-clustering of N_r predicts co-memberships of mates in N_{ds} (*i.e.*, count how many pairs of elements that are mates in $C_t(ds)$ are also mates according to the centers of $C_t(r)$).

Formally, the measure computed in step (ii) is obtained as follows. Given t, clusterings $C_t(ds)$ and $C_t(r)$, and elements e_i and e_j belonging to N_{ds}, let $D[i,j]=1$ if e_i and e_j are mates according to both $C_t(ds)$ and $C_t(r)$, otherwise $D[i,j]=0$. Let $C_t(ds)=\{C_{t,1}(ds),\ldots, C_{t,t}(ds)\}$, then the *prediction strength* $PS(t)$ of $C_t(ds)$ is defined as:

$$PS(t) = \min_{1 \le \ell \le t} \frac{1}{\#\, pairs\ in\ C_{t,\ell}(ds)} \sum_{\substack{i,j \in C_{t,\ell}(ds)\\ i<j}} D[i,j], \tag{5}$$

where the number of pairs in $C_{t,\ell}(ds)$ is given by its binomial coefficient over 2. In other words, $PS(t)$ is the minimum fraction of pairs, among all clusters in $C_t(ds)$, that are mates according to both clusterings, hence $PS(t)$ is a worst case measure. The above outlined procedure terminates at the largest value of t such that $PS(t)$ is above a given threshold, setting k equal to such t.

In our implementation, we first run the FPF algorithm on Nr up to t= η, storing all the computed centers c1,…,cη. Such preprocessing costs O(η |Nr|) = O(η 2). We then run k-center with input set Nds. Suppose at step t we have computed the clusters

Ct,1(ds), ..., Ct,t(ds) and suppose, for each $e \in N_{ds}$, we keep the index i(e,t) of its closest center among c1,…,ct. Such index can be updated in constant time by comparing d(e,ci(e,t-1)) with d(e,ct), i.e., the distance of e from the "current" center and that to the new center c_t. Now, for each $C_{t,\ell}(ds)$, $\ell = 1,...,t$, we can easily count in time $O(|C_{t,\ell}(ds)|)$ the number of elements that are closest to the same center c_j, j=1,…, t, and finally compute the summations in formula (5) in time $O(|N_{ds}|)$. After the last iteration, we obtain the clustering of N by simply associating the points $c_1,\ldots,c_\eta$ to their closest centers in $C_k(ds)$. The overall cost of our modified procedure is $O(\eta^2 + k(n-\eta) + k\eta) = O(kn)$ for $\eta = O(n^{1/2})$. Note that, differently from the original technique, we stop this procedure at the first value of t such that $PS(t) < PS(t-1)$ and set $k=t-1$. We empirically demonstrate that our choice of termination condition gives good results.

4 Experiments

Evaluation. We performed experiments on yeast datasets and other species. We compare our algorithm with some of the most used and valid clustering algorithms for microarray gene expression data, namely CLICK, *k*-means, SOM, and the basic FPF algorithm. CLICK, *k*-means and SOM algorithms have been run under EXPANDER[2] (EXpression Analyzer and DisplayER) [18], a java-based tool for analysis of gene expression data, that is capable of clustering and visualizing the correspondent results. Among the clustering algorithms used for comparison, CLICK is the only one that does not need to know the number of clusters in advance, while both *k*-means and the basic FPF need *k* à priori, and SOM requires the grid dimension (not always corresponding to the required number of clusters[3]). Since FPF-SB and CLICK usually suggest a different *k*, we run *k*-means, SOM and FPF with both values of *k*.

In order to assess the validity of our method we evaluate the clusterings by means of the *z*-score computed by ClusterJudge tool [10], also available on line[4]. This tool scores yeast (Saccharomyces cerevisiae) genes clusterings by evaluating the mutual information between a gene's membership in a cluster, and the attributes it possesses, as annotated by the Saccharomyces Genome Database (SGD)[5] and the Gene Ontology Consortium[6]. In particular, ClusterJudge first determines a set of gene attributes among those provided by Gene Ontology, that are independent and significant; then it computes the mutual information of the proposed clustering and that of a reference random clustering. Finally it returns the *z*-score $(MI_{real} - MI_{random})/\sigma_{random}$, where MI_{random} is the mean of the mutual information score for the random clustering used, and σ_{random} is the standard deviation. The higher the *z*-score the better the clustering. Given the randomized nature of the test, different runs produce slightly different

---

[2] http://www.cs.tau.ac.il/~rshamir/expander.

[3] Note, for example, that in Table 2 the grid for Spellman et al. dataset was set to 9x3 resulting in only 18 clusters instead of 27.

[4] http://llama.med.harvard.edu/cgi/ClusterJudge/cluster_judge.pl.

[5] http://www.yeastgenome.org.

[6] http://www.geneontology.org.

numerical values, however the ranking of the method is stable and consistent across different applications of the evaluation tool. For this reason, for each dataset used we repeated three times the evaluation of the output of all the different algorithms, here reporting the average *z*-score. Even if the ClusterJudge methodology is available only for yeast genes, it is independent of both the algorithm and the metric used to produce the clustering, and thus is in effect validating both choices.

We tested our algorithm on some of the most used yeast datasets in literature [4,6,19] and our results show that, on the average, we achieve a better score than that obtained by the other clustering algorithms, while using far fewer resources, especially time.

Datasets. The algorithm was tested on three well studied yeast datasets. The first is the yeast cell cycle dataset described by Cho et al. in [4]. In their work the authors monitored the expression levels of 6,218 Saccharomyces cerevisiae putative gene transcripts (ORFs). Probes were collected at 17 time points taken at 10 min intervals (160 minutes), covering nearly two full cell cycles[7]. The second dataset, described by Spellman et al. in [19], is a comprehensive catalog of 6178 yeast genes whose transcript levels vary periodically within the cell cycle (for a total of 82 conditions) [8]. The third dataset, described by Eisen et al. [6], contains 2467 probes under 79 conditions, and consists of an aggregation of data from experiments on the budding yeast Saccharomyces cerevisiae (including time courses of the mitotic cell division cycle, sporulation, and the diauxic shift)[9].

Experimental results. The results reported here have been obtained on an 1.4 GHz AMD ATHLON XP workstation with 1.5 GB RAM running Linux kernel 2.6.11.4.

Table 1 summarizes the properties (number of probes and number of conditions) of the three datasets, while experimental results are reported in Tables 2 and 3. For each experiment we report the number of clusters *k*, the computation time in seconds and the value of the *z*-score.

Note that CLICK is the only algorithm, among those that we have tested, that sometimes produces singletons in its clustering (136 in Cho et al. dataset, 17 in Spellman et al. dataset and none in Eisen et al. dataset) and put them into a single cluster labeled cluster zero. Hence, to correctly evaluate CLICK results we created a new cluster for each of these singletons.

Concerning Table 2, we observe that FPF-SB achieves a better *z*-score on all three datasets using far less computation time. The first two rows of the table show the results obtained with FPF-SB and CLICK, the two algorithms that do not need *k* as input. It can be seen that the two algorithms make a very different choice of *k*, except for the Eisen et al. dataset. The results show that FPF-SB outperforms CLICK, both in terms of actual computational time (in seconds)[10] and *z*-score, even when the guess of *k* is the same. FPF-SB outperforms all the other algorithms when fed with CLICK's guess of *k*.

---

[7] The complete dataset, now containing 6601 Saccharomyces cerevisiae putative gene transcripts (ORFs), is available at http://genomics.stanford.edu/yeast_cell_cycle/cellcycle.html.
[8] The complete dataset and description are available at http://cellcycle-www.stanford.edu.
[9] The data is fully downloadable from http://genome-www.stanford.edu/clustering.
[10] As expected, since CLICK asymptotically runs in $O(n^3)$.

Table 1. Summary of dataset properties

Dataset	**Cho et al.**	**Spellman et al.**	**Eisen et al.**
Probes	6601	6178	2467
Conditions	17	82	79

Table 2. Experimental results comparing FPF-SB with algorithms that need *k* as input parameter (using CLICK's guess of *k*)

Dataset	**Cho et al.**			**Spellman et al.**			**Eisen et al.**		
Algorithm	***k***	**Time**	***z*-score**	***k***	**Time**	***z*-score**	***k***	**Time**	***z*-score**
FPF-SB	8	12	77	16	94	61.6	8	15	62.9
CLICK	30	660	52.3	27	4200	52.9	8	240	38
***k*-means**	30	720	66.3	27	1140	52.5	8	120	34.2
SOM	29	300	54.9	18	240	49.1	8	60	59
FPF	30	22	46.8	27	80	47.9	8	7	62.9

Table 3. Experimental results comparing FPF-SB with algorithms that need *k* as input parameter (using FPF-SB's guess of *k*)

Dataset	**Cho et al.**			**Spellman et al.**			**Eisen et al.**		
Algorithm	***k***	**Time**	***z*-score**	***k***	**Time**	***z*-score**	***k***	**Time**	***z*-score**
FPF-SB	8	12	77	16	94	61.6	8	15	62.9
***k*-means**	8	420	92.5	16	900	63	8	120	34.2
SOM	8	180	75.9	16	420	52.1	8	60	59
FPF	8	5	77	16	43	61.6	8	7	62.9

The results reported in Table 3 show that, with the number of clusters suggested by FPF-SB, the quality of the clusterings computed by *k*-means and SOM, as judged by the *z*-score, is always better than in the previous case. This suggests that our implementation of the stability-based method produces a good guess of *k* in a computationally efficient way. Note also that we can evaluate the time spent by FPF-SB for the computation of *k* by comparing its running time with that of the basic FPF algorithm when computing the same number of clusters. Empirically, we evaluated this time as 55% of the total time on the average. However, the overall computation time is still much less than that of all the other algorithms.

The results also show the scalability of FPF and FPF-SB in terms of actual performance. Note that the size of a dataset is correctly measured as the product $n \cdot m$ (the dimension of the point space where the clusterings are computed). The actual running time is thus proportional, to first order, to the product $m \cdot n \cdot k$. It can be easily seen that the increase in running time of the Spellman et al. and Eisen et. al. datasets over the Cho et al. (the smallest one) can be quite accurately predicted using the ratio of the corresponding asymptotic costs, plus the time needed to compute *k* (the stability-based part). For instance, FPF-SB prediction for Spellman et al. dataset is 96 seconds[11].

Moreover, our experiments give us a way to estimate the multiplicative constant hidden by the big-O notation and, thus, make a reliable prediction of the running time of FPF and FPF-SB in actual seconds, for any given dataset. This constant can be obtained by taking the maximum ratio, among all datasets, between the measured time (expressed in μsecs) and $m \cdot n \cdot k$. We obtained the constants 13.4 for FPF-SB and 6.5 for FPF. Hence, given any dataset consisting of n genes and m conditions, we can estimate (an upper bound to) the time needed by FPF-SB (resp. FPF) to compute a k-clustering in $13.4 \times 10^{-6} \times m \cdot n \cdot k$ seconds (resp. $6.5\ 10^{-6} \times m \cdot n \cdot k$ seconds) [11].

5 Conclusions

Efficient and effective analysis of large datasets from microarray gene expression data is one of the keys to time-critical personalized medicine. The issue we address here is the scalability of the data processing software for clustering gene expression data into groups with homogeneous expression profile. In this paper we propose *FPF-SB*, a new clustering algorithm that efficiently applies to microarray data analysis, being scalable to large datasets without sacrificing output quality.

In order to validate both the choice of the algorithm and the metric used to produce the clusterings we used ClusterJudge, an independent tool that only scores yeast (Saccharomyces cerevisiae) genes clusterings. Therefore, one of our future tasks will be to find methodologies for the evaluation of clusterings of gene datasets from other species (human, mouse, rat, etc.).

References

[1] Alon, U., et al.: Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Proc. Natl Acad Sci USA 96(12), 6745–6750 (1999)

[2] Belacel, N., Cuperlovic-Culf, M., Laflamme, M., Ouellette, R.: Fuzzy J-Means and VNS methods for clustering genes from microarray data. Bioinf. 20(11), 1690–1701 (2004)

[3] Ben-Dor, A., Shamir, R., Yakhini, Z.: Clustering gene expression patterns. J Comput Biol. 6(3-4), 281–297 (1999)

[4] Cho, R.J., et al.: A genome-wide transcriptional analysis of the mitotic cell cycle. Mol Cell. 2(1), 65–73 (1988)

[5] Clarkson, K.L.: Nearest-neighbor searching and metric space dimensions. In: Shakhnarovich, G., Darrell, T., Indyk, P. (eds.) Nearest-Neighbor Methods for Learning and Vision: Theory and Practice, pp. 15–59. MIT Press, Cambridge (2006)

[6] Eisen, M.B., Spellman, P.T., Browndagger, P.O., Botstein, D.: Cluster analysis and display of genome-wide expression patterns. PNAS 95(25), 14863–14868 (1998)

[7] Ernst, J., Naur, G.J., Bar-Joseph, Z.: Clustering short time series gene expression. Bioinf. 21(1), i159–i168 (2005)

[11] Note that these are our workstation estimation.

[8] Feder, T., Greene, D.H.: Optimal algortihms for approximate clustering. In: Proc. of 20th ACM Symposium on Theory of Computing, pp. 434–444 (1988)

[9] Geraci, F., Pellegrini, M., Sebastiani, F., Pisati, P.: A Scalable Algorithm for High-Quality Clustering of Web Snippets. In: Proc. of 21st ACM Symposium on Applied Computing 2006 (2006)

[10] Gibbons, F.D., Roth, F.P.: Judging the Quality of Gene Expression-Based Clustering Methods Using Gene Annotation. Genome Research 12, 1574–1581 (2000)

[11] Gonzalez, T.: Clustering to minimize the maximum intercluster distance. Theoretical Computer Science, pp. 293–306 (1985)

[12] Hastie, T., et al.: Gene shaving' as a method for identifying distinct sets of genes with similar expression patterns. Genome Biol. 1(2) (2000)

[13] Holloway, A.J., et al.: Options available - from start to finish - for obtaining data from DNA microarrays II. Nature Gen. Suppl. 32, 481–489 (2002)

[14] Huang, D., Pan, W.: Incorporating biological knowledge into distance-based clustering analysis of microarray gene expression data. Bioinf. 22(10), 1259–1268 (2006)

[15] Jiang, D., Tang, C., Zhang, A.: Cluster Analysis for Gene Expression Data: A Survey. IEEE Trans. on Knowledge and Data Eng. 16(11), 1370–1386 (2004)

[16] Ramoni, M.F., Sebastiani, P., Kohane, I.S.: Cluster analysis of gene expression dynamics. Proc. Nat Acad Sci USA 99(14), 9121–9126 (2002)

[17] Schadt, E.E., et al.: A comprehensive transcript index of the human genome generated using microarrays and computational approaches. Genome Biology 5(10), R73 (2004)

[18] Sharan, R., Maron-Katz, A., Shamir, R.: CLICK and EXPANDER: A System for Clustering and Visualizing Gene Expression Data. Bioinf. 19(14), 1787–1799 (2003)

[19] Spellman, P.T., et al.: Comprehensive Identification of Cell Cycle-regulated Genes of the Yeast Saccharomyces cerevisiae by Microarray Hybridization. Mol Biol Cell. 9, 3273–3297 (1998)

[20] Tamayo, P., et al.: Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. Proc. Natl Acad Sci USA 96(6), 2907–2912 (1999)

[21] Tavazoie, S., Hughes, J.D., Campbell, M.J., Cho, R.J., Church, G.M.: Systematic determination of genetic network architecture. Nature Genetics 22, 281–285 (1999)

[22] Tibshirani, R., Walther, G., Botstein, D., Brown, P.: Cluster validation by prediction strength. Journal of Computational & Graphical Statistics 14, 511–528 (2005)

[23] Trent, J.M., Bexevanis, A.D.: Chipping away at genomic medicine. Nature Genetics (Suppl), p. 426 (2002)

[24] Wen, X., et al.: Large-scale temporal gene expression mapping of central nervous system development. Proc. Natl Acad Sci USA 95(1), 334–349 (1988)

[25] Xing, E.P., Karp, R.M.: CLIFF: clustering of high-dimensional microarray data via iterative feature filtering using normalized cuts. Bioinf. 17(1), 306–315 (2001)

[26] Yeung, K.Y., Haynor, D.R., Ruzzo, W.L.: Validating clustering for gene expression data. Bioinf. 17(4), 309–318 (2001)

[27] WWW, Personalized Medicine Coalition, The case for Personalised Medicine http://www.personalizedmedicinecoalition.org

[28] WWW, The Royal Society, Personalised medicines: hopes and realities http://www.royalsoc.ac.uk

A Finite Element 3D Model of in Vivo Human Knee Joint Based on MRI for the Tibiofemoral Joint Contact Analysis

Zhixiu Hao, Dewen Jin, Yu Zhang, and Jichuan Zhang

Division of Intelligent and Biomechanical System, State Key Laboratory of Tribology, Tsinghua University, Beijing, P.R. China, 100084
haozx@tsinghua.edu.cn

Abstract. The contact behaviors of in vivo knee during weight bearing were quantified using MRI-based 3D knee model. The model included the cortical and trabecular bone of the femur and tibia, cartilage of the femoral condyles and tibial plateau, both the medial and lateral menisci with their horn attachments. Explicit dynamic nonlinear finite element techniques were discussed to simulate biomechanical features of tibio-femeral joint under different loads condition, considering the relative slide and friction existing in the knee. The simulating results show that the contact area increased while the loads increased. Both the ratio of the contact area and the ratio of contact force of medial and laterial tibial plateau were almost constant under the different loads along the axis of the tibia during the supporting phase in a gait, and yet the contact points of the compressive force were changed.

Keywords: tibiofemoral joint, contact; finite element.

1 Introduction

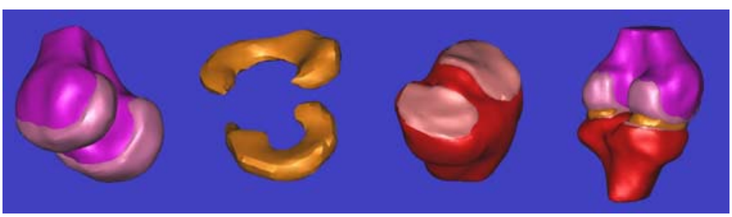
Human knee joint is a diarthrodial joint and supports high level of mechanical load for several decades during ones lifetime . This often causes degenerative arthritis in the knee. The research finding of biomechanics of the human knee is a reference to not only deeply understand the behavior of joints, protect joints consciously against injury, but also to know the causes of degenerative arthritis in the knee. Some recent studies have implemented 3D models of human knee to predict the biomechanical behavior of the joint. Louis E[1] used 3D MRI-based knee joint to analyze the in vivo tibio-femoral contact analysis. Tammy L[2] constructed a finite element model of the human knee joint of a cadaveric specimen and analyzed the tibiofemoral contact pressure using implicit dynamic finite element method techniques under ones body weight. Guo an[3] constructed a 3D finite element model using MRI of a knee to evaluate the joint contact stress analysis and the effect of changes in cartilage thickness on contact stress analysis was investigated too. William J[4] described a method, which combines high-speed biplane radiographic image data and three-dimensional bone surface information derived from CT, to identify the regions of close contact during motion. No previous studies have used the 3D finite model of

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 616–622, 2007.

knee in vivo with both the medial and lateral menisci to quantify tibiofemoral cartilage contact stress and contact points with different loading conditions. The object of this study was to analyze the contact behaviors during in vivo knee flexion using 3D MRI-based computational models by explicit dynamic nonlinear finite element techniques.

2 Methods

A single right knee of a healthy subject (female, 36) in the sagittal plane was scanned by ESAOTE/Artoscan-C (GE Company). A surface coil was used to generate parasigittal images, and forty-three slices were obtained at knee of 25^0 flexion while the subject was relaxed and full extension as 0^0 flexion respectively. The subject lay supine on the MRI scanner. The total scan time was approximately 40min. The MRI scan had a resolution of 256×256 pixels. Sagittal plane images spaced 0.70mm apart were manually traced in a solid modeling software[5] and then used to construct 3D solid models of the knee including both the cortical and trabecular bone of the femur and tibia, articular cartilage of the femoral condyles and tibial plateau, the medial and lateral menisci in Fig.1. Both of the femoral cartilage and the tibial cartilage were constructed as two pieces, namely, medial and lateral pieces. Considering the previous studies concluded that finite element solution for rigid versus deformable bones indicated that none of the contact variables changed by more than 2% when femur and tibia were treated as rigid[2], bones were assumed to be rigid in this knee joint finite element model. Only the surfaces of the bones were needed to demonstrate the relative position of the tibia and femur to specify external loads. The material behaviors of both the tibial and femoral cartilage layers were modeled as linear elastic, isotropic, and homogeneous materials. The elastic modulus was 15MPa and Poisson ratio was 0.475 under short loading time [2].



(a) (b) (c) (d)

Fig. 1. 3D geometrical model of human tibiofemoral joint (a) cartilage and femur (b) material and lateral menisci (c) cartilage and tibial plateau (d) tibiofemoral joint

The menisci were modeled as linearly elastic, transversely isotropic with moduli of 20MPa in the radial and axial directions, and 140MPa in the circumferential direction. The Poisson ratio in the plane of isotropy was 0.2,and the out-of-plane poisson ratio was 0.3. The values for the moduli and Poisson ratio were taken from the literature [2]. It was assumed that each of the horns of the menisci was attached to the tibial plateau by linear springs. Considering the width of the horns of the menisci, ten linear springs, each with a stiffness of 200N/mm were used to simulate horn attachment [2].

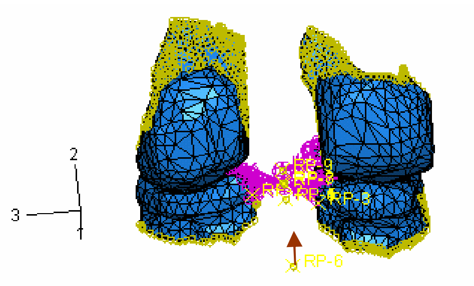


Fig. 2. 3D finite element model of human knee joint for tibia-femoral contact analysis with menisci

There were contacts between the femur and menisci, the menisci and tibia, and femur and tibia for both the lateral and medial hemijoints, resulting in six contact pairs in the model. The finite sliding of each contact pairs were involved in the model. Because the surfaces of the contact pairs were curved and deformable, and the contacting conditions of each contact pairs were uncertain, obtaining the high-quality computational solutions under these complex conditions represented a significant challenge. In previous studies the friction in each of contact pairs were ignored[2,6]. In this study, friction was involved for the friction exists in human knee joints, especially in the joints with arthropathy. In this study, both the finite sliding and friction of each contact pairs were involved.

The goal this study was to investigate the contact behaviors of the tibia-femoral joint under different load conditions along the axis of the tibia during the supporting phase in a gait. The knee flexion angle is about 25 degree, at which the load on the knee joint is about six times of body's weight[7]. It is the time when knees bear the maximum loads during a gait cycle. It was supposed that the knee had no other movement in this position except relative slide along the tibia.

The numerical problems have been described in the literatures when using implicit finite element techniques to simulate problems that involve contact and large displacements. The explicit finite element techniques can be used to solve the quasistatic problems. The stability of solution and low computational cost are the two main advantages of the method over classic implicit techniques. One limitation of the explicit finite element technique is that the method is only conditionally stable. For explicit finite element technique uses the central-difference operator, the duration between two integration instances must be smaller than a critical time step, which is relate to the minimal element size and the material properties. Refinement of the mesh is more penalizing in explicit than in implicit techniques since it also effects the minimum time step.

The general finite element code ABAQUS/ Explicit was used to obtain solution to the tibia femoral contact. The cartilages and menisci were meshed as 10-node modified second-order tetrahedral elements which are designed to be used in complex contact simulation. It can calculate the contact pressure accurately. In order to get high quality solution, the contact analysis was done in each of three mesh size. The initial mesh size was 3.0mm, then 2.5mm, final 2.0mm, and the number of the elements was 16,782, 25,733, and 45,832 respectively. Decreasing the mesh from

2.5mm to 2.0mm changed the contact variables by 3% at most. Therefore, the 2.5mm was enough fine for the simulation[5]. The final simulating model is shown in fig.2. In order to reduce the effect of the inertia force in the simulation using ABAQUS/ Explicit, the selected loading curve was smooth in both first derivative and second derivative and simulating time was 0.02s. The criteria used to judge whether the simulating result was quasi-static was that the ratio of kinetic energy to inner energy of the computational model was less than 0.5%, and the variation of the inner energy drove to stabilization.

3 Result

Cosidering the load supporting phase is from 10% to 15% of a gait cycle, and when the flexion angle of a knee is about 25^0, the maximum load reach about 6 times body weight[7], The biomechanical behaviors at this position were studied.

Table 1. Medial and lateral contact variables at different loads

Load(N)	Contact Area (mm^2)			Contact force (N)		
	M	L	M/L (%)	M	L	M/L (%)
1×600	300	232	56.4	345	245	58.5
2×600	413	303	57.7	650	510	56.0
3×600	555	360	60.7	1020	739	58.0
4×600	627	445	58.5	1400	945	59.7
5×600	696	492	58.6	1800	1140	61.2
6×600	725	540	57.3	2150	1380	61.0

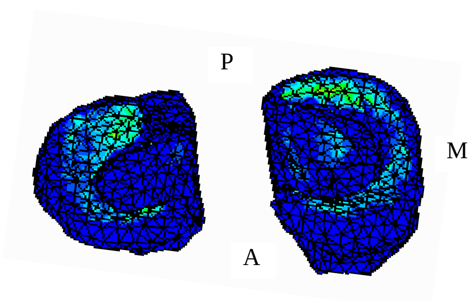


Fig. 3. Contact pressure distribution in menisci and articular cartilage of tibial plateau under 3 times body weight at knee 25^0 flexion

These results indicated the contact area between each contact pairs is increased with the load increasing, and the ratio of the contact area of medial and lateral tibial plateau was almost constant(0.582±0.01470). So was the ratio of the contact force of medial and lateral tibial plateau(0.591±0.01978) under different loads(table1).

Table 2. Position of the Medial and lateral tibia contact center at different loads

Load(N)	Anterior-posterior(mm)		Media-lateral(mm)	
	Medial condyle	Lateral condyle	Medial condyle	Lateral condyle
1×600	108.0	109.0	-32.7	8.0
3×600	108.7	109.2	-32.4	8.6
6×600	108.0	109.4	-30.8	7.7

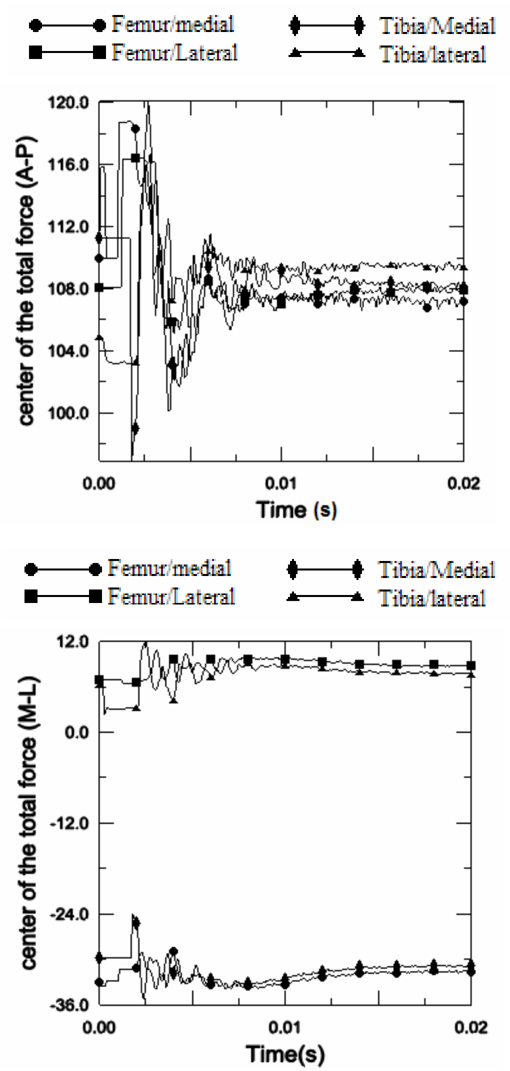


Fig. 4. Center of the total force due to contact pressure and friction stress under 6 times body weight at 25^0 flexion

Therefore, the average contact pressure of medial and lateral tibial plateau were almost equal during the loading phase in a gait. It can keep the medial and lateral tibial plateau have same life period.

As to the load transmission, when the extra load was 600N(1BW), almost all of load was transmitted by menisci. When the load was increased to 1800N, 90 percent load was transfered by menisci, and menisci carried 70 percent load when extra load reached 3600N. The map of the contact stress distribution from the tibial plateau under 1800N compressive force at 25 degree flexion was shown in fig.3.

More ever, both medial and lateral contact points of compressive force moved laterally, slightly anterior when loading to 3 times body weight. Increasing to 6 times body weight, the contact point on the medial condyle moved laterally, slightly posterior. On the contrary, the contact point on the lateral condyle moved medially, slightly anterior. Table 2 lists the contact position in detail.

Fig.4 was the simulation history of the center of the total force due to contact pressure and friction stress while bearing 6 times body weight. The contact point moved more obviously in Media-lateral direction than in Anterior-posterior direction.

4 Discussion

The 3D finite element model of in vivo knee was constructed using MR images, and the model was used to compute the contact behaviors at 25 degree flexion of knee during the supporting phase of human gait cycle. The explicit dynamic nonlinear finite element techniques can be used to solve the quasi-static problem and have a high efficiency especially in solving the problem of contacts and large displacements.

For a direct validation of the cartilage contact area and contact stress of in vivo knee is unrealized, a qualitative comparision with literatures was done. Fukubayashi and Kurosawa[8] measured contact area of the knee joint using a casing technique. Under an axial compressive load of 1000N, the contact area at the medial tibial plateau of the knee was 300±80mm^2 (mean±S.D. n=7). The simulation in this study showed a contact area of 300mm^2 under 600N and 413 mm^2 under 1200N at medial tibial plateau. Kurosawa et al[9] estimated average surface pressure over the knee joint using the same casting technique. The average surface pressure over the medial and lateral sides of the knee was 1.97±0.56MPa under 1000N. The average surface pressures of the medial sides of the knee were 1.15MPa under 600N and 1.57MPa under 1200N. The average surface pressures of the lateral sides of the knee were 1.06MPa under 600N and 1.68MPa under 1200N. The calculation data in our study are within the range of experimental measurement.

In our present study, only the contact behaviors at 25 degree flexion were analyzed. The future study should compute the contact behaviors during the whole gait cycle.

Acknowledgments. The authors thank Mingxue Sun and Xin Wang for their assistant in magnetic resonance imaging taken.

References

1. Defrate, L.E, Sun, H.: In Vivo Tibiofemoral Contact Analysis Using 3D MRI-based Knee Models. Journal of Biomechanics 37, 1499–1504 (2004)
2. Haut Donahue, T.L., Hull, M.L., Rashid, M.M., et al.: A Finite Element Model of the Human Knee Joint for the Study of Tibio-femoral Contact. Journal of Biomechanical Engineering 124, 273–280 (2002)
3. Guoan, L., Orlando, L., Harry, R.: Variability of a three-Dimensional Finite Element Model Constructed Using Magnetic Resonance Images of a Knee for Joint Contact Stress Analysis. Journal of Biomechanical Engineering 123, 341–346 (2001)
4. Anderst, W.J., Tashman, S.A.: Method to Estimate in vivo Dynamic Articular Surface Interaction. Journal of Biomechanics 36, 1291–1299 (2003)
5. Zhixiu, H.: Study of the in vivo Human Knee Model with Menisci and Investigation of Biomechanical Behaviors of Tibiofemoral Join. Ph.D.Thesis. Beijing, Tsinghua University (2006)
6. Meakin, J.R., Shrive, N.G., Frank, C.B., et al.: Finite Element Analysis of the Meniscus: the Influence of Geometry and Material Properties on its Behavior. The knee 10, 33–41 (2003)
7. Yicong, W.: Basis and Clinical of Knee Joint Surgery. People Sanitation Press, Beijing, China (1999)
8. Fukabayasi, T., Kurosawa, H.: The Contact Area and Pressure Distribution Patten of the Knee. Acta Orthopaedica Scandinavica 51, 871–879 (1980)
9. Kurosawa, H., Fukabayasi, T., Nakajima, H.: Load-Bearing Model of the Knee Joint, Physical Behavior of the Knee with or without Menisci. Clinical Orthopedics 149, 283–290 (1980)

Advances in Visible Human Based Virtual Medicine

Pheng Ann Heng[1,2]

[1] Department of Computer Science and Engineering,
The Chinese University of Hong Kong, Hong Kong
[2] Shenzhen Institute of Advanced Integration Technology,
Chinese Academy of Sciences/The Chinese University of Hong Kong, Shenzhen
pheng@cse.cuhk.edu.hk

Abstract. In this paper, important virtual medicine research related to visible human data sets would be outlined. In particular, latest visible human related research carried out in the Visualization, Virtual Reality and Imaging Research Centre of the Chinese University of Hong Kong (CUHK) will be discussed in detail, while related applications such as virtual acupuncture, virtual anatomy, and virtual orthopedics training briefly shaped our current and future research and development on visible human-related virtual medicine. Our latest advancement in segmentation, imaging, user-interface design as well as surgical education and training will be discussed.

Keywords: Visible Human data, visualization, surgical simulation, virtual reality.

1 Introduction

As the importance of digital images in medical research increases, construction of standardized human body datasets has been indispensable. Being proposed by the National Library of Medicine of United States, the Visible Human Project (VHP) was carried out at University of Colorado, Denver, Colorado [1]. Complete digitized datasets for normal male and female, which consist of cryosectional photographic images as well as the corresponding computerized tomography (CT) and magnetic resonance images (MRI), were collected in 1994 and 1995 respectively. This accomplishment symbolizes a new era in virtual medicine with an in-depth understanding of human anatomy and systematic application of computational techniques in clinical medicine and biomedical research.

Motivated by VHP, the Chinese Visible Human (CVH) Project was started in 2001 while the first male dataset was collected in 2002 at the Third Military Medical University, Chongqing, China [2]. Up to now, five CVH datasets have been successfully collected, representing a diversity of Chinese population [3]. Compared to the US visible human, the CVH data features a higher resolution of images; this provides a clearer identification of blood vessels, for example. The Visible Korean Human (VKH) project was started in 2001 at Ajou University, Suwon, Korea, which is another visible human representing East Asian [4]. This dataset contains CT, MRI and cryosectional photographs of a complete male cadaver.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 623–632, 2007.

One natural application of visible human data is virtual anatomy where the rich anatomical information contained in the VH datasets enables the construction of informative anatomical atlases for anatomy education [7], this gives a convenient way to view structures from different directions. Voxel-Man project built photorealistic anatomical atlases for the brain, internal organs [8], and small organs like the hand [9]. Virtual anatomy complements the traditional cadaver dissection by constructing 3D models using the cryosectional images such as the visible human data, and it serves as a powerful and flexible way to learn human anatomy [10].

A concept extended from “visible human” is “virtual human”, which integrates static anatomical data with dynamic physiological properties. Once such a virtual human is constructed, it can serve as a test-base for the simulating and validating human anatomical functions. The Oak Ridge National Laboratory (ORNL) proposed the Integrated Respiratory System Model [13] that uses the virtual human to simulate the generation and transmission of lung sounds. Another direct application of virtual human is surgery simulation, which simulates surgery to estimate the effect on a patient [14].

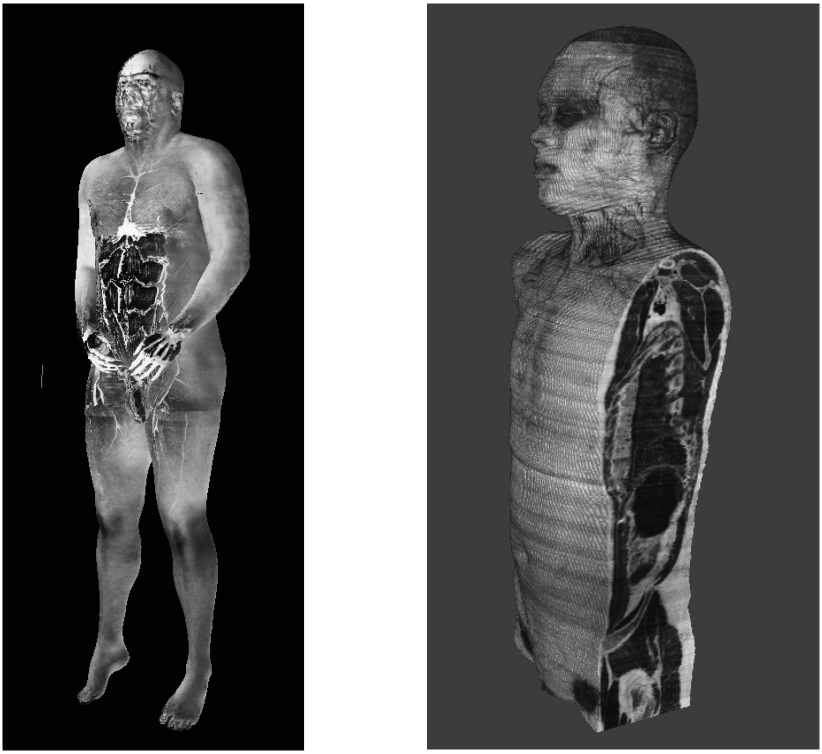
As a repository of medical images, the visible human dataset can be used to test different image processing and analysis methods. The Biomedical Imaging Resource (BIR) at the Mayo Clinic have been using the visible human dataset extensively for testing image visualization, segmentation, registration, classification, and modeling algorithms [5]. The Anatquest project aims at building a universal platform for visual and textual queries to the VH dataset and investigating the image segmentation techniques tailored for the VH cryosectional images [6].

There has been many virtual medicine applications implemented using visible human data, including surgical planning and rehearsal, as well as virtual endoscopy. To assist surgical planning and intraoperational navigation, 3D organ models built from the VH data are registered with the patient MRI data. The visible human data have also been used to build phantoms for radiation treatment planning [11]. The Monte Carlo method is used to calculate the maximum radiation dose for a tumor while minimizing the exposure of normal tissue. Virtual endoscopy broadens the accessibility of endoscopic examination and minimizes the uncomfortable experience of patients by generating simulated visualizations of patient anatomy. Visible human datasets have been used to produce simulated endoscopic views of the stomach, colon, spinal column, esophagus, trachea, and heart [12]. By mapping textures onto the cryosectional images of visible human datasets, the simulation of endoscopic views can be vividly generated.

In this paper, some of the latest research and development projects carried out in our research centre, the Virtual Reality, Visualization and Imaging Centre, CUHK will be discussed in detail. The rest of this paper is organized in this way. First, the visible human visualization will be discussed. The use of Visible Human Project data as well as the Chinese Visible Human data on virtual medicine research, for example, in the application of virtual acupuncture, virtual anatomy shall be discussed next. Finally, conclusion will be drawn.

2 Visible Human Visualization

Visualization, Virtual Reality and Imaging Research Centre has started medicine-related research on the visible human project since 1999. **Fig. 1** (a) shows our



(a) (b)

Fig. 1. Screenshots of the real-time visible human visualization are shown. (a) The VHP data, having segmented into different tissue layers, can be interactively viewed. (b) The highest resolution Chinese Visible Human data can be visualized with multi-planar reconstruction.

real-time rendered shots on a multi-layer VH visualization. Since the Third Military Medical University in Chongching started the CVH project in 2002, we have been collaborating with the CVH project team in the development of 3-D visualization techniques [15]. A depiction of the early work on CVH data visualization can be observed in Fig. 1 (b).

Gelatin is injected into the body in order to prevent the collapse of the respiratory system or the digestive system at the abdominal region. The milling process is carried out inside a cold chamber in order to avoid dropping of smaller structures. Compared to the American visible human data, the CVH dataset can be said to be more complete in terms of the anatomical system being preserved [16]. CUHK, as one of the important collaborators in the CVH project, has been devoted into the imaging, visualization as well as virtual surgery research on the CVH data in these few years. Innovative techniques in these areas have been proposed in order to handle these precise but huge data sets.

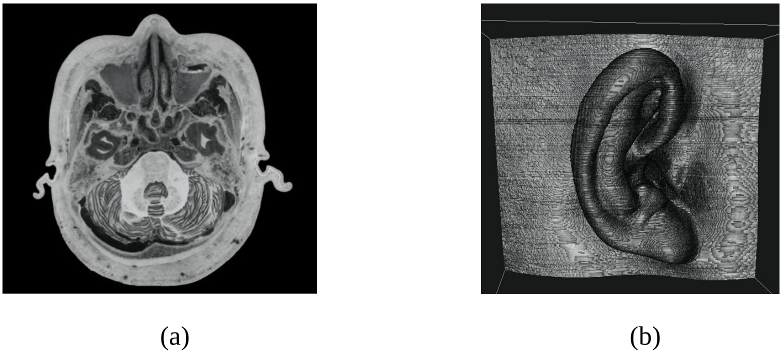


Fig. 2. (a) A processed Chinese Visible Human cryosectional slice is shown. Preciseness of the tiny details, such as blood vessel or nerve can be observed in the dataset. (b) The volume rendered view of the ear from the second female Chinese Visible Human dataset is shown.

Slices pre-processing is one of the essential process before actual visualization can be done. Since the overall milling process has been carried out for a number of weeks, the photo capturing condition is altering from slice to slice, color or contrast balancing as well as accurate alignment between multiple slices is thus necessary. One of the processed CVH slices is shown in **Fig. 2** (a). Having processed every slice, the slice reconstruction comes next.

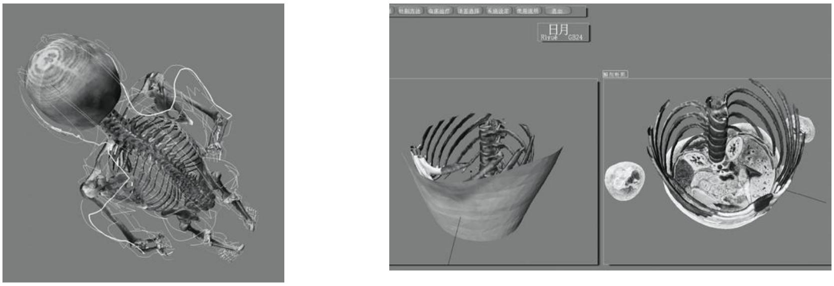
Slice reconstruction is relatively more straight-forward, slices in different axial order are re-arranged and then the volume texture is being compressed in way which conforms to the graphics pipeline in personal computer grade video accelerator. In this sense, the reconstructed volume can be visualized by deploying advanced shader-based rendering tactics. **Fig. 2** (b) shows the screen shots of our CVH visualization application, which incorporates the second female CVH data set which is composed of 131G data size in total. Such dataset can be presented in real-time stereoscopic rendering too. This application has won the Credit of Merits in RSNA'04 [17] and is now being shown as a permanent exhibition in the Hong Kong Science Museum.

3 Virtual Medicine with Visible Human datasets

In this section, we introduce some of the visible human based virtual medicine projects being conducted in our research centre. These projects include virtual acupuncture, virtual anatomy and virtual orthopedics training.

3.1 Virtual Acupuncture

Virtual Acupuncture is one of our early works based on VHP data. As a key component of traditional Chinese medicine, the therapeutic effectiveness of acupuncture on various diseases by stimulating acupoints using needle has long been appreciated. However, incorrect needle penetrations may lead to serious complications, in this sense, extensive training is essential for acupuncturists to build up their skill, virtual



(a) (b)

Fig. 3. (a) The meridian system can be shown interactively. The brighter white line is showing the highlighted meridian when other dimmer lines are showing the rest of the system. (b) The user interface for displaying every acupuncture point is shown. In addition to the external skin surface, bony tissue as well as the cross-sectional anatomical details can be observed simultaneously with the selected acupoint. The line is showing the location of the selected acupoint.

reality training is an elegant choice. CUHK has collaborated with the Nanjing Chinese Medical School of China in the virtual acupuncture project aiming at delivering a user-friendly environment for learning acupuncture therapeutic concept and knowledge as well as a virtual puncturing training system for acquiring needle lifting and thrusting operational skills.

Fig. 3 (a) shows our virtual acupuncture application. Based on the Chinese Visible Human dataset, the data is first semi-automatically segmented into different tissue: the skin layer (including the dermis and hypodermis), the muscle and the bone [13]. The skin and bone is crucial for locating acupuncture points while hypodermis and muscle contributes to a successful realistic tactile modeling of needle puncturing simulation. **Fig. 3** (b) shows how our system can assist the learning of the 3-D location of acupoints and meridians. Acupuncture novice can acquire the knowledge of acupuncture points or meridians, while practitioners can further research the correlation between these acupoints and underlying anatomical structures. In other words, our intuitive learning platform can foster scientific research on traditional acupuncture.

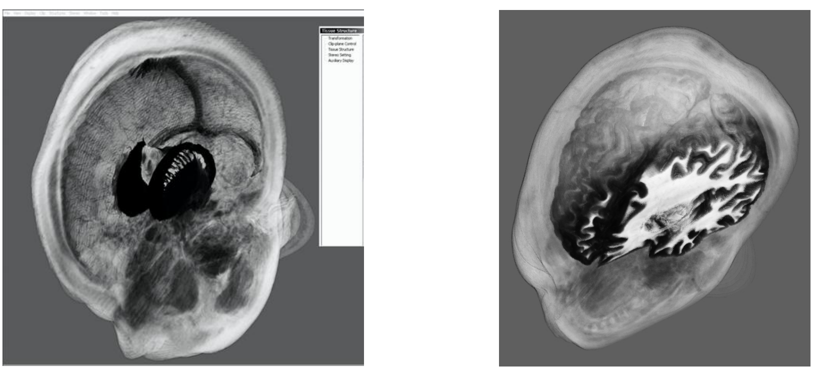
Fig. 4 shows the haptic needle puncture training system. Based on a multi-layer model obtained through segment visible human, physical tissue properties have been incorporated in order to constitute a bi-directional needle thrust and uplifting force model for realistic haptic rendering [14]. With such a tactile simulation, users can practice their needle manipulation skill including thrusting depth and frequency as if they were puncturing real patients.

3.2 Virtual Anatomy

The CVH dataset have been utilized as the basis of our virtual anatomy application, different regions of interest are segmented for different anatomical education platforms. In our virtual anatomy platform, interactive user control such as multi-planar

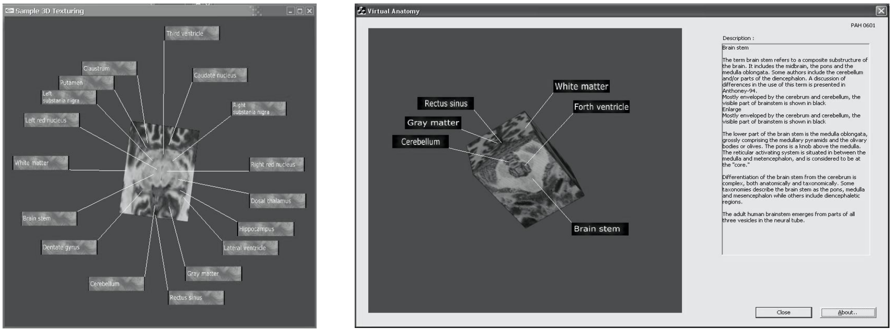
Fig. 4. The user interface for needle puncturing training can be observed. Reach-in display system is deployed in order to provide a more immersive environment. Users can freely practice needle thrusting or uplifting within this virtual environment.

reconstruction (MPR), translucency visualization, and pseudo-color highlights can be applied to the structures in a real-time manner. 3-dimensional stereoscopic visualization, informative anatomical descriptions, selective functional structure display can help medical student shortening the learning curve in acquiring anatomical knowledge. The preciseness of the data set equipped with interactive visualization delivers a completely new learning experience on anatomy. One of our virtual anatomy series, virtual neuroanatomy, central nervous system structures in the head and neck region for example, cerebrum, cerebellum, ventricles, diencephlons, bascal ganglia and nerves etc are identified. This application has been shown in the RSNA'05 InfoRAD Exhibits [18].



(a) (b)

Fig. 5. (a) The user interface of the virtual neuroanatomy is shown. Different tissue can be visualized selectively, for example, the cerebrum has been set totally transparent so that the limbic system can be observed clearly. (b) Multi-Planer Reconstruction is one of essential visualization mode in virtual anatomy applications.



(a) (b)

Fig. 6. (a) the automatic anatomic annotation result (b) detailed anatomical information will be shown when user clicks on one of the labels

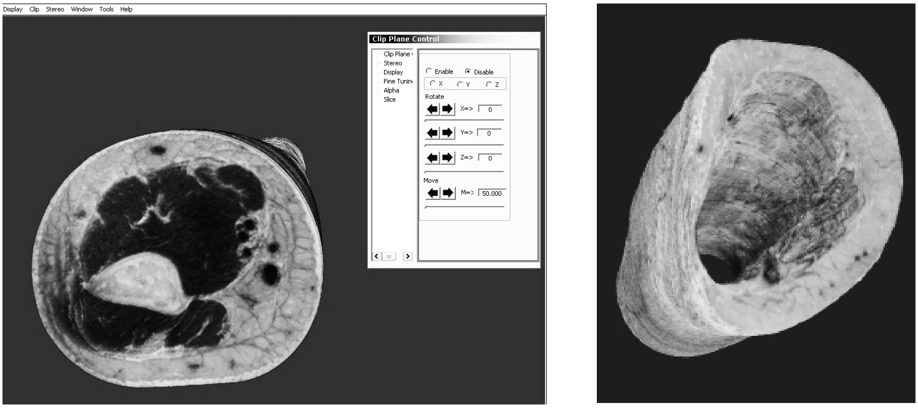
Fig. 5 (a) shows the user interface of the virtual neuroanatomy application. Translucent visualization and pseudo-color highlight are supported. Interactive planar reconstruction (**Fig. 5** b) and level of details are two other important features of the software. In addition, a novel computing tool for annotating anatomical structures in virtual anatomy applications has been proposed. The automatic labeling algorithm can also be applied to other related applications, e.g. virtual acupuncture, which involves similar anatomic annotation requirements. A real-time interactive application which will automatically generate tissues labels was developed. The user can gain further description of a tissue by clicking on the corresponding label and also drag the labels to desired positions.

Although the data size is huge with respect to normal personal computer, for example, the total data size of the virtual brain is over 4G bytes, shader-accelerated dependent texture lookup is being exploited in order to achieve interactive user manipulation of such a huge volume texture.

3.3 Virtual Orthopedics Training

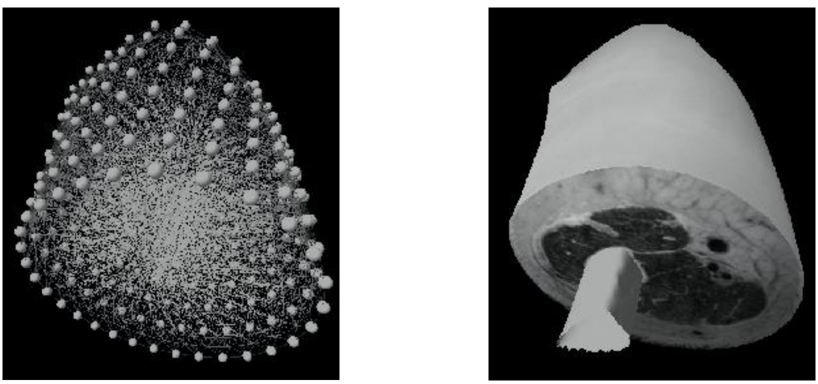
The Orthopaedics trainer is another visible human based surgical education and training system we developed under collaboration with the Department of Orthopaedics and Traumatology at the Prince of Wales Hospital, Hong Kong [19]. The upper limb region of the selected CVH data is segmented into dermis, hypodermis, fascia, deltoid muscle, brachial muscle, triceps brachii muscle, biceps brachii muscle, coracobrachialis muscle, extensor carpi radialis longus muscle, artery, vein, nerve, bone and bone marrow. **Fig. 7** (a) shows the user interface of the orthopaedics training application while another selected shot is shown in **Fig. 7** (b).

Realistic dermal deformation has been researched so that the simulation environment copes with orthopaedics training. A relatively new multi-core hardware accelerator called Physics-Processing Unit (PPU) has been exploited for efficient modeling of volumetric mass-spring system. Microscopic computational parameters of the non-linear spring and damper model are learned through a simulated annealing engine



(a) (b)

Fig. 7. (a) The user interface of the orthopaedics training application is shown. The upper limb region obtained from the second Chinese Visible Human dataset is utilized, (b) The subcutaneous fatty tissue is shown.



(a) (b)

Fig. 8. (a) The mass spring model of the upper limb model (b) the deformed limb

through inputting the macroscopic biomechanics properties of various dermal tissue, like epidermis, dermis, hypodermis. In this sense, real-time physiological deformation of skin tissue can be achieved for assisting surgical education and training.

Fig. 8 shows the mass-spring volumetric model being generated for realistic soft tissue simulation. The multilayer topologic structure can be automatically extracted from segmented CVH datasets. Interactive deformation can be computed using the PPU mass-spring framework while visualization can be done at interactive frame-rate.

4 Conclusion

In this paper, important virtual medicine researches related to visible human data sets have been outlined. In particular, the Visible Human related research in carried out in CUHK has been discussed in detail, while related applications such as virtual knee arthroscopy, virtual acupuncture as well as virtual anatomy briefly shaped our current and future research and development on visible human-related virtual medicine.

Acknowledgments. The work described in this paper was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CUHK4223/04E). Authors would like to thank Dr. Zhang Shaoxiang and his research team at the Third Military Medical University, Chongqing, China for providing the CVH data.

References

1. Spitzer, V., Ackerman, M.J., Scherzinger, A.L., Whitlock, D.: The Visible Human Male: A technical report. J Am. Med. Inform Assoc 3, 118–130 (1996)
2. Zhang, S.X., Heng, P.A., Liu, Z.J.: Chinese Visible Human Project: the Present and Future. Clinical Anatomy 19(3), 204–215 (2006)
3. Zhang, S.X., Heng, P.A., Liu, Z.J., Tan, L.W., Qiu, M.G., et al.: Creation of the Chinese Visible Human data set. Anat Rec (Part B: New Anat) 275, 190–195 (2003)
4. Park, J.S., Chung, M.S., Hwang, S.B., Lee, Y.S., Har, D.H., Park, H.S.: Visible Korean Human: Improved Serially Sectioned images of the entire body. IEEE Trans Med. Imaging 2005 24, 352–260 (2005)
5. Robb, R.A.: The biomedical imaging resource at Mayo Clinic. IEEE Transaction on Medical Imaging 2001 20, 854–867 (2001)
6. http://archive.nlm.nih.gov/pubs/reports/bsc2003-AQ/BSC2003-AQ.pdf
7. Nowinski, W.L., Yang, G.L., Yeo, T.T.: Computer-aided stereotactic functional neurosurgery enhanced by the use of the multiple brain atlas databases. IEEE Trans Med Imaging 19, 62–69 (2000)
8. Höhne, K.H. (ed.): VOXEL-MAN, Part 1: Brain and Skull, Version 1.1. Springer-Verlag Electronic Media, Heidelberg (1996)
9. Gehrmann, S., Höhne, K.H., Linhart, W., Pflesser, B., Pommert, A., Riemer, M., Tiede, U., Windolf, J., Schumacher, U., Rueger, J.M.: A novel interactive anatomic atlas of the hand. Clinical Anatomy 3, 258–266 (2006)
10. Celina, I., Pat, M.: Incorporating 3D Virtual Anatomy into the Medical Curriculum. Communications of the ACM 2, 49–54 (2005)
11. http://www.ornl.gov/info/ornlreview/v33_1_00/virtual.htm
12. Robb, R.A.: Virtual endoscopy: development and evaluation using the Visible Human datasets. Comput Med Imaging Graph. 24(3), 133–151 (2000)
13. Heng, P.A., Xie, Y., Wang, X., Chui, Y.P., Wong, T.T.: Virtual Acupuncture Human based on Chinese Visible Human Dataset. In: Proceedings of the Medicine Meets Virtual Reality 14 Conference, Long Beach, CA, USA (January 2006)
14. Heng, P.A., Wong, T.T., Leung, K.M., Chui, Y.P., Sun, H.: A Haptic Needle Manipulation Simulator for Chinese Acupuncture Learning and Training. International Journal of Image and Graphics 6(2), 205–230 (2006)

15. Heng, P.A., Zhang, S.X., Xie, Y.M., Wong, T.T., Chui, Y.P., Cheng, C.Y.: Photorealistic Virtual Anatomy Based on Chinese Visible Human Data. Clinical Anatomy 19(3), 232–239 (2006)
16. Zhang, S.X., Heng, P.A., et al.: The Chinese Visible Human (CVH) Datasets Incorporate Technical and Imaging Advances on Earlier Digital Humans. Journal of Anatomy 204(3), 165–173 (2004)
17. Heng, P.A., Zhang, S.X., Xie, Y.M., Wong, T.T., Chui, Y.P., Cheng, C.Y.: Virtual Anatomy Based on Chinese Visible Human Data, RSNA'04 InfoRAD Exhibits
18. Heng, P.A., Zhang, S.X., Xie, Y.M., Wong, T.T., Chui, Y.P., Cheng, C.Y.: Virtual Neuroanatomy, RSNA'05 InfoRAD Exhibits
19. Heng, P.A., Cheng, C.Y., Wong, T.T., Wu, W., Xu, Y.S., Xie, Y.M., Chui, Y.P., Chan, K.M., Leung, K.S.: Virtual Reality Techniques - Application to Anatomic Visualization and Orthopaedics Training. Clinical Orthopaedics and Related Research, No. 442, 5–12 (2006)

Finite Element Analysis of a Six-Component Force Sensor for the Trans-Femoral Prosthesis

Xiao Hu, Rencheng Wang, Fangfang Wu, Dewen Jin, Xiaohong Jia, Jichuan Zhang, Fuwen Cai, and Shuangxi Zheng

Division of Intelligent and Biomechanical System
State Key Laboratory of Tribology, Tsinghua University Beijing, 100084, China
wangrc@mail.tsinghua.edu.cn

Abstract. It is significant to detect and analyze its mechanical property for the design of the artificial knee joint, especially for the design of an osseointegrated prosthetic limb. Since normal six-component force sensors are unsuitable for detecting the mechanical property of the lower limb prosthesis, a novel sensor is presented in this paper. It can be easily fixed between the artificial knee joint and the stump of the patient to detect the load condition during walking. The mathematic model of the sensor is analyzed, and strength check, stiffness design and the linearity of the sensor were studied with FEA. Finally, the Transmission Matrix is calculated. This kind of sensor can help us to get academic foundations for the designment and the evaluation of the limb prosthesis and its performance.

Keywords: Six-component force sensor, Prosthesis, FEA.

1 Introduction

Due to the complex load conditions of the lower limb prosthesis during walking, it is significant to detect and analyze its mechanical property for the design of the artificial knee joint, especially for the design of an osseointegrated prosthetic limb, which is implanted directly to the bone of the stump without socket [1][2]. The sensor used to detect the mechanical stress for prosthetic limb necessarily requires relatively great measuring range and should be easy to link to the artificial joint and the stump. However, the present six-component sensors have not enough range [3][4], or are too complicated to be utilized to detect the load of the lower prosthesis limb [5][6]. Moreover, most of these sensors are typically used in robotics and other industrial applications where a large mass is often not as much of a concern as in prosthetics applications [7]. Hence a special six-component force sensor is designed in this paper to meet such requirements.

2 Sensor Structure and Mathematic Model

The arrangement of strain gages on the elastic body of the sensor is shown in the Fig. 1. Six pairs of strain gages are stuck to the outer surface of the cylinder

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 633–639, 2007.

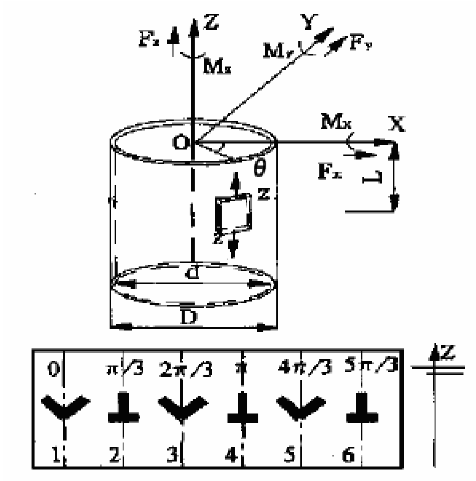


Fig. 1. Arrangement of strain gages on the elastic body

equidistantly. Each pair of strain gages is composed of two single perpendicular gages. The 1st, 3rd and 5th pairs of gages are used to detect the shearing strain. And the 2nd, 4th and 6th pairs are used to detect the normal strain. When the elastic body is deforming, according to the principle of superposition, we can derive the relationship between stress and each component of 6-axis forces.

The relationship of normal stress and relevant components of 6-axis forces is:

$$\sigma_{Z1} = -\left(F_X L \cos\theta\right) / W_X \tag{1}$$

$$\sigma_{Z2} = -\left(F_Y L \sin\theta\right) / W_X \tag{2}$$

$$\sigma_{Z3} = \left(M_X L \sin\theta\right) / W_X \tag{3}$$

$$\sigma_{Z4} = -\left(M_Y L \cos\theta\right) / W_X \tag{4}$$

$$\sigma_{Z5} = -F_Z / S \tag{5}$$

Where $W_X = \pi\left(D^4 - d^4\right)/32D$ is the section modulus in bending, $S = \pi\left(D^2 - d^2\right)/4$ is the cross-section area of the elastic body, D is the outer diameter of the elastic body and d is the inner diameter, L and θ are shown in the Fig. 1.

The shearing stress of the element can be caused by F_x, F_y, M_z, and the relationships are shown below:

$$\tau_{Fy} = \left(F_Y \cos^2\theta\right)/S \tag{6}$$

$$\tau_{Fx} = \left(F_X \sin^2\theta\right)/S \tag{7}$$

$$\tau_{Fy} = M_Z / B \tag{8}$$

Where $B = S\left(D + d\right)/2$.

From (1) ~ (8), it could be derived the relationship between the output voltage of the six pairs of strain gages and the 6-axis forces:

$$U=KC'F=CF \tag{9}$$

Where $U=\left[U_1, U_2, U_3, U_4, U_5, U_6\right]^T$ is output voltage of the six pairs of strain gages, $F=\left[F_1, F_2, F_3, M_1, M_2, M_3\right]^T$ is the 6-axis forces, K is the gage and bridge gain, which is a diagonal matrix, and C is the Transmission Matrix of the sensor, which is defined as (10) following:

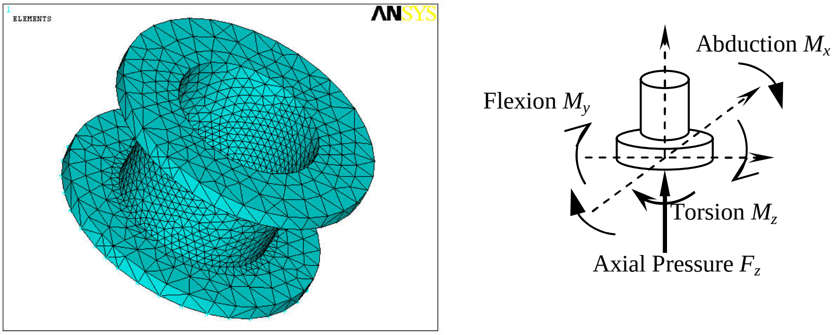
$$C=\begin{bmatrix} 0 & \frac{K_1}{SG} & 0 & 0 & 0 & \frac{K_1}{BG} \\ -\frac{K_2L}{2W_xE} & -\frac{\sqrt{3}K_2L}{2W_xE} & -\frac{K_2}{SE} & \frac{\sqrt{3}K_2L}{2W_xE} & -\frac{K_2L}{2W_xE} & 0 \\ -\frac{3K_3}{4SG} & -\frac{K_3}{4SG} & 0 & 0 & 0 & \frac{K_3}{BG} \\ \frac{K_4L}{W_xE} & 0 & -\frac{K_4}{SE} & 0 & \frac{K_4}{W_xE} & 0 \\ \frac{3K_5}{4SG} & -\frac{K_5}{4SG} & 0 & 0 & 0 & \frac{K_5}{BG} \\ -\frac{K_6L}{2W_xE} & \frac{\sqrt{3}K_6L}{2W_xE} & -\frac{K_6}{SE} & -\frac{\sqrt{3}K_6L}{2W_xE} & -\frac{K_6L}{2W_xE} & 0 \end{bmatrix} \tag{10}$$

Where E is the modulus of elasticity, and G is the shearing modulus.

3 Finite Element Analysis

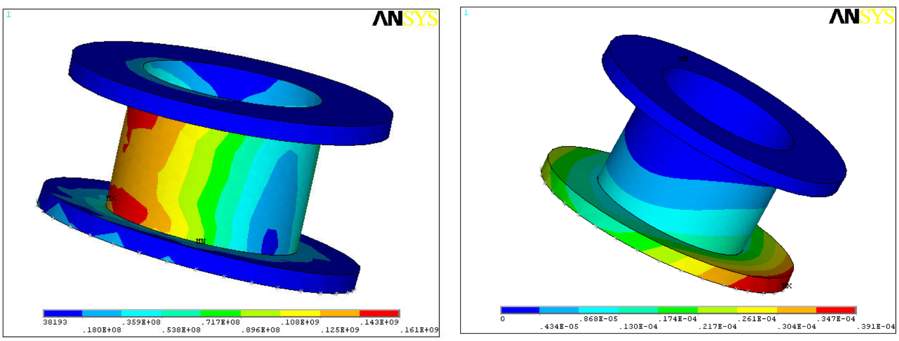
3.1 Sensor Performance

The sensor was modeled as commercially 40Cr with elastic modulus of 210GPa and a Poisson's ratio of 0.28. The wall thickness of the elastic body is 1mm and its length is



(a) (b)

Fig. 2. Sensor model and applied loads



(a) (b)

Fig. 3. Sensor contour maps of von Mises stress and deflection

20mm. The ANSYS FE package was used for stress and strain analysis. Simplified three-dimensional model of the sensor was developed using eight-noded axisymmetric harmonic elements (SOLID45), shown in Fig. 2(a). The model was free meshed. A total number of 6294 solid elements were generated to simulate the structure of the sensor. As shown in Fig. 2(b), a combination of axial compressive forces (F_y=3750N), torsion (M_z=20Nm), flexion (M_x=60Nm) and abduction (M_y=40Nm) bending moments load cases were applied to the models for the following investigation [8].

Fig. 3(a) shows the von Mises stress of the elastic body of sensor under the applied forces. The maximal stress of the model is 161MPa, which is less than the allowable stress of 40Cr (230MPa). Fig. 3(b) also shows the deflection of the sensor under the same condition. The maximal deflection is 39.1um, which has no negative effects on the sensor structure, so that the sensor can be safely used on the prosthesis. Therefore, it can be concluded that the material 40Cr is qualified in the sensor design.

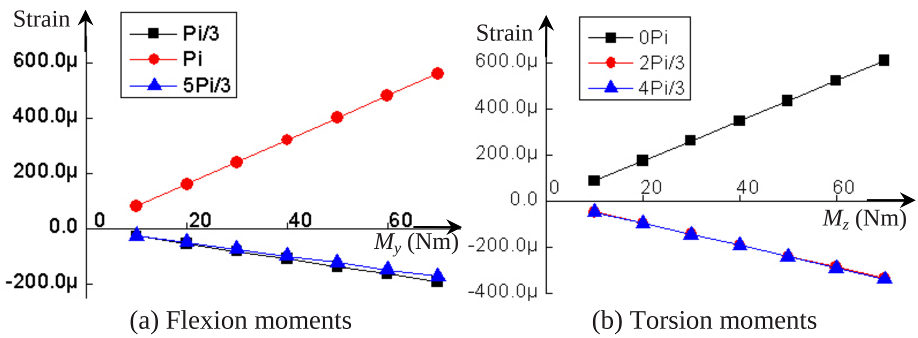


Fig. 4. Sensor linearity under flexion and torsion moments

The strain gages would be stuck on the middle of the sensor cylinder, 10mm from each end. With ANSYS, on those positions the gages would be stuck, the linearity of the sensor during the full measure range could be studied. In the full range of flexion and torsion moment (0~70Nm), we apply seven flexion moments to the sensor, each with an increment of 10Nm. The results are shown in the Fig. 4. The gage poison is shown with angle, Pi/3, Pi, etc. Other applied conditions also show the same results, but can not be indicated here due to the abstract length limitation.

3.2 Transmission Matrix of the Sensor

From (9), it could be derived:

$$F=C^{-1}U=C^{*}U \tag{11}$$

Where C^* is called the Transmission Matrix of the sensor.

According to (9), it can be derived that:

$$C^{*}=\left(C'\right)^{-1}K^{-1} \tag{12}$$

So, (11) could be transformed as:

$$F=C^{*}U=\left(C'\right)^{-1}K^{-1}U=\left(C'\right)^{-1}S=C_1'S \tag{13}$$

The Eq. (9) could be transformed as:

$$S=C'F \tag{14}$$

Where S is the strains of the key points where the six pairs of gages are stuck. Eq. (13) and (14) indicates the relationship between the strain of gages and the forces applied to the sensor.

Since $C^{*}=\left(C'\right)^{-1}K^{-1}=C_1'K^{-1}$ and $C=C'K$, the Transmission Matrix and Calibration Matrix could be studied preliminarily with C'. With ANSYS, six loads were respectively applied to the model of the sensor (Table 1). The strains of the key point could be obtained with ANSYS simulation.

Table 1. Six loads applied to the sensor model

Shear Force (F_x, N)	Shear Force (F_y, N)	Axial Pressure (F_z, N)	Abduction (M_x, Nm)	Flexion (M_y, Nm)	Torsion (M_z, Nm)
1000	1000	2000	50	50	50

With the simulation results, the C' could be obtained:

$$C'=1.0e\text{-}5\bullet\begin{bmatrix} 0.0000 & 0.0262 & 0.0000 & -0.0037 & 0.0174 & 0.9294 \\ -0.0048 & -0.0082 & -0.0053 & 0.6283 & -0.3721 & 0.0011 \\ -0.0109 & -0.0064 & -0.0000 & -0.0013 & -0.0003 & 0.9452 \\ 0.0093 & 0.0000 & -0.0049 & -0.0057 & 0.7026 & 0.0007 \\ 0.0118 & -0.0070 & 0.0000 & 0.0003 & 0.0058 & 0.9356 \\ -0.0047 & 0.0081 & -0.0049 & -0.6177 & -0.3436 & -0.0009 \end{bmatrix} \tag{15}$$

Compared with the theoretic Eq. (10), the main features are the same. Every column relates to one component of 6-axis forces applied to the sensor. For example, the first column relates to F_x, the fourth column relates to M_x, etc. Then, the Transmission Matrix could be obtained:

$$C^*=(C')^{-1}K^{-1}=1.0e6\bullet\begin{bmatrix} 0.0821 & 0.0232 & -4.4398 & -0.0103 & 4.4040 & 0.0351 \\ 3.0519 & 0.0289 & -1.5902 & -0.0423 & -1.4253 & 0.0142 \\ -0.0008 & -6.5710 & -0.0023 & -6.7181 & 0.0094 & -6.6208 \\ 0.0398 & 0.0775 & -0.0199 & -0.0005 & -0.0196 & -0.0833 \\ -0.0009 & -0.0459 & 0.0586 & 0.0952 & -0.0583 & -0.0477 \\ 0.0218 & 0.0004 & 0.0437 & -0.0005 & 0.0411 & 0.0002 \end{bmatrix}\bullet K^{-1} \tag{16}$$

When calibrating the sensor, the obtained theoretical Transmission Matrix could be used as a reference. However, due to the fabricating error and error caused by sticking the gages on the sensor, the practically observed value of C^* would deviate a little from the theoretical calculated value.

4 Conclusion

The sensor proposed in this paper is easy to link to the prosthesis, and strength check, stiffness design and the linearity of the sensor were studied with FEA. The results indicate the sensor could meet the requirements of the design and application.

Acknowledgement. This study is supported by a grant from the Chinese NSFC (№50475017).

References

1. Xu, W., Crocombe, A.D., Hughes, S.C.: Finite Element Analysis of Bone Stress and Strain around a Distal Osseointegrated Implant for Prosthetic Limb Attachment. Journal of Engineering in Medicine 6, 595–602 (2000)
2. Hsieh, M.C., Fang, Y.K., Ju, M.-S., Chen, G.-S., Ho, J.-J., et al.: A Contact-Type Piezoresistive Micro-Shear Stress Sensor for Above-Knee Prosthesis Application. IEEE Journal of Microelectromechanical Systems 1, 121–127 (2001)
3. Kim, G.-S.: The Design of a Six-Component Force/Moment Sensor and Evaluation of its Uncertainty. Meas. Sci. Technol. 9, 1445–1455 (2001)
4. Chao, L.-P., Chen, K.-T.: Shape Optimal Design and Force Sensitivity Evaluation of Six-Axis Force Sensors. Sens Actuators A Phys 2, 105–112 (1997)
5. Nishiwaki, K., Murakami, Y., Kagami, S., Kuniyoshi, Y., Inaba, M. et al.: A Six-Axis Force Sensor with Parallel Support Mechanism to Measure the Ground Reaction Force of Humanoid Robot. In: Proc IEEE Int Conf Rob Autom, pp. 2277–2282 (2002)
6. Liu, S.A., Tzo, H.L.: A novel six-component force sensor of good measurement isotropy and sensitivities. Sens Actuators A Phys 2-3, 223–230 (2002)
7. Sanders, J.E., Miller, R.A., Berglund, D.N., Zachariah, S.G.: Modular Six-Directional Force Sensor for Prosthetic Assessment: A Technical Note. J Rehabil Res. Dev. 2, 195–202 (1997)
8. Baumann, J.U., Schar, A., Meier, G.: Reaction Forces and Moments at Hip and Knee. In: Orthopade 1, 29–34 (1992)

The Influence of Shoe-Heel Height on Knee Muscle Activity of Transtibial Amputees During Standing

Xiaohong Jia[1], Jichuan Zhang[1], Rencheng Wang[1], Lidan Fang[1], Dewen Jin[1] and Ming Zhang[2]

[1] Division of Intelligent and Biomechanical System, State Key Laboratory of Tribology, Tsinghua University, Beijing 100084, China
[2] Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong, China
jiaxh@mail.tsinghua.edu.cn

Abstract. In order to access the effects of shoe-heel height on knee muscle activity for transtibial amputees during standing, five male subjects volunteered for the study. Three pairs of shoes with zero, 20 mm and 40 mm heel height were used during normal standing. Surface EMG of 10 muscles was recorded by the Noraxon surface EMG collection system. EMG-MAV of the medial and lateral gastrocnemius of the sound leg almost change double with increase in heel height from zero to 40 mm, and EMG-MAV of the rectus fomris, vastus lateralis and vastus medialis of prosthetic side became larger to different extent. The finding in this paper suggested that an alignment change was necessary to accommodate the heel height change and the prostheses users should be cautious to choose shoes in daily life.

Keywords: Shoe-heel height, Surface EMG, Prosthesis, Standing.

1 Introduction

Artificial limb is a main way to recover the activity and appearance for amputees. The performance of the prosthesis influences the level of rehabilitation for the transtibial (TT) amputees greatly. The alignment, defined as the relative position and orientation of the prosthetic components, has been recognized as one main factor that determines the success of a prosthesis for TT amputees [1]. Misalignment always results in deviation of standing pattern and gait, increase of energy cost, pain or even tissue breakdown at the socket/stump interface.

Lots of work has been done to obtain a more quantitative technique and optimization criteria for TT prosthetic alignment [2-5], which has theoretical merit to provide quantitative information to the practitioners that is difficult to obtain through their direct subjective observation, and also to enable rehabilitation outcomes to serve as valuable evidence of treatment efficacy. However, even though the prosthesis was adjusted optimally for one amputee, misalignment would still happen when prosthesis was not used properly. However, even though the prosthesis is aligned optimally for one amputee, the alignment would be changed if the prosthesis is not used properly. It

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 640–645, 2007.

is well accepted that persons who use lower-limb prostheses are often restricted to the choice of a narrow range of heel heights [6]. But the change of shoes with different heel heights is necessary in human daily life which may destroy the fitted alignment setting by the prosthetist. At the same time, the static standing status is a main and basic activity of the amputee's daily life, it is important to study how much the alignment influences the static standing biomechanics. The objective of this study is to look into the effect of shoe-heel height on knee muscle activity of below-knee amputee during standing, and to provide quantitative information for a biomechanically optimal prosthetic alignment.

2 Methods

2.1 Subjects and Prostheses

Five males with unilateral TT amputation participated in this work. Table 1 outlined the subject characteristics. Each subject signed an approved human subject's consent form. From each subject, a certified prosthetist duplicated his existing PTB socket with 5-mm-thick polyurethane liner, and fitted it with new shank and SACH foot with two Otto Bock adapters 4R21 (Otto Bock Orthopadische GmbH & Co., Duderstadt, Germany). Each component can be angularly adjusted relative to each other in the antero-posterior and medio-lateral planes. The five subjects had no other concomitant disabilities and skin complications. This experienced prosthetist also joined in this study to ensure the function of each participant's prosthesis.

Table 1. Subjects data

Subject number	Age (years)	Body mass (Kg)	Height (cm)	Prosthetic side	Years since amputation	Cause	Activity level
1	23	57	170	left	6	trauma	high
2	58	70	168	right	12	trauma	moderate
3	32	75	177	left	3	infection	high
4	44	65	172	left	11	trauma	moderate
5	28	69	174	left	2	trauma	high

2.2 Experiment and Data Collection

Three levels of shoe-heel height were investigated in this study: shoes with zero, 20 mm and 40 mm heel height. All shoes used here were commercially available. Surface EMG signals were recorded to investigate the effect of shoe-heel height on knee muscle activity, at a sampling rate of 1500 Hz using pre-amplified bipolar, surface electrodes (Noraxon, Scottsdale, AZ, USA), which were placed on the clean-shaven skin overlying biceps femoris, rectuc femoris, vastus lateralis, and vastus medialis of both legs and the medial and lateral gastrocnemius of the sound leg. The ground electrode was placed overlying the tuberosity of the tibia. Since the shoes with 20 mm heel height are normal choice according to subjects' responses, the prostheses were aligned to be optimal for the subjects with this height before experiment. The alignment was kept the same when the heel height was changed. Every subject was asked to load the prosthesis to 40% of body weight as measured by the Otto Bock L.A.S.A.R. Posture alignment system for each trial. The experimental set up was shown in Fig. 1. Loading line of prosthetic side, plantar foot pressure were also simultaneously recorded to reveal the effects of heel height.

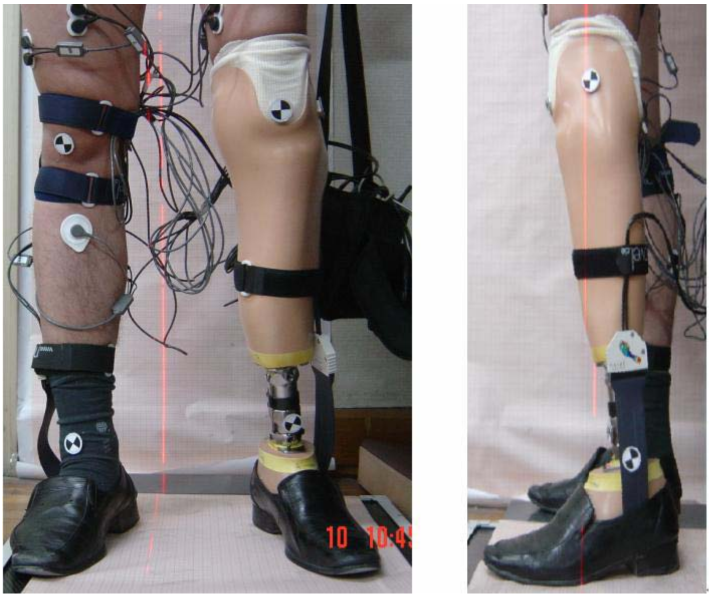


Fig. 1. Experiment system to study effect of heel height (anterior and lateral view)

The same experiment repeated twice. The shoe-heel height was changed in a random turn, and after a short break the data were re-measured. All data was recorded for 5 seconds when the subjects were asked to keep the standing posture for about 10 seconds. After each change, the subjects were asked to walk with close observation of the prosthetist to ensure the safety of the prostheses with different shoes.

2.3 Data Analysis

The EMG signal is widely used in analyzing the activity of muscles. Mean absolute value (MAV) of EMG was used to reveal the influence of the changes of the shoe-heel height, which was defined as equation 1. Here x_i is signal and N is the number of points.

$$\mathrm{MAV} = \frac{1}{N}\sum_{i=1}^{N} |x_i| \tag{1}$$

3 Results

Fig. 2 illustrated the comparison of typical muscle activity between sound and prosthetic side (these data were from subject 1, the others showed the same trend). In order to keep standing balance, muscles of prosthetic side were more active than the same ones of sound side. Especially for biceps femoris and vastus medialis of prosthetic side, the EMG-MAV was almost doubled comparing with sound side. This may be due to subjective adjustment. Biceps femoris and vastus medialis, which are responsible for keeping stability of knee joint, must contract more to prevent injury from hyperextension or flexion.

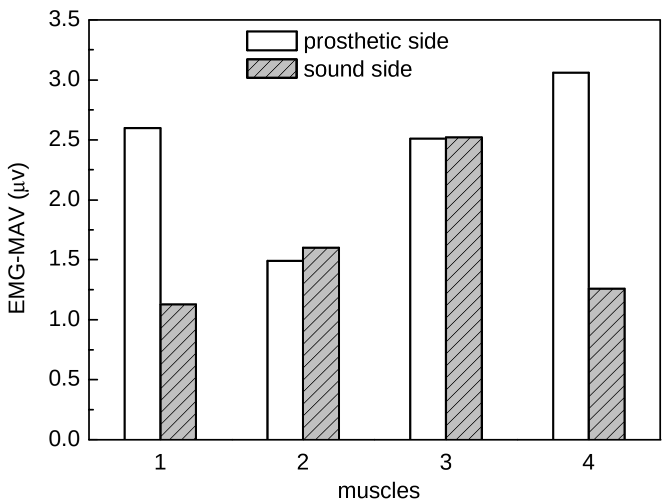


Fig. 2. The comparison of mean absolute value (MAV) of EMG of muscles of both sides, including 4 muscles (Subject 1). 1- biceps femoris, 2 –rectus femoris, 3 - vastus lateralis, 4 - vastus medialis.

Fig. 3 illustrated the typical muscle activity for sound and prosthetic side related to different shoe-heel height (these data were from subject 1 too, the others showed the same trend). During static standing, the EMG signals of knee muscles of the sound leg did not change a lot when the heel height was varied, whereas EMG-MAV of the medial and lateral gastrocnemius of the same leg was almost doubled when the heel-height increased from zero to 40 mm. This was because the two posterior caff muscles, responsible for ankle plantarflexion, played an important role in standing balance (Buckley, 2002). In contrast, the EMG of the knee muscles of the prosthetic side was systematically affected by the heel height change. When the heel height changed from zero to 40 mm, the knee flexion angle increased as a result. In order to keep the standing balance and knee stability, the muscles for knee extension must be more active. In Fig. 3, it can be found that EMG-MAV of the rectus fomris, vastus lateralis and vastus medialis became larger with higher shoe-heel to different extents. At the same time, the EMG-MAV of the biceps femoris reduced slightly.

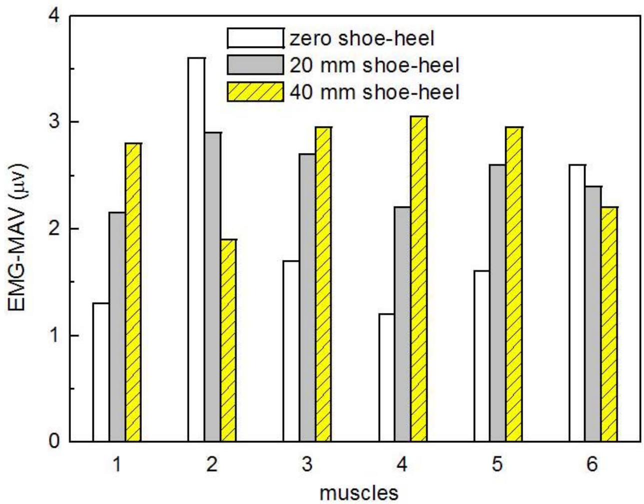


Fig. 3. Mean absolute value (MAV) of EMG of muscles change with the three heel heights, including 6 muscles (Subject 1). Sound side: 1- medial gastrocnemius, 2 - lateral gastrocnemius; Prosthetic side: 3 - rectus femoris, 4 - vastus medialis, 5 - vastus lateralis, 6 - biceps femoris.

4 Discussion

The importance of prosthetic alignment to the success of prosthetic fitting is well known, but work remains to be done in this area. Although it was reckoned that an alignment change is necessary to accommodate higher or lower heel height when shoe-heel height changes beyond a small range [6], only the effects of shoe-heel height on the rollover characteristics of the biologic ankle–foot system for non-disabled adult females were studied. In our primary work, the effect of shoe-heel height on the knee muscle activity was studied. Since the heel-height change for the TT amputee during standing actually caused a change in orientation angle in the sagittal plane of the socket relative to the ground, our results could be explained by other reports about sagittal alignment in the literature.

When the heel height was reduced to zero, the knee joint was hyper-extended. The rectus fomris, vastus lateralis, and vastus medialis, which are responsible to extend the knee joint, provided less force when heel height was zero because the EMG amplitude was approximately proportional to muscle force when the EMG was measured under isometric conditions. On the other hand, when heel height was changed from zero to 40 mm, an external flexion moment happened to the knee joint. To maintain the standing balance and joint stability, muscle forces of the rectus fomris, vastus lateralis, and vastus medialis had to increase. Although the posterior knee ligament relaxed, more energy consumption in the muscles could increase fatigue in the amputee. With long-term standing, some damage might occur in the knee joint.

Standing alignment for the TT amputee has been defined as biomechanically optimal when the anatomical knee center is 15 mm posterior to the individual loading line of the prosthetic side [7]. In this study, the prosthesis was adjusted to optimal alignment when the heel height was 20 mm. Under this condition, the loading line of the prosthetic side is 6~23 mm anterior to the knee center for five subjects, agreeing with the above optimal criterion by Blumentritt to some extent. The distance between the load line of prosthetic side and knee joint center was defined as *L*, which was positive when the load line was anterior to the knee center. It was found that *L* was 58 mm for zero shoe-heel and -32 mm for 40 mm shoe-heel (for subject 1).

In this study, the effects of shoe-heel height of transtibial amputees during standing were investigated based on knee muscle activity. Because the subjects would adjust the standing posture automatically to compensate the changes of the heel height, it may be difficult to find some directly practical criterion for prosthetic alignment. This study tried to emphasize the importance of proper use of prosthesis. The finding in this paper suggested that an alignment change was necessary to accommodate the heel height change and the prostheses users should be cautious to choose their shoes in daily life.

Acknowledgments. This work was supported by National Science Funding of China, No. 50575122. The all authors would like to thank Xiaobing Li and Yong Luo of National Rehabilitation Center of China.

References

1. Yang, L., Solomonidis, S.E., Spence, W.D.: The Influence of Limb Alignment on the Gait of Above-knee Amputees. J. Biomech. 11, 981–997 (1991)
2. Andres, R.O., Stimmel, S.K.: Prosthetic Alignment Effects on Gait Symmetry: a Case Study. Clin. Biomech. 5, 88–96 (1990)
3. Winter, D.A., Sienko, S.E.: Biomechanics of Below-knee Amputee Gait. J. Biomech. 21, 361–367 (1988)
4. Geil, M.D., Lay, A.: Plantar Foot Pressure Responses to Change during Dynamic Trans-tibial Prosthetic Alignment in a Clinical Setting. Prosthet. Orthot. Int. 28, 105–114 (2004)
5. Hannah, R.E., Morrison, J.B., Chapman, A.E.: Prosthesis Alignment: Effect on Gait of Persons with Below-knee Amputations. Arch. Phys. Med. Rehabil. 65, 159–162 (1984)
6. Hansen, A.H., Childress, D.S.: Effects of Shoe Heel Height on Biologic Rollover Characteristics during Walking. J. Rehab. Res. Dev. 4, 547–553 (2004)
7. Blumentritt, S., Schmalz, T., Jarasch, R., Schneider, M.: Effects of sagittal plane prosthetic alignment on standing trans-tibial amputee knee loads. Prosthet. Orthot. Int. 23, 231–238 (1999)

3D Modeling of the Vessels from X-Ray Angiography

Na-Young Lee[1], Gye-Young Kim[1], and Hyung-Il Choi[2]

[1] School of Computing, Soongsil University, Seoul, Korea
nayoung4493@ssu.ac.kr, gykim11@ssu.ac.kr
[2] School of Media, Soongsil University, Seoul, Korea
hic@ssu.ac.kr

Abstract. Coronary artery diseases are usually revealed using X-ray angiographies. Such images are complex to analyze because they provide a 2D projection of a 3D object. Medical diagnosis suffers from inter- and intra-clinician variability. Therefore, reliable software for the 3D modeling of the coronary tree is strongly desired. This paper focuses on the automatic 3D modeling of the vessels from X-ray angiography. Our approach is based on a 3D model of standard vessels. The model is projected because it is difficult to suitably transform standard into individual vessels on 3D space. The modeling process is carried out in two steps. The first step consists of selecting automatically two sets of corresponding control points between standard and individual vessels. In the second step, 3D model of individual vessels is performed by warping with corresponding control points. Accurate descriptions of 3D model would be useful for quantitative diagnosis of atherosclerosis, for surgical or treatment planning, for monitoring disease progress or remission, and for comparing efficacies of treatments.

Index term: Coronary Angiography, Control point, Image Transformation, 3D Modeling.

1 Introduction

Coronary arteries are a key subject in cardiac diagnosis and therapy. Several medical imaging modalities such as X-ray angiography, computer tomography, and magnetic resonance imaging are available to assess the condition of the coronary vasculature. Applying domain knowledge in the form of a coronary artery model can considerably enhance associated algorithms for medical image processing, visualization, and reporting. With respect to image processing tasks such as the segmentation of the coronary arteries [1], an estimate of the expected course of the coronary arteries within the imaged anatomy is of large interest. This information can for example help to bridge gaps caused by either imaging errors (e.g. motion artifacts) or by true vascular properties (e.g. a stenosis).

This paper describes a new method for fully automatic 3D modeling of the vessels from X-ray angiography. 3D model provides many important anatomical

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 646–654, 2007.

measurements that neither are available, nor can be accurately measured in 2D. For example, the projected length of a vessel is shorter in the projected views. Torque and the curvature of vessels are virtually impossible to estimate from 2D views. The 3D model of vessels provides patients with better and cleaner visualization without extensive training to understand vessels geometry. It saves reviewing time for physicians since 3D model may be performed by a trained technician, and may also help visualize dynamics of the vessels.

The structure of the paper is as follows. In Section 2, we describe the two major stages of our algorithm. Experimental results obtained for clinical datasets are discussed in Section 3. Finally, we discuss conclusion in Section 4.

2 Methodology

We propose a new approach for fully automatic 3D modeling of the vessels from X-ray angiography. The modeling method consists of two steps. The first step consists of selecting automatically two sets of corresponding control points between standard vessels and individual vessels. In the second step, 3D model of individual vessels is performed by warping with corresponding control points.

Overall system configuration is as shown in Fig. 1.

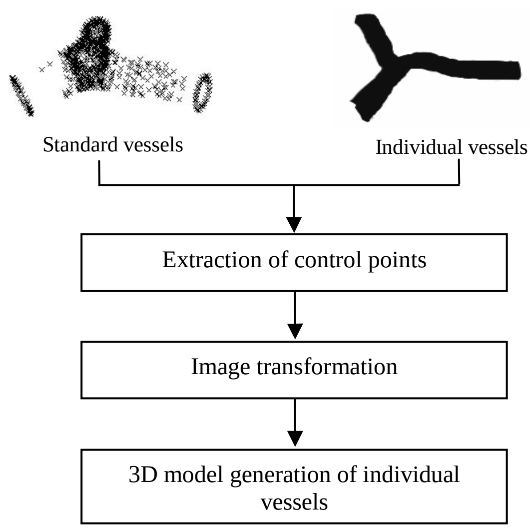


Fig. 1. Overall system configuration

2.1 Extraction of Control Points

To generate 3D model of individual vessels, individual vessel is modified and adjusted standard model based on extracted vessel area in angiogram.

We used the Harris corner detector to extract the control points from the standard and individual vessels. The Harris corner detector is a popular interest point detector due to its strong invariance to [2]: rotation, scale, illumination variation and image noise. The Harris corner detector is based on the local auto-correlation function of a signal; where the local auto-correlation function measures the local changes of the signal with patches shifted by a small amount in different directions. A discrete predecessor of the Harris detector was presented by Moravec [3]; where the discreteness refers to the shifting of the patches.

Given a shift $(\Delta x, \Delta y)$ and a point (x, y), the auto-correlation function is defined as,

$$c(x,y) = \sum_{W} [\, I(x_i\, , y_i) - I(x_i + \Delta x\, , y_i + \Delta y)\,]^2 \tag{1}$$

Where I denotes the image function and (x_i, y_i) are the points in the window W centered on (x, y).

The shifted image is approximated by a Taylor expansion truncated to the first order terms,

$$I(x_i + \Delta x\, , y_i + \Delta y) \approx I(x_i\, , y_i) + [\, I_x(x_i\, , y_i) \;\; I_y(x_i\, , y_i)] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \tag{2}$$

Where I_x and I_y denote the partial derivatives in x and y, respectively.

Substituting approximation equation (2) into equation (1) yields,

$$\begin{aligned} c(x,y) &= \sum_{W} [\, I(x_i\, , y_i) - I(x_i + \Delta x\, , y_i + \Delta y)\,]^2 \\ &= [\, \Delta x \; \Delta y\,]\, C(x,y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \end{aligned} \tag{3}$$

Where matrix $C(x,y)$ captures the intensity structure of the local neighborhood.

Let λ_1, λ_2 be the eigenvalues of matrix $C(x,y)$. The eigenvalues form a rotationally invariant description. If both λ_1, λ_2 are small, so that the local auto-correlation function is flat (i.e., little change in $c(x,y)$ any direction), the windowed image region is of approximately constant intensity. If one eigenvalue is high and the other low, so the local auto-correlation function is ridge shaped, then only local shifts in one direction (along the ridge) cause little change in $c(x,y)$ and significant change in the orthogonal direction; this indicates an edge. If both eigenvalues are

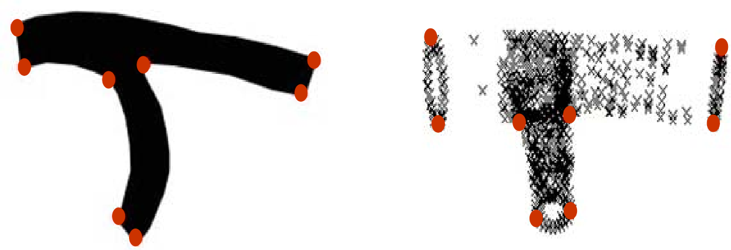


Fig. 2. Result of extracted control points

high, so the local auto-correlation function is sharply peaked, then shifts in any direction will result in a significant increase; this indicates a corner.

We extracted automatically two sets of corresponding control points between standard and individual vessels.

The Fig.2 shows the result of extracting the control points by applying the Harris corner detector.

2.2 Image Transformation

We have warped the standard vessels with respect to the individual vessels. Given the two sets of corresponding control points, $S = \{s_{1,} s_2, \ldots s_m\}$ and $I = \{i_{1,} i_2, \ldots i_m\}$, 3D model of vessels is performed by warping of the standard vessels. Here, S is a set of control points in standard vessels and I is a set of one in individual vessels.

In this work, standard vessels warping is performed by applying the elastic TPS(Thin-Plate-Spline) interpolation function[5] on the two sets of feature points.

The TPS are interpolating functions, representing the distortion at each feature point exactly, and defining a minimum curvature surface between control points. A TPS function is a flexible transformation that allows rotation, translation, scaling, and skewing. It also allows lines to bend according to the TPS model. Therefore, a large number of deformations can be characterized by the TPS model.

The TPS interpolation function can be written as

$$h(x) = Ax + t + \sum_{i=1}^{m} W_i K(\| x - x_i \|) \tag{4}$$

Where A are t are the affine transformation parameters matrices, W_i are the weights of the non-linear radial interpolation function K, and x_i are the control points. The function $K(r)$ is the solution of the biharmonic equation $(\Delta^2 K = 0)$ that satisfies the condition of bending energy minimization, namely $K(r) = r^2 \log(r^2)$.

To solve for the x, y components, the transformation h is defined from two sets of m corresponding control point pairs as

$$x' = h_x(x,y) = a_{11}x + a_{12}y + t_x + \sum_{i=1}^{m} W_{xi} K(\| (x_i, y_i) - (x,y) \|)$$
$$y' = h_y(x,y) = a_{21}x + a_{22}y + t_y + \sum_{i=1}^{m} W_{yi} K(\| (x_i, y_i) - (x,y) \|) \quad (5)$$

With parameters $a_{11}, a_{12}, a_{21}, a_{22}, t_x,$ and t_y representing the linear affine transformation, and with parameters W_{xi} and W_{yi} representing the weights of the function K.

The warping of the standard vessels with respect to the individual vessels is accomplished by carrying out the following steps:

1. Given the two sets S and I of m corresponding control point pairs, the affine transformation parameters A and t are obtained using the following equation (6):

$$y = Mz \quad (6)$$

$$where, \quad y = \begin{bmatrix} x_1' \\ y_1' \\ x_2' \\ y_2' \\ M \\ x_m' \\ y_m' \end{bmatrix}, \quad M = \begin{bmatrix} x_1 & y_1 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_1 & y_1 & 0 & 1 \\ x_1 & y_1 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_1 & y_1 & 0 & 1 \\ M & M & M & M & M & M \\ x_m & y_m & 0 & 0 & 1 & 0 \\ 0 & 0 & x_m & y_m & 0 & 1 \end{bmatrix}, \text{ and } \quad z = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \\ t_x \\ t_y \end{bmatrix}$$

$a_{11}, a_{12}, a_{21}, a_{22}, t_x,$ and t_y are the affine transformation parameters and are found by taking the pseudo-inverse of matrix M

$$z = (M^T M)^{-1} (M^T y) \quad (7)$$

Where M^T is the transpose of M and $(M^T M)^{-1}$ denotes the inverse of $M^T M$.

2. The two sets of corresponding control point pairs are then used to calculate the weights W_{xi} and W_{yi} using the equations (8):

$$\begin{bmatrix} x_1' - a_{11}x_1 - a_{12}y_1 - t_x \\ x_2' - a_{11}x_2 - a_{12}y_2 - t_x \\ M \\ x_m' - a_{11}x_m - a_{12}y_m - t_x \end{bmatrix} = \begin{bmatrix} 0 & K(r_{12}) & K & K(r_{1m}) \\ K(r_{21}) & 0 & K & K(r_{2m}) \\ M & M & M & M \\ K(r_{m1}) & K(r_{m2}) & K & 0 \end{bmatrix} \begin{bmatrix} W_{x1} \\ W_{x2} \\ M \\ W_{xm} \end{bmatrix}$$

$$\begin{bmatrix} y_1' - a_{21}x_1 - a_{22}y_1 - t_y \\ y_2' - a_{21}x_2 - a_{22}y_2 - t_y \\ M \\ y_m' - a_{21}x_m - a_{22}y_m - t_y \end{bmatrix} = \begin{bmatrix} 0 & K(r_{12}) & K & K(r_{1m}) \\ K(r_{21}) & 0 & K & K(r_{2m}) \\ M & M & M & M \\ K(r_{m1}) & K(r_{m2}) & K & 0 \end{bmatrix} \begin{bmatrix} W_{y1} \\ W_{y2} \\ M \\ W_{ym} \end{bmatrix} \quad (8)$$

Where $r_{ij} = \| (x_i , y_i) - (x_j' , y_j') \|$ and $K(r) = r^2 \log (r^2)$.

3. The complete set of parameters, defining the interpolating registration transformation is then used to transform the standard vessels. It should be noted that in order to be able to carry out the warping of the standard vessels with respect to the individual vessels, it is required to have a complete description of the TPS interpolation function.

The Fig. 3 shows the results of modifying the standard vessels to suit the model of individual vessels.

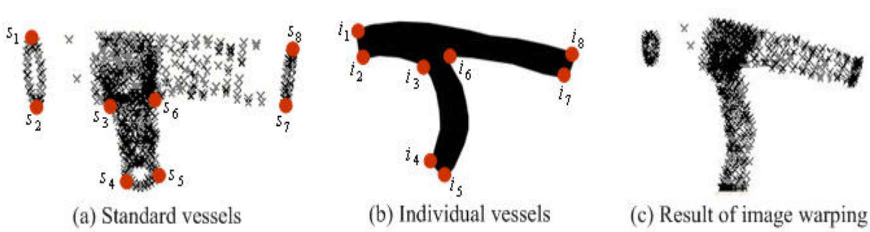


Fig. 3. Result of image transformation

3 Experimental Results

We simulated the system environment that is Microsoft Windows XP on a PentiumⅣ 3GHz, Intel Corp. and the compiler used was VC++ 6.0. The image used for experimentation was 512×512. Each image has a gray-value resolution of 8 bits, i.e., 256 gray levels.

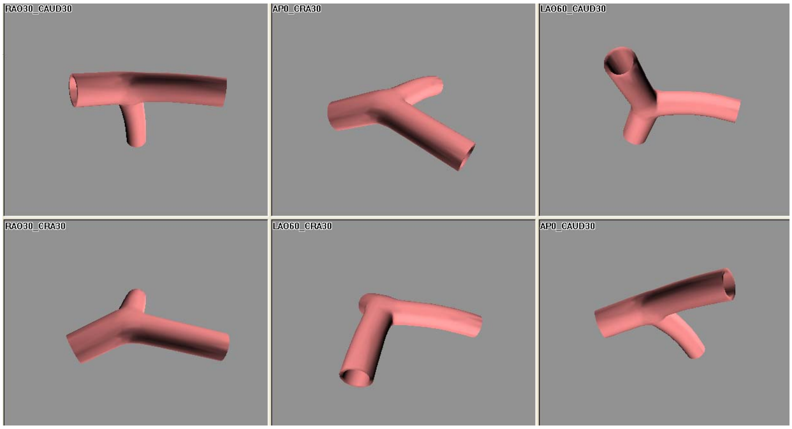


Fig. 4. 3D model of standard vessels in six views

The Fig. 4 shows the results of generating a 3D model of standard vessels from six different angiographic views.

The angiogram shows the different shape of images according to individual vessel. The 3D model of standard vessels is projected onto a 2D plane because angiogram is in 2D. The Fig. 5 shows the result of extracting the control points by applying the Harris corner detector to the standard vessels.

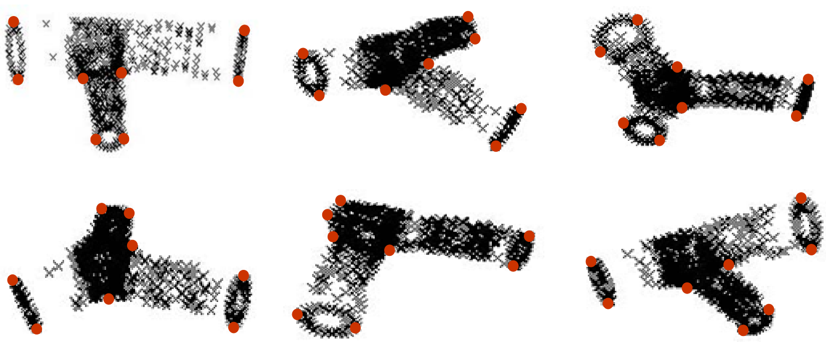


Fig. 5. Control points automatically extracted from standard vessels

The Fig. 6 shows the result of extracting the control points by applying the Harris corner detector to the vessels extracted from the angiogram.

The Fig. 7 shows the results of modifying the standard vessels to suit the model of individual vessels using the TPS interpolation function. The Fig.7(a) indicates the result of extracted eight control points from the standard and individual vessels. And the Fig.7(b) shows the result of generating an 3D model of individualized vessels by inverse projection onto the modified 2D standard model.

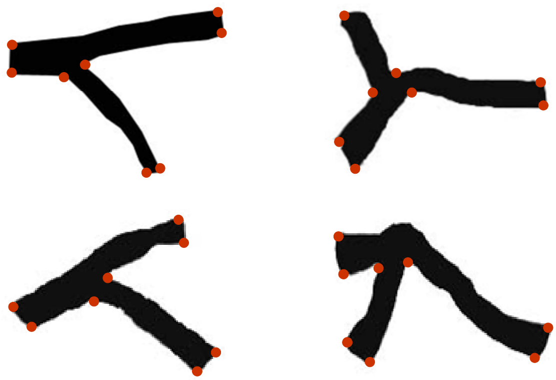
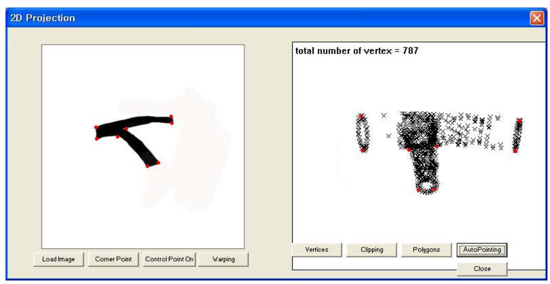
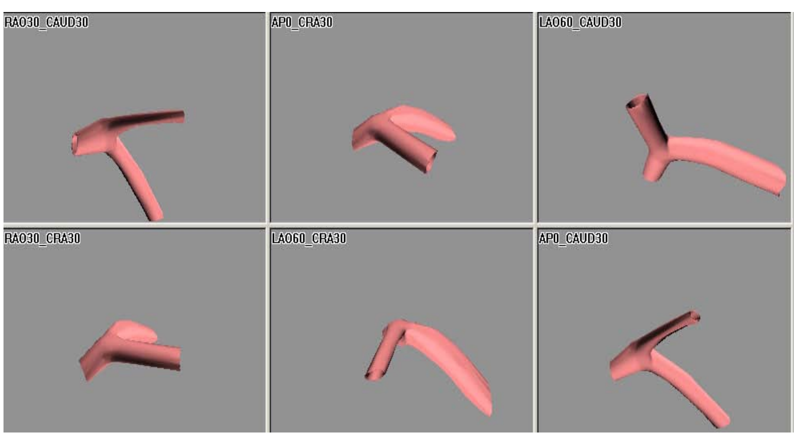


Fig. 6. Control points automatically extracted from individual vessels



(a) Extraction of corresponding control points from two vessels



(b) 3D model of individual vessels in six views

Fig. 7. Automatic generation result of individual vessels to 3D model

4 Conclusion

We have developed a fully automatic algorithm to perform the 3D model of individual vessels from six angiograms, taken from a single rotational acquisition. As demonstrated in the previous section, this approach can be used to recover the geometry of the main arteries. The 3D model of vessels enables patients to visualize their progress and improvement. Such a model should not only enhance the level of reliability but also provides speedy and accurate identification. In order words, this method can expect to reduce the number of misdiagnosed cases.

Acknowledgement. This work was supported by the Korea Research Foundation (KRF-2006-005-J03801).

References

1. Lorenz, C., Renisch, S., Schlatholter, S., Bulow, T.: SPIE, Simultaneous Segmentation and Tree Reconstruction of the Coronary Arteries in MSCT Images. In: Int. Symposium Medical Imaging, Proc. SPIE, San Diego, vol. 5032, pp. 167–177 (2003)
2. Schmid, C., Mohr, R., Bauckhage, C.: Evaluation of interest point detectors. International Journal of Computer Vision, 151–172 (2000)
3. Moravec, H.: Obstacle avoidance and navigation in the real world by a seeing robot rover, Technical Report CMU-RI-TR-3, Carnegie-Mellon University, Robotics Institute (1980)
4. Tomasi, C., Kanade, T.: Detection and Tracking of Point Features, Carnegie Mellon University Technical Report CMU-CS-91-132 (1991)
5. Bentoutou, Y., et al.: An invariant approach for image registration in digital subtraction angiography. Pattern Recognition, 34–48 (2002)
6. Brown, B.G., Bolson, E., Frimer, M., Dodge, H.: Quantitative coronary arteriography estimation of dimensions, hemodynamic resistance, and atheroma mass of coronary artery lesions using the arteriogram and digital computation. Circulation 55, 329–337 (1977)
7. de Feyter, P., Vos, J., Reiber, J., Serruys, P.: Value and limitations of quantitative coronary angiography to assess progression and regression of coronary atherosclerosis. In: Advances in Quantitative Coronary Arteriography, 255–271 (1993)
8. Ross, J., et al.: Guidelines for coronary angiography. Circulation 76, 963A–977A (1987)
9. Bookstein, F.L.: Principal warps: thin-plate splines and the decomposition of deformations. IEEE-PAMI 11, 567–585 (1989)
10. Flusser, J., Suk, T.: Degraded image analysis: an invariant approach. IEEE Trans. Pattern Anal. Mach. Intell, 590–603 (1998)
11. Venot, A., Lebruchec, J.F., Roucayrol, J.C.: A new class of similarity measures for robust image registration. Comput. Vision Graph. Image Process, 176–184 (19984)
12. Shi, J., Tomasi, C.: Good features to track. In: IEEE Conference on CVPR Seattle, 593–600 (1994)

Finite Element Modeling to Aid in Refining the Rehabilitation of Amputees Using Osseointegrated Prostheses

Winson Lee[1], Laurent Frossard[1], Nicola Cairns[1], Rickard Branemark[2], John Evans[1], Clayton Adam[1], and Mark Pearcy[1]

[1] Institute of Health and Biomedical Innovation, Queensland University of Technology, Australia

[2] Centre of Orthopaedic Osseointegration, Sahlgrenska University Hospital, Sweden

Abstract. The direct anchorage of a lower-limb prosthesis to the bone has been shown to be an excellent alternative for amputees experiencing complications in using a conventional prosthetic socket. After surgical implantation, amputees have to go through a weight bearing exercise program to prepare the bone to tolerate forces and promote bone-remodeling. Currently, the load magnitude prescribed by the clinician is measured by a weight scale which reports only the axial force in the limb. Previous study using a load transducer revealed that in addition to the axial force there were other forces and moments. This study develops a FE model and utilizes our load data to investigate the stress distribution at the bone-implant interface. The model shows that the stress distribution could be highly non-uniform during the exercise. Bone-implant interface stress has certain implications in pain adaptation and bone-remodeling, and a good understanding of it can assist in future attempts to refine and shorten the period of rehabilitation exercise.

Keywords: transfemoral amputation, bone anchorage prosthetics, osseointegration, weight bearing exercise, finite element modeling.

1 Introduction

Pain and tissue breakdown are sometimes experienced by amputees using conventional prostheses because of the high pressure applied from the socket to the residual limb. In addition, fitting problems are produced because of the fluctuations of the residual limb volume and the residual limb being too short. In an attempt to avoid these problems, a surgical approach for connecting a prosthesis into the femur through a titanium threaded-implant (osseointegration) was developed [1]. The implant is connected to the external prosthetic components through an abutment, which protrudes through the soft tissues and skin.

Trans-femoral amputees being fitted with osseointegrated prostheses have to undergo some specific rehabilitation exercises prior to full weight-bearing. One exercise is to apply a prescribed static load to the abutment, approximately along the axis of the residual femur using a conventional weigh-scale as force transducer.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 655–658, 2007.

Initially 20 kg of load is prescribed, and the loading is increased incrementally until full standing weight can be borne safely without pain. The weight is decreased to a non-painful level if pain is perceived. Full weight bearing is usually achieved between 3 and 6 months after the beginning of the static load bearing exercises.

A domestic weigh scale displays the magnitude of the vertical force only and, if the femur is vertical this corresponds to the axial force it experiences. If the limb is not vertical or is otherwise subjected to forces in the horizontal plane, these forces and the moments generated are not measured. Previous studies using a 6-channel load transducer have revealed that in addition to the axial forces there were some forces acting on the other axes and corresponding moments [2,3]. The effects of those additional loads on bone-implant interface stresses are of interest as they impact on the process of rehabilitation.

The stresses at the bone-implant interface have been calculated in this study using a finite element (FE) model with our previous load data as input. Comparisons of the stresses are made between two loading cases: 1) A load applied on the long axis of the limb only which corresponds to the load clinically prescribed (monitored with the weigh scale); and 2) A load applied on the three axes corresponding to the "true" load (measured simultaneously by a 6-channel load transducer).

2 Methods

The forces generated during the weight bearing exercise of one trans-femoral amputee fitted with an osseointegrated fixation were measured simultaneously using a weigh scale and a 6-channel load transducer. Detailed measurement techniques have been described in [3]. Utilizing the measured load data, finite element (FE) analysis was performed in Abaqus 6.6 to investigate the bone-implant interface stresses (Figure 1). Details of the model are as follows:

Geometry: The dimensions of the commercial implant (approximately 100mm long and 20mm in diameter) developed by Dr. R. Branemark were used. Three-dimensional bone geometry was downloaded from the BEL Repository. The amputated bone had a length of 230mm measuring from the greater trochanter to the distal cut end. The implant and the bone were assumed to be tied together.

Material properties: Young's moduli of 15GPa (bone) and 110GPa (implant) were assigned. Poisson's ratios of 0.3 were used for the bone and the implant.

Loadings: Two load cases were applied to at the distal end of the abutment. A prescribed axial force was 900N (Loading A), which represented nearly the full body weight of the subject; and the corresponding "true" load (Loading B) recorded by the 6-channel transducer which suggested there were lateral and posterior forces of 112.1N and 40.4N, respectively, in addition to the axial force, and with a lateral rotational moment of 13.4Nm. The femoral head was fully fixed.

Mesh: The number of 4-node tetrahedral elements reached approximately 100,000. A fine mesh was used at the interface between the bone and the implant.

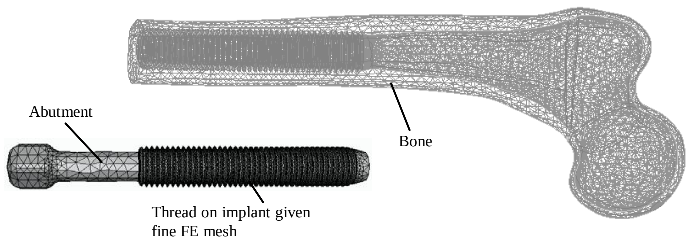


Fig. 1. FE meshes for bone and implant-abutment fixation

3 Results

Loading A: The FE model shows that the von Mises stress at the implanted site of the bone is distributed quite evenly with the peak von Mises stress of 1.5MPa when the axial-only force is applied. Immediately above the proximal end of the implant the von Mises stress becomes higher reaching 3MPa.

Loading B: The stress distribution is highly non-uniform when the forces and moments are applied on the three axes (Figure 2). The peak von Mises stress at the implanted site of the bone becomes significantly higher (7.7MPa) than the axial-only loading. Some areas along the implanted site experienced very low stresses (<0.5MPa). The von Mises stresses reached approximately 9MPa immediately above the proximal end of the implant.

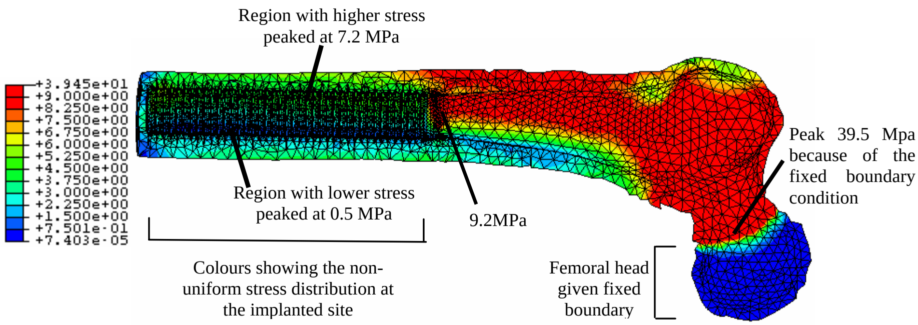


Fig. 2. Section view of the bone showing the von Mises stress distribution at Loading B condition

4 Discussion and Conclusion

This model demonstrated that the stress distribution at the bone-implant interface induced by the actual 3-dimensional loadings measured during the weight bearing exercise can be significantly different from that created by an axial-only force.

Bone-implant stress distribution could have certain implications in pain adaptation and bone-remodeling. Particularly high stresses applied to the bone should be avoided as it can produce pain. On the other hand, suitable amounts of stresses should be applied to the appropriate regions of the bone so as to prepare the amputees to walk without perception of pain and to promote bone-remodeling. Future attempts will be made to refine and shorten the period of rehabilitation exercise after the insertion of the implant and the abutment by re-distributing the stresses that favor bone-remodeling and improve bone tolerance to pain.

References

1. Branemark, R., Branemark, P.I., Rydevik, B., Myers, R.R: Osseointegration in skeletal reconstruction and rehabilitation: A review. J Rehabil Res. Dev. 38, 175–181 (2001)
2. Cairns, N., Frossard, L., Hagberg, K., Branemark, R.: Static load bearing during early rehabilitation of transfemoral amputees using osseointegrated fixation. Proceedings of the 30th Annual Scientific of ISPO Australia-Australia, pp. 51–52 (2006)
3. Frossard, L., Beck, J., Dillon, M., Evans, J.: Development and preliminary testing of a device for the direct measurement of forces and moments in the prosthetic limb of transfemoral amputees during activities of daily living. J Prosthet Orthot 15, 135–142 (2003)
4. Day, R.: 3rd generation composite femur, From: The BEL Repository, http://www.tecno.ior.it/VRLAB

Actions of an External Electrical Shock on Human Atrial Excitation – A Computer Model Study

Jihong Liu[1], A.V. Holden[2], and Henggui Zhang[3]

[1] College of Information Science and Engineering, Northeastern University, Shenyang, P.R. China,110004
liujihong@ise.neu.edu.cn
[2] School of Biomedical Sciences, The University of Leeds, Leeds, LS2 9JT, UK
[3] Biological Physics Group, School of Physics & Astronomy, The University of Manchester, Manchester, M60 1QD, UK

Abstract. Atrial fibrillation (AF) is one of the most common cardiac diseases that cause morbidity and mortality. One of the most frequently used clinical treatments of AF is to use a large and brief external electrical shock to reset atrial tissue from a disordered fibrillation state to a quiescent state, then the pacemaker of the heart resumes its control of atrial excitation rhythm and thus a defibrillation is achieved. Though widely used in practice, the mechanisms underlying the success of an electrical shock in defibrillation is incompletely understood. In this study, we developed a computer model of human atrial tissue to investigate the actions of an external electrical shock on atrial excitations. Using the model, we computed the defibrillation threshold of the human atrium. We found that due to the supernormal excitability of human atrium, the computed successful defibrillation threshold is much less than the excitation threshold of human atrium in resting state. This study provides some new insights to understand the mechanisms underlying AF defibrillation.

Keywords: External electrical stimulation, defibrillation threshold, bi-domain model of cardiac tissue, reentrant excitation.

1 Introduction

Atrial fibrillation (AF) is the most common cardiac diseases and can cause high risks of sudden cardiac death and disability (Nattel, 2002). Typically AF is characterised by spatio-temporal irregular excitation waves propagating at a high rate (300-500 /min). One of the most frequently used clinical treatments of AF is electrical cardioversion, which uses a brief (0.01-5 ms) and large (0.1-100 KV) external electrical shock to reset atrial tissue from disordered fibrillation to a quiescent state, and then the pacemaker of the heart can resume its control over the rhythm of atrial excitations. This technique is widely used in clinical practice and has dramatically saved the lives of patients with life-threatening cardiac arrhythmias including AF. However, its practical uses are limited by the lack of understanding the underlying mechanisms. A better understanding of these mechanisms will help to refine the design of defibrillation electrodes and the characteristics of stimulation pulse, and to improve the efficiency of defibrillation.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 659–667, 2007.

In this study, we developed a biophysically detailed bi-domain computer model of electrical activity of human atrial tissue. The model considered spatial effects of an external electrical field on a single cell (Biktashev et al., 1997; Zhang & Holden, 2004). Using the model, we studied the excitability and the activation of the atrial tissue by an external stimulation pulse. We found that the human atrial tissue has a feature of supernormal excitability. As a consequence the successful defibrillation threshold was much less than that of human atrial excitation threshold at a resting state and is also less than that predicted by the "upper limit of vulnerability" defibrillation theory (Pumir and Krinsky; 1996; Krinsky and Pumir, 1998).

2 Bi-domain Model of Electrical Activity of Human Atrial Myocytes

In this study, we used the Nygren et al. (1998) model to simulate the electrical action potential (AP) of human atrial myocytes. This is a biophysically detailed model, which was constructed based on a series of voltage clamp experiments on the kinetics of ionic channel of human atrial myocytes. We modified the model to incorporate the actions of an external electrical field, subjected to which a cardiac cell should be taken as a spatially extended object and modelled by a bi-domain model (Biktashev et al., 1997; Zhang & Holden, 2004). In general the modified model takes the form as Equation (1) to Equation (6).

Equation (1) describes the time-dependent rate of membrane potential V(t) of an atrial cell with a total cell membrane capacitance C_m. In the equation, i_j (V,t) is the j_{th} membrane ionic current, $P_j(C_i^m,C_i^n,C_o^m,C_o^n)$ the j_{th} pump/exchanger current, and C_i^m, C_o^m internal and external concentrations of the m_{th} ionic species as described by Nygren et al. (1998). *E* is the external field magnitude, s the index of spatial position of a membrane site with respect to the external field. Equations (3)-(5) specify the general form of voltage and time dependent membrane ionic channel currents.

$$\frac{dv}{dt} = -\frac{1}{C_m}\int (\sum_j i_j^s(V^s,t) + \sum_j P_j^s(C_i^m(s), C_i^n(s), C_o^m(s), C_o^n(s), V^s)K(s)sds \tag{1}$$

$$V^s = V + sE(t) \tag{2}$$

$$i_j^s(V^s,t) = (V^s - E_j)G_j^{\max} m_j^{p_j}(V^s)h_j^{q_j} \tag{3}$$

$$\frac{dm_j}{dt} = \alpha_j^m(V^s)(1-m_j(V^s)) - \beta_j^m m_j(V^s) \tag{4}$$

$$\frac{dh_j}{dt} = \alpha_j^h(V^s)(1-h_j(V^s)) - \beta_j^h h_j(V^s) \tag{5}$$

$$\frac{dC_i^j(s)}{dt} = -\frac{\sum_j i_j^{tot}(V^s)}{z_j F V_i} \tag{6}$$

K(s) is the spatial dependent weighting function, for which we have Equation (7) as following.

$$\int_s K(s)ds = 1 \tag{7}$$

Equations (1)-(7) can be simplified by using a two-compartment approximation of the integral $\int$ operator following the approaches developed by Biktashev et al. (1997) and Zhang & Holden (2004). Under the action of an external electrical field, a cardiac cell is nearly iso-potential, but it is depolarised in one part, and hyperpolarised in the other part, both by the same amplitude of external stimulation current i_{ext}. These two iso-potential parts are coupled through the intracellular resistance. If we denote the depolarised part by +, and the hyperpolarised part by -, the intracellular conductance by α, then equation (1)-(7) becomes equation (8) -(16).

$$\frac{dV}{dt} = -\frac{1}{2C_m} i_{tot} \tag{8}$$

$$i_j^+(V^+) = (V^+ - E_j)G_j^{\max} m_j^+(V^+)^{p_j} h_j^+(V^+)^{q_j} \tag{9}$$

$$\frac{dm_j^+}{dt} = \alpha_j^m (V^+)(1-m_j^+) - \beta_j^m (V^+) m_j^+ \tag{10}$$

$$\frac{dh_j^+}{dt} = \alpha_j^h (V^+)(1-h_j^+) - \beta_j^m (V^+) h_j^+ \tag{11}$$

$$i_j^-(V^-) = (V^- - E_j)G_j^{\max} m_j^-(V^-)^{p_j} h_j^-(V^-)^{q_j} \tag{12}$$

$$\frac{dm_j^-}{dt} = \alpha_j^m (V^-)(1-m_j^-) - \beta_j^m (V^-) m_j^- \tag{13}$$

$$\frac{dh_j^-}{dt} = \alpha_j^h (V^-)(1-h_j^-) - \beta_j^m (V^-) h_j^- \tag{14}$$

$$\frac{dC_i^+}{dt} = -\sum_j i_j^{tot^+}(V^+)/(z_j F V_i) \tag{15}$$

$$\frac{dC_i^-}{dt} = -\sum_j i_j^{tot^-}(V^-)/(z_j F V_i) \tag{16}$$

where

$$i_{tot} = \sum_j (i_j^+(V^+) + i_j^-(V^-)) + \sum_j (p_j(C_i^{m^+}, C_i^{n^+}, C_o^{m^+}, C_o^{n^+}) + p_j(C_i^{m^+}, C_i^{n^+}, C_o^{m^+}, C_o^{n^+})) \quad (17)$$

$$V^+ = V + \frac{1}{2\alpha} i_{ext} \quad (18)$$

$$V^- = V - \frac{1}{2\alpha} i_{ext} \quad (19)$$

3 Multi-cellular Model of Human Atrial Tissue

A bi-domain model of human atrial tissue is given by a parabolic partial differential equation (PDE) of reaction diffusion type based on the modified Nygren et al. equations, which takes the form as equation (20).

$$\frac{dV}{dt} = \frac{1}{C_m}(-i_{tot} + D\Delta V) \quad (20)$$

D is the diffusion coefficient modelling the extracellular conductance. D is different from the intracellular conductance α. D scales the conduction velocity of a solitary travelling wave solution. Δ is a Laplacian operator. x and y represents spatial coordinates in two dimensions (in *mm*). In the model, the diffusion coefficient D was set to 0.3125cm^2 s^{-1} that gave a plane wave velocity of 32 cm s^{-1}. To solve the equation, we used the explicit Euler method with a 3-node approximation of Laplacian operator for 1D, and 5-node approximation of Laplacian operator for 2D. In numerical simulations, the time step Δt is set to 0.001 ms, and space step (Δx = Δy) is 0.32*mm*. The space step of 0.32 mm is about 1/3 of the space constant of the model and corresponds to the long axis of about 4 cells (the length of a cardiac cell is about 0.08mm). It is sufficiently small for stability of the numerical solution. External electrical stimulation is modelled by a square pulse i_{ext} with a variable magnitude and constant duration of 2 ms. The intracellular conductance α was set to 10 S.

4 Results

4.1 Measuring Excitability of Cardiac Tissue

The excitability of human atria was computed from a 1D bi-domain model of human atrial strand (with a zero-flux boundary condition). The strand is 96 mm in length and discretised by 300 nodes. To measure the excitability of the tissue, the standard S1-S2 stimulus protocol was used. Firstly, a sequence of 5 supra-threshold S1 stimuli was applied to the node-1 of the strand. Each of the stimulus pulse had an amplitude 135 nA and duration 6 ms. The time interval between two successive stimuli is 450 ms, which is large enough for the tissue to recover from its previous excitation before a new stimulation was delivered. Each of the 5 stimuli evoked an excitation wave propagating from node-1 to node-300. After the 5th stimulus, a test S2 stimulus was delivered to node-100, a time delay δt after the wavefront of the 5th excitation arrived this point (the time defined as dV/dt of the node reaching its maximal value). The test stimulus has a fixed duration 6 ms, but a changeable magnitude of i_{ext}. The excitability of cardiac tissue was determined as the minimal strength of the test stimulus that re-excites the node-100 at a time interval δt after its depolarisation and produces an action potential with an overshoot over -20 mV. Depending on the time interval δt, the S2-evoked excitation can either propagates bi-directionally if δt is sufficiently large such that atrial tissue has recovered its excitability from previous excitation (Figure 1A), or propagates uni-directionally if δt is small such that the excitability of atrial tissue is recovered only in the retrograde direction of previous excitation (Figure 1B), or fails to propagate at all if δt is too small such that the atrial tissue has not yet recovered its excitability from previous excitation (Figure 1C).

The measured excitability of the atrial strand is δt -dependent as shown in Figure 1D. In the figure, the measured minimal stimulus strength is plotted against the time interval δt. With decrease of δt, it is expected that the excitation threshold increases as more energy is required to re-excite less recovered tissue in normal excitable cardiac tissue. However, in the human atrial model, decrease in δt reduces, rather than increases, the excitation threshold when δt is in the range of 220-450 ms. Around δt =220 ms, the tissue reaches a maximal excitability. The excitability of the tissue during the course of repolarisation or immediately after action potential repolarisation is greater than the excitability of the tissue in resting state. This is a typical feature of excitable tissue with supernormal refractory period (Boyett and Fedida, 1984).

In figure 1D, the δt-dependent dynamical behaviour of the S2 evoked excitation wave was also plotted. With decrease of δt, the S2-evoked excitation wave changed from bi-directional conduction (•) to unidirectional conduction (Δ) and then failed to propagate at all (□). The time window during which the S2-evoked excitation wave propagates uni-directionally (the time window marked by Δ) indexes the vulnerable window of the tissue (Starmer at al., 1993; Zhang & Holden, 1997).

4.2 Defibrillation Threshold

Defibrillation threshold can be estimated from the strength-interval curve shown in the figure 1D. According to the theory of “upper limit of vulnerability” (Pumir and Krinsky; 1996; Krinsky and Pumir, 1998), for a successful defibrillation, it is necessary to apply a stimulus pulse large enough to re-excite the part of tissue in its refractory period and the evoked excitation can propagates uni-directionally only in the retrograde direction, not in the anterograde direction of its previous excitation wave. The re-excitation wavefront annihilates with the previous excitation wavefront when they meet leading the tissue quiescent. According to this theory, the minimal strength of an external stimulation for a successful defibrillation should reaches at least the level indicated by the □ in the figure 1D, which is about 155 nA.

4.3 Defibrillation Simulation in a 1D Model of Atrial Strand

Numerical simulations of actions of an external electrical stimulus on the excitation wave of human atria were carried out with a 1D bi-domain model of a ring fibre with circumference of 96 mm. The re-entrant excitation wave was initiated on the strand by a supra-threshold stimulus. After 5 rotations of excitation waves when a steady stable re-entrant excitation was established, a uniform electrical stimulus was applied to the ring strand. The stimulus was with a changeable amplitude i_{ext}, but a fixed duration of 6 ms.

Actions of an external stimulation pulse on re-entrant excitation are shown in Figure 2. In the figure, the propagating action potentials are represented horizontally in space, vertically in time. The time runs from present at the front to the future at the back with a time interval of 10 ms. Stimulation time is indicated by the arrow.

Panel (A) shows the action of an external stimulation pulse with i_{ext}=90 nA. This stimulus is well below the excitation threshold of the tissue in its resting state (110 nA) and the most excitable tissue in its refractory period. This stimulus failed to produce re-excitation in the tissue, and has no effect on the on-going excitation wave. Panel (B) shows the action of an external stimulus with i_{ext}= 102 nA. This stimulus is smaller than the excitation threshold of the tissue in its resting state (110 nA), but is larger than the excitation threshold of the tissue in its most excitable refractory window (101 nA). When this stimulus was applied, most of the ring was not affected, but there were two parts, indexed by P1 and P2, responded. In the part of P1, a pair of excitation waves were generated which propagated in both the retrograde and the anterograde directions of the previous excitation wave. The wave propagating in the retrograde direction annihilated with the wavefront of the previous excitation. The wave propagating in the anterograde direction collides with the wave evoked in the part P2, which propagated only in the retrograde direction. As a result, existing excitation wave dies and a successful defibrillation is achieved. Panel (C) shows the action of an external stimulus with a strength i_{ext}=160 nA. This stimulus is above the defibrillation threshold predicted by the defibrillation theory. This stimulus re-excited most of the tissue and the evoked excitation wavefront did not expand, but shrunk leading to a successful defibrillation. Due to the supernormal excitability of human virtual atria tissue, the successful defibrillation threshold (102 nA) is smaller than the excitation threshold of the tissue in resting state (110 nA) and remarkably less than that predicted by the defibrillation theory (155 nA) of “upper limit of vulnerability” (Pumir and Krinsky; 1996; Krinsky and Pumir, 1998).

4.4 Defibrillation Simulation in a 2D Model of Human Atrial Tissue

In 2D simulations, an external stimulus with a strength of i_{ext}=102 nA (duration of 6 ms) also successfully defibrillated a re-entrant excitation. This is shown in figure 3. In the figure the snapshots of the re-entrant activity before and after application of the pulse stimulation are displayed. Figure 3(A) shows the snapshot 5 ms just before stimulation, figure 3(B) shows the snapshot 10 ms after the stimulation, and figure 3(C) shows the snapshot 100 ms after the stimulation. In figure 3(B), the subthreshold external stimulation generated non-propagating excitation in the parts of tissue in its refractory period. The non-propagating excitation formed a conduction block to the wavefront of the existing re-entrant excitation, which leads it to termination of re-entrant activity as shown in figure 3(C).

5 Conclusions

In this study we have developed a bi-domain model of human atrial electrical activity based on modification of the Nygren et al. (1998) model to include the actions of an external electrical filed, subjected to which cell membrane potential shows spatial gradient distribution. Using the model, we have quantified the rate-dependent excitability and investigated the effects of a large and rapid external electrical stimulus pulse on human atrial excitation. The significance of this study is that due to the supernormal excitability of human atrial tissue, the successful defibrillation threshold is remarkably lower than the excitation threshold of atrial tissue in a resting state and is also less than that estimated by the "upper limit of vulnerability" theory of defibrillation (Pumir and Krinsky; 1996; Krinsky and Pumir, 1998). This provides new insights to understand the mechanisms underlying the defibrillation by a short and rapid electrical shock.

References

1. Biktashev, V.N., Holden, A.V., Zhang, H.: A model for the action of external current onto excitable tissue. International Journal of Bifurcation and Chaos 7(2), 477–485 (1997)
2. Boyett, M.R., Fedida, D.: Changes in the electrical activity of dog cardiac Purkinje-fibres at high heart-rates. Journal of Physiology, London 350, 361–391 (1984)
3. Krinsky, V., Pumir, A.: Models of defibrillation of cardiac tissue. Chaos. 8(1), 188–203 (1998)
4. Nattel, S.: New ideas about atrial fibrillation 50 years on. Nature 415, 219–226 (2002)
5. Nygren, A., Firek, L., Fiset, C., Clark, J.W., Linblad, D.S., Clark, R.B., Giles, W.R: Mathemaical model of an adult human atrial cell: the role of K+ currents in repolarisation. Circulation Research 82, 63–81 (1998)
6. Pumir, A., Krinsky, V.: How does an electric current defibrillate cardiac muscle. Physica D. 91, 205–219 (1996)
7. Starmer, C.F., Biktashev, V.N., Romashko, D.N., Stepanov, M.R., Makarova, O.N., Krinsky, V.I.: Vulnerability in an excitable medium: analytical and numerical studies of initiating unidirectional propagation. Biophysical Journal. 65, 1775–1787 (1993)
8. Zhang, H., Holden, A.V.: One-dimensional modeling of the vulnerability to re-entry of homogeneous atrial tissue. Journal of Theoretical Biology 184, 119–124 (1997)
9. Zhang, H., Holden, A.V.: Defibrillation threshold computed from normal and supernormal excitable cardiac tissue. Mathematical Biosciences 188, 175–190 (2004)

Appendix

Figure legends

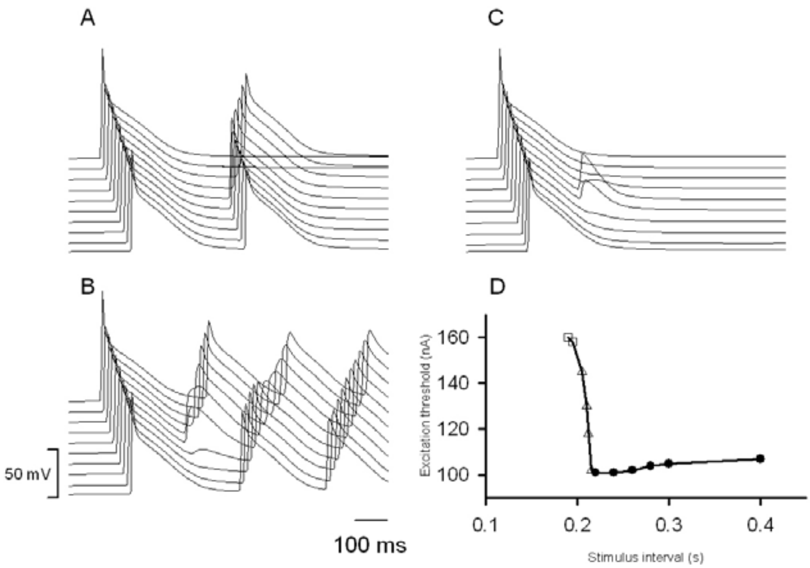


Fig. 1. Time interval (t)-dependent dynamical behaviours of excitation wave evoked by S2 stimulus and the estimated excitation threshold (nA) of human atrial tissue. (A) The re-excitation propagates in both retro- and anterograde directions of a condition wave. (B) The re-excitation propagates only in the retrograde direction of a condition wave. (C) The re-excitation fails to propagate. (D) The estimated excitation threshold against □t, superimposed with the dynamical behaviours of re-excitation (bi-directional conduction: •; unidirectional conduction: Δ; non-propagation: □). Supernormal excitability is illustrated by the biphasic relationship between the excitation threshold and the time interval.

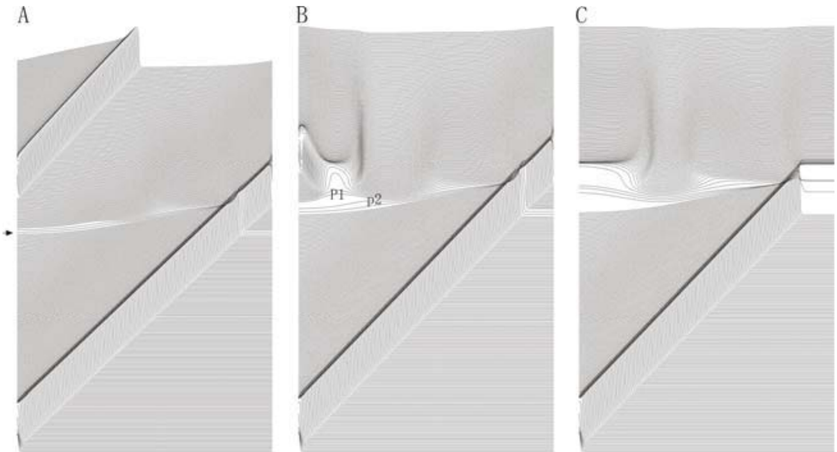


Fig. 2. Actions of an external stimulus on the re-entrant excitation wave in 1D ring fibres. The external stimulus is with different stimulus strengths and a fixed duration (6 ms). Due to the supernormal excitability, the successful defibrillation threshold is much less than that of the normal excitation threshold and that predicted by the defibrillation theory. (A) i_{ext}=90 nA, an unsuccessful defibrillation; (B) i_{ext}=102 nA, a successful defibrillation with an external stimulus strength much less than that predicted by defibrillation theory. (C) i_{ext}=160 nA. A successful defibrillation with an external stimulus strength over the value predicted by defibrillation theory of upper limit of vulnerability.

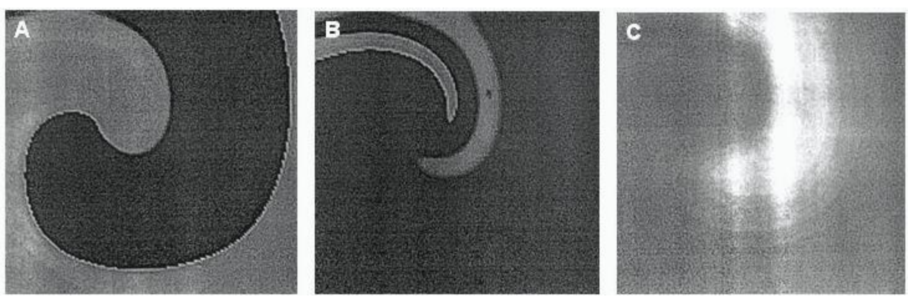


Fig. 3. Defibrillation in a 2D model of human atrium with i_{ext}=102 nA (duration of 6 ms). The part of the tissue with the greatest excitability responds. The evoked excitation propagated in the retrograde direction and collides with the existing reentry wavefront leading to a successful defibrillation. The stimulation strength is 102 nA, which is less than that of a resting tissue and the value predicted by the defibrillation theory of upper limit of vulnerability.

Study and Application of Medical Image Visualization Technology

Jihong Liu[1], Weina Ma[1], Fei Liu[1], Ying Hu[2], Jinzhu Yang[1], and Xinhe Xu[1]

[1] College of Information Science and Engineering, Northeastern University. 110004 Shenyang, China
[2] Research Center of Automation, 116026 Dalian, China
{liujihong,xuxinhe}@ise.neu.edn.cn

Abstract. In medical imaging, many applications require visualization and analysis of three-dimensional (3D) objects. Visualization is the process of exploring, transforming, and view data as images to gain understanding and insight into the data, which requires fast interactive speed and high image quality. In this paper, we describe the key techniques in medical image visualization. In order to improve ray casting rendering speed, a synthetically accelerated algorithm is proposed. Firstly, rendering algorithms are fully studied and compared. Secondly, proximity clouds algorithm has been selected and extended to continuous ray casting. Finally, the accelerated algorithm based on ray coherence has been realized. The experimental results on 3D medical image reconstruction are given, which show the medical image visualization technology has provided a powerful technology base for computer-aided diagnosis, virtual surgery and e-learning in medicine field.

Keywords: Visualization, 3D reconstruction, Proximity Clouds, Volume Rendering.

1 Introduction

Medical imaging technique is becoming a major component of clinical decision, which is making in a trend towards evidence-based medicine that relies heavily on objective findings for diagnostic tasks. Consequently, the need for image visualization and analysis has extended beyond the traditional diagnostic arena in radiology departments and is becoming common practice in all medical subspecialty. Digital and 3D techniques for imaging diagnostic devices have made a dramatic progress, and Computed Tomography (CT) and Magnetic Resonance (MR) imaging are common in this decade. Visualization is the process of exploring, transforming, and view data as images to gain understanding and insight into the data. CT and MR scanners can be used to create a volume by reconstructing a series of slicing images. The flow of visualization is shown in Fig.1. The visualization technology makes it possible to provide a multimedia tool for medical instruction, diagnosis and research, in which anatomic structure and functions of human organs are explored interactively. To effectively visualize volumes, it is important to be able to image them from different

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 668–677, 2007.

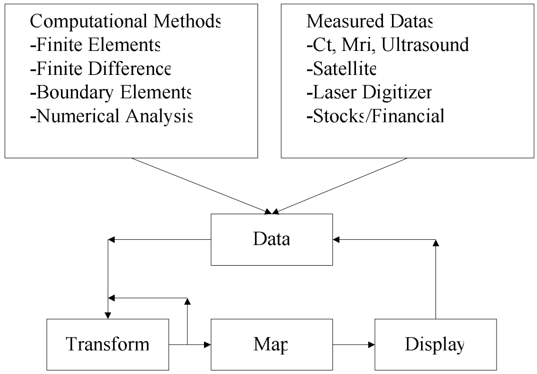


Fig. 1. The flowchart of visualization

viewpoints, and to shade them in a manner which brings out volume display and subtle variations in density or opacity.

Besides the key technologies of realizing 3D reconstruction, image collection and image processing are also the essential prior work. Though there has been astonishing progress in the speed of image collection techniques recently, the accurate 3D reconstruction still relies on availability of high quality images that represent the feature and structure of an interested object better. Before 3D reconstruction, some image procession methods are used to improve the quality of images. Most 3D reconstruction takes the contour of object as the essential data, which may be obtained directly from the image by extracting interested boundary. Therefore, the image processing technologies, which form the basis of medical image analysis for various purposes and of 3D visualization, include image enhancement, characteristic extraction and segmentation. The current research focuses on the algorithms completing segmentation and 3D reconstruction.

Many approaches to reducing rendering cost without affecting image quality by exploiting coherence in the dataset have been proposed. These approaches rely on spatial data structures that encode the presence or absence of high-opacity voxels so that computation can be omitted. The data structures are built during a preprocessing step on a classified volume on which an opacity transfer function is applied. The opacity transfer function and shading functions classify the volume and throw off the uninterested parts.

2 Key Techniques Before 3D Reconstruction

2.1 Image Registration

Registration is an important technology to computer-aided diagnosis, virtual surgery and e-learning in medicine field and has been one of the hot research fields [1],[2],[3]. By choosing a transformation, it produces the best matching between the components of the objects in the images to be matched. Depending on the combination of images to be matched, various transformations can be used, such as rigid transformation, nonrigid transformation, and projective transformation.

Recent studies focus on that all of the input image information is utilized directly in evaluating the similarity between the images, without extraction of features. There is an increasing need from the viewpoint of diagnostic support to display multiple examination image data with registration (but no additional information extraction is required). The processing is simple in these methods, since no preprocessing such as feature extraction from the image is required. It is also an advantage that error is not accumulated in the course of processing. In order to evaluate the similarity between images, various similarity indices have been considered, such as correlation of intensity and correlation in Fourier space. Among these, the maximization of the mutual information (MI), which was introduced by Maes and colleagues [4], is currently considered the most interesting approach.

2.2 Distance Transform

The intensity of medical images corresponds mostly to tissue information, and does not correspond directly to morphological information. For this reason, when interpolation based on intensity is applied to data with bigger slice spacing, the desired shape of the region may not be obtained. If the region to be interpolated can be binarized within the slice, the 2D signed distance transform (the object region is defined as positive and the background region as negative) is applied to each slice. The signed distance is used in the interpolation between slices. By extracting the voxels with negative distance from the interpolated signed distance image, a region with high resolution interpolation is obtained. The distance transform is an operation in which the shortest distance is determined from the voxels in the object region (with value 1) of the binary image to the background region (with value 0). As the distance, the Chamfer distance [5] and others, which can be determined with high speed using only local operations or integer operations, were used in the past, but the Euclidean distance is often used at present due to improvements in computer performance and the development of fast computational techniques [6].

2.3 Image Segmentation

One of the basic problems in medical image analysis is to precisely segment interested structures from a huge dataset, accurately represent them, efficiently visualize them and perform measurements appropriate for diagnosis, surgery and therapy guidance, and other applications [7], [8]. Most current segmentation algorithms applied to medical imaging problems only detect the rough boundaries of the structures in two dimensions (2D), and as such do not satisfy the requirement of high accuracy required of many medical applications. The continuing evolution of computer-aided diagnosis, image-guided and robotically-assisted surgery, improves the development of efficient, accurate three dimensional segmentation procedures.

Segmentation techniques can be divided into classes in many ways, according to different classification schemes, model-based and region-based techniques represent the two main groups. Model-based procedures include “snake” algorithms (deformable model, active contours etc.) and level set methods (as well as fast marching methods). These techniques are based on deforming an initial contour or surface towards the boundary of the object to be detected. The deformation is obtained by minimizing a customized energy function such that its local minima are reached at the boundary of the desired structure. Region-based algorithms include region growing [9], morphological reconstruction [10] and watershed [11]. Since these procedures are generally based on neighborhood operations, and examine each pixel during the evolution of the edge, the outcomes are usually highly accurate.

Virtual surgery is one of the hot research fields in recent years. For instance, the common orthopedic surgery training can be a typical example. In the virtual environment, the surgeon will deal with the virtual models of bones, surgical tools, and implants. Visualization is one of its key technologies. In order to meet the high accurate requirements in virtual surgery, there are several segmentation algorithms described in the literature to facilitate medical image visualization and manipulation [12], [13]. An algorithm of serial images segmentation was proposed based on the combination of Live Wire algorithm and contour interpolation algorithm in our paper, refer to [14]. Using the ideas of active contour model, this algorithm combined with the ternary algorithm, modified traditional contour interpolation and Live Wire algorithm. Combining with the ideas of active contour model□the images are self-shrinking on the basis of actual images local character in contour of rebuild. The experiment shows the algorithm obtains the boundary of ROI from a series of medical images quickly and reliably. The result of extracted boundary is shown in Fig. 2. The more detail algorithm has been proposed in [14].

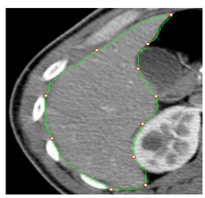


Fig. 2. The result of extracting the boundary

A new 3D hybrid segmentation algorithm has been proposed in [15], which works in a multistage manner to perform segmentations rapidly and precisely. In the first stage, it employed a morphological recursive erosion to reduce the connectivity between the object and its neighboring tissues. Then a fast marching method is employed in the second stage to greatly accelerate the initial propagation from the user-defined seed structure to the vicinity of the boundary of object. In the third stage, it employed a morphological reconstruction algorithm to refine the propagation to achieve a highly accurate boundary. At the end of the segmentation procedure, morphological recursive dilation is implemented, using the same number of iterations as recorded in stage one, to recover regions lost during the initial erosion, but at the same time avoiding re-connection to the surrounding structures. The segmented result may be presented to the user by volume rendering (ray casting) or surface rendering (matching cube) methods. It developed a segmentation environment, “TkSegment”, based on the visualization toolkit (VTK) and the Python language, into which the multistage hybrid segmentation algorithm was integrated.

3 The Main Rendering Algorithms

3.1 Rendering Algorithms

According to the different models, the 3D reconstruction methods can be divided into two different categories. One is the surface rendering, in which only the surface of the object display or the boundary between tissues is displayed, the other one is the volume rendering, in which radiation and absorption of the tissue is simulated.

The surface expression is the most essential method to present a 3D object. It may provide comprehensive information about the shape of 3D object. Specifically there are two different approaches: boundary contour line expression and superficial curved surface expression. Boundary contour line expression is comparatively simple and data involved in computation is little, but its result is not realistic for the loss of information between contours. Researchers have proposed arithmetic of using triangles or quadrilateral meshes as a rendering primitive. Shape and topological differences between object contours in adjacent sections cause severe difficulties in the reconstruction process. A way to approach this problem is using the skeleton to create intermediate sections that represent the place where the ramifications occur. Several authors have proposed the use of some type of skeleton to face the problem illustrate, and have gotten the excellent performance of the methods in especially difficult cases [16].

When we render an object using surface rendering techniques, we mathematically model the object with a surface description such as points, lines, trianglesm polygoins, or 2D and 3D splines. The interior of the object is not described, or only implicitly represented from the surface representation. Although techniques do exist that allow us to make the surface transparent or translucent, there are still many phenomena that cannot be simulated using surface rendering techniques alone.

The volume rendering is a technique for visualizing sampled functions of three spatial dimensions by computing 2D projections of a colored semitransparent volume, which can show the inhomogeneity inside objects. According to the different methods of data processing, volume rendering algorithm can be divided into space domain method and transform domain method. The former is that processing and displaying the original data directly; the latter is that transforming the data into transform domain firstly and then displaying the processing results.

The methods of space domain include: ray casting, splatting, shear-warp and 3D texture mapping based on hardware.

The main rendering algorithms are compared in Table 1.

3.2 Ray Casting

In ray casting method, the image is composed by distributing opacity and light intensity for every volume. The main problem of ray casting is unable to store the volume data by physical sequence. Any change relating to viewpoint can induce complete resample of ray casting. Another problem of ray casting is that the voxel values of the 3D volume must be read along the view line, which lowers memory access efficiency.

Table 1. The comparison of algorithm charectures

<table>
<tr><th colspan="3">Volume rendering algorithm</th><th>Image quality</th><th>Rending velocity</th><th colspan="2">Algorithm characteristic</th></tr>
<tr><td rowspan="4">Spatial domain</td><td colspan="2">Ray casting</td><td>best</td><td>slow</td><td rowspan="4">may make use of an opacity to get the overall hiberarchy without segmentat ion</td><td>Occupied memory is big</td></tr>
<tr><td colspan="2">Footprint</td><td>good</td><td>medium</td><td>Occupied memory is small, may gradually display</td></tr>
<tr><td colspan="2">Shear-warp</td><td>medium</td><td>fastest</td><td>Occupied memory is small</td></tr>
<tr><td colspan="2">3D texture mapping</td><td>bad</td><td>fast</td><td>The hardware accelerates, image quality are dependent on frame of memory place resolution ratio</td></tr>
<tr><td rowspan="3">Transform domain</td><td colspan="2">Volume rending of frequency domain</td><td>better</td><td>fast</td><td colspan="2">the effect of X－ray picture, makes use of fft, the algorithm is concise</td></tr>
<tr><td rowspan="2">Wavelet domain</td><td>Ray casting</td><td>good</td><td>slow</td><td colspan="2">may make use of an opacity to get the overall hiberarchy, Occupied memory is big</td></tr>
<tr><td>Footprint</td><td>better</td><td>faster</td><td colspan="2">the effect of X－ray picture, may gradually display , part detail added</td></tr>
</table>

3.3 Splatting

The splatting (or footprint) method [17], proposed by Westover, is a form of volume rendering in which the projection patterns of the voxels (footprints) are superposed on the image plane from farther (or nearer) voxels, considering the opacity (called alpha-blending). In contrast to ray casting, which is image-based, the splatting method is considered to be object-based. Voxels which are completely transparent or pixels onto which no voxel is projected can be ignored in processing, which improves the speed compared to the ray casting method. In particular, in the case of parallel projection, the footprint always has the same shape for any voxel. By preparing a table for this shape, the computation speed can be further improved.

3.4 Shear-Warp

Shear-warp volume rendering [18] has the advantages of a moderate image quality and a fast rendering speed. It is faster than others methods except for those which uses graphic hardware. Its realization includes two steps: decomposing the projection transformation of 3D discrete data field into shear transform of 3D data field and warp of 2D image, then changing the resample process of 3D space into that of two-dimension. However, it has not been preferable in recent medical applications because the image quality is worse than in ray casting. Lots of new methods that improve image quality of shear-warp rendering without noticeable performance loss has been proposed [19]. It transforms the volume into sheared object space. In the sheared object space, the viewing ray is adjusted to be vertical to the data slices, and then the projection from volume to the image plane is simplified, provides an effective method for the solution of acceleration problem on medical visualization.

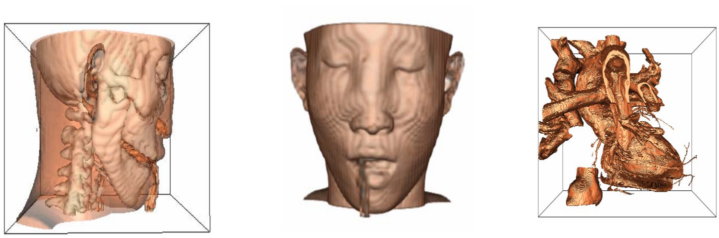
3.5 3D Texture Mapping

The 3D texture mapping based on hardware realizes interpolation of resampling and image combination of opacity in texture space, which is completed by hardware. So the speed of calculation has been greatly improved. This method has been widely applied to the shading volume rendering at present.

3.6 The Volume Rendering of Transform Domain

Transform domain rendering methods includes: Fourier volume rendering (FVR) and Fourier-wavelet volume rendering (FWVR). FVR, which makes use of frequency domain techniques, is based upon the Fourier slice theorem [20],[21],[22]. FWVR includes two methods: Fourier-wavelet volume rendering based on the X-ray transform, which is the realization of ray rendering in wavelet space with some advantages of ray rendering and wavelet splatting, which is a modification of the standard splatting algorithm through the use of wavelets as reconstruction filters. A disadvantage of FWVR is that it requires resampling of a slice in Fourier space at full resolution in order to perform a 2-D wavelet decomposition.

The traditional methods of 3D reconstruction for medical images are completed in this paper. Some experimental results gained by the method of Surface Shaded Display (SSD) are shown in Fig.3.



(a) (b) (c)

Fig. 3. Surface Shaded Display (SSD) (a) Marching Cubes (b) Ray Casting (c) Serial 2D Contours

4 The Accelerated Algorithm Based on Ray Coherence

The accelerated algorithm based on ray coherence falls into: (1) pixel-space coherency (2) object-space coherence (3) inter-ray coherency (4) space-leaping (5) sequence coherence. The purpose of 3D medical visualization is to assist medical diagnosis, so it needs a high requirement in reconstruction quality. The technology of space-leaping is an exact accelerated algorithm, which improves the speed of rendering dramatically, while having no effect on the image quality.

The accelerated ray casting algorithm based on space-leaping has been proposed in our paper [23], which leaps the peripheral disturbance of the region of interest(ROI) using bounding box technology,then completes the volume rendering by improving the proximity clouds algorithm. Fig.4 is the results of the improved accelerated algorithm.

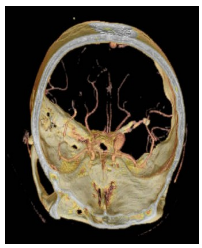


Fig. 3. The results of accelerated algorithm

Compared with standard ray casting algorithm, our algorithm is better. This accelerated algorithm not only improves the rendering speed for nearly ten times but also brings no compressing on image quality. At 512×512 image resolution, a nearly real time interactive speed has been gotten. It gives a more capability method for the application of medical image 3D reconstruction.

5 Conclusion

In this paper, some key technologies of medical image visualization are discussed. By combining these technologies into computer-aided diagnosis system, the remarkable progress would be achieved in the automated diagnosis technology. Some experimental results on 3D medical image reconstruction are given, which show the medical image visualization technology has provided a powerful technology base for computer-aided diagnosis, virtual surgery and e-learning in medicine field.

Reference

1. Maintz, J.B.A., Viergever, M.A.: A survey of medical image registration. Med. Image Anal. 2, 1–36 (1998)
2. Ashburner, J., Friston, K.J.: Nonlinear spatial normalization using basis function. Human Brain Mapping 7, 254–266 (1999)
3. Audette, M.A., Ferrie, F.P., Peters, T.M.: An alogorithmic overview of surface registration techniques for medical imaging. Med. Image Anal. 4, 201–217 (2000)

4. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P.: Multimodality image registration by maxi-mization of mutual information. IEEE Trans Med. Imaging 16, 187–198 (1997)
5. Borgefors, G.: Distance transformations in digital images. Comput Vis. Graph Image Process 34, 344–371 (1986)
6. Saito, T., Toriwaki, J.: New algorithms for Euclidian distance transformation of an N-dimensional digital picture with applications. Pattern Recognition 27, 1551–1565 (1994)
7. Bresson, X., Vandergheynst, P., Thiran, J.-P.: A Variational Model for Object Segmentation Using Boundary Information and Shape Prior Driven by the Mumford-Shah Functional. International Journal of Computer Vision 68(2), 145–162 (2006)
8. Heinonent, T., Dastidar, P., Kauppinen, P., Malmivuo, J., Eskola, H.: Semi-automatic tool for segmentation and volumetric analysis of medical images. Medical and Biological Engineering and Computing 36(3), 291–296 (1998)
9. Logeswaran, R., Eswaran, C.: Discontinuous Region Growing Scheme for Preliminary Detection of Tumor in MRCP Images. Journal of Medical Systems 30(4), 317–324 (2006)
10. Capowiez, Y., Pierret, A., Moran, C.J.: Characterisation of the three-dimensional structure of earthworm burrow systems using image analysis and mathematical morphology. Biology and Fertility of Soils 38(5), 301–310 (2003)
11. Heede, B.H.: Deteriorated watersheds can be restored: A case study. Environmental Management 3(3), 271–281 (1979)
12. Wismüller, A., Lange, O., Dersch, D.R., Leinsinger, G.L., Hahn, K., Pütz, B., Auer, D.: Cluster Analysis of Biomedical Image Time-Series. International Journal of Computer Vision 46(2), 103–128 (2002)
13. Salah, Z., Kastner, M., Dammann, F., Schwaderer, E., Maassen, M.M., Bartz, D., Straßer, W.: Preoperative planning of a complete mastoidectomy: semiautomatic segmentation and evaluation. International Journal of Computer Assisted Radiology and Surgery 1(4), 213–222 (2006)
14. Gu, L., Peters, T.: 3D segmentation of medical images using a fast multistage hybrid algorithm. Int. J CARS 1, 23–31 (2006)
15. Jin-zhu, Y., Shu-ying, Z., Ying, H., Ji-hong, L., Xin-he, X.: Algorithm of segmentation of medical series image 3D reconstructs. Journal of system simulation 17(12), 2896–2900 (2005)
16. Amargós, J.P., Mancho, R.A.: Reconstruction of Surfaces from Cross Sections Using Skeleton Information. In: Sanfeliu, A., Ruiz-Shulcloper, J. (eds.) CIARP 2003. LNCS, vol. 2905, pp. 180–187. Springer, Heidelberg (2003)
17. Westover, L.: Footprint evaluation for volume rendering. In: Proceedings of SIGGRAPH '90. vol. 1009, pp. 367–376, ACM Press, New York (1990)
18. Lacroute, P.: Real-time volume rendering on shared memory multiprocessors using the shear-warp factorization. Parallel Rendering Symposium, pp. 15–21 (1995)
19. Kye, H., Oh, K.: High-Quality Shear-Warp Volume Rendering Using Efficient Supersampling and Pre-integration Technique. In: Pan, Z., Cheok, A., Haller, M., Lau, R.W.H., Saito, H., Liang, R. (eds.) ICAT 2006. LNCS, vol. 4282, pp. 624–632. Springer, Heidelberg (2006)
20. Malzbender, T.: Fourier volume rendering. ACM Transactions on Graphics 12(3), 233–250 (1993)

21. Napel, S., Dunne, S., Rutt, B.K.: Fast Fourier projection for MR angiography. Magnetic Resonance in Medicine 19, 393–405 (1991)
22. Yagel, R.: Towards real time volume rendering [A]. In: Proceedings of GRA PH ICON ’ 96 [C], Saint Petersburg, Russia, pp. 230–241 (1996)
23. Ying, H., Xinhe, X.: Accelerate Ray Casting Algorithm Base on Ray Coherence. Journal of Image and Graphics 9(2) (2004)

A Method for Gene Identification by Dynamic Feature Choosing

Jia-Wei Luo, Li Yang, and Xi-Zhen Zhang

Computer Science and Communication College, Hunan University, Hunan 410083, China
jt_ljw@hnu.cn, yoyoer@tom.com, chunagp@hotmail.com

Abstract. With growth of sequenced genome, a number of algorithms for gene identification were created. These algorithms use fixed gene features which are chosen based on observation or experience. These features may not be major features of a genome. In this paper, we illustrate several candidate features and propose a dynamic feature choosing algorithm to determine the major features. We describe nucleotide sequence by feature vector and use Discriminant analysis to them to make decision on coding/non-coding. To test the algorithm, we apply the algorithm to the S.cerevisiae genome and achieve accuracy of above 98%.

Keywords: Gene identification, Discriminant analysis, three cross-validation tests.

1 Introduction

Due to the underway of genome sequencing and the development of Human Genome Project, more and more species have been sequenced. However, genome annotation is not keeping pace with overflow of raw sequence data. There is urgent need for accurate and fast tools to analyze these sequences. A DNA sequence consists of many regions. DNA coding region, i.e., gene, carries the genetic information of species. It's the only region that can be translated into RNA, which participates in the protein synthesis. Thus, recognizing gene region is one of the most important problems in genome research.

There are many algorithms for gene identification such as codon bias index [1], YZ score [2], period-3 features of a nucleotide sequence in the coding regions [3], Spectrum rotation [4]. These algorithms use fixed gene features which are chosen based on observation or experience. These features may not be major features of a genome. If simply combine all the known features, the dimension of the feature vector will become large, moreover, some features may interfere with others, which will lower the efficiency in the calculation. In this paper, we illustrate several candidate features and propose a dynamic feature choosing algorithm to determine the major features. Describe nucleotide sequence by feature vector and use Discriminant analysis to them to make decision on coding/non-coding. We test the method on S.cerevisiae genome and find that the accuracy reaches above 98%. The comparison between our methods and YZ score [2] and Wang's [3] methods is also presented. It

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 678–683, 2007.

suggests ours has better performance. The advantage of the new methods is that it can choose major features dynamically so that the minor features are deserted, further more, lower the dimension of feature vector.

2 Database and Methods

2.1 Database

S.cerevisiae is the firstly sequenced eukaryotic single-cell species. We use S.cerevisiae genome to build database in this paper. There are 16 chromosomes in the genome and are of 12.16Mbp length. The database can be obtained from http://mips.gsf.de.

2.2 Composition Feature

Nucleotide compositon is a fundamental feature of gene. We firstly consider 1-mer , 2-mer, 3-mer and 6-mer composition of sequence. For k-length words, there are 4^k probabilities of combination.

Describe the 1-mer composition bias by define a 4 element vector $(f_{A,} f_{T,} f_{G,} f_C)$, where $f_{A,} f_{T,} f_{G,} f_C$ denote the frequencies of A, T, G and C, respectively. Similarly, define the 2-mer, 3-mer usage to be 16 elements, 64 elements vector.

For 6-mer frequency, there are 4096 (4^6) possible words. So we use the Shannon entropy H to lower the vector dimension. See equation 1.

$$H = -\sum_{w=1}^{4096} p(w) \log_2 p(w) \tag{1}$$

where $p(w)$ denote frequencies of every possible hexamer w.

G+C content evaluate the significance of the coding scores. Define the ratio of G+C content to A+T content to describe this feature.

So for 1-mer, 2-mer, 3-mer, 6-mer and G+C content feature, we get 4 elements, 16 elements, 64 elements, 1 element and 1 element vector, respectively.

2.3 Codon Prototype Feature

A number of investigators have noticed prototypical properties of codons, e.g. that they are often of the form RNY or WWS, or that a certain base is more common in one position than in another [5]. Name the subsequence with nucleotides in position 1,4,7,… as subsequence-1, the nucleotides in position 2,5,8,… as subsequence-2, the nucleotides in position 3,6,9,… as subsequence-3. Denote the frequencies of purine (R), amino (M), strong hydrogen (S) in subsequence-1 by $r(1), m(1), s(1)$, in subsequence-2 by $r(2), m(2), s(2)$, in subsequenc-3 by $r(3), m(3), s(3)$.

Then for Codon prototype feature, we get a 9 elements vector.

2.4 Termination Codon Feature

The termination codons are very strong signals in DNA sequences [4]. In coding region, contents of TAA, TAG or TGA are usually very low, whereas in non-coding region, contents of termination codon are higher.

We describe the termination codon feature by the frequencies of TAA, TAG or TGA in a sequence. Then, for termination codon feature, we get 1 element vector.

2.5 Algorithm for Feature Choosing

We choose 7 candidate features for sequences. They are 1-mer, 2-mer, 3-mer, 6-mer, G+C content, Codon prototype, and Termination codon frequency. Each feature's contribution to the coding decision is different. Some features may contribute the similar information of gene. Some one may interfere with the result. Thus, it's essential to choose the candidate features. Procedure of feature choosing can improve the calculation efficiency, get rid of the disturbing information and improve the accuracy.

We need two sets to complete the algorithm. One is set of positive samples consist of coding sequence; another is a set of negative samples consist of non-coding sequence. Each sequence is represented by a vector **u** with components of above obtained features. The sense of feature can be evaluated by squared Mahalonobis distance D^2 of two sets. If D^2 value is high, we can come to conclusion that the feature is significant. The calculation of D^2 is described here. Name coding set as group 1 and non-coding set as group 2. Calculate the Arithmetic mean $\bar{U}_k^g$ for each group,

$$\bar{U}_k^g = (\bar{u}_1^g, \bar{u}_2^g, ..., \bar{u}_m^g)^T, g = 1,2 \tag{2}$$

Where m denote the size of the vector that contains k features, and

$$\bar{u}_j^g = \frac{1}{n_g}\sum_{l=1}^{n_g} u_{jl}^g, g = 1,2; j = 1,2,...,m \tag{3}$$

Denoting by $S^k = \left(s_{ij}^k\right)$ the sum of the covariance matrices of two groups, we have

$$s_{ij}^k = \sum_{g=1}^{2}\sum_{l=1}^{n_g}(u_{il}^g - \bar{u}_i^g)(u_{jl}^g - \bar{u}_j^g), i, j = 1,2,...,m \tag{4}$$

D_k^2 of k features will be calculated by following equation

$$D_k^2 = (\bar{U}_k^1 - \bar{U}_k^2)^T \left(S^k\right)^{-1} (\bar{U}_k^1 - \bar{U}_k^2) \tag{5}$$

We use our method to choose ultimate features. The detailed procedure is described below:

(1) Calculate D_1^2 of each of the 7 candidate features. Sort D_1^2 in descending order to establish candidate queue Q. Select the feature which has the highest D_1^2 value and add it to the feature set A.
(2) Add to feature set: assume that we have obtained k features in set A. For the head candidate feature α in queue Q, calculate D_{k+1}^2 of $A+\alpha$. If $D_{k+1}^2 - D_k^2 > d_0$ (d_0 is a threshold determined during the calculation), move α to A and go to (3). Or else, choose the next feature in queue Q and go to (2). If the queue Q is empty, the algorithm ends.
(3) Delete feature: suppose that in step (2) we have add a new feature to A, for any other feature x in A, calculate D_k^2 of $A-x$. If $D_{k+1}^2 - D_k^2 < d_0$, delete x from A. Go to (2).

After above choosing procedure, we get 3 features in set A: Codon prototype, 6-mer and Termination codon frequency. They are described by 9 elements, 1 element and 1 element vector. Incorporate them as a 11 elements vector **u**: $\mathrm{u} = (u_1, u_2, ..., u_{11})^T$.

2.6 Discriminant Analysis

Now, each ORF is represented by a vector **u** in the 11-D space. Apply the Fisher discriminant algorithm to the two training groups defined above. Then some coefficient vector **c** and some threshold c_0 is obtained. The decision of coding/non-coding is performed by the criterion of $\mathrm{c} \bullet \mathrm{u} > c_0 \,/\, \mathrm{c} \bullet \mathrm{u} < c_0$.

Method for deriving such a coefficient vector **c** is described below. One begins by defining the covariance matrices S, as follows: let u_{jk}^g be the j th component of the vector of the k th sample in the g group, where g=1,2; j=1,2,…,11; and k=1,2,…, $\mathrm{n_g}$ ($\mathrm{n_1}$=$\mathrm{n_2}$). Let $\bar{u}_j^g$ be the mean of u_{jk}^g over k, and $\bar{U}_g$ the vector of $\bar{u}_j^g$, i.e., $\bar{U}_g = (\bar{u}_1^g, \bar{u}_2^g, ..., \bar{u}_{11}^g)^T$, where $g = 1, 2$. Then the element in the i th row and j th column of S is $s_{ij} = \sum_{g=1}^{2}\sum_{k=1}^{n_g}(u_{ik}^g - \bar{u}_i^g)(u_{jk}^g - \bar{u}_j^g)$, where $i, j = 1, 2, ..., 11$.

The vector **c** is simply determined by the following equation

$$\mathrm{c} = S^{-1}(\bar{U}_1 - \bar{U}_2) \tag{6}$$

Where S^{-1} is the inverse of the matrix S. See the detailed explanation on these equations in Mardia et al.[6].

Based on the data in the training set, the threshold c_0 is chosen so that the fraction of errors on the coding is equal to the fraction of errors on the non-coding. Finally, the Fisher coefficients $c_1, c_2, ..., c_{11}$ and threshold c_0 is applied to the test set. An ORF is thought to be coding if and only if $\mathrm{c} \bullet \mathrm{u} > c_0$, where $\mathrm{c} = (c_1, c_2, ..., c_{11})$ and $\mathrm{u} = (u_1, u_2, ..., u_{11})^T$.

3 Results and Discussion

In the version of MIPS database, the total number of S.cerevisiae ORFs is 6449, and was classified into 6 categories; include known proteins, no similarity, questionable ORFs, similarity or weak similarity to known proteins, similarity to unknown proteins and strong similarity to known proteins. Each contain 3410, 516, 471, 820, 1003 and 229 sequences. Choose the first 3410 ORFs as positive sets.

Construct negative sets in following method: (1) note down the starting position of every ORF section; (2) pick up the intergenic sequences with length longer than 300bp from the 16 chromosomes; (3) In the DNA sequences with length longer than 300bp, search for codon ‘ATG’ start from the first base; then search for termination codon {TAA, TGA, TAG} till find the first termination codon. Randomly select 3410 sequences to form the negative sets. To test the new algorithm, we use Three-fold cross-validation test. Divide the positive and negative sets into 3 partitions, respectively. Each of the three partitions was tested after training on the other two.

We denote by TP the number of correctly predicted coding nucleotides; TN, the number of correctly predicted non-coding nucleotides; FP, the number of non-coding nucleotides predicted to be coding; FN, the number of missed coding nucleotides. Sensitivity Sn=TP/(TP+FN) is the fraction of correctly predicted coding sequences among all coding ORFs. Specificity Sp=TN/(TN+FP) is the fraction of intergenic sequences that have been correctly predicted as non-coding sequences. The accuracy AC is defined as the average of Sn and Sp.

Table 1 lists the results of the identification. It shows that the 3 test sets are with accuracies of higher than 98%. The average accuracy is as high as 98.7%, which shows superior performance of our approach.

As comparison, results of the YZ score [2] and Wang’s [3] methods are reproduced. YZ score make use of 1-mer and Codon prototype features of gene. Wang integrates several features of gene. To make all the methods comparable, we rewrite the programs for both methods and apply the same database mentioned above to them.

Table 2 shows the results of the comparison. It can be observed that our method has higher accuracy than the other two. Our new method need not consider the major features of a genome, for the features are dynamically chosen. It can be applied to many species, which can make the identification procedure more automatic and more reliable.

Table 1. Identification results for three different test sets

Test set	Sn(%)	Sp(%)	AC(%)
1	98.6	98.7	98.7
2	98.7	98.7	98.7
3	98.9	98.7	98.8
average	98.7	98.7	98.7

Table 2. Identification results based on different methods

	Sn(%)	Sp(%)	AC(%)
YZ score	95.2	96.3	95.7
Wang	97.2	97.8	97.5
Our work	98.7	98.7	98.7

4 Conclusion

Although great progress has been made in the development of gene identification in recent years, there is still in need of method to make more correct annotations from uncharacterized DNA sequences. We demonstrated here the method for gene identification by dynamic feature choosing. It chooses the major features in a scientific and rational way. So the accuracy of gene identification is improved.

Acknowledgments. We thank the help of Chun-Ting Zhang and Bo Liao. This work is supported by the Item Foundation of Hunan Provincial Department of Finance([2005]90) and Hunan Provincial National Science Foundation of China(06JJ4076).

References

1. Bennetzen, J.L., Benjamin, D.H.: J.Biol.Chem 257, 3026–3031 (1982)
2. Zhang, C., Wang, J: Recognition of protein coding genes in the yeast genome at better than 95% accuracy based on the Z curve. Nucleic Acid Res.2000 28(14), 2804–2814 (2000)
3. Wang, Y.H, Zhang, C.-T., Dong, P.X.: Recognizing shorter coding regions of human genes based on the statistics of stop codons[J]. Biopolymers 63, 207–216 (2002)
4. Kotlar, D., Lavner, Y.: Gene Prediction by Spectral Rotation Measure: A New Method for Identigying Protein-Coding Regions. Genome Res. 13, 1930–1937 (2003)
5. Fickett, J.W., Tung, C.S.: Assessment of protein coding measures. Nucleic Acids Res. 20, 6441–6450 (1992)
6. Mardia, K.V., Kent, J.T., Bibby, J.M.: Multivariate Analysis. Academic Press, London, UK (1979)

Latest Development of an Interventional Radiology Training Simulation System: *NeuroCath*

Xin Ma[1,2]

[1] Biomedical Imaging Lab, Singapore Bioimaging Consortium, Agency for Science, Technology and Research
30 Biopolis Stree, #07-01 Matrix, Singapore 138671
[2] Centre for Human-Computer Interaction
Shenzhen Institute of Advanced Integration Technology,
Chinese Academy of Sciences/The Chinese University of Hong Kong
F3 Building A, Nanshan Mega-Surplus Medical Equipments Industry Park,
Gongye 3 Road, Skekou, Shenzhen, China, 518067
maxin@sbic.a-star.edu.sg

Abstract. We describe the latest development of a computer-based virtual reality environment for training interventional neuroradiology procedures. The system, *NeuroCath (*Neuroradiology Catheterization Simulator), includes extraction and construction of a vascular model from different imaging modalities that represents the anatomy of patient in a computationally efficient manner, and a finite element method (FEM) based physical model that simulates the interaction between the devices and neuro-vasculature. A realistic visual interface with multiple, synchronized windows and plenty of video control functions are developed. The latest version is also equipped with haptic feedback module that gives the sense of touch in real-time, and customizable vascular model so that trainer can understand the importance of vascular variations and practice. According to the validation in several clinical centers, 70%-75% of training features have been realized which makes the system well suitable for training of interventional neuroradiologists.

Keywords: interventional neuroradiology, simulation, augmented reality, modeling.

1 Introduction

Generally, minimally invasive surgical (MIS) procedures, such as interventional radiological procedures, are utilized by physicians to accomplish tasks that would otherwise require a patient to undergo open surgery [1-2]. For example, an angioplasty-balloon procedure may be utilized to open and eliminate blockages in a blood vessel without subjecting a patient to open heart surgery. Briefly, in the actual angioplasty-balloon procedure, a variety of changeable guidewires, catheters and sheaths are inserted into a patient and manipulated through the patient's arterial network until reaching the point where a blockage occurs. Navigation of these

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 684–693, 2007.

components through the arterial network is aided by a fluoroscope display showing the positions of these radiopaque instruments within the arterial network. Upon reaching the blockage point a contrasting fluid is injected into the patient permitting the blockage to be viewed on the fluoroscope display. The catheter is changed to an angioplasty catheter with a balloon disposed at its distal end which is centered in the blockage region and inflated to compress the blockage material on the artery walls to open the blood passageway. Balloon inflation is viewed on the display to confirm that the balloon is appropriately inflated to eliminate the blockage without rupturing the artery walls.

Performing MIS procedures requires great skill to avoid complications that may cause serious injury to a patient and/or requires the patient to undergo open surgery. Thus, physicians need to acquire the necessary skill and experience to perform MIS procedures to ensure successful performance. Although practicing MIS procedures on live patients provides excellent training, the procedure may only be performed once on a particular live patient and typically requires the presence of a skilled physician to supervise and oversee the procedure to avoid serious injury to the patient. Further, training physicians or other medical professionals in MIS procedures on live patients requires the use of proper facilities and equipment (e.g., hospital facilities and equipment), thereby incurring substantial costs and limiting procedure practice to a particular time and location. Moreover, since only one physician is able to practice one procedure on a particular live patient, the quantity of physicians who may practice or perform MIS procedures is severely restricted, thereby limiting the quantity of physicians that can acquire sufficient experience to perform these types of procedures. In addition, this traditional training scheme – learning on a patient, is becoming unacceptable because it is risky and unfair for the patient.

To overcome the above described disadvantages, employing simulation techniques to train physicians or other medical professionals could be an ideal solution. Consequently, in August 2004, the FDA (Food and Drug Administration) mandated that physicians performing carotid stenting must complete a multi-step educational program including simulator-based training to proficiency [3]. Several groups have reported work in this field. HT Medical, Inc. has developed a *Dawson-Kaufman* simulator for practicing angioplasty and other procedures [4]. This simulator is for catheter navigation in abdominal aorta. Physical modeling-based computation analysis of the catheter movement requires a little longer time than the desired real-time response. Another simulator is *CathSim*, and its applications include peripheral intravenous catheterization [5] and more realistic training of nursing students in performing venipuncture [6]. *ICTS* is an interventional cardiology training system [7]. Another training system for interventional radiologists [8] is for filter placement in the inferior vena cava.

Our group has been pioneering the development of interventional radiology simulators from the beginning of the 90s. The first development was *daVinci* [9], an interventional radiology simulator for catheter navigation. *daVinci* provides a basic framework of human vascular system registered with synthesized fluoroscopic views to create a virtual patient for radiologists' training. A potential field supported finite element method [9] [10] was developed for real-time analysis of catheter navigation. The next simulator *ICard* [11], [12] has been developed for interventional cardiology. These simulators have been demonstrated as commercial or scientific exhibits at

(a) (b) (c)

Fig. 1. Interventional radiology simulators: a) *daVinci* at RSNA 1995 and SCVIR 1997, b) *ICard* at MICCAI 1999, and c) *NeuroCath* at ECR 2000

several meetings: *daVinci* as a technical exhibit at the Radiological Society of North America RSNA 1995, *ICard* at MICCAI 1999, and *NeuroCath* at the European Congress of Radiology ECR 2000, Figure 1.

The next prototype was *NeuroCath*, which is focused on neuron-interventional procedures. The system was demonstrated at ECR 2000 [13], RSNA 2001 (award winner) [14], 2004 (award winner) [16], 2005 [17] and 2006 [18], and MMVR 2002 [15]. The latest development of *NeuroCath* including improvement of system design, interface, haptic feedback, modeling on vascular variations is reported in this paper.

2 Interventional Neuororadiology Simulator

2.1 System Overview

Figure 2 shows the architecture of *NeuroCath*.

The user interacts with the system through visual and haptic interfaces. The patient data in the format of VHD (Visible Human Data), XRA (X-ray Angiography), CTA (Computerized Tomography Angiography) or MRA (Magnetic Resource

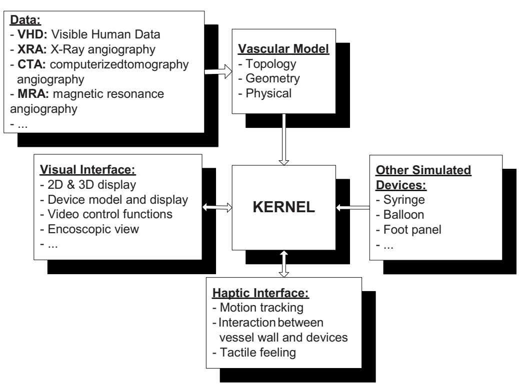


Fig. 2. Architecture of *NeuroCath*

Angiography) are input and subsequently the vascular model is extracted from them encompassing aspects of topology and geometry. Physical properties can be gathered from other resources, such as literature.

The visual interface provides 2D and 3D displays of the patient-specific data, devices, and intervention situation. A 3D volume and/or surface rendered display of the vascular model is provided. The location of the device on axial, coronal, and sagittal cross-sections can be displayed. *NeuroCath* also provides plenty of video control functions such as zoom in/out, rotation, shifting, shuttle and recording. In addition, the interior view of blood vessels that simulates endoscopic imaging is available.

Motion information of devices (catheter and guidewire) is detected by the motion tracking structure in the motion tracking and force feedback box (MFB). The information is used to calculate the interaction between the vessel wall and devices. It allows the user to deftly push, pull, and twist the navigation devices.

Other devices, including syringe and balloon device are simulated as the inputs for the simulator.

Details of hardware and software components are described as following.

2.2 Hardware Components

2.2.1 Motion Tracking and Haptic Feedback Box (MFB)

A library of catheters, guidewires, stents, and coils, from various manufacturers has been built into *NeuroCath* for selection. Then, the user sets the insertion/entry point. The navigating directions of catheter and/or guidewire have to be set towards or away from the target. The MFB, an electromechanical device, is used to track these movements [19]. MFB is an input device that has been specially constructed for simulators of interventional radiology procedures. It has a close resemblance to the real catheter control mechanism and allows the user to manipulate the translation and rotation of the catheter the way he or she experiences in a real interventional procedure. The principle and physical set is shown in Figure 3.

As shown in Figure 3(a), two encoders attached to wheels A and B record rotation and translation motion separately. Units 9 and 10, controlled by a servo motor, realize tactile feedback on the device. When they move towards the catheter, force feedback is provided and the force increased; when move backwards, force is reduced and the device is released.

In real procedures, in some dangerous cases, such as tending to puncture a vessel wall, the device should be prevented from being pushed further, but free to be pulled back. To simulate this process, a self-locking structure is designed. The structure consists of unit 5 - a string and unit 7. When friction force added by unit 9 on unit 3 reaches a customized critical level, the catheter will be “locked” so that pushing motion is prevented and only pulling is allowed.

2.2.2 Other Simulated Devices

Contrast Injector: A syringe is connected to a special control panel to simulate the work of a contrast injector. A user is able to set the injection volume and rate. The appearance of contrast filling vessels is modeled using opacity which reflects the rate

Fig. 3. (a): Mechanical structure of the MFB. The motion tracking module processes the translation and rotation motion of the catheter/guidewire. (b): Physical set of MFB.

Fig. 4. (a): Simulated contrast injection with a syringe; (b): Fluoroscopic images with real time catheter tracking; (c): roadmap image, after contrast media is injected.

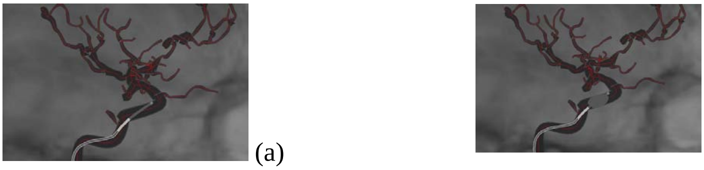
of injection relative to the blood flow rate in vessels. This provides a realistic impression of the injection, Figure 4. In addition, the washout of contrast through the vasculature is simulated.

Ballooning Inflation Device: NeuroCath is able to simulate ballooning and stent placement. The balloon is inflated in a cerebral artery using a customized simulated balloon inflation device, Figure 5. Compared with commonly used balloon inflation devices, this specially designed one provides an electronic output signal that is useful for simulation purposes.

Figure 5(a) shows the handheld balloon inflation device for simulation purposes. Figure 5(b) illustrates the simulation process using the device. Figure 6 illustrates inflation of the angioplasty in the human cerebral vasculature. Figure 7 shows a stent deployed by balloon inflation at the middle portion of the same cerebral artery. This provides a realistic simulation environment for the user to practice or for pretreatment planning of pertaneous transluminal coronary angioplasty procedures.

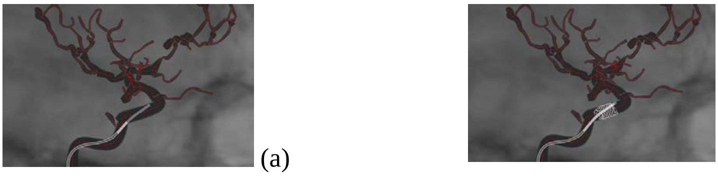
Fig. 5. (a) Balloon inflation device. (b) Simulating balloon inflation and deflation

This new desktop-based automatic catheter balloon inflation device can be used not only in the simulation system, but also in clinical procedures. The manipulation of the device allows the physician to complete the balloon inflation procedure using one hand. The accuracy of the system in recording and displaying inflation pressures is



(a) (b)

Fig. 6. Balloon inflation on a fluoroscopic image and the surface rendered model of vessels. (a): before inflation; (b) after inflation.



(a) (b)

Fig. 7. Stent deployment on fluoroscopic image and surface rendered model of vessels. (a): before deployment; (b): after deployment.

very high [20]. Finally, physicians can control the clinical procedure according to the pressure display on the handheld remote controller and be provided with additional balloon inflation safety alarm features.

2.3 Software Components

2.3.1 Vascular Modeling

The vascular model, based on centre-line format, uses a hierarchical structure (a tree) to represent the topological connectivity of the centre-line segments. Each segment is a skeleton section of the corresponding blood vessel and it contains a sequence of sampling nodes. Each node contains a group of basic properties defined on the cross section in the corresponding MRA image, including spatial coordinates of the central point, vessel radius, and orientation of the blood flow. The centre-line model can supply the effective and sufficient information to build the 3D vasculature model. Details were described in [21].

2.3.2 FEM Modeling

Beneath our simulator are the physical models that abstract the behavior of the various components in the simulation process. These physical based models define the interaction between physiology, anatomy and instrumentation through the governing physics represented in the physical modeling. It dictates how different levels of simulations (anatomy, physiology, device or physics) interact with each other. We take into account the functional nature of human vasculature. For instance, damaging the vessel wall from excessive pushing (physical phenomenon) may have an effect on the blood pressure. On the other hand, the development of tumorous lesions (physiological phenomenon) modifies locally the tissue mechanical properties.

2.3.2 Customization Vascular Models

The human vascular anatomy is unique to each patient, individual variations of vascular anatomy are common, and each individual patient's angiographic appearance can be very different from each other and from those featured in anatomy textbook or atlases. Some of these variations may be challenging for less experienced interventionists to manipulate the interventional devices, or may even require the use of unique interventional device shapes and sizes. Also, these variations are not easily to be found and used as training cases.

Till now, almost all the interventional radiology simulators can only provide limited, pre-generated vascular models, which are clearly not sufficient for students to understand and practice on variable models. So we took our efforts to develop the new function of providing customizable vascular models so that the applicability of the simulator can be improved significantly.

We firstly develop a tool which is called Vascular Editor. This tool can load in initial vascular model. In the Editor, 3D coordinates of each node and corresponding radius can be modified. More advanced functions like dragging nodes/segments in 3D space, adding/removing branches are also developed. After the vascular structure is customized, new model is integrated into simulator and available for training. An example about four major types of variations of ICA (internal carotid artery) is shown in Figure 8.

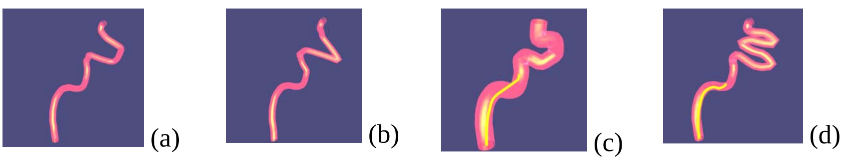


Fig. 8. Variations of ICA (internal carotid artery). (a): U-shape; (b): V-shape; (c): Omega shape with a interventional device; (d): Double Siphon shape with a interventional device**.**

The initial testing shows that:

- When manipulating similar size of device in these variations, double siphon is the most difficult situation, omega shape is much simpler, U-shape and V-shape are in the middle;
- Properties of the device influence the simulation result seriously. For example, if the material parameter is more flexible, and the size is smaller, the manipulation is much easier to be successful.

These results are as reasonable as we can imagine. It is encouraging that it shows the value of this feature but more clinical validations are necessary.

3 Current Results

The physical setup of *NeuroCath* (version 4.0) is shown in Figure 9.

NeuroCath is firstly validated in Johns Hopkins School of Medicine, then licenced to several hospitals and research centres including Clinic of Diagnostic and

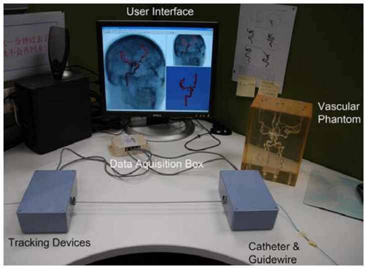


Fig. 9. Physical setup of *NeuroCath* 4.0

Interventional Neuroradiology (University of the Saarland, Germany), Xuanwu Hospital (Beijing, China), German National Research Center for Information Technology and so on.

Table 1 summarizes the feedback from these centers.

Table 1. Summary of clinical feedback

Criteria		Clinical feedback
Training Acceptance	To Entry Level Training	Good
	To Higher Level Training	Not sufficient, more clinical knowledge need to be integrated.
Visual Interface		Good.
Device Tracking		Good.
Training Feature Fulfilled		70-75%
Visual Reality		Needs improvement, particularly on the vascular model rendering.
Haptic Reality		Needs improvement.
Improvement Needed		1).The accuracy of the tactile feeling; 2).Vascular extraction: faster segmentation and processing method 3).Vascular modelling: deformable and smoother model;

4 Discussion

NeuroCath provides 2D and 3D views of patient-specific data and interventional devices, quantitative information of segmented vasculature along with pathology, realistic haptic interfacing environment for manipulating navigation devices, realistic deformation analysis of interaction between the devices and the blood vessels, and support for key interventional neuroradiology procedures. These features are important and helpful in planning and training interventional neuroradiology procedures.

From the clinical feedback, we are encouraged that the system can fulfil 70%-75%of training requirements and entry-level trainees find it especially useful. But there need a lot of improvements.

- Acceptance for higher level training. The feedback about key problem is that training cases are limited and too simple to higher level trainees.

Customizable vascular models can be a good approach, but more clinical meaningful cases are needed. Then more validation work on the vascular modeling and FEM analysis method will be necessary to make sure the simulation results on these variations can represent the real situations accurately.

- Deformable vascular model and memodynamic model for blood flow. They are critical in improving the visual reality and haptic reality of the system. As the ultimate goal of *NeuroCath* is to provide interventional neuroradiologists and neuroradiology residents with an accurate, realistic interventional neuroradiology simulator, we need to enhance the 3D deformable model that facilitates simulation of interventional neuroradiology procedures more realistically in both visual and haptics senses, while a hemodynamic model to credibly simulate the blood flow and its interaction with the vessels in the presence of vascular malformations is desirable.
- Future development direction. It will be reasonable and more valuable to extend the training system into a pre-planning system and even a surgical assistant system. We are in the process of developing a more accurate and faster extraction approach for our current vascular extraction and reconstruction procedure. We are also exploring to use the system for telemedicine of interventional radiology procedure. [22]

Acknowledgements. Support of this research and development by Agency for Science, Technology and Research of Singapore is gratefully acknowledged**.**

References

1. Verel, D.: Cardiac Catheterization and Angiocardiography, 3rd edn. Longman, New York, USA (1978)
2. Ring, E.J., McLean, G.K.: Interventional Radiology: Principles and Techniques, Little, Brown, Boston, USA (1981)
3. Cotin, S., Luboz, V., Pergoraro, V., Neuman, P.F., Wu, X., Dawson, L.S.: High-fidelity simulation of interventional neuroradiology procedures. In: Proc. American Society of Neuroradiology ASNR May 23-27, 2005, Toronto, Canada, 443 (2005)
4. Meglan, D.: Making Surgical Simulation Real, Computer Graphics, 30(4), 37–39 (1996)
5. Ursino, M., Tasto, J.L., Nguyen, B.H., Cunningham, R., Merril, G.L.: Cathsim: An Intravascular Catheterization Simulator on a PC. Studies in Health Technology and Informatics 62, 360–366 (1999)
6. Barker, V.L.: CathSim. Studies in Health Technology and Informatics 62, 36–37 (1999)
7. Cotin, S., Dawson, S.L., Meglan, D., Shaffer, D.W., Ferrell, M.A., Bardsley, R.S., Morgan, F.M., Nagano, T., Nikom, J., Sherman, P., Walterman, M.T., Wendlandt, J.: ICTS, an Interventional Cardiology Training System. Studies in Health Technology and Informatics 70, 59–65 (2000)
8. Hahn, J.K., Kaufman, R., Winick, A.B., Carleton, T., Park, Y., Lindeman, R., Oh, K.M., Al-Ghreimil, N., Walsh, R.J., Loew, M., Gerber, J., Sankar, S.: Training Environment for Inferior Vena Caval Filter Placement. Studies in Health Technology and Informatics 50, 291–297 (1998)

9. Wang, Y., Chui, C., Cai, Y. et al.: Potential Field Supported Method for Contact Calculation in FEM Analysis of Catheter Navigation. In: 19th International Congress of Theoretical and Applied Mechanics, Kyoto, Japan 1996 (1996)
10. Chui, C., Nguyen, H., Wang, Y., Mullick, R., Raghavan, R., Anderson, J.: Potential field and anatomy vasculature for real-time computation in daVinci. In: 1st Visible Human Conference, Abstract:P113, Bethesda, MD, USA, October 1996 (1996)
11. Wang, Y., Chui, C., Lim, H., Cai, Y., Mak, K.: Real-Time Interactive Surgical Simulator for Catheterization Procedures. Journal of Computer Aided Surgery 3(5), 211–227 (1999)
12. Chui, C., Wang, Y., Cai, Y., Lim, H., Ooi, Y., Mak, K.: ICard: An Interventional Cardiology Simulator for Percutaneous Coronary Revascularisation. In: Computer Assisted Radiology and Surgery (CAR'98), Tokyo, June 24-27, 1998 (1998)
13. Nowinski, W.L., Wang, Y.P., Srinivasan, R., Chui, C.K., Guo, H.H., Pang, P., Xu, M., Thiru, A.: Model-enhanced radiology. European Radiology 10(2) Supplement 1, 444 (2000)
14. Anderson, J.H., Chui, C.K., Li, Z.R., Ma, X., Cai, Y.Y., Wang, Y.P., Murrphy, K., Venbrux, A., Nowinski, W.L.: PC based simulator for cardiovascular catheterization and image guided therapy. In: 87th Radiological Society of North America Annual Meeting 2001, Chicago, USA (2001)
15. Chui, C.K., Li, Z.R., Anderson, J.H., Murphy, K., Venbrux, A., Ma, X., Wang, Z., Gailloud, P., Cai, Y.Y., Wang, Y., Nowinski, W.L.: Training and Planning of Interventional Neuroradiology Procedures - Initial Clinical Validation. In: Medicine Meets Virtual Reality (MMVR), Jan 23-26, 2002, vol. 85, pp. 96–102, IOS Press, CA (2002)
16. Ma, X., Zhao, L., Zheng, W.L., Volkau, I., Aziz, A., Nowinski, W.L.: Computer assisted interventional neuroradiology procedures: optimal path guidance and remote operation system. In: The 90th Radiological Society of North America, Chicago, USA, 2004, p. 831(2004)
17. Ma, X., Zhao, L., Aziz, A., Zaoli, Z., Jeremy, L.O.H., Nowinski, W.L.: The use of impedance sensors for intravascular navigation based on landmarks and a virtual map. In: Program 91st Radiological Society of North America Scientific 2005, Chicago, Illinois, USA, 27 November - 2 December 2005, p. 864 (2005)
18. Xin, M., Jun, T.D., Jiming, L., Volkau, L., Marchenko, Y., Nowinski, W.: A customizable training simulator for interventional radiology. In: Program 92st Radiological Society of North America Scientific 2006, Chicago, Illinois, USA, p. 791(2006)
19. Ma, X., Chui, C., Wang, Y., Nowinski, W.L.: Haptic Interfacing Device in Computer Simulator for Interventional Procedure. In: IEEE-EMBS Asia Pacific Conference on Biomedical Engineering, Hangzhou, China, September 2000 (2000)
20. Ma, X., Zhong, F., Chui, C., Cai, Y., Anderson, J., Nowinski, W.L.: Digital Balloon Inflation Device for Interventional Radiology. In: Niessen, W.J., Viergever, M.A. (eds.) MICCAI 2001. LNCS, vol. 2208, pp. 335–342. Springer, Heidelberg (2001)
21. Volkau, I., Zheng, W., Baimouratov, R., Aziz, A., Nowinski, W.L.: Geometric modeling of the human normal cerebral arterial system. IEEE Trans Med. Imaging 24(4), 529–539 (2005)
22. Xin, M., Lei, Z., Volkau, I., Weili, Z., Aziz, A., Ang Jr, M.H., Nowinski, W.L.: A virtual reality simulator for interventional tele-radiology: concept, design, and initial validation. IEEE Biomedical Engineering 53(8), 1696–1700 (2006)

Methodologies to Evaluate Simulations of Cardiac Tissue Abnormalities at a Cellular Level

Nicos Maglaveras and Ioanna Chouvarda

Aristotle University, The Medical School, Lab of Medical Informatics,
Box 323, 54124, Thessaloniki, Greece
{nicmag,ioanna}@med.auth.gr

Abstract. Computer Simulations of cardiac wave propagation may be used as a tool towards understanding the mechanisms of cardiac conduction, the nature of various heart diseases, as well as the effect of drugs in cardiac function. Such simulations depend on the ionuc current model adopted, and various such models have been proposed. The exact propagation wavefront thus depends on the ionic model and the tissue properties, being homogeneous or heterogeneous. The latter case, which corresponds to infracted cardiac tissue, is the focus in this work. The ionic current properties and the sodium kinetics on a two-dimensional grid where wavefront rotations around barriers at bifurcation sites take place, are examined in detail and differences in propagation characteristics elicited by using fast or slow fast inward current kinetics such as can be found in the Beeler-Reuter, Luo-Rudy and Ebihara-Johnson models are elaborated.

Keywords: ionic current models, cardiac propagation, conduction abnormalities.

1 Introduction

Cardiac tissue has received lots of attention in previous years regarding modeling and simulation of its function in normal and abnormal situations. Cardiac tissue has a number of different types of cells related with functionally different parts, such as the atrium, the AV node, the ventricles and the SA node to name the most well known ones.

In most cases the modeling and simulation of the function of the cardiac tissue at the cellular level depends on detailed ionic current based models such as the Luo-Rudy (LR), Ebihara-Johnson (EJ), Beeler Reuter (BR) and DiFrancesco and Noble (DFN) to name a few of the most well known ones [1-3]. Studies on the underlying physiological characteristics of conduction and their effect to the measured electrical field either at the surface of the heart or at the body surface are found in the literature. Usually the most addressed topics are the ones linked with the fast upstroke characteristics or the ones linked with the dispersion of refractoriness and the global repolarisation properties of the cardiac muscle.

The type of the ionic current models used and the kinetics properties may have a profound effect even in the basic propagation characteristics in the cardiac muscle. In

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 694–702, 2007.

this study we investigate the effects of different ionic current models used when simulating propagation in myocardial infracted cardiac tissue linked with zig-zag type pathways, especially at the bifurcation sites as can occur in infracted tissue.

Fast and Kleber [4] presented a review regarding the theoretical background relating impulse propagation to wavefront curvature and analysed the role of wavefront curvature in electrical stimulation, formation of conduction block, and the dynamic behavior of spiral waves. The effects of tortuosity, created by spatial distribution of dead cell barriers, on 2-D propagation have been also studied by Maglaveras et al [5].

In this work, considering a rotating propagation wave, due to dead cell barriers, we take a closer look at the model kinetics in the area of rotation. The behavior of three ionic current models in such areas is compared. We show that the changes in ionic current spatial properties largely depend on the type of model of the fast inward current used, something that can be linked as well with the extracellular fields calculated at these sites, which may effect the correct identification of the activation instant when analyzing the electrograms.

2 Methods

2.1 The Ionic Current Models

The ionic current models used were the Beeler-Reuter (BR), Luo-Rudy (LR) and Ebihara-Johnson (EJ) [1-3]. The BR and LR in the fast inward current (I_{na}) formulation exhibit one activation gating variable (m) and two inactivation gating variables (h,j), while the EJ model exhibits one activation (m) and one inactivation (h) variable. The kinetics of the BR model are much slower when compared with the LR and EJ models, and this may play a role in the behaviour of the propagating wavefront especially in regions of high anatomical and functional complexity such as the bifurcation sites in an infracted cardiac tissue.

2.2 Data and Experimental Setup

Simulated data have been used for this work, in order to test and compare different methods for a structure of known complexity and behavior. Simulated experiments also allowed us to control the number of extracellular measurements to be used during the estimation procedure. The 2D monodomain model of cardiac muscle was used [5]-[6]. Signals are extracellular electrograms recorded at the surface of the tissue. A rectangular sheet of cells was simulated with a grid of 40x200 elements, which corresponds to 40x50 cells. Each cell consists of 4 longitudinal elements (3 cytoplasmic and 1 junctional). Elements were coupled by resistances. Beeler-Reuter Ebihara-Johnson and Luo-Rudy models have been used for ionic kinetics. These models express different kinetics, which lead to different frequency content. Regions of scar tissue were defined in order to introduce heterogeneity (see Figure 1). Simulation of propagation was implemented using Gauss Seidel method. The grid intervals were 10 μm transversely and 25 μm longitudinally and time step was 1μsec.

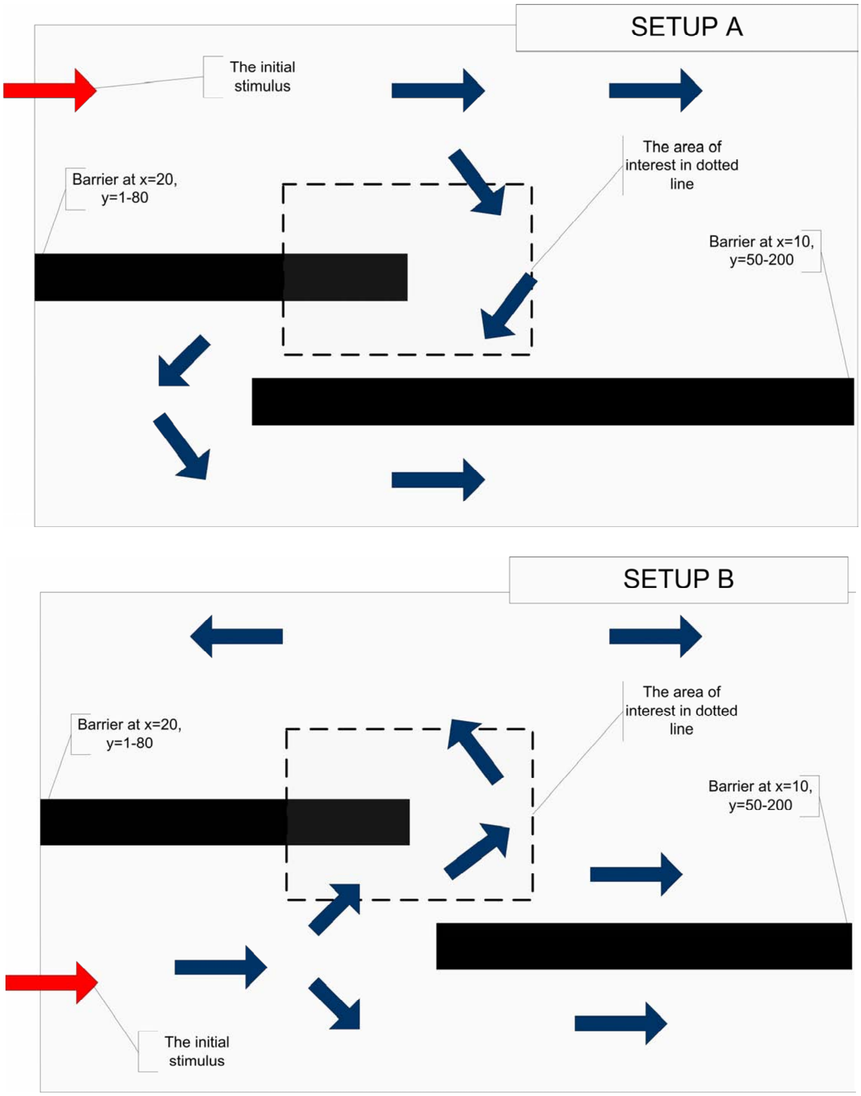


Fig. 1. The two experimental setups used. Both contain areas of heterogeneity

During simulation, transmembrane current was stored for certain grid-points. Activation times were also calculated using transmembrane current criteria analysis. Current waveforms and activation maps were used for comparison with the ones estimated by the methods described in this work.

The two settings used in this paper are shown in Figure 1. Two barriers of scar tissue, cause wavefront to follow a zigzag pathway. Within an area of interest where both areas of normal propagation and velocity changes are present, due to the barriers of scar tissue, specifically in the rectangle defined by $15<x<25$ and $67<y<103$, thick recording take place, in a grid, with 10 μm transverse interval and 100μm longitudinal interval, corresponding to one recording per cell.

Regarding implementation issues, simulations were performed with an in-house application written in standard C code, while and post analysis was performed in Matlab.

2.3 Analysis and Features

Having available the calculated data of the simulation setups within the predefined grid, the features depicted for further analysis included the ionic current and the fast kinetic currents, specifically:

- The absolute minimum value of the ionic current, Ion_{min} as a basic morphological feature also strongly affecting the membrane current and the excitability, which is supposed to change due to heterogeneity, and specifically decrease in the infarcted regions.
- Considering the relation between *m*, *h* and *j* fast current kinetic functions, the underlying surfaces of the m-h and h-j plots, i.e. S_{mh} and S_{hj}, as described in Equations 1 and 2 respectively.
- In the same relation, feature D_{mh} describes the minimum distance between m-h curve and (1,1) point, as shown in Figure 2. A similar definition stands for D_{hj} feature (see Equation 3 and 4 for these descriptions).

$$S_{mh} = \frac{\sum_{i=1:N} m(i) \cdot h(i)}{N} \tag{1}$$

$$S_{hj} = \frac{\sum_{i=1:N} h(i) \cdot j(i)}{N} \tag{2}$$

$$D_{mh} = \min\{\sqrt{(m(i)-1)^2 + (h(i)-1)^2}\} \tag{3}$$

$$D_{hj} = \min\{\sqrt{(h(i)-1)^2 + (j(i)-1)^2}\} \tag{4}$$

The reason for selecting these features is to have a close examination at the detailed behavior of each model, especially in heterogeneous areas, instead of the extracellular potentials morphological analysis, usually followed for a more macrospopic analysis and activation times determination.

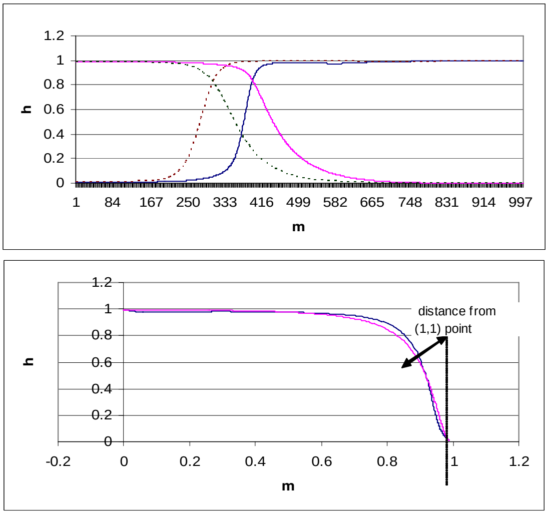


Fig. 2. (top) The voltage-dependent inactivation gate of the fast sodium channel - the h gate and the .fast_sodium_current_m_gate functions. Solid lines BR, dashed lines LR. (down) m-h plots and the distance feature.

3 Results

3.1 Analysis of the Ionic Current

A close examination at the variation of Ion_{min} in the area of interest, which corresponds to an area where the wave is initially planar, then starts rotating until it gets again planar, reveals differences in the behavior of the three models. Near the turning point (20,80), in BR model Ion_{min} decreases, while LR and EJ have an abrupt increase close to the turning point (see Figure 3). Specifically, in setup A, the maximum Ion_{min} value for BR model is at point (25, 62),), at the beginning of the rotation, while point (20,82), just in front of the barrier, holds the maximum Ion_{min} value. In setup B, point (17, 62), at the beginning of the rotation, has the maximum Ion_{min} value, while point (21, 82), just beside the barrier holds the maximum Ion_{min} for LR and EJ models. As shown in Figure 4, the behavior of Ion_{min} in LR-EJ models is quite similar, but quite distinct from that of BR model.

Another way to view possible differences in the area of wavefront rotation, is a representation in polar coordinates (see Figure 5). The different behavior in the areas of rotation is quite obvious. In both setups, the Ion_{min} changes in LR and EJ models in the rotating area seem to be more dependent on the angle. Of course, as expected, in bigger distances from the rotating point, these phenomena decrease.

In order to show the different dependence of Ion_{min} on distance and angle from the rotation point, we have computed a regression model. as shown in Equation 5. The model parameters, as calculated based on the two setups, for the three different models are depicted in Table 1. It has to be noted that the two setups create a

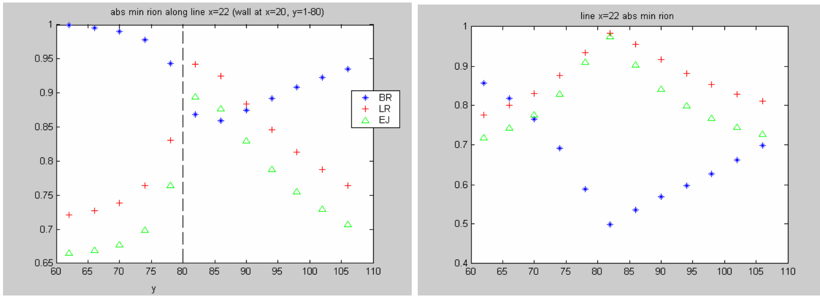


Fig. 3. Along a longitudinal lines (x=22), near the barrier, the evolution of Ion_{min}, normalized to the maximum value within the area of interest, for the 3 models. (left) setup A; (right) setup B.

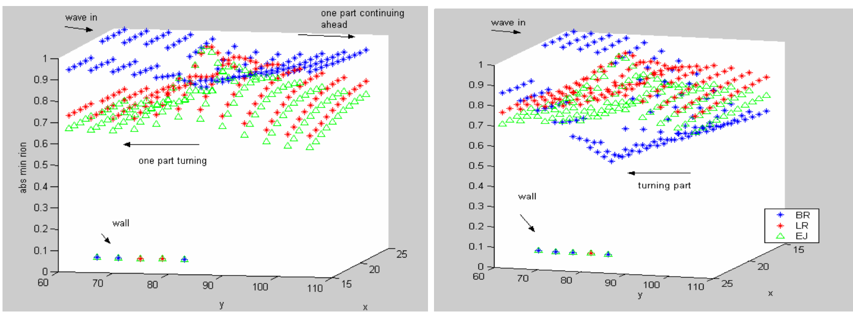


Fig. 4. Normalised Ion_{min} values in the area of interest, in cartesian coordinates, for the three models. (left) setup A, (right) setup B.

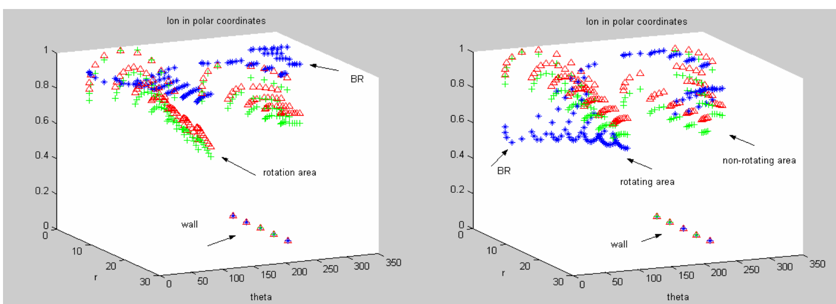


Fig. 5. Normalised Ion_{min} values in the area of interest, in polar coordinates, for the three models. (left) setup A, (right) setup B.

curvature of the opposite side, since propagation follows the opposite direction in the two cases. The difference in the parameter values among the BR model and the other two models are quite obvious.

$$Rion\_min=F(r,theta)= a*r+b*theta+c. \tag{5}$$

Table 1. Linear regression parameters for the relating ionic current with the polar coordinates, for the three models and the two setups

	BR		LR		EJ	
	Setup 1	Setup 2	Setup 1	Setup 2	Setup 1	Setup 2
A (radius)	-0.2330	3.3101	-3.5677	-4.2512	-3.4527	-8.7566
B (theta)	-0.2399	0.6967	-1.6461	-2.6526	-2.2098	-3.1777
C (const)	-115.5088-	236.9449	-224.002	-168.779	-323.035	-193.41

3.2 Analysis of m-h-j Kinetic Functions

m-h Fast Gates Relation

In BR model, D_{mh} increases towards the rotation area, while it decreases again after the rotation towards the normal propagation (see Figure 6). Lowering of Dmh corresponding to more stiff m,h functions, more 0-1 behavior, with less broad transitional area. D_{mh} has the opposite behavior in LR, EJ models, which in general show less variation in D_{mh}. This might imply that LR and EJ kinetics do exhibit the variation expected in varying propagation conditions.

Regarding the surface S_{mh}, again in the turning point there is a significant increase in the BR model, coming back to normal when propagation gets normal, as depicted in Figure 7. There is a similar behavior in LR and EJ models, but with with much lower values.

h-j Inactivation Gates Relation

BR and LR models are quite close, while EJ model has a different behavior, regarding Shj and Dhj. Dhj has much larger values in EJ model (around 0.5) than in BR and LR models. Regarding S_{hj} feature, there are some interesting observations: a) the three models have a distinct behavior, but in general decrease of S_{hj} near the rotation area, and b) there is no axial symmetry as can be seen when comparing S_{hj} in line 22 in the two setups, where the wave is rotating in the opposite direction, as can be seen in Figure 8 (down). This is in contrast with S_{mh}, where there seems to be symmetry.

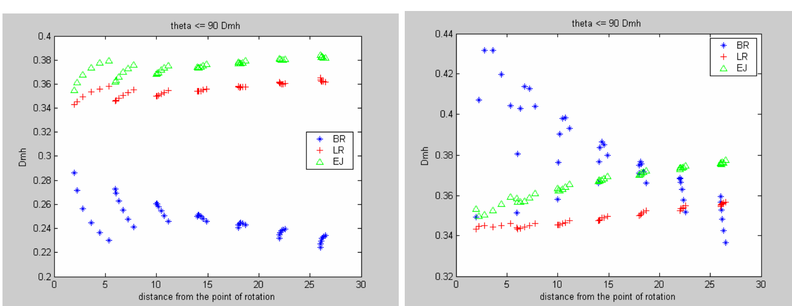


Fig. 6. The D_{mh} feature with distance, for points within the angles 0<theta<90, i.e. in the rotating area. Left setup A, right setup B.

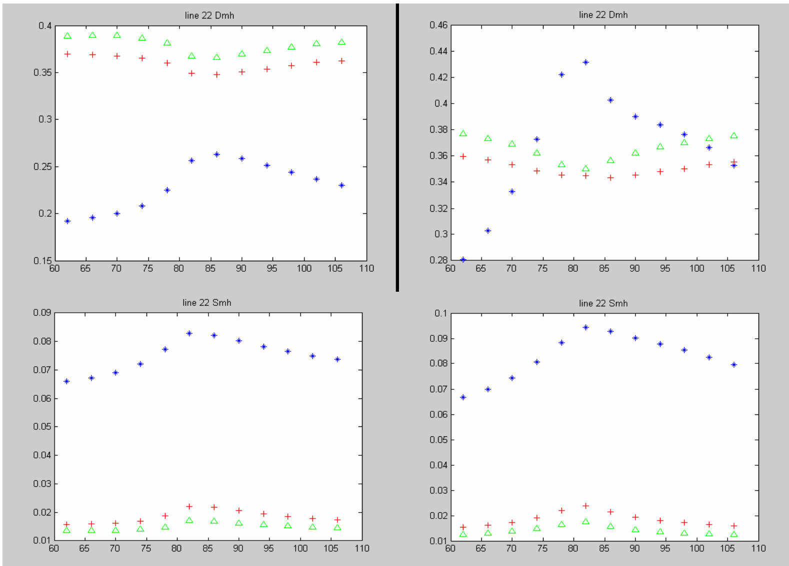


Fig. 7. (Top) The D_{mh} along the longitudinal line x=22. (Down) S_{mh} along the same line. Left setup A, right setup B.

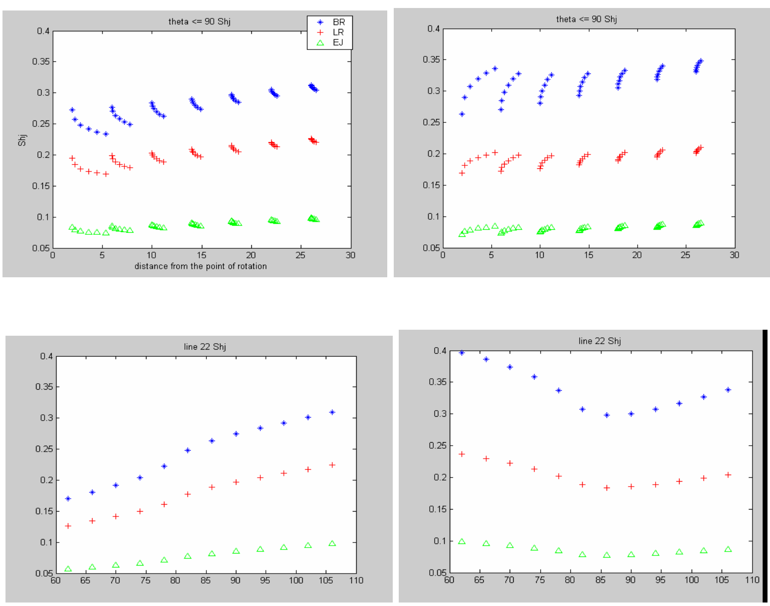


Fig. 8. (Top) The S_{mh} for all point with angle < 90 degrees from th rotating point. (Down) S_{mh} along the longitudinal line x=22. Left setup A, right setup B.

5 Discussion

In the present wok, three different models are used in a 2-D tissue and the differences in the sodium channel kinetics and the ionic currents are studied, for these models and for areas of heterogeneity. The minimum value of the ionic current is a feature found to have different behavior in the BR model than in LR and EJ models. Specifically, in the former it is decreasing, while in the latter it is abruptly increasing in the vicinity of a propagation barrier. In computer simulations of the infarcted myocardial tissue by others [7], LR model has been used, and it is suggested that reduction of the current density of I_{to} found in failing myocytes of human hearts does not seem to contribute significantly to the APD prolongation in heart failure, which instead is mainly due to the enhanced activity of the Na+-Ca2+ exchanger.

Furthermore, regarding the sodium kinetics, and specifically the m-h and the h-j relations in the three models, it was derived that while LR and EJ have a very similar m-h fast kinetics behavior, however quite opposite from the one in BR model, the three models exhibit a distinct behavior as far as h-j relation is concerned, although a reduction in the channel availability is reported in all models. Sodium channel availability, as expressed by the product $h\ j$ of the two inactivation gates (h=fast, j=slow) of I_{Na}., has been studied by [8] and a significant reduction has been reported in the transition zone as compared to the homogeneous segments.

The differences of these models should be checked against experimental evidence, and under different conditions of propagation, in order to understand and evaluate the conditions favorable .for the adoption of each model.

References

1. Beeler, G.W., Reuter, H.: Reconstruction of the action potential of ventricular myocardial fibres. Journal of Physiology 268, 177–210 (1977)
2. Luo, C.H., Rudy, Y.: A dynamic model of the cardiac ventricular action potential. I. Simulation of ionic currents and concentration changes. Circ. Res. 74, 1071 (1994)
3. Ebihara, L., Johnson, E.A.: Fast sodium current in cardiac muscle. A quantitative description. Biophysical Journal 32, 779–790 (1980)
4. Fast, V.G., Kleber, A.G.: Role of wavefront curvature in propagation of cardiac impulse. Cardiovasc Res. 33(2), 258–271 (1997)
5. Maglaveras, N., Offner, F., van Capelle, F.J., Allessie, M.A., Sahakian, A.V.: Effects of barriers on propagation of action potentials in two-dimensional cardiac tissue. A computer simulation study. J Electrocardiol 28(1), 17–31 (1995)
6. Henriquez, C.S., Papazoglou, A.: Using Computer Models to Understand the Roles of Tissue and Membrane Dynamics in Arrhythmogenesis. Proc. IEEE 84(3), 334–354 (1996)
7. Priebe, L., Beuckelmann, D.J.: Simulation study of cellular electric properties in heart failure. Circ Res. 82(11), 1206–1223 (1998)
8. Wang, Y., Rudy, Y.: Action potential propagation in inhomogeneous cardiac tissue: safety factor considerations and ionic mechanism. Am. J Physiol Heart Circ. Physiol 278(4), H1019–H1029 (2000)

Multi-level Analysis and Information Extraction Considerations for Validating 4D Models of Human Function

Kostas Marias[1], Dimitra D. Dionysiou[2], Georgios S. Stamatakos[2], Fotini Zacharopoulou[3], Eleni Georgiadi[1,2], Thanasis Margaritis[4], Thomas G. Maris[3], and Ioannis G. Tollis[1]

[1] Biomedical Informatics Laboratory, ICS-FORTH, Vassilika Vouton, P.O. Box 1385, 71110 Heraklion, Crete, Greece
[2] In Silico Oncology Group
Laboratory of Microwaves and Fiber Optics Institute of Communication and Computer Systems School of Electrical and Computer Engineering National Technical University of Athens Iroon Polytechniou 9, GR-157 80 Zografos, Greece
[3] Medical Physics Department, University of Crete, Faculty of Medicine, Heraklion, Greece
[4] Institute of Molecular Biology and Biotechnology, IMBB -FORTH, Vassilika Vouton, P.O. Box 1385, 71110 Heraklion, Crete, Greece
kmarias@ics.forth.gr

Abstract. Recent research trends focus on how multiscale biomedical information can be modeled and transformed into knowledge, in order to lead to a less interfering but also more individualized diagnosis and therapy. In order to assess the clinical importance of models of human pathology (e.g. cancer), it is necessary to validate them with prior and post treatment clinical data which in turn requires the determination of the tumor size and shape with high resolution, accuracy and precision, as well as structural and physiological information. This paper discusses some of the most important image analysis challenges in order to define an optimal method for extracting more accurate and precise anatomical and functional information related to the underlying pathology, which can be used for initializing and validating models of pathophysiology as well as simulations/predictions of the response to therapeutical regimes.

Keywords: Virtual Physiological Human, biomedical data analysis, modeling.

1 Introduction

Modern scientific advances promise more efficient ways to optimally extract, analyze and model biomedical information in order to model pathophysiology. However, the integration of models from different levels remains an open issue. This paper addresses an important facet of this problem; the challenge of efficient and accurate extraction of biomedical information from different levels (from molecular to tissue/organ) in order to shed light in the computation of 4D maps of pathophysiological properties of tumors for model development and validation.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 703–709, 2007.

Multi-level data analysis and information extraction are necessary steps for developing and validating mathematical models of human physiology and pathology.

Despite the recent advances in medical, molecular/genetic imaging, the robust extraction of physiological parameters remains an open issue since computing them from measurements (e.g. pixel values), isn't a trivial task. Although, pathophysiological measurements are essential for modeling human processes, the physics of imaging modalities (e.g. MRI) does not absolutely generalize; it is the physics of MRI electromagnetism or the ionizing radiation in CT that produces a representation of anatomy and physiology for interpretation. Also, the mechanism by which the image is formed inherently emphasizes certain aspects and provides less sensitivity to others.

In addition, traditional diagnostic imaging, no matter how sophisticated, is only a representation of the underlying anatomical and (patho) physiological characteristics of disease. Even the most sophisticated MRI and CT techniques can only achieve an order of magnitude below millimeter spatial resolution [1,2]. Therefore the information captured in the image is only an average of what is contained within the sampling range of the instrument. Given the local complexity and “multi-scale” nature of human physiology, any attempt to use imaging to capture detailed physiological processes will understandably generalize the underlying phenomena. Thus, care must be taken when interpreting any kind of physiological parameter from imaging. This is particularly true in contrast-enhanced (CE) MRI analysis where the choice of a different analysis models may lead to a completely different result regarding the true spatial extent of a cancer.

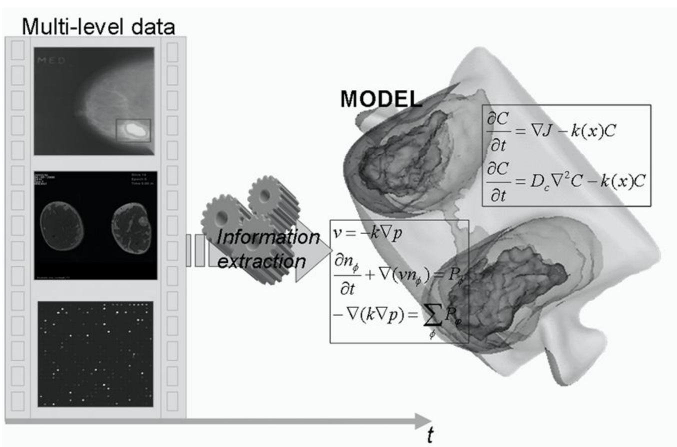


Fig. 1. Extraction of temporal pathophysiological information is essential for developing and validating multi-level models

The above considerations are very important for the development of computational frameworks for multi-level (from molecular/genetic to tissue/organ) modeling and simulation of human pathophysiology. Multi-level measurements (e.g. molecular/genetic, diagnostic imaging, etc.), should be properly interpreted and combined in order to be used in simulation models of human function. Additionally, this process needs to be repeated in several time instances in order to assess the validity of each model temporally. This concept is schematically illustrated in Fig. 1 where information extraction of temporal multi-level data is driven into the corresponding multi-level model of breast cancer. To reach this goal, it is essential to be able to optimally extract robust temporal pathophysiological information from multi-level measurements (e.g. microarrays, CE MRI, etc.).

Medical Imaging has focused in providing anatomical information, mainly imaging human bones, dense tissue and arteries. Recent advances especially in PET and functional MRI allowed the study of various pathological processes via radio-labelled tracers (PET) or pharmaco-kinetic models in contrast enhanced MRI. The whole field of molecular medicine and molecular imaging is opening up new possibilities for targeted assessment of disease and disease mechanisms. Also, microarray imaging has created exiting possibilities for measuring gene differential expression and defining new disease biomarkers. Concluding, from the imaging standpoint it is essential to consider the following points relevant to human pathophysiology modeling:

a) There is a need for a holistic understanding of pathophysiology and this clearly implies a multidisciplinary approach. To this end, molecular and genetic imaging offer unique opportunities to better understand pathophysiology in smaller scales and built multilevel models. As an example, one can think of the crucial role of imaging the genetic profile changes in cancer modeling as well as global changes in tumour size, density, etc. by analysing temporal medical imaging data.

b) It is necessary to pre-process biomedical data at all possible scales (e.g. medical images, microarray scans) in order to extract all the information that is needed for a given model. This way, multiscale information extraction aims to ‘individualise’ a given model.

The next section focuses on specific issues regarding the previous points.

2 Problems Related to Multi-level Physiological Information Extraction

2.1 Limitations in Extracting Pathophysiological Information from Medical Images

It is always essential to use of some kind of process (e.g. segmentation) to identify important structures and features in medical images (e.g. tumours can be segmented using a pharmacokinetic model of gadolinium uptake with contrast-enhanced MRI, while microarray spots can be segmented by combining the two different information channels i.e. Cy3 and Cy 5 [7]). As mentioned before, the intrinsic limitations of imaging technologies often are responsible for problems related to the assessment of subtle pathophysiological changes in micro-structure. The main problem is that in order to characterize the function of an organ it is essential to acquire multi-modality imaging data, since each imaging modality produces a representation that has different parametric properties. For example, the active volume of a tumor detected with X-ray mammography is less accurate than that using with MRI.

In cancer imaging, perfusion is a complex process that encompasses the delivery of nutrients (primarily oxygen and glucose), their diffusion and convection into the tumor parenchyma, and the removal of waste products. Therefore, any perfusion information is crucial for characterizing the tumor microenvironment. The perfusion of tumors can be interrogated with time-dependent delivery of exogenous contrast, using contrast agents. Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is in most cases applied by the clinicians in order to decide upon the state of the tumor. This is performed after the administration of intravenous contrast medium in order to noninvasively assess tumor vascular characteristics. Kinetic parameters can be then correlated with immunohistochemical surrogates of tumor angiogenesis, including microvessel density, and with pathologic tumor grade. DCE-MRI is being applied to monitor the clinical effectiveness of a variety of treatments, including antiangiogenic drugs. Kinetic parameter changes following treatment have been correlated with histopathological outcome and patient survival [8].

However, when no DCE-MRI image data could be acquired a standard CE-MRI should be used to trace the tumor. The success of this technique depends on its ability to demonstrate quantitative differences of contrast medium behavior in a variety of tissues. In particular, malignant tumors exhibit an increased vascularity, since they begin to grow their own blood supply network. For this reason when the contrast agent is distributed, malignant masses enhance faster. This led to the development of models of contrast uptake as is illustrated in Figure 2.

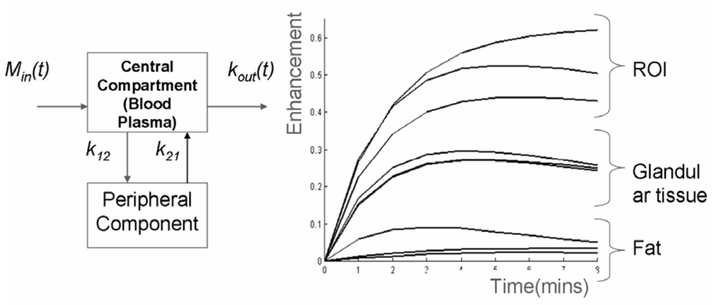


Fig. 2. A two-compartment pharmacokinetic model with typical contrast curves for fat, parenchymal (glandular) tissue and enhancing regions of interest. M_{in} is the mass of contrast injected into the blood stream with respect to time. k_{12} and k_{21} are inter-compartment exchange rates and k_{out} is the leaving contrast rate.

This technique can be misleading on several conditions. It is common knowledge that the contrast media perfuse over time in tissue. There is no evidence that after a specific time interval the contrast agent perfuse by tumor tissue exclusively. In addition, it has been observed that tumor tissue 'liquifies' after treatment due to destruction of the protein structure. Quantitative MRI (qMRI) techniques can actually quantify this alteration of the tissue structure by measuring its physical parameters however this technique isn't routinely used in clinical practice.

In conclusion, there are several limitations in extracting actual pathophysiological information from medical images. Therefore, in order to build robust models it is essential to carefully chose the imaging modalities and use as much complementary information as possible. However, the information extraction process is difficult to generalize and specific tools should be combined when different modalities are used (e.g. the two-compartment pharmacokinetic model for segmentation shown in Fig. 2). In the case of cancer modeling, several scientists have suggested the inclusion of qMRI techniques in the clinical protocols associated with the validation of insilico models since the alteration of the tissue structure, detected by MRI, could shed light in the accurate delineation of the tumor especially in areas where DCE technique is unreliable (e.g. the periphery of the tumor). As an example of possible multimodal information fusion, the combination of DCE-MRI and parametric qMRI could provide more accurate and precise tumor information, which would serve as a basis for validating insilico models within a higher level confidence.

2.2 Geometrical Normalization

It is necessary to ensure that a point correspondence is computed between different images. This is a typical problem in the case of breast imaging owing to the differences in breast shape/compression but also in newer applications such as molecular imaging and microarray imaging. Non-rigid alignment or registration is required to compensate for such differences. This problem does not only pertain to the multi-modal scenaria, as a temporal acquisition of the same modality will still likely involve registration in order to facilitate comparison. Several registration frameworks have been proposed (see [3-5] for selected publications in the field), traditionally for medical imaging applications but more recently also for correcting time dependent geometries in 2D molecular optical imaging studies [6].

2.3 Intensity Normalization

As stated above each biomedical measurement in essence disguises the true physiological property due to the image formation process. The non-linearities introduced by varying imaging conditions may alter significantly the image-intensity profile and reduce the efficiency of generic analysis algorithms. An interesting example is the model of Highnam and Brady [9] for mammogram image normalisation that eliminates variations related to imaging conditions (e.g. tube voltage, time of exposure, etc). Highnam and Brady’s method estimates – in millimetres – the amount of interesting tissue in each pixel column. This effectively provides objective quantitative information about the breast anatomy. If, for example, the separation between the Lucite plates is 6.5cm, the amount of interesting tissue at a location (x,y) might be 4.75cm, implying 1.75cm of fat. This way, the algorithm estimates, and then eliminates the effects of, the particular parameters that were used to form the mammographic image providing true anatomical information.

This problem also exists in microarray imaging technologies where several non-linearities in the experimental process render the measured expression values prone to variability and often, to poor reproducibility. To achieve normalization, one has to adjust the sensitivity of detection (photomultiplier voltage with fluorescence or exposure time with radioactivity) so that the measurements occupy the same dynamic range in the detector and exploit the fact that the gene expression values (e.g. from the

Cy3 and Cy5 matrices), should ideally follow a linear trend [10]. The later can be performed to all the genes or to a ‘ground truth’ subset that is known a priori to be the same in both channels (Cy3 and Cy5).

2.4 Visualisation

This is a fundamental aspect of biomedical data information fusion that is typically less well addressed in the literature, but which can dramatically increase the clinical utility of a solution if implemented intelligently. The effectiveness of visualisation depends very strongly on how clearly different indicators can be extracted from data and therefore segmentation is of utmost importance. A great deal of effort has been made in this research to produce visualizations of the results of temporal and multi-modal image fusion that optimizes the presentation of available clinical information (an example is illustrated in Figure 3).

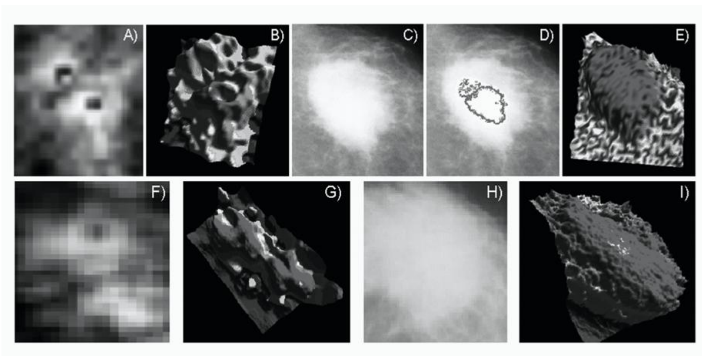


Fig. 3. A-D and F-I: Multimodal (MRI and X-ray) 3D visualization of 2D medical images better highlights the necrotic centers of each tumour

3 Conclusions

In order to approach the vision of the Virtual Physiological Human, it will be essential to develop and validate individualized, multi-level models taking into consideration pathophysiological information at all scales. As discussed in this paper, the extraction of useful anatomical and physiological information from biomedical measurements isn’t a trivial task due to the complex physical interactions involved in each acquisition as well as several systematic and random errors involved in the process. In many cases, the problems that arise in different scales are common (e.g. geometrical inconsistencies over time) and is therefore important to develop generic tools for multi-scale temporal analysis in order to robustly extract and visualize pathophysiological information over time. Such information is crucial for initializing (i.e. in the case of in silico models of cancer, 3D voxels should be classified as ‘proliferating’, ‘necrotic’, etc.), inspiring and validating 4D models of human function.

References

1. Behrenbruch, C.P., Marias, K., Armitage, P.A., Yam, M., Moore, N.R., English, R.E., Clarke, P.J., Leong, F.J., Sir Brady, J.M.: Fusion of contrast-enhanced breast MR and mammographic imaging data. British Journal of Radiology 77, 201–208 (2004)
2. Behrenbruch, C.P., Marias, K., Armitage, P.A., Yam, M., Moore, N., English, R.E., Clarke, P.J., Brady, M.: Fusion of Contrast-Enhanced Breast MR and Mammographic Imaging Data. In: Medical Image Analysis (MedIA), vol. 7(3), pp. 311–340. Elsevier, Amsterdam (2003)
3. Guimond, A., Roche, A., Ayache, N., Meunier, J.: Multimodal Brain Warping Using the Demons Algorithm and Adaptative Intensity Corrections. IEEE Transaction on Medical Imaging 20(1), 58–69 (2001)
4. McInerney, T., Terzopolous, D.: Deformable models in medical image analysis: a survey. In: Medical Image Analysis, pp. 91–108. Oxford University Press, New York (1996)
5. Penney, G.P., Weese, J., Little, J.A., Hill, D.L.G., Hawkes, D.J.: A comparison of similarity measures for use in 2D-3D medical image registrtation. IEEE Transactions in Medical Imaging 17, 586–595 (1998)
6. Marias, K., Ripoll, J., Meyer, H., Ntziachristos, V., Orphanoudakis, S.: Image Analysis for Assessing Molecular Activity Changes in Time-Dependent Geometries. IEEE Transactions on Medical Imaging, Special issue on Molecular Imaging 24(7) (July 2005)
7. Margaritis, T., Marias, K., Kafetzopoulos, D.: Improved Microarray Spot Segmentation by Combining two Information Channels. In: proceedings of the 2006 IEEE Engineering in Medicine and Biology Society (EMBS) Annual International Conference, IEEE. 2006, New York, USA (2006)
8. Bhujwalla, Z.M., Artemov, D., Glockner, J.: Tumor angiogenesis, vascularization,and contrast-enhanced magnetic resonance imaging. Top Magn Reson Imaging 10, 92–103 (1999)
9. Highnam, R.P., Brady, J.M.: Mammographic Image Analysis. Kluwer Academic Publishers, Dordrecht (1999)
10. Chen, Y., Dougherty, E.R., Bittner, M.L.: Ratio-based decisions and the quantitative analysis of cDNA microarray images. J. Biomed. Optics 2, 364–374 (1997)

Clinical Patient Safety—Achieving High Reliability in a Complex System

Kathryn Rapala and Julie Cowan Novak

Purdue University, School of Nursing, 502 N. University,
West Lafayette, IN 47907 USA
{Krapala,novakj}@purdue.edu

Abstract. Since the 2001 Institute of Medicine Report which estimated that 44,000 to 98,000 patients die each year as a result of healthcare error. This report in effect launched a global patient safety movement, with many proposed regulatory, research and administrative solutions. Patient safety areas of focus such as work complexity, teamwork and communication, technology, and evidence based practice provide a basis for understanding healthcare error. Reliability concepts are the goal of healthcare organizations; and applications such as simulation theory provide means to achieve this status. The translation of research into practice is the foundation of organizational patient safety. Understanding and awareness of patient safety issues has increased; however, significant work to improve patient care outcomes remains.

Keywords: patient safety, culture, work complexity, human factors, education, root cause analysis, failure mode event analysis, simulation, high reliability organization.

1 Introduction

The landmark 2001 Institute of Medicine report *To Err Is Human: Building a Safer Health System*, estimated that from 44,000 to 98,000 patient die every year due to healthcare error. (1) This attention has driven a variety of important patient safety related solutions that include regulatory requirements, error reporting systems, and technology. Five years after this report, improvements, particularly in the area of regulatory standards, have been achieved. However, significant barriers remain. These barriers include complexity, individual, professional autonomy, fear, hierarchical authority structure, diffuse accountability, lack of leadership, difficulty of measuring patient safety and reimbursement structure. (2) This paper will explore the current state of patient safety knowledge, and identify challenges and a direction for future multidisciplinary patient safety research, design, and collaboration.

2 Patient Safety Areas of Practice

To create a reliable healthcare environment, an understanding of theoretical patient safety is necessary to understand, interpret and formulate patient safety solutions. Patient safety areas of practice may be grouped in the following broad categories.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 710–716, 2007.

2.1 Work Complexity

Understanding the work environment is a key to understanding patient care error. Work complexity, which is really the study of the healthcare environment from the frontline perspective, is related to human factors engineering. The healthcare environment is highly complex, involving multiple people, products, equipment, interruptions and patients. The needs and conditions of patients may change dramatically within a short period of time. Researchers in nursing and other disciplines have determined that this complex, rapidly changing and unpredictable environment causes a series of small latent failures or gaps, which may cause or contribute to patient care error. (3) For example, a surge in patient admissions increases the workload of a nurse. Unless the surge is addressed by staffing, the nurse must juggle additional information and a critical piece of information may be missed. Human factors concepts, such as the impact of labeling on healthcare provider performance, are an integral part of understanding the work environment.

2.2 Technology

While technology has provided many solutions to patient safety, it may also provide unintended consequences. (4) Technology is only as good as the culture in which the technology is used. Technology must be developed, implemented and analyzed with the assistance of the frontline staff that uses the particular technology. The work environment must be evaluated prior to technology implementation.

2.3 Communication and Teamwork

Collaborative relationships between healthcare providers are key to understanding error providing good patient outcomes. The JCAHO Sentinel Event database identifies communication as a root cause in most serious reported healthcare error regardless of the type of error. The Health Work Environment Standards by AACN are evidence based interventions based on the “Silence Kills” study. This study showed that in most healthcare environments, staff members do not speak out regarding patient safety issues for a variety of reasons. (5)

2.4 Evidence Based Practice

Evidence based practice is defined as the conscientious use of current best evidence in making decisions about patient care”. (6) Many decisions made in healthcare are not based on evidence, even when evidence exists. (7) This failure to translate research into practice may be the result of difficulty translating evidence into practice, or the result of leadership style. Evidence based practice should be required for every healthcare decision.

3 Reliability and Healthcare

Reliability is the measurable capability of a process or procedure to perform its intended function in the required time under commonly and uncommonly occurring conditions (8). High reliability organizations are those that, even with highly complex and dangerous activities, have very few errors. Examples of high reliability organizations include nuclear power plants, aircraft carriers, and air traffic control systems. High reliability organizations exhibit harmony, balance and a sense of purpose with all aspects of the given process or product to produce a reliable result.

Healthcare is even more complex than the examples listed above. An aircraft carrier focuses on landing one aircraft at a time. Manufacturers are able to standardize parts and processes on the assembly line. In contrast, healthcare practitioners manage the arrivals, transfers, and departures of highly individual patients with very different needs on any given day. When the individual practitioner practice, process, and products are factored into patient needs, the result is fantastically complex care. The challenge for all of us, as healthcare professionals, is to convert healthcare into high reliability organizations.

It is one thing to study high reliability organizations, but quite another to introduce reliability into practice. There are four elements present in every high reliability organization. These elements are:

- Systems, structures, and procedures conducive to safety and reliability are in place
- A culture of safety permeates the organization
- Safety and reliability are examined prospectively for all the organization's activities; organizational learning by retrospective analysis of accidents and incidents is aggressively pursued
- Intensive training of personnel and teams takes place during routine operations, drills and simulations (9)Aspects of all these elements are present

in every organization, service and academic. Healthcare organizations wishing to decrease error need to address every element in depth, implementing the latest evidence in each area.

3.1 Systems, Structures, and Procedures Conducive to Safety and Reliability Are in Place

To provide a safe environment of care for both staff and patient, certain factors must be present. Systems, such as an electronic medical record or the way drugs are distributed in pharmacy, are necessary to run the hospital efficiently. Physical structures, such as the needed equipment, the physical environment or software, are components of the reliable workplace. Additionally, the providers' infrastructure, or organizational structure, is another example of structure needed. Finally policies and procedures for safety—from policies on the actual procedure to regulatory policies are necessary. Some regulatory bodies, such as the Joint Commission for Healthcare Accreditation (JCAHO), as well as government agencies, require baseline system, structures and procedure to ensure patient safety. The JCAHO National Patient Safety Goals are an example of a regulatory required procedure. (10)

3.2 A Culture of Safety Permeates the Organization

A "culture of safety" is a term used to describe a working environment where staff members feel comfortable reporting errors and mistakes without fear of retribution, where staff may question those in higher authority when a patient safety question arises, and where management values and acts on patient safety information from any source. In a culture of safety, rather than blame the individual practitioner who may have given the wrong medication or procedure, patient care error is viewed as a systems issue. Many in health care fear that non-punitive reporting allows a substandard staff member to stay in the system without taking responsibility. As a result, researchers such as James Reason have suggested the concept of a just culture, where the error is analyzed from both a systems and a personal accountability perspective. There are several surveys, instruments and tools designed to measure a culture of safety.

3.3 Safety and Reliability Are Examined Prospectively for All the Organization's Activities; Organizational Learning by Retrospective Analysis of Accidents and Incidents is Aggressively Pursued

Borrowing from the world of industry, tools such as failure modes event analysis (FMEA) and root cause analysis are used by healthcare providers to learn from error. Although required by some accreditation organizations such as JCAHO, there is a great degree of variability in the application of these tools to the healthcare setting and to what degree the healthcare organization uses these tools. Additionally, the learning from healthcare error is often difficult to disseminate within the organization and to other healthcare providers. The JCAHO Sentinel Event program and various state reporting systems are methods that accrediting bodies and the government have chosen to assist healthcare organizations in sharing and disseminating healthcare error information. (11)

3.4 Intensive Training of Personnel and Teams Takes Place During Routine Operations, Drills and Simulations

Training and simulation has a strong evidence base in healthcare. Like the airline industries that put their pilots and crews through teamwork and simulation exercises, similar training and simulation exercises have been developed in some regions of the nation. There are three types of simulation and training. The first is using equipment or props to perform a task. Examples are anchoring a foley catheter, mock codes or starting an intravenous line. A second type of simulation is a simulation of a particular procedure or scenario, using high tech simulated patients. These patients may be used in a variety of circumstances and scenarios, such as performing a cardiac catheterization or simulating a process such as anesthesia in the operating room. This type of simulation may be done individually or with teams. Lastly, simulations may be computer based. Much like video games, these simulations can teach the operation of equipment or any number of processes or procedures.

Simulation and training is important to both professionals and students. Moving from the old model of "see one, do one", simulation and training provides learning and practice on virtual patients rather than very real patients. Additionally, the concept of teamwork training is a concept very important to healthcare. Sentinel event data, both internal and external, indicates that communication is a root cause of almost every adverse event. Learning how to communicate and work in teams not only raises awareness, but also teaches essential skills. Simulation and teamwork training also can mimic work complexity. Traditional training teaches a perfect procedure in a perfect environment, which happens infrequently in the real environment.

Simulation and training provides a basis for service and academic research. Simulation and training were adopted primarily through comparisons to military and airline training—high-risk technical operations. Although simulation is used by academia and service to increase performance, there is not a great deal of evidence in the literature to show outcome measures.

4 Designing and Implementing Patient Safety Programs

4.1 Leadership

Leadership is a requirement of an effective patient safety program. Diffuse accountability has been cited as a barrier to patient safety solutions. Programs such as the Institute for Healthcare Improvement 5 Million Lives Campaign have targeted governance as a key element in the success of patient safety.

An excellent example of nursing leadership that produces quality outcomes is the American Nurses Credentialing Center (ANCC) Magnet Recognition Program®. This program was developed by to recognize health care organizations that provide nursing excellence. A Magnet facility is marked by a professional environment guided by a strong visionary nursing leader who, through shared governance with the staff, creates an environment that embodies quality and patient safety. (12)

4.2 Framework

Each healthcare organization or should have a framework or plan that not only describes patient safety in that organization, but also provides measurable goals, objectives, measurement and evaluation. The patient safety framework or plan should include not only regulatory goals, but also areas of priority for that particular organization. Due to a lack of standardization and many patient safety approaches, not every patient safety plan will look the same and some organizations do not have such a plan. However, to make patient safety improvements, the identification of baseline status and direction is essential.

4.3 Performance Improvement and Outcome Measurement

Healthcare organizations and providers may choose several different methods, most based on industrial applications to execute patient safety performance improvement activities. Common performance improvement tools include the Plan Do Study Act method, lean 6 sigma and 6 sigma. Culture assessment is another common way to measure safety.

Outcome measurement is a challenge for healthcare organizations. Benchmarking or comparison data is not readily available due to litigation concerns in regard to patient error. Thus, there is no standard for a medication error rate, for example. Organizations use statistical analysis such as statistical control charts to track error internally.

Denominators are frequently an issue in obtaining rates in patient care error. Rates per 1000 patient days are commonly used as some indication of unit or hospital activity.

Benchmarks and comparative data are emerging. The ANCC Magnet Program has comparison data available for falls and skin problems; regional and professional collaborative also have comparison data. There are some studies in the area of staffing and patient care error. Electronic medical records are providing information and simultaneous data collection for some patient safety issues.

5 Education

The IOM noted in 2001 that educational institutions, among other stakeholders, must improve the environment to prepare the workforce for a complex health environment, and recommended a multidisciplinary summit review the state of education in light of the current complex health environment. (IOM (2001) p.5) The *Health Professions Education: a Bridge to Quality* was published in 2003. The report noted "[a]lthough there is a spotlight on the serious mismatch between what we know to be good quality care and the care that is actually delivered, students and health professionals have few opportunities to avail themselves of coursework and other educational interventions that would aide them in analyzing the root causes of errors and other quality problems and in designing system-wide fixes." (13) The Purdue vision was a practice doctoral program that would address the complexity of the health-care system; the information, knowledge and technology explosion; spiraling costs; and the need for a systems approach to create new models of care and solve existing health-care dilemmas. Launched in 2003, the Purdue Doctor of Nursing Practice Program is an example of educational design that has incorporated patient safety and systems thinking into the curriculum.

6 Challenges for the Future

Although much progress has been made in patient safety, challenges remain. Following is a summary of significant challenges:

- Utilizing patient safety concepts, such as work complexity and human factors, to plan, design and redesign healthcare and healthcare products, equipment and education.
- Working with industries including but not limited to healthcare in order to provide quick implementation of safety methods and research.
- Providing, in conjunction with healthcare and industry, tools and methods that can be standardized to measure patient safety outcomes.
- Disseminating evidence based practice and research quickly and efficiently.
- Utilizing evidence based practice to make educational, administrative, and technical healthcare decisions.
- Integrate high reliability concepts into healthcare administration, education and product design.

References

1. Institute of Medicine, Crossing the quality channel—a new health system for the 21st century, National Academy Press, Washington, DC (2001)
2. Leape, L., Berwick, D.: Five years after To Err is Human: What have we learned? Journal of the American Medical Association 293(19), 2384–2390 (2005)
3. Ebright, P., Patterson, E., Chalko, B., Render, C.: Understanding the complexity of registered nurse work in acute care settings. Journal of Nursing Administration 33(12), 630–638 (2003)
4. Ash, J.S., Berg, M., Coiera, E.: Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. Journal of American Medical Informatics Association 11, 104–112 (2004)
5. American Association of Critical Care Nurses. Healthy Work Environment (2007) (Accessed February 28, 2007) Available at: http://www.aacn.org/AACN/ hwe.nsf/vwdoc/ HWEHomePage
6. Melnyk, B., Fineout-Overholt, E.: 20050.Evidence-based practice in nursing & healthcare: A guide to best practice. Lippencott, Philadelphia
7. Pfeffer, J., Sutton, R.: Evidence based management. Harvard Business Review, 63–74 (2006)
8. Berwick, D., Nolan, T.: High reliability healthcare (2003) (Accessed February 28, 2007) Available at: http://www.ihi.org/IHI/Topics/Reliability/
9. Gaba, D.: Safety first: Ensuring quality care in the intensely productive environment—The HRO model. Anesthesia Patient Safety Foundation Newsletter 18(1), 1–16 (2003)
10. The Joint Commission. National Patient Safety Goals (2007) (Accessed February 28, 2007) Available at: http://www.jointcommission.org/ PatientSafety/ NationalPatient Safety Goals/
11. The Joint Commission. Sentinel Event (2007) (Accessed February 28, 2007) Available at: http://www.jointcommission.org/SentinelEvents/
12. American Nurses Credentialing Center (ANCC) Magnet Recognition Program (2007) (Accessed February 28, 2007) Available at: http://www.nursingworld.org/ancc/
13. Institute of Medicine. Greiner, A., Knebel, E.: Health professions education: a bridge to quality, National Academy Press, Washington, DC (2003)

Novel Methods for Human-Computer Interaction in Multimodal and Multidimensional Noninvasive Medical Imaging

Tomasz Soltysinski

Institute of Precision and Biomedical Engineering, Warsaw University of Technology, Sw. A. Boboli 8, 02-525 Warsaw, Poland
solek@mchtr.pw.edu.pl

Abstract. The newly developed method for medical noisy data segmentation for the purpose of presentation and supporting the diagnostics is introduced. It also allows for morphometry and visualization of medical multimodal and dynamical data. A general mathematical framework is presented and characterized together with numerous applications. As this tool is designed to support human-computer interaction by means of involving the sense of sight, and suspected to be worthy in the virtual environment sensitive to the sense of touch, the discussion is supported with numerous examples of visualizations and multimodal and multidimensional applications of proposed method.

Keywords: Digital human modeling, Bayesian inference, spectral method, multidimensional imaging, ultrasound, CT, MRI, PET, histopathology.

1 Introduction

Medical imaging is one of the most important fields of present days science. This field is based on numerous achievements of mankind technology and science done in recent decades. It requires fast computers, modern high quality materials and specialized electronic circuits of high speed together with sophisticated software. However, the human factor is still the crucial point of the diagnostic process. Even if the equipment is sophisticated, the decision that physician makes is based on his experience and knowledge as well as the information which he gains during the medical imaging.

Very effective and fast imaging equipment provide enormous amount of data, often in distributed manner. This data covers dynamical time series of signals, complimentary spatial data of images and volumes, as well as some other data that is related to individual patient history and condition. All the data is stored by computers and these machines are responsible for its presentation to the physicians, researchers and the patients as well. This indicates the strong need for methods of data presentation, analysis and visualization.

Throughout this study, some new trends are presented and methods applied to different modalities of medical imaging. This is followed by discussion of the

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 717–726, 2007.

possible ways of data combination and presentation to the end users of modern complex and synergistic medical systems.

2 Methods

The main point is to focus on multimodal imaging and visualization of data in 3-, 4-, or N-dimensional space. Combining this methodology with high level math and data presentation technics is believed to broaden the view on medical diagnostic process.
A special attention is paid to combination of multimodal data by means of Bayesian inference. Medical imaging is a process that transforms the information about the body state into some virtual numerical space. This transformation depends on many circumstances and parameters, hence the final result is known only with some probability dependent on these parameters. Some models may also be adapted into the imaging, however their incorporation requires to use the terms of probability as well. Then, the imaging is just the manipulation of probabilities and the search for a strategy that maximizes the probability that the correct and the most relevant information is obtained.

The proposed method is based on Bayesian inference that provides the information about edges of the investigated structure or organ. This information is fuzzy and uncertain and forms radial function of initial guess on edge position, called edge map. This edge map is used as initial data for spectral method that reveals the final radial function describing the most probable contour of the structure. Important point of proposed method is the incorporation of priors. The a priori knowledge is provided by patterns or probability distributions into the inference. The priors are constructed of denoised and decomposed data. The decomposition is done in multiscale fashion, by means of a trous, B3-spline wavelet based decomposition of images. The priors are also constructed from analytical functions when necessary, if these functions may be guessed or approximated. Application of Bayesian inference and the usage of priors handles with the cases that are hard to analyze and delineate, like those with high level of noise or distortion.

Multidimensional Segmentation

The proposed method is a new approach to image segmentation that combines the speed of spectral method in contour detection [5] and the Bayesian inference that allows for the most probable estimation on initial edge map. The method has been successfully applied to ultrasonic data and CT brain data with aneurysm and is described in details in [3,4]. The real contour of investigated and segmented organ [4] is extracted in the iterative process of solving the nonlinear partial differential equation (PDE). This step is realized in Fourier space by fast spectral method. PDE is approaching the function that reveals the real edge of the object and starts from an initially guessed edge map. This iterative process is controlled by two parameters that describe fidelity of reconstruction, one, μ is steering the PDE and the other is responsible for slight smoothing of the resolved subsequent partial approximations of the final solution.

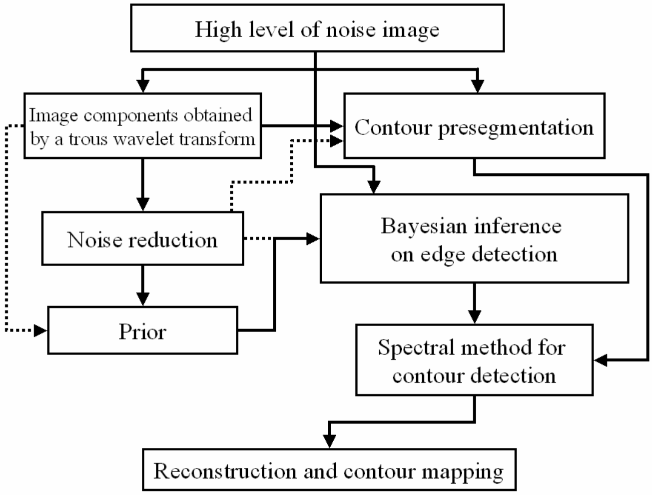


Fig. 1. Flow diagram for proposed Bayesian constrained spectral method. Solid lines mark necessary steps in the algorithm while dotted lines optional paths.

Contouring may be expressed as partial differential equation (PDE). This is the approach commonly found in methods based on active contours that are iteratively approaching the final contour. Following the work presented in [5] the final contour f may be found by solving the elliptic equation of Helmholtz type

$$\nabla^2 f - \mu(f - g) = 0 \tag{1}$$

This equation uses known variable which is initially guessed edge map g. It is solved in spherical coordinates. Moving the g term to right hand side and relating it to a previously found value of f, called f_n the PDE can be expressed in linearized form

$$\alpha\nabla^2 f_{n+1} - f_{n+1} = g_{f_n} \tag{2}$$

Such an equation is further easily solved by the fast spectral method. Applying $\alpha=1/\mu$ the solution may be controlled by value of μ.

Bayesian Inference in Tissue and Shape Selection

The edge map g is determined by Bayesian inference on edge placement in image data. Let $P(E_i/I)$ denote the required probability of the most appropriate edge in our existing data set. This is the conditional probability as it depends on the contents of I. $P(E_i/I)$ is the probability of the fact that the I point belongs to the edge class E_i, knowing the value of intensity of this point. Let $P(I/E_i)$ be a probability of how much the value or intensity of a point is depending on edge class E_i. This term serves as a kernel. $P(E_i)$ is simply the probability of existence of the edge class E_i among all other detected edge classes. Edge class is a set of some subsequent pixel intensity values. Then the required probability can be found by solving the Bayes rule:

$$P(E_i / I) = \frac{P(I / E_i)P(E_i)}{P(I)} = \frac{P(I / E_i)P(E_i)}{\sum_i P(I / E_i)P(E_i)} \tag{3}$$

P(I) is a sum of all probabilities $P(I/E_i)$ weighted by $P(E_i)$ and thus remaining constant. *P(I)* is only a normalizing term and can be excluded from further analysis. The standard way of solving Bayes equation is the maximization of the right side over the parameter E_i (maximum likelihood, *ML* step) and then maximization of the found solution over all accessible data (maximum-a-posteriori, *MAP* step). The $P(E_i)$ is a prior and contains a priori knowledge.

In practice the $P(I/E_i)$ is estimated from the histogram of *I* along given radius from centroid to a point of circumference of some bounding box selecting the region of interest (*ROI*). The histogram is shrank in such a way that each bin is equal to the edge class size assuming that each expected edge class covers the same number of intensity levels. Calculating the normalized probability over all edges E_i, *ML* step is performed and the most probable edge in *I* is estimated. Then the *MAP* is done by searching for maximum over the data itself, and usually the first maximum in $P(I/E_i)$ is detected as an edge. $P(E_i)$ may be a constant value if we assume all edge classes as equally probable or may be distributed uniquely according to the prior knowledge. From *P(E/I)* the polar edge map, $g(\theta)$ is derived.

Spectral Methods

The linearized equation 2 is solved by a spectral method [5]. Adapting Cheong's method [11] the Laplacian operator ∇^2 on the unit circle is expressed as follows:

$$\nabla^2 \frac{1}{\sin\theta}\frac{\delta}{\delta\theta}\left(\sin\theta\frac{\delta}{\delta\theta}\right) \tag{4}$$

Both functions, *f* and *g* are defined on the computational grid (θ_i), where $\theta_i = \pi(j+0.5)/J$. *J* is the number of points along the longitude of unit circle's circumference high enough to engage all points covered by *g*. Each point in *g* may be now expressed by its discrete cosine transform (DCT) representation yielding

$$g(\theta_i) = \sum_{n=0}^{J-1} g_n \cos n\theta_i \tag{5}$$

with g_n being simply the coefficients of discrete cosine transform. Applying equation 4 into equation 1, it may be written as an ordinary differential equation (ODE):

$$\frac{1}{\sin\theta}\frac{\delta}{\delta\theta}\left(\sin\theta\frac{\delta}{\delta\theta}f(\theta)\right) = \mu[f(\theta) - g(\theta)] \tag{6}$$

which yields an algebraic system of equations in Fourier space:

$$p_{n-2}f_{n-2} - p_n f_n + p_{n+2}f_{n+2} = \mu\left[\frac{1}{4}g_{n-2} - \frac{1}{2}g_n + \frac{1}{4}g_{n+2}\right] \tag{7}$$

where

$$p_{n-2} = \frac{(n-1)(n-2)+\mu}{4},$$
$$p_n = \frac{n^2+\mu}{2}, \qquad p_{n+2} = \frac{(n+1)(n+2)+\mu}{4} \tag{8}$$

after substitution of eq. 5 into eq. 6 and expressing f in the same way as g. The index $n=1,3,\ldots J-1$ for odd n and $n=0,2,..,J-2$ for even n. The system of equation 7 may be now expressed as a double matrix equation:

$$\mathbf{B}_e\hat{h}_e=\mathbf{A}_e\hat{g}_{e,} \qquad \mathbf{B}_o\hat{h}_o= \mathbf{A}_o\hat{g}_o \tag{9}$$

with subscripts e for even and o for odd n, $\hat{h}$ and $\hat{g}$ denote the column vector of expansion coefficients of $f(\theta)$ and $g(\theta)$, respectively. **B** is a tridiagonal matrix containing the left side of equation 7 and **A** is tridiagonal matrix with constant coefficients along each diagonal corresponding to right side of eq. 7.

The calculated set of expansion coefficients f_{n+1} serves for the reconstruction of f_i the representation of g on the certain level of approximation i. Smoothing and summing all partial functions f_i the required smooth approximation to g is recovered revealing the most probable edge map.

Wavelet *a trous* Analysis

It may be summarized as follows:

1. Initialize j, the scale index, to 0, starting with an image $c_{j,k}$ where k ranges over all pixels.
2. Carry out a discrete convolution of the data $c_{j,k}$ using a wavelet filter, yielding $c_{j+1,k}$. The convolution is an interlaced one, where the filter's pixel values have a gap (growing with level j) between them of 2^j pixels, giving rise to the name *a trous* – with holes.
3. From this smoothing it is obtained the discrete wavelet transform, $w_{j+1,k}=c_{j,k}-c_{j+1,k}$.
4. If j is less than the number J of resolution levels wanted, then increment j and return to step 2.

The original image is reconstructed by the summation of all $w_{j,k}$ and the smoothed array $c_{J,k}$, where J is the maximum that j may reach.

For the purpose of this study, further in the paper, the reversed notation is used that, as it is believed, makes the analysis more intuitive and convenient. The smoothed component (originally $c_{J,k}$) has index 0 and all subsequent components, with increasing level of details have growing indexes. Hence, level $w_{1,k}$ (originally $w_{J,k}$)

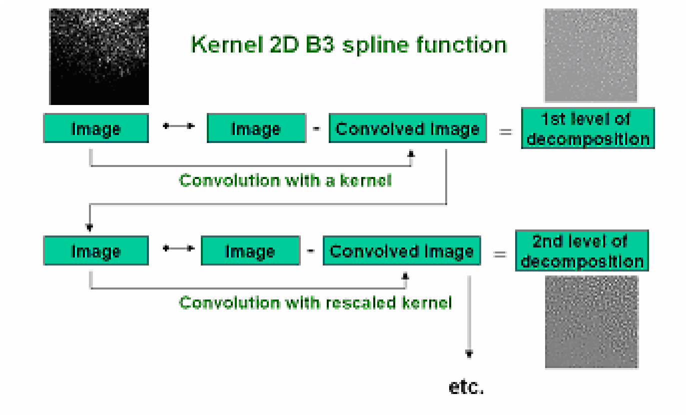


Fig. 2. Scheme that describes *a trous* wavelet decomposition algorithm

with index 1 has low details but of higher resolution than the base level with index 0, level $w_{2,k}$ contains details with higher resolution than previous level but lower than these ones at next level, etc.

3 Practical Applications for Human-Computer Interaction

Hierarchical, Dynamical and Structural Modeling

The method has been tested in many scales from cell level through internal organs to whole body. This makes it a universal tool for segmentation of a human tissue and organs of body in hierarchical or dynamical digital models. Another advantage of the method is arbitrary scaling of surface of visualized object and compression of information which is necessary for object's description. Applications of this approach are shown for delineation of heart left ventricle walls in dynamic ultrasonic imaging at high level of speckle noise (figs. 3,4) together with derivation of unique proposal for dynamical analysis based on time-contour function plane (fig. 5) covering the total information about dynamics of shape of heart left ventricle; morphometry and visualization of upper part of a thigh bone (figs. 6,7) compared to method based on pixel counting; delineation of active metabolic region of liver in PET-CT multimodal

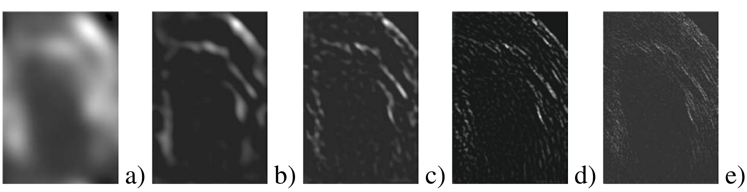


Fig. 3. Multiscale *a trous* decomposition of heart left ventricle ultrasonic image of one frame of cardiac cycle. Scales of decomposition, from grain level (a) to finer scales (b,c,d,e).

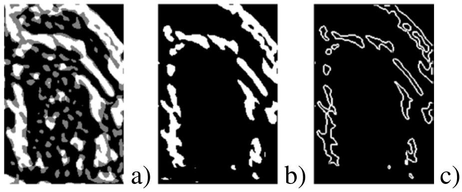


Fig. 4. Construction of enhanced prior for improved strategy in Bayesian inference based on multiresolution support (MRS), a technics based on *a trous* multiscale decomposition. Original image of MRS (a), filtered image (b) and the same image with edges extracted by Sobel filtration (c), the final prior used in inference.

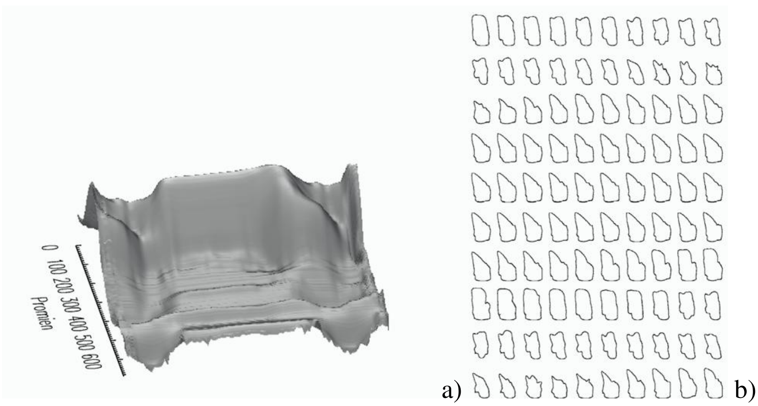


Fig. 5. Time-contour radial function space (a). The dynamics of mitral valve region is visible. All the cycle contours derived in fully automatic manner from raw, ultrasonic image data (b).

Fig. 6. Comparison of contours in 3D structure of a piece of thigh bone obtained by Sobel edge detection (a) and proposed method based on Bayesian inference (b)

parallel imaging (figs. 8,9) with the applications of complex priors for Bayesian inference; morphometry of structural and functional regions of brain tumor tissue in subsequent MRI-PET whole head scans (fig. 10) and in cells morphometry for angiogenesis research (fig. 11).

Fig. 7. Examples of the bone shown on fig. 6 reconstructed by proposed method and visualized in IDL environment

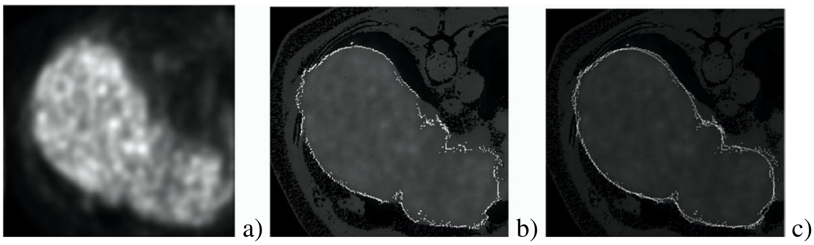


Fig. 8. The prior for Bayesian inference on the region of PET activity in human liver (a). Initial edge map revealed by Bayesian inference (b) and the resulting contour based on this map, returned by spectral method (c).

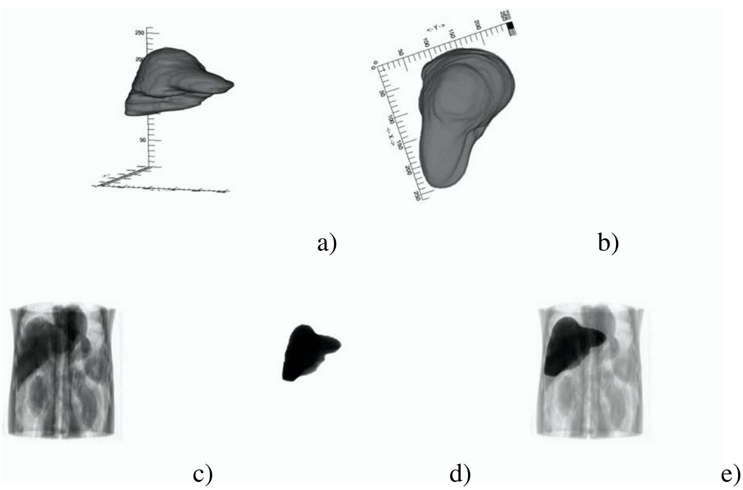


Fig. 9. Visualizations of 3D structure PET active region in human liver determined by proposed method (a,b). Structural 3D projection based on the CT volume (c), functional 3D projection of PET active region of liver obtained by proposed method (d) and the fusion of two projections (e).

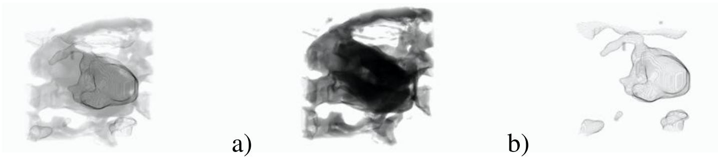


Fig. 10. Fusion (a) of structural MRI (b) and functional PET 3D volumes (c) limited to the tissue of brain cancer. The data was registered first by a method based on mutual information estimation.

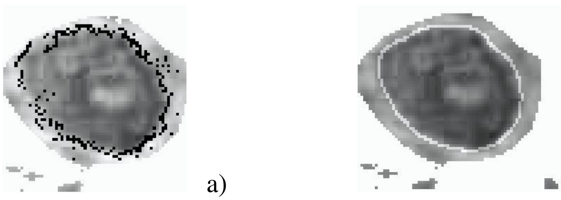


Fig. 11. Cell morphometry for angiogenesis research. Proposed method is applied to find out the edge map of a cell (a) and the final contour (b). Multiscale priors were used.

4 Discussion

The method is well suited for the analysis of noisy and disturbed data like that obtained by ultrasonography. Applications of the method to ultrasonic data is an alternative for existing methods [13][14] especially in the case of dynamical modeling. Moreover, it offers a unique opportunity for missing information reconstruction. This is the main advantage of the method in applications to multimodal data as well. High ability of mutual information recovering supports the medical diagnostic process based on human-computer interaction.

5 Conclusion

Presented work may serve as a general framework to develop useful and suitable graphical interface for supporting the diagnostic process, it helps to manage huge amount of information about human body and its functions, provides tools for virtual reality and digital human modeling.

Acknowledgments. This scientific work has been funded by resources of Science and High Education Ministry of Poland for years 2006-2007, in the frame of research grant number N518 023 31/1387.

References

1. Rogowska, J.: Overview of fundamentals of medical image segmentation. In: Bankman, w I.N. (ed.) Handbook of Medical Imaging, Processing and Analysis, pp. 69–86. Academic Press, San Diego (2000)
2. Bankman, I.N. (ed.): Handbook of Medical Imaging, Processing and Analysis. Academic Press, San Diego (2000)

3. Soltysinski, T.: Novel algorithm for active surface reconstruction by Bayesian constrained spectral method. In: Hozman, J., Kneppo, P. (eds.): IFMBE Proceedings, vol. 11. Prague: IFMBE, pp. 4191–4195 (2005)
4. Soltysinski, T., Kałużynski, K., Pałko, T.: Cardiac ventricle contour reconstruction in ultrasonographic images using Bayesian constrained spectral method. Lecture Notes in Artificial Intelligence Series. Springer, Heidelberg (in press, 2006)
5. Li, J., Hero, A.O.: A fast spectral method for active 3D shape reconstruction. Journal of Mathematical Imaging and Vision 20, 73–87 (2004)
6. Holschneider, M., Kronland-Martinet, R., Morlet, J., Tchamitchian, P.: A real time algorithm for signal analysis with the help of the wavelet transform. In: w Wavelets: Time-Frequency Methods and Phase-Space, pp. 286–297. Springer, Heidelberg (1989)
7. Soltysinski, T., Rodatus, A., Golnik, N., Pałko, T.: Comparison of multiscale entropy of healthy and cancer tissue imaged by optical polarimeter. In: IFMBE Proceedings, vol. 14 (2006)
8. Soltysinski, T.: Speckle Noise Removal By A Trous Decomposition And Threshold-limited Synthesis In Ultrasonic Cardiac Images. In: Proceedings of the conference Biosignal 2006, Brno Technical University, Czech Republic, vol. 18 (2006)
9. Gravel, P., Beaudoin, G., De Guise, J.A.: A method for modeling noise in medical images. IEEE Transactions on medical imaging 23(10), 1221–1232 (2004)
10. Starck, J.-L., Murtagh, F.: Astronomical image and data analysis. Springer, Berlin (2002)
11. Cheong, H.: Double Fourier series on a sphere: Applications to elliptic and vorticity equations. Journal of Computational Physics 157(1), 327–349 (2000)
12. Soltysinski, T.: Influence of multiscale denoising on Bayesian constrained spectral method in segmentation of noisy medical images. In: IFMBE Proceedings, vol. 14 (2006)
13. Dydenko, I., Friboulet, D., Gorce, J.M., D'Hooge, J., Bijnens, B., Magnin, I.E.: Towards ultrasound cardiac image segmentation based on the radiofrequency signal. Medical Image Analysis 7, 353–367 (2003)
14. Hammoude, A.: Endocardial border identification in two-dimensional echocardiographic images: review of methods. Computers in Medical Imaging and Graphics 32, 181–193 (1998)

A Hybrid AB-RBF Classifier for Surface Electromyography Classification

Rencheng Wang[1], Yiyong Yang[2], Xiao Hu[1], Fangfang Wu[1], Dewen Jin[1], Xiaohong Jia[1], Fang Li[1], and Jichuan Zhang[1]

[1] Division of Intelligent and Biomechanical System, State Key Laboratory of Tribology, Tsinghua University, Beijing, 100084, China
[2] School of Engineering and Technology, China University of Geosciences (Beijing), Beijing 100083, China
wangrc@mail.tsinghua.edu.cn

Abstract. In this paper, we aim to classify surface electromyography (sEMG) by using Attribute Bagging-Radial Basis Function (AB-RBF) hybrid classifier. Eight normally-limbed individuals were recruited to participate in the experiments. Each subject was instructed to perform six kinds of finger movements and each movement was repeated 50 times. Features were extracted using wavelet transform and used to train the RBF classifier and the AB-RBF hybrid classifier. The experiment results showed that AB-RBF hybrid classifier achieved higher discrimination accuracy and stability than single RBF classifier. It proves that integrating classifiers using random feature subsets is an effective method to improve the performance of the pattern recognition system.

Keywords: Attribute Bagging Algorithm, RBFN, sEMG, Prosthesis control.

1 Introduction

It is an ideal and bio-mimetic control method that multi-DOFs artificial limb is controlled by Surface Electromyography (sEMG). Many researches have been conducted at home and abroad [1][2]. Because of instability of the electromyography, while classifying sEMG, it is difficult to acquire satisfying classification accuracy by only one method; however, the shortage of one method can be compensated by another one. So we can decrease classification mistakes and strengthen classification robustness through integrating different classifiers. Therefore integrated classifiers have been the focuses of an increasing number of studies in recent years.

From this point, this paper proposes an AB-RBF hybrid classifier system based on Attribute-Bagging Algorithm (AB Algorithm), which is used to classify 6 kinds of finger movements.

2 AB Algorithm

AB Algorithm integrates classifiers using random attribute subsets [3], which is conceptually similar to Bagging Algorithm. In AB Algorithm, firstly it is given an original data set S, which consists of N-eigenvectors and *M* samples. Then the

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 727–735, 2007.

original data set S randomly yields a subset of n'-eigenvector ($n' < n$) S_n^1 used to train the ensemble of the classifiers. Repeat this process for T times independently and get T independent subsets of n'-eigenvector $S_n = \{S_n^1, S_n^2, \cdots, S_n^T\}$. The obtained subsets are used to train classifiers independently, and then rank the classifiers by the classification accuracy of the training data. Only the classifiers constructed on the highest ranking subsets (the number is L, $L<<T$) participate in the ensemble voting. While classifying, the L selected classifiers are used to vote and classify the samples respectively, then make a final decision according to the principle of plurality.

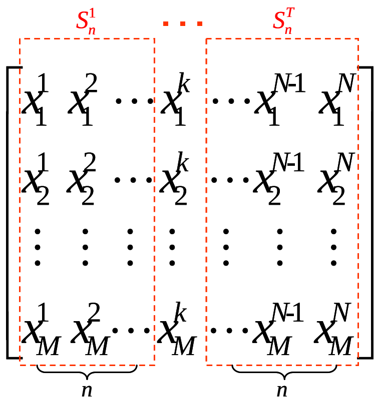


Fig. 1. The diagram of eigenvector subsets of AB algorithm

Using attribute subsets to integrate classifiers has many advantages [3][4]:

(1) Reduction of the dimensionality in the input sample, which lessens the impacts of the “curse of dimensionality” and raises the efficiency of training classifiers.

(2) The correlation among the classifiers is reduced by training them with different features and, as a result, the classification performance of the integrated classifiers is also improved.

(3) Because attribute subsets are the projection of the original feature and the total number of the training samples does not change, the recognition performance of each trained classifiers to the related kinds of samples is identical. In contrast with the AB Algorithm, the Bagging Algorithm only selects parts of the samples randomly, so the number of the trained samples may be disproportional.

3 Fulfillment of the sEMG Classification System Based on AB-RBF

3.1 Background

This sEMG identifying system aims to control an HIT skillful bionic hand developed by Harbin Institute of Technology (HIT), China. HIT hand has five fingers. The thumb has 2 joints; the other four fingers have 3 joints respectively. The thumb, the index finger and the middle finger can move independently and they are all driven by a electromotor respectively. The ring finger and the little finger are not driven by electromotors, but they can move along with middle finger through the connection rods and springs. In all, the hand has 14 joints, 3 independent DOFs, and the others are passive DOFs.

According to the function of human hands in daily life and the mechanism features of the HIT skillful hand, six patterns of the finger movements were discriminated in the experiment. These movements were Thumb Extension/Thumb Flexion (TE/TF), Two Fingers Pinch/ Two Fingers Extension (TFP/TEF), Power Grasp (PG) and Flatten Hand (FH).

sEMG data were collected by four active electrodes, after filtered and magnified, then transmitted to A/D DAQ card (UA303-DAQ). The frequency of data acquisition was 2,400 Hz. Four electrodes were applied to lightly abraded, degreased skin over the short extensor muscle of thumb, extensor muscle of fingers, superficial fleaor muscle of fingers and ulnar flexor muscle of wrist, which are most involved when executing the six patterns of finger movements.

Eight healthy and normal volunteers were recruited in the experiments, including four males and four females. Each subject was required to perform six movements aforementioned, and each movement lasted one minute and was repeated fifty times, at last we got three hundred groups of data. The data obtained from the earlier 30 times are used to train and the data obtained from the later 20 times are used to test. The results of the experiment represented the best performance that an amputee can achieve, because all the subjects are healthy.

3.2 Extraction of Eigenvectors of EMG

Wavelet transform has high resolution in time domain and frequency domain so that can perfectly represent the instability of EMG data. The processes of picking features of the wavelet transform are: firstly, get the continuous wavelet transforming coefficient of N points of EMG at measure a=1,…,5, as a result, we can get a N×5 coefficient matrix; then calculate the singular value of the coefficient matrix; finally integrate the singular values of the four-channel coefficient matrixes to get the wavelet decomposition eigenvectors of the N points of EMG.

According to Ref. [5], using four-order Coiflet Mother Wavelet to decompose wavelet has better precision than other orders and other kinds of wavelet. So we used Coif4 Mother Wavelet to decompose the continuous wavelet in the experiment. The length of EMG (N) was 256.

The definition of Cepstrum is following:

$$c(n) = Z^{-1}[\log|Z[x(n)]|] = Z^{-1}[\log|X(z)|] \tag{1}$$

When z takes a value on the unit circle, according to Ref. [6], the first P orders of Cepstrum coefficient can be got through iteration of the P orders of AR model coefficient:

$$\begin{cases} c_i = -a_i & i = 1 \\ c_i = -a_i - \sum_{n=1}^{i-1}(1-\frac{n}{i})a_n c_{i-n} & 1 < i \le p \end{cases} \tag{2}$$

In Eq. (2), p means the order of AR model, $a_i,\ i=1,\cdots,p$ are the coefficients of AR model, and $c_i,\ i=1,\cdots,p$ are the first p orders of cepstrum coefficients.

In this paper, first of all, we got the first p orders of cepstrum coefficients of the signals, and then integrated cepstrum coefficients of all channels to get a cepstrum eigenvector. In the experiment, it was set p=4. The single cepstrum eigenvector was obtained as following:

$$\mathbf{x}_{CEP}=[c_1^1,c_2^1,c_3^1,c_4^1,\cdots,c_1^4,c_2^4,c_3^4,c_4^4] \quad (3)$$

As Ref. [7] pointed out, the features of cepstrum coefficient reflect the spectrum difference of low frequency of signals. So, when compared with the feature of AR model, it has better performances of decomposition.

3.3 Fulfillment Steps of AB-RBF Hybrid Classifiers

The AB Algorithm is realized through two steps:

(1) Determine an optimal dimensionality n^* according to the accuracy of classifiers, which is trained by eigenvectors subsets of different dimensionality. n^* varies as situation changes, because it is affected by redundancy and correlation of the original eigenvectors.

(2) The L most accurate classifiers that are selected in T classifiers based on $S_{n^*}=\{S_{n^*}^1,S_{n^*}^2,\cdots,S_{n^*}^T\}$ participate in voting. The other less accurate classifiers are abandoned.

The operating stages are as following:

The first stage: Fix an appropriate dimensionality n^*.

(1) The follow conditions are given: the training sample set $\mathbf{x}_i$ and the corresponding sort label $\mathbf{y}$, ,the test sample set $\mathbf{d}_j$ and the corresponding sort label O_j;the number of training sample M, the dimensionality of a single eigenvector N, and the number of sampling T'.The original value of T' is 0, the original dimensionality of attribute subset is 1;

(2) Select n-dimension eigenvectors from the original training set randomly to form a training attribute subset S_n^t, the sequence number of which is t;

(3) Use the obtained subsets to train the ensemble of classifiers;

(4) Get the predicting function h_n^t of the ensemble of classifiers, the sequence number of which is t;

(5) If $t<T'$, $t=t+1$;turn to (2), otherwise, turn to (6) ;

(6) Coordinate all predicting functions h_n^t （$t=1,2,...,T'$; to get the final predicting function h_n^* complying with the principle of plurality. Test the accuracy of h_n^* in the testing sample set $\{\mathbf{d}_j\}$, and all the T' subsets participate in voting. To get

improve the precise of curve, repeat testing for K times and compute the average, which is the accuracy (r_n) of the predicting function.

(7) If $n<N$, $n = n+1$, turn to(2), otherwise, turn to(8);

(8) Set $\{n^* = q \mid r_q = \max_{p=1,\dots,N}(r_p)\}$.

The second stage: Optimizing classifiers.

(1) The follow conditions are given: the training sample set $\mathbf{x}_i$ and corresponding sort label $\mathbf{y}_i$, the number of training sample M, the dimensionality of a single eigenvector N, and the number of sampling T. The original value of T is 0, the original dimensionality of attribute subset is n^*;

(2) Select n*-dimension eigenvectors from the original training set randomly to form a training attribute subset $S_{n^*}^t$, the sequence number of which is t;

(3) Use the obtained subsets to train the sub-classifiers;

(4) Get the predict function $h_{n^*}^t$ of the sub-classifier, the sequence number of which is t;

(5) Test the accuracy of $h_{n^*}^t$ in the testing sample set. Repeat the test for K times and calculate the average, which is the accuracy ($r_{n^*}^t$) of the predicting function;

(6) If $t<T$, $t = t+1$, turn to (2), otherwise, turn to (6);

(7) Rank $r_{n^*}^t (t = 1,\cdots,T)$ by accuracy, and integrate the L most accurate predicting functions to get the final predicting list $\{h_{n^*}^{c_1},\cdots,h_{n^*}^{c_L}\}$, $c_i \in [1,\cdots,T]$, which will participate in voting in classification. The other (T-L) predicting functions will be abandoned.

4 Experimental Results and Analysis

4.1 Analysis of Affection of the Sub-eigenvector's Different Sizes on the AB Algorithm

In the experiment of analyzing AB algorithm, RBF Network was used as the predicting function, and the original features were composed of WT feature and CEP feature (Table 1). The number of training samples was 180, and the number of testing samples was 120.

To determine the optimal dimensionality of the sub-eigenvector, sampling were repeated for 25 times (T'=25), and then we obtained 25 sub-classifiers which all

Table 1. Meaning of Feature Component

Feature Component	Meaning	Feature Component	Meaning
X1-X20	WT feature	X21-X36	CEP feature

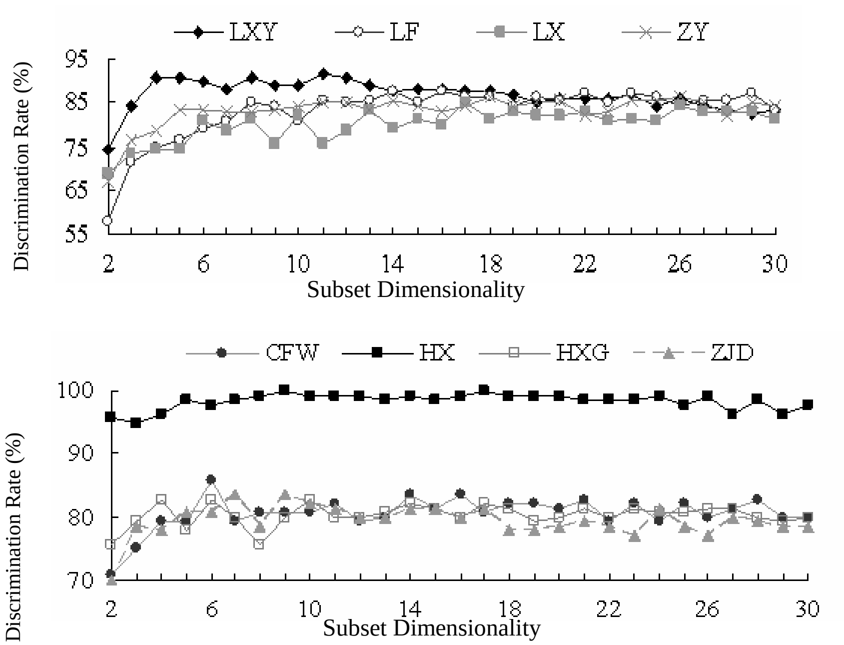


Fig. 2. The classifying accuracy curve of AB algorithm with different subset dimensionality

participated in voting. The dimension of the sub-eigenvector was from 2 to 30. Fig. 2 illustrates the relationship between the dimension of sub-eigenvector and the classifying accuracy of AB algorithm. The most appropriate dimensionalities of sub-eigenvectors of different experimental subjects are in Table 2. In Table 2, it can be seen that, for the situations researched in this paper, the most appropriate dimensionality of the sub-eigenvector is usually in the range of 1/4 ~ 1/2 of the original dimensionality.

After obtained the optimal dimensionality set of different sub-eigenvectors belonging to different experimental objects, take the minimum as the most appropriate value. Sample from the original attribute sets for 25 times and get 25 classifiers, which will participate in voting. In Table 3, we compared the obtained experimental result with the classifying results of a single RBF classifier with WT feature、CEP feature and most appropriate sub-feature(a training attribute subset in AB). Accuracy of the single RBF Network classifier with most appropriate sub-feature is in the "Single Classifier" row.

In Table 3, it can be seen that the hybrid RBF classifier integrated by AB algorithm is more accurate than other single classifiers, only except for subject ZJD. Especially when the dimensionality is small, the advantage of AB algorithm is much more obvious. As the result of subject CFW showing, the dimensionalities of eigenvector in

Table 2. The most Optimal Sub-feature Dimensionality of Different Subjects

Subjects	LF	LX	LXY	ZY	CFW	HX	HXG	ZJD
The most Optimal Sub-feature Dimensionality	14/16	17	11	18/26	6	9/17	4/6/10	7/9

Table 3. Discrimination Rate Comparison of Different Algorithm

Algorithm	LF	LX	LXY	ZY	CFW	HX	HXG	ZJD
AB Algorithm	87.86	87.5	91.67	86.43	85.71	100.00	82.86	83.57
Single RBF	79.17	84.17	84.17	81.67	68.33	98.33	82.50	70.00
WT Feature	81.67	84.17	89.17	85.83	79.17	99.17	73.33	85.83
CEP Feature	80.83	84.17	85.00	71.67	70.00	99.17	81.67	83.33

both “AB algorithm” and “single classifier” are all 6, but the recognition accuracy of the former is higher than the later by 17.38%.

4.2 Analysis of Affection of the Number of Sub-classifiers and Voting Classifiers on AB Algorithm

Keeping a generality, in the experiment of selecting optimized classifiers on the second stage, we compared and analyzed the data of subject LX, whose accuracy improved moderately with AB algorithm. The number of training samples was 180 and that of testing samples was 120. The dimensionality of attribute subset was 17. If we have not point out specially, “classification accuracy” means the classification accuracy of testing samples. When the number of voting classifiers L was 10 and 25, and that of sub-classifiers T is 25, 50 and 100, the accuracy of AB Algorithm was given in the Table 4. When the number of sub-classifiers is 25, 50 and 100, and the proportion of voting classifiers was changing, the accuracy of AB Algorithm was given in Fig. 3.

Table 4. Discrimination Rate of Same Number of classifiers with different number of sub-classifiers

	T=25	T=50	T=100
L=10	86.67	83.33	85.83
L=25	85	87.5	86.67

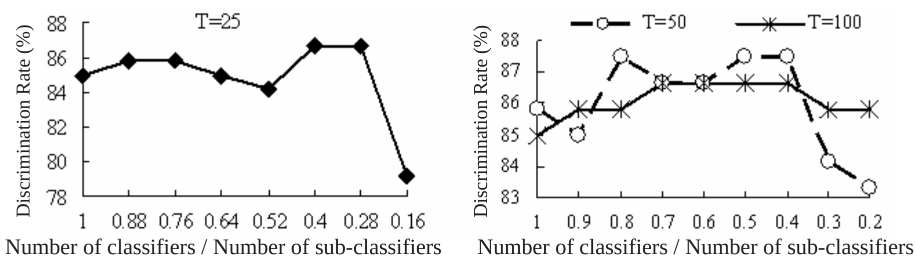


Fig. 3. Discrimination rate comparison of same number of classifiers with different number of sub-classifiers

From the Fig. 3 and the Table 4, it can be seen that when the most optimal dimensionality is determined, the highest accuracy is determined not only by the number of voting classifiers and sub-classifiers, but also largely by the ratio of the former to the later. Usually, the ratio, 1/5 ~ 1/2, is the most optimal.

5 Conclusion

Results showed that AB-RBF hybrid classifier achieved higher discrimination accuracy and stability than single RBF classifier. This proves that integrating classifiers with random feature subsets is an effective method to improve the performance of the pattern recognition system.

Moreover, with small number of dimensionality, AB-RBF hybrid classifier can achieve higher recognition accuracy and stability than single RBF classifier which can lessen the impact of the “curse of dimensionality”.

Acknowledgements. This study is supported by a grant from the Chinese NSFC (№50435040).

References

1. Wang, J.Z., Li, F., Wang, R.C., Zhang, S.: Research Development in the Discrimination of sEMG Signals with Artificial Neural Networks. Chinese Journa1 of Rehabilitation Medicine(in Chinese) 01, 81–83 (2006)
2. Li, F., Wang, R.C.: The Development Trend of the Discrimination Methods on EMG and Its Movement Patterns. Chinese Journa1 of Rehabilitation Medicine(in Chinese) 07, 492–493 (2005)
3. Bryll, R., Gutierrez-Osuna, R., Quek, F.: Attribute Bagging: Improving Accuracy of Classifier Ensembles by Using Random Feature Subsets. J. Pattern Recognition 36, 1291–1302 (2003)
4. Tumer, K., Oza, N.C.: Decimated Input Ensembles for Improved Generalization. In: Proceedings of the International Joint Conference on Neural Networks, Washington DC, 1999, vol. 05, pp. 3069–3074 (1999)

5. Daubechies, I.: Ten Lectures on Wavelets. In: CBMS-NSF Regional Conference Series in Applied Mathematics, 1992, vol. 61 (1992)
6. Atal, B.S.: Effectiveness of Linear Prediction Characteristics of the Speech Wave for Automatic Speaker Identification and Verification. The Journal of the Acoustical Society of America 55, 1304–1312 (1974)
7. Kang, W.J., Shiu, J.R., Cheng, C.K., et al.: The application of cepstral coefficients and maximum likelihood method in EMG pattern recognition [movements classification]. IEEE Transactions on Biomedical Engineering 42, 777–785 (1995)

An Epileptic Seizure Prediction Algorithm from Scalp EEG Based on Morphological Filter and Kolmogorov Complexity

Guanghua Xu[1], Jing Wang[2], Qing Zhang[2], and Junming Zhu[3]

[1] State Key Laboratory for Manufacturing Systems Engineering, Xi′an Jiaotong University, Xi'an, Shaanxi, 710049, P.R. China
[2] Department of Instrument Science and Technology, Xi′an Jiaotong University, Xi'an, Shaanxi, 710049, P.R. China
[3] Department of Neurosurgery, Zhejiang Provincial People's Hospital, Hangzhou, Zhejiang, 310014, P.R. China
xugh@mail.xjtu.edu.cn, aii@mail.xjtu.edu.cn, Junming_zhu@hotmail.com

Abstract. Epilepsy is the most common neurological disorder in the world, second only to stroke. There are nearly 15 million patients suffer from refractory epilepsy, with no available therapy. Although most seizures are not life threatening, they are an unpredictable source of annoyance and embarrassment, which will result in unconfident and fear. Prediction of epileptic seizures has a profound effect in understanding the mechanism of seizure, improving the rehabilitation possibilities and thereby the quality of life for epilepsy patients. A seizure prediction system can help refractory patients rehabilitate psychologically. In this paper, we introduce an epilepsy seizure prediction algorithm from scalp EEG based on morphological filter and Kolmogorov complexity. Firstly, a complex filter is constructed to remove the artifacts in scalp EEG, in which a morphological filter with optimized structure elements is proposed to eliminate the ocular artifact. Then, the improved Kolmogorov complexity is applied to describe the non-linear dynamic transition of brains. Results show that only the Kolmogorov complexity of electrodes near the epileptogenic focus reduces significantly before seizures. Through the analysis of 7 long-term scalp EEG recordings from 5 epilepsy patients, the average prediction time is 8.5 minutes, the mean sensitivity is 74.0% and specificity is 33.6%.

Keywords: Scalp EEG, Epileptic seizure prediction, Kolmogorov complexity, Morphological filter, Artifact removal.

1 Introduction

Epilepsy is one of the most common neurological disorders and affects almost 60 million people worldwide. About 75% of them can be controlled by medications or curable by surgery. This leaves one-quarter of the patients with refractory epilepsy.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 736–746, 2007.

Over the past decade, researchers have gradually noticed that seizures do not begin abruptly but develop several minutes to hours before clinical symptoms. This discovery leads to the possibility of predicting epileptic seizures, which can greatly improve the quality of life for these patients. For instance, patients can be forewarned to take timely preventive steps such as interrupting hazardous activities or contacting the physician for help. Also, therapeutic concepts could move from preventive strategies towards an on-demand therapy by electrical stimulation in an attempt to reset brain dynamics to a state that will no longer develop into a seizure.

EEG is an effective tool for clinical evaluation of brain activity. Majority of studies on epileptic seizure prediction are based on intracranial EEG [0~0]. However, it requires intensive surgical operations to implant electrodes inside brain, which are hazardous to the patient. Recently, researchers begin to pay their attention to scalp EEG because of its flexibility for outpatient and ambulatory applications [0,0]. Since scalp EEG is often contaminated with noises, a preprocessing of artifact removal is necessary. Corsini [0] presents a seizure prediction approach from scalp EEG. But ocular artifact, one of the most common contaminations, is not concerned.

So, the goal of this paper is to propose an integrated seizure prediction method based on scalp EEG signals, including artifacts elimination, seizure anticipation and analysis of clinical cases. The content of the paper is laid out in 5 sections with this section as the introduction. Section 2 describes the method of removing artifacts. A prediction method of epileptic seizure is presented and the results of applying this algorithm to 7 long-term scalp EEG recordings are given in section 3 and 4, respectively. Finally, conclusions are given in section 5.

2 Artifacts Removal for Scalp EEG

There exist a lot of artifacts in scalp EEG, which may possibly affect seizure prediction results. Due to the different characteristics of these artifacts, it is difficult to remove them with a single filter. So, a complex method is proposed as follows.

2.1 Baseline Excursion and 50Hz Frequency Component

The artifact of baseline excursion often occurs on temporal and frontal lobe with frequency of less than 1Hz because of sweat. In addition, most recordings present a 50Hz frequency component on all electrodes. Therefore, a band-pass filter with cutoff frequency of 0.5 and 45 Hz is utilized to eliminate these artifacts, as shown in Fig. 1.

2.2 Ocular Artifact

Eye blinks may cause ocular artifact on Fp1 and Fp2 electrodes. Owing to the frequency overlap between eye blinks and basic rhythm in scalp EEG, a band-pass filter will eliminate not only the artifacts, but also some useful information which affects seizure prediction result. Morphological filter can decompose EEG signal into several physical parts according to geometric characteristics, by using certain

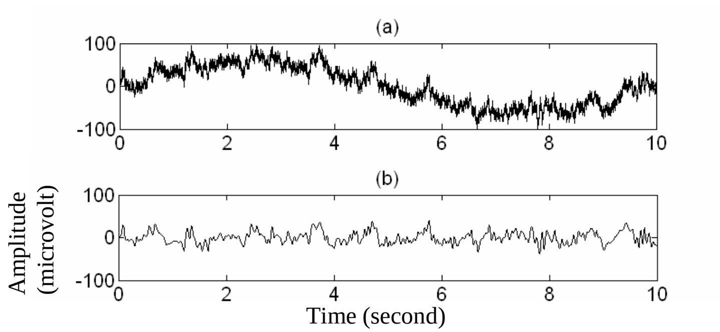


Fig. 1. Artifact removal: (a) Scalp EEG with artifacts of baseline excursion and 50Hz frequency component; (b) processed data with band-pass Butterworth filter

structure elements. Maragos [0] put forward four basic morphological operations shown in equation (1)~(4):

Erosion:

$$(f \Theta g^s) = \min_{\tau \in D}\{f(\tau) - g(-(t-\tau))\} \tag{1}$$

Dilation:

$$(f \oplus g^s) = \max_{\tau \in D}\{f(\tau) + g(-(t-\tau))\} \tag{2}$$

Opening:

$$(f \circ g)(t) = [(f \Theta g^s) \oplus g](t) \tag{3}$$

Closing:

$$(f \bullet g)(t) = [(f \oplus g^s) \Theta g](t) \tag{4}$$

where $f(t)$ is the original time series, and $g(t)$ is a structure element function. $g^s(t)$ points to reflection of $g(t)$, which is defined as $g^s(t) = g(-t)$. D means the set of real number. According to the characteristics of ocular artifact, clos-opening operation is employed to eliminate ocular artifact, as shown in equation (5):

$$CO(f(t)) = f(t) \bullet g_1(t) \circ g_2(t) \tag{5}$$

In order to separate the ocular artifact and background EEG, structure elements, which can insert into shape of background EEG but not into spike waves, are constructed by $g_1(t)$ and $g_2(t)$. The structure element pair is determined in (6) and shown in Fig. 2.

$$g_i(t) = a_i t^2 + b_i, \qquad i = 1,2 \tag{6}$$

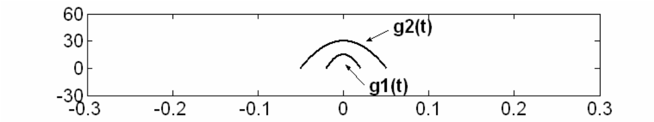


Fig. 2. Structure elements

where a_i accounts for the width and b_i stands for the center amplitude of the structure elements. Since eye-blink signals of different patients may have various amplitude and frequency, the structure elements should be adjusted to proper size where the eye-blink component can be best extracted. Suppose that $s(t)$ is the original EEG data and $x(t)$ is calculated with $x(t) = s(t) - CO(s(t))$, the structure elements are optimized as follows:

$$K = I_f / R_{pz} \tag{7}$$

for which

$$I_f = \hat{x} / \bar{x} \tag{8}$$

and

$$R_{pz} = N_{pz} / N \tag{9}$$

where I_f is the pulse index for the input signal x and R_{pz} points to the zero-pass rate of signal x. Thereinto $\hat{x}$ is the peak value of x obtained by $\hat{x} = \max\{|x(t)|\}$, $\bar{x}$ is the average amplitude of x calculated with $\bar{x} = \int_{-\infty}^{+\infty} |x| \, p(x) dx$. N is total data length of x and N_{pz} is the number of zero-pass points in signal x determined by

$$N_{pz} = \sum_{n=1}^{N-1} \mu[x(n) * x(n+1)] \tag{10}$$

assuming

$$\mu(x) = \begin{cases} 0 & x > 0 \\ 1 & x \le 0 \end{cases} \tag{11}$$

I_f is sensitive to peak and valley values in eye-blink signals and R_{pz} reflects the degree of restraining the background activities. So, a larger K shows that eye-blink component in EEG signals is better extracted and background activities are better restrained. In the present paper, the amplitude and width of the structure elements are optimized respectively with the mentioned K criterion, as described in Fig. 3.

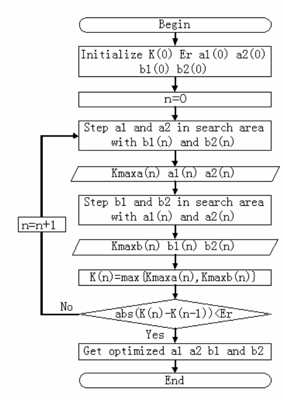


Fig. 3. Optimization of structure elements

Giving a is a typical EEG signal with ocular artifact in Fig. 4(a), the process of ocular artifact elimination is as follows:

(i) Firstly, the search area on amplitude and width of structure elements are decided: $K(0)=0$, $a_1(0)=-1.23\times10^4$, $a_2(0)=-4.80\times10^3$, $b_1(0)=100\mu V$, $b_2(0)=300\mu V$. There into, $b_1(0)$ and $b_2(0)$ indicate the probable amplitude of ocular artifact. Since the sample rate of EEG signal is 256 Hz , $a_1(0)$ and $a_2(0)$ determine that the initialized width of the structure elements are 0.18 s and 0.5 s , which covers the frequency range of ocular artifact;

(ii) With a certain structure element pair in the search area, clos-opening operation is implemented followed by its subtraction with original EEG signal: $x(t)=f(t)-CO(f(t))$;

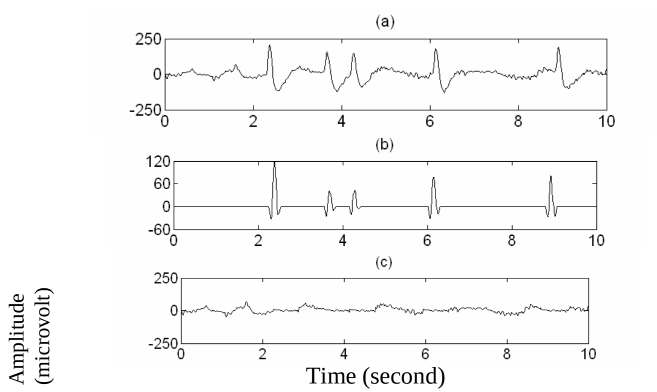


Fig. 4. Ocular artifact removal: (a) EEG signal with ocular artifact; (b) Detected ocular artifacts; (c) Artifact removal result

(iii) Parameter K for $x(t)$ is calculated with equation (7);

(iv) Repeat step (ii) and (iii) to get the K matrix;

(v) The maximum of K corresponds to the optimized structure elements: $a_1 = -3.13\times10^3$, $a_2 = -2.33\times10^3$, $b_1 = 180\mu V$, $b_2 = 210\mu V$;

(vi) Background activity is fully restrained, using optimized strcture elements. With a $10\mu V$ threshold, the ocular artifact is detected in Fig. 4(b);

(vii) Ocular artifact is replaced with normal background activity in Fig. 4(c).

3 Prediction Algorithm Based on Kolmogorov Complexity

EEG signals result from the complex processes inside the brain. Since epileptic neuronal networks are essentially complex nonlinear structures, the theory of complexity can be used to quantify its dynamics.

3.1 Improved Kolmogorov Complexity

In 1965, Kolmogorov suggested that complexity of a given string with zeros and ones was decided by the length of the shortest computer program which can generate this string. An appropriate measure of Kolmogorov complexity was put forward by Lempel and Ziv [0]. Usually, the symbol ‘0’ or ‘1’ is assigned to each sample by comparing it with the average of the signal. However, this transformation may neglect a lot of useful information in the original signal. Many modifications have been proposed to overcome the disadvantage. In [0], spikes with high amplitude, a typical epileptic activity in EEG signal, are reserved as ‘1’ after binary symbolization, while others are set to ‘0’. This means that many important epileptic characteristics, such as slow-amplitude spikes and spike-slow complex, are ignored. Here, we present a new method to extract all the epileptic activities in EEG, which is described as follows.

(i) Suppose $x(n) = \{x_1, x_2, ..., x_N\}$ is the original EEG signal and the average value of the time series is $x_p = [\sum_{i=1}^{N} x_i]/N$.

(ii) Calculate the absolute value of the subtraction between $x(n)$ and $x(p)$. $y(n) = \{y_1, y_2, ..., y_N\}$, where $y_i = |x_i - x_p|$, and the average value of $y(n)$ is obtained by $y_p = [\sum_{i=1}^{N} y_i]/N$.

(iii) Absolute difference of $x(n)$ is defined as $derx(n) = \{derx_1, derx_2, ...derx_N\}$, where $derx_j = \begin{cases} 0, & j = 1 \\ |x_j - x_{j-1}|, & otherwise \end{cases}$.

(iv) The average value of $derx(m)$ is acquired with $derx_p = [\sum_{i=1}^{N} derx_i]/N$.

(v) Binary symbolization of $x(n)$ is denoted as $bs(n) = \{bs_1, bs_2, \ldots bs_N\}$, which is decided by $bs_i = \begin{cases} 1, & y_i > y_p \;\; or \;\; derx_i > derx_p \\ 0, & otherwise \end{cases}$.

Thus, background activities in the original EEG signal are symbolized with '0', and epileptic characteristics are expressed with '1', which constitute a new series $bs(n)$. Then, Kolmogorov complexity of $bs(n)$ is calculated with the mentioned method in [0] to evaluate the dynamic features of $x(n)$.

3.2 Seizure Prediction Algorithm

The basic idea of our seizure prediction algorithm is to identify the change of brain activities from interictal state to preictal state. The decreases of Kolmogorov complexity at the electrodes near the epileptogenic focus are considered as the main identifier for preictal state. The algorithm can be described in detail as follows:

(i) Artifacts removal. Band-pass filter and morphological filter are employed to eliminate 50Hz frequency component and ocular artifact mentioned in section 2.

(ii) Electrode selection. Only electrodes near epileptogenic focus shows apparent changes when seizures are coming [0]. These electrodes are obtained knowledge through clinical diagnosis, shown in Table 1. The detailed information of the clinical EEG data for the patients will be explained in section 4.

(iii) Kolmogorov complexity computation. The Kolmogorov complexities on focal region electrodes are calculated according to the method proposed in section 3.1, with a sliding window of 40s and moving step of 3s. For comparison, Kolmogorov complexity of an electrode in remote focal area is also calculated. We take an EEG segment (SEG1) of patient 1 for example and the result is shown in Fig. 5.

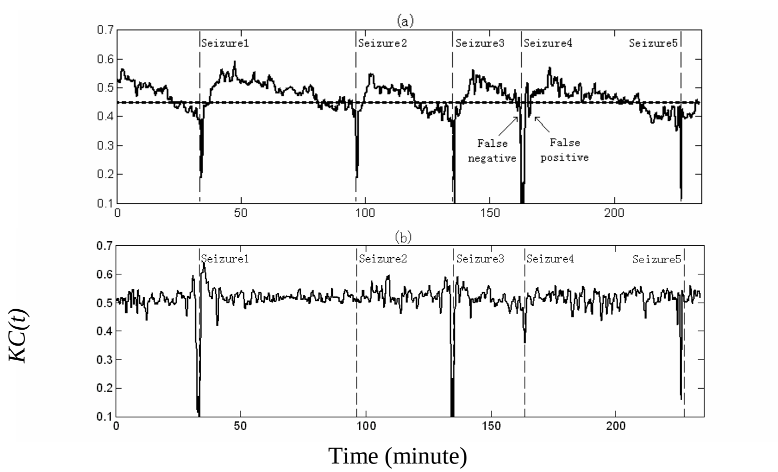


Fig. 5. Changes of Kolmogorov complexity: (a) Epileptogenic region (Fp2 and F4); (b) Remote region (C3)

Table 1. Electrodes selection

Patient	Epileptic focus	Electrodes for seizure prediction
1	Right frontal lobe	Fp2 F4
2	Right temporal lobe	F8 T4 T6
3	Left temporal lobe	T3 T5
4	Left frontal lobe	Fp1 F3
5	Left frontal lobe	Fp1 F3 F7

The change of Kolmogorov complexity in epileptogenic region is revealed in Fig. 5(a). Seizures are denoted as Seizure1, Seizure2, … and the Kolmogorov complexity at time t is denoted as $KC(t)$. The $KC(t)$ in Fig. 5(a) is an average value of Kolmogorov complexity of electrodes in focal area (Fp2, F4). We can see an obvious decrease of $KC(t)$ from several to tens of minutes before seizure. The curve reaches minimum during seizure and re-increases gradually after seizure. This phenomenon cannot be found in that of remote region (C3), as shown in Fig. 5(b).

(iv) Warning criteria. By comparing the difference of $KC(t)$ between epileptic region and remote region, we have the following criteria on seizure prediction: A warning alarm will be triggered when the average $KC(t)$ in recent 1 minute is lower than a preset threshold. In this paper, the threshold is set to 0.45.

4 Clinical Results

4.1 Data Acquisition

The clinical scalp EEG data were collected from five patients of Zhejiang Provincial People's Hospital, China, with detailed information shown in Table 2. Data acquisition system is Phoenix Unique Ambulatory EEG of EMS Handelsges.mbH company, Austria. The EEG data are amplified with band-pass filter of 0.15 - 60 Hz and the sampling rate is set to 256 Hz. Exploring cup electrodes were fixed according to the International 10 - 20 System. Video recordings are acquired synchronously with EEG, by which an experienced physician confirms the clinical onset.

Table 2. Summary of clinical data

Patient	Sex	Segment	Duration of recordings (min)	Number of seizures
1	Male	SEG1	236	5
		SEG2	119	2
2	Male	SEG3	120	2
3	Female	SEG4	60	1
		SEG5	120	3
4	Male	SEG6	179	7
5	Female	SEG7	180	9

4.2 Definition of Evaluation Criterion

An alarm can be either true or false, depending on whether it is followed by a clinical seizure or not. So, it is necessary to define a prediction range that points to the period within which a seizure is expected after an alarm. If a seizure arises after alarm within prediction range, it is defined as " true positive (TP)", otherwise regarded as "false positive (FP)". Also, if a seizure is not preceded by an alarm within prediction range, this will be thought of as "false negative (FN)". In this paper, the prediction range is determined to 15 minutes. Typical FP and FN can be found in Fig. 5(a).

Most studies have reported specificity in false prediction per hour, i.e., count all FPs and divide this number by the total duration of the recording. This definition ignores a fact that no FP will be announced in prediction range once an alarm arises. So, we employ a revised definition of specificity rates proposed in [0], which can be described as follows.

$$SPF = \frac{N_{FP} \times T_{PRE}}{T_A - N_{TP+FN} \times T_{PRE}} \tag{12}$$

where N_{FP} is the number of FPs, T_{PRE} means prediction range, T_A points to the total duration of the recording, and N_{TP+FN} denotes the number of clinical seizures. SPF indicates the portion of time from the interictal period during which a patient is not in the state of falsely waiting for a seizure.

Another assessing criterion of prediction algorithm performance is sensitivity shown in (13).

$$SEN = \frac{N_{TP}}{N_{TP+FN}} \tag{13}$$

Where N_{TP} is the number of TPs.

Finally, "Mean Prediction Time (MPT)" is utilized to evaluate the average forewarning time of the algorithm.

4.3 Statistics of Clinical Application

Using the proposed prediction algorithm in section 3, the prediction results for the EEG recordings shown in Table 2 are listed as follows.

Table 3. Prediction results of clinical data

Patient	Segment	Number of seizures	TPs	FPs	SPF (%)	SEN (%)	MPT (s)
1	SEG1	5	4	1	16.9%	80%	613
	SEG2	2	1	1	9.3%	50%	400
2	SEG3	2	1	0	0%	50%	365
3	SEG4	1	1	1	13.4%	100%	488
	SEG5	3	3	1	100%	100%	515
4	SEG6	7	5	1	20.3%	71.4%	562
5	SEG7	9	6	2	75%	66.7%	620
Total		29	21	7	33.6%	74.0%	509

The specificity is affected by factors such as duration of recordings, number of seizures and true positive. Once a patient undergoes several seizures in short time and a *FP* is announced, the interictal period decreases largely and therefore, *SPF* increases exceedingly. The result of average *SPF* indicates that the patients spend one third of their intericatal time on falsely awaiting a seizure.On the other hand, we get an average sensitivity of 74.0% and mean prediction time of more than 8 minutes, which demonstrates the validity of the proposed seizure prediction method.

5 Conclusion

In this paper, a method for epileptic seizure prediction from scalp EEG using morphological filter and Kolmogorov complexity is proposed. Firstly, the artifacts in the original signals are removed by a complex method. Baseline excursion and 50Hz frequency component are eliminated with a band-pass filter, while ocular artifact on Fp1 and Fp2 are removed with morphological filter. Thereinto, two aspects are focused on: a) clos-opening operation is utilized to extract ocular artifact in EEG data; b) the structure elements of the morphological filter are decided by two parabolas according to the geometric characteristics of ocular signals, and a new criterion is put forward to optimize the width and center amplitude of the structure elements. Secondly, a forewarning system based on Kolmogorov complexity is constructed. We present a new symbolization method to extract epileptic activities including spikes with different amplitude and slow waves, which is reasonable in clinical analysis. The proposed method is applied to 7 EEG segment from 5 epilepsy patients. Statistical results show that the mean prediction time (MPT) is 8.5 minutes, the mean sensitivity is 74.0% and specificity is 33.6%. In the future research work, we will focus on achieving sufficient sensitivity and specificity of seizure prediction techniques, and ultimately, establish an on-line prediction system based on scalp EEG.

References

1. Iasemidis, L.D., Shiau, D.S., Pardalos, P.M., Chaovalitwongse, W., Narayanan, K., Prasad, A., Tsakalis, K., Carney, P.R., Sackellares, J.C.: Long-term prospective on-line real-time seizure prediction. Clinical Neurophysiology 116, 532–544 (2005)
2. Alessandro, M.D., Esteller, R., Vachtsevanos, G., Hinson, A., Echauz, J., Litt, B.: Epileptic seizure prediction using hybrid feature selection over multiple intracranial EEG electrode contacts: a report of four patients. IEEE Trans. on biomedical and engineering 50, 603–615 (2003)
3. Corsini, J., Shoker, L., Sanei, S., Alareon, G.: Epileptic seizure predictability from scalp EEG incorporating constrained blind source separation. IEEE Trans. on biomedical and engineering 53, 790–799 (2006)
4. Hively, L.M., Protopopescu, V.A., Munro, N.B.: Enhancements in epilepsy forewarning via phas-space dissimilarity. Journal of clinical neurophysiology 22, 402–409 (2005)
5. Maragos, P., Schafer, W.: Morphological filters—part 1:Their set-theoretic analysis and relations to linear shift-invariant filters. IEEE Trans. on ASSP 35, 1153–1169 (1987)

6. Lempel, A., Ziv, J.: On the complexity of finite sequences. IEEE Trans. on information theory 22, 75–81 (1976)
7. Jia, W.Y., Kong, N., Li, F., Gao, X.R., Gao, S.K., Zhang, G.J., Wang, Y.P., Yang, F.S.: An epileptic seizure prediction algorithm based on second-order complexity measure. Physiological Measurement 26, 609–625 (2005)
8. Mormann, F., Elger, C.E., Lehnertz, K.: Seizure anticipation: from algorithm to clinical practice. Current Opinion in Neurology 19, 187–193 (2006)

A New Virtual Dynamic Dentomaxillofacial System for Analyzing Mandibular Movement, Occlusal Contact, and TMJ Condition

Chi Zhang[1,*], Lei Chen[2,*], Fengjun Zhang[1], Hao Zhang[2], Hailan Feng[2], and Guozhong Dai[1]

[1] Intelligence Engineering Laboratory, Institute of Software, Chinese Academy of Sciences
[2] Department of Prosthodontics, Peking University School & Hospital of Stomatology
{zhangchi,zfj,dgz}@iel.iscas.ac.cn, {chenlei,zhanghao,kqfenghl}@bjmu.edu.cn

Abstract. This paper describes a new virtual dynamic dentomaxillofacial system. Mechanical articulators have been used to reproduce mandibular movements and analyze occlusal contacts. With the help of VR and visualization technologies, virtual articulator systems can provide dynamic simulation and quantitative information visualization, enhance the functionality, and extend the system to additional application areas. We integrate mandibular movement simulation, occlusal analysis and TMJ analysis into our system, and design new algorithms to improve the results of analysis. This system is helpful to the education, the research, and the clinic in dentistry. An evaluation is conducted to prove the functionality and usability of the system.

Keywords: Visualization, Movement Simulation, Occlusion, Inter-articular Space.

1 Introduction

To simulate human mandibular movement is of great importance in dentistry. It can be used in diagnosing and evaluating the madibular movements, occlusal characteristics, the TMJ (tempromandibular joint) condition and other functions of the stomatognathic system.

Since introduced, mechanical articulators have been used to reproduce mandibular movements and analyze occlusal contacts. However, these mechanical devices have the following main disadvantages: 1. they cannot simulate the physiological mandibular movements; 2. they cannot provide time resolved information on the jaw movement; 3. they cannot visualize the real occlusal contact situation quantitatively and dynamically; 4. they cannot represent real dynamic condition of TMJ.

* C. Zhang and L. Chen contribute equally to the work.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 747–756, 2007.

In recent years, paramount progress is made in the area of computer graphics. With the help of VR and visualization technologies, it comes possible to build a more powerful mandibular movement simulation system. The computerization provides quantitative results dynamically and visually, enhances the functionality, and extends the system to additional application areas.

In this paper, we discuss a novel dynamic dentomaxillofacial system which satisfactorily integrates visualization, measurements, jaw movement simulation, occlusal analysis and TMJ analysis. Section 2 provides a brief survey of related research and previously developed systems. Section 3 discusses the process of data acquisition. Section 4 and Section 5 explain the main algorithms used in this system. Section 6 describes the system architecture and the user interface. We present our results and evaluations in Section 7.

2 Related Works

The computerized system for simulating the mandibular movements has been the focus of many studies. Hayashi et al firstly presented a computerized system to simulate mandibular movement and to visualize and analyze occlusion by measuring tooth shape with a 3D laser scanner and recording mandibular movement optically[1]. They made use of discrete 2D height-field data of tooth surfaces, and calculated minimal distances between points on opposing surfaces to generate occlusion. Okamoto et al reported a similar system that combined an optoelectronic tracking device for mandibular movements with a 3D contact-type digitizer for tooth shape[2]. This system could simulate dynamic occlusal situation and estimate occlusal contact area by exhaustively checking the distances from every point of the height-field to the points of the opposing surface. Kordass and Gartner developed the VR-articulator software tool—DentCAM using a 3D laser scanner and an ultrasonic measurement system[3]. This software allowed users to analyze static and dynamic occlusion as well as jaw relation. In addition, the movement paths of bilateral condylar points were simultaneously visualized which were in a sagittal and transversal view. However, all of the three systems did not mention the penetration between upper and lower teeth during the simulation. Because the skull model was not incorporated into these systems, the dynamic situation of the whole mandible (including the TMJ condition) cannot be visualized.

Enciso et al introduced computer methods to precisely visualize the movement of the whole mandible in 3-dimensions using CT data, digital models produced by destructive scanning, and mandibular motion capture[4]. Gallo et al presented a method of reconstruction and animation of the TMJ function using MRI data and jaw-tracking technology[5]. Dynamic intra-articulator space of the TMJ was computed and visualized in this system. Nonetheless, these two systems concentrated merely on analyze the dynamic condition of the mandible or the TMJ, rather than analyzing occlusal situation.

Based on the research noted above, we developed a novel dynamic dentomaxillofacial system, including the following functions: 1. simulating the movement of the whole skull, and providing time resolved information on the jaw movement; 2. producing various movement paths of arbitrary points on the mandible;

3. analyzing dynamic occlusal contact situation and the TMJ condition using a new method; 4. providing various 2D and 3D measurement tools.

3 Data Acquisition

The digital skull models are necessary to visualize the jaw movements on the screen. We create the 3D CT skull models from CT data. The skull of the patient undergoes CT scanning, and the digital CT data are transferred from the CT scanner to the computer using a CCD. The 3D skull model is reconstructed using the Marching Cubes algorithm. Thus the upper and lower jaw models are segmented, and stored in STL format. Afterwards, we reconstruct the digital dentitions by scanning the plaster models with the optical 3D measuring system. The upper and lower casts aligned in the relationship of ICP (intercuspal position) are scanned, and the digital upper and lower dentitions are then registered in ICP. Finally, the digital bone models are registered into the digital dentitions using an iterative closest point algorithm. The composite skull models with high detailed dental crowns are created, and exported for computerized manipulation.

Furthermore, we must measure the mandibular movement data for simulation of the patient's jaw motion. The ultrasonic-based system (ARCUS digma) is used to recording various jaw movements with six degrees of freedom[6]. This system measures the position of three sensors using four receivers according to the run times of ultrasonic impulses. In this system, the 3D positions of three virtual points (an incisal point, and both condylar points) in relation to the standard coordinate system are computed and stored sequentially with respect to time. In addition, the bitefork with the upper cast in its silicone occlusal registration is 3D scanned. By registering the composite skull models into the bitefork, the digital models and the motion data are combined in the standard coordinate system.

4 Visualization and Measurements

We utilize the Visualization ToolKit (VTK) and OpenGL to provide 3D data visualization. The system reads in the 3D models generated from the scan data, and provides various tools for users to browse, measure and inspect the whole skull conveniently. These tools include 2D/3D measurement tools, clipping planes and specific plane visualization, as shown in Fig. 1.

Distance and angle measurements are provided both in 2D and 3D, clearly distinguished from each other by different colors and notations. Users can simultaneously perform multiple measurements on the models. The measurements keep updating during the motion playback, providing a dynamic view of the statistics which users are concerned about.

Also users can easily create a clipping plane along arbitrary direction, so as to inspect and measure concealed anatomical structures. The system enables users to freely perform measurements on the clipping plane, thus a large number of anthropometric parameters of the occlusion, anatomy, orthodontics, and orthognathic surgery can be measured, which are applicable in operation planning. In addition,

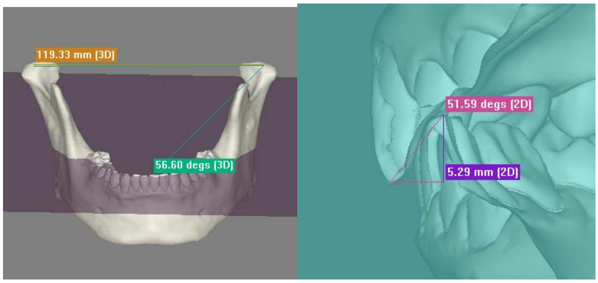


Fig. 1. 2D/3D measurement tools, the clipping plane (cyan) and specific plane visualization (magenta). Different colors are used to distinguish different measurements.

users can define and visualize a specific plane by picking three points that are supposed to be on the expecting plane. It can be used to visualize imaginary planes, such as occlusal plane and Frankfort plane. These planes can be acted as reference planes during the measurements.

5 Jaw Movements and Occlusions

5.1 Movement Simulation

We build a parser to read in the text data file exported by ARCUS Digma. The parser separates the motion data of each movement, and groups them distinctively for further processing. Each group contains coordinate sequences of two condylar points and the incisal point during one movement. The system converts the coordinate sequences into transformation matrix sequences, which are later used for movement simulation.

Users are allowed to select from seven separate movements, pause or continue the playback, and freely navigate throughout the motion playback using a video control panel (the toolbar below the menu in Fig. 2). Several time-related parameters are visualized by the playback, such as the chewing cycle, the motion velocity, and the timing of the mastication.

5.2 Motion Paths

The system provides users with various movement paths of arbitrary mandibular points, including the kinematic center points, and any points that users designate on the mandible (Fig. 3). The system calculates the movements of the points throughout the recorded motions, and generates a movement path in 3D for each of them. These paths can be browsed, measured and monitored as the models can, and the system automatically keeps the consistency between the paths and the models in real time. By measuring these paths, many distances and angles were provided, such as ranges of movement, opening angle, shift angle. All of these parameters can be used to analyze the motion of the condyles and specialized points.

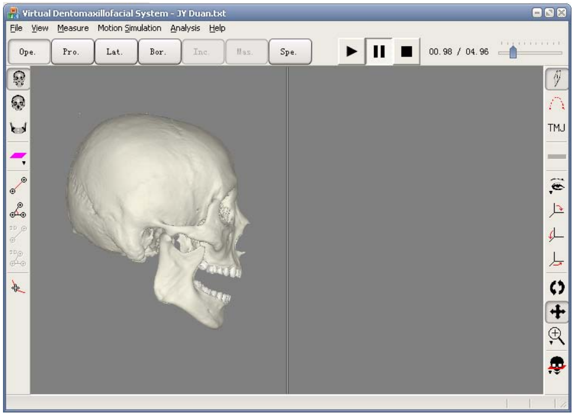


Fig. 2. The user interface of the system. Video control panel is positioned below the menu.

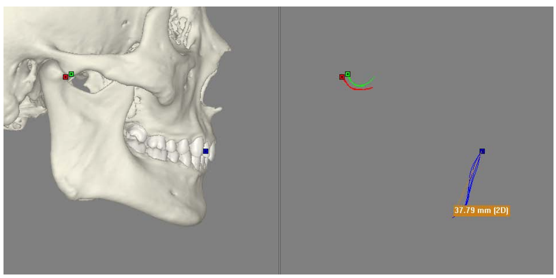


Fig. 3. Motion path visualization. The points tracked are indicated by the square markers (on both sides). Users are also enabled to perform measurements on the tracks.

5.3 Real-Time Occlusal Analysis

We improve Hayashi's method to perform real-time occlusal analysis[7]. The system resamples tooth surfaces into uniform 2D height-field grids, and calculates minimal distances between grid points on opposing surfaces to generate occlusion (Fig. 4). In addition, we check whether the opposing point falls below the current point, which indicates an occurrence of penetration. By this means all penetrations are checked and can be precisely reported, allowing users to make the decision how to treat the error introduced by the penetrations. The grid resolution can be adjusted to take the tradeoff between precision and the response time, which allows users to monitor the changes of the occlusion while browsing through the motion playback, and analyze the order and the timing of the occlusal contacts.

5.4 Accurate Occlusal Analysis and TMJ Analysis

As a result of the compromise, the real-time occlusal analysis algorithm gives only less-detailed version of the occlusion, and has distinct limitation because of its 2D

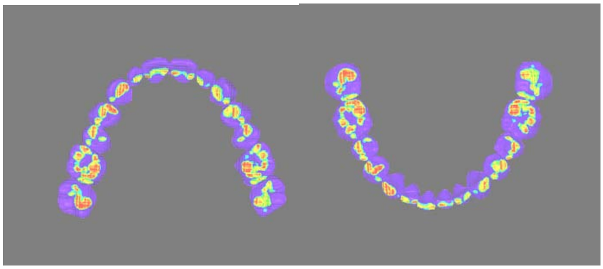
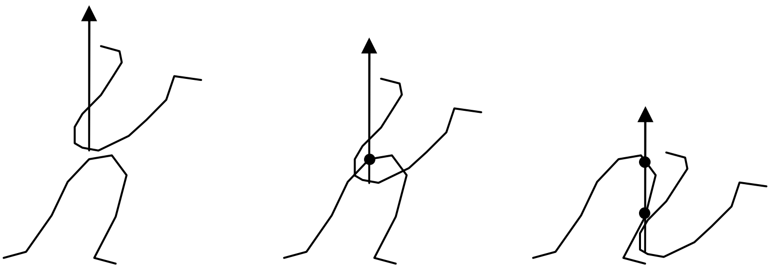


Fig. 4. Real-time occlusal analysis results. Inter-surface distances are indicated by different colors.



(a) No intersection. No penetration.

(b) One intersection. Penetration.

(c) Two intersections. No penetration.

Fig. 5. Penetration checking algorithm. Different situations are illustrated by (a), (b) and (c).

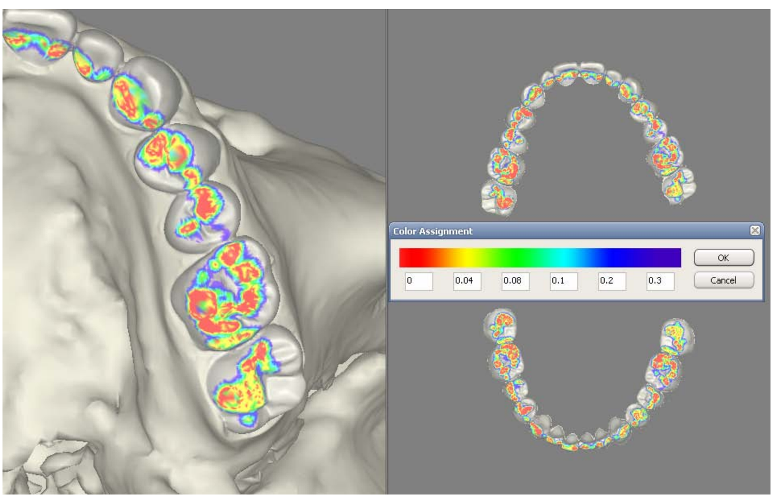


Fig. 6. The results of accurate occlusal analysis and the color assignment dialog

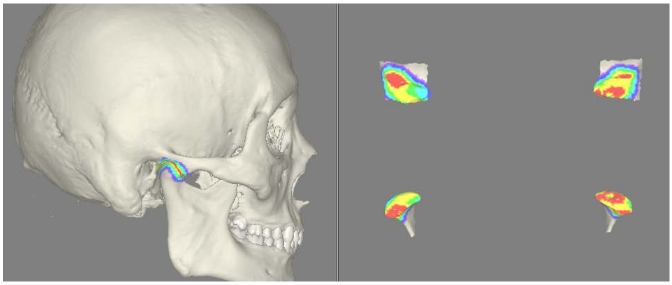


Fig. 7. TMJ analysis results

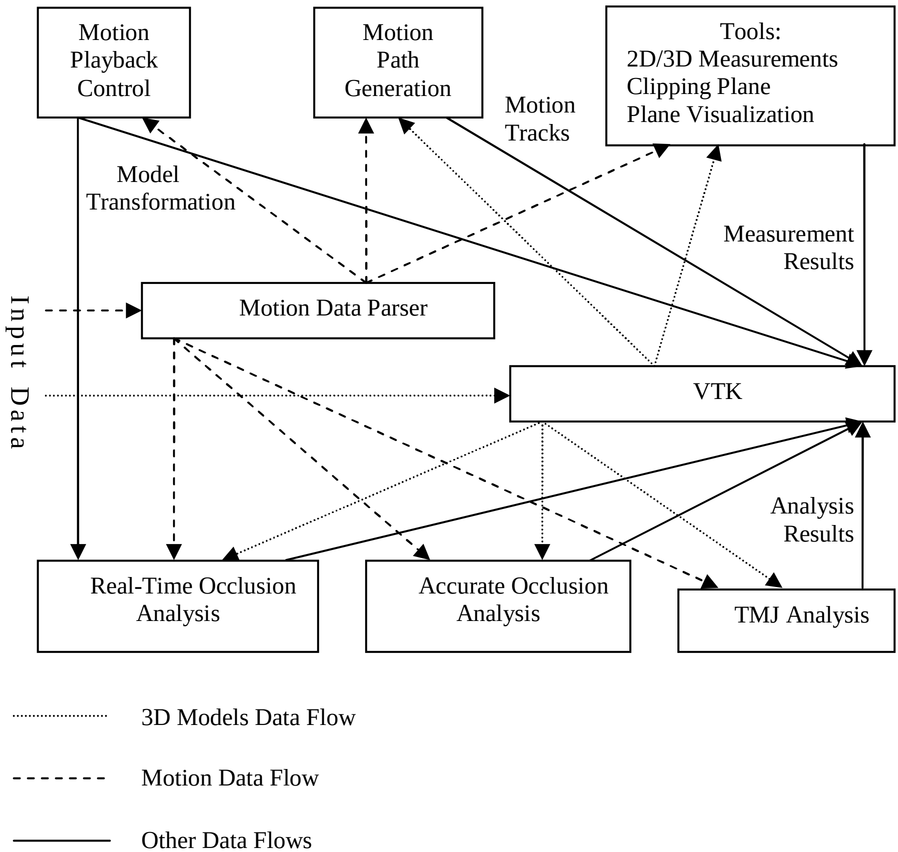


Fig. 8. System architecture

height-field assumption[7]. To provide users with more comprehensive and detailed information, we design another occlusal analysis algorithm that is more time-consuming but gives better results.

The algorithm consists of two stages: the penetration checking stage and the distance calculation stage.

In the first stage, we traverse all vertices of the teeth models and determine whether a vertex penetrates into the opposing model. This is done via the following steps:

1. Shoot a ray from the vertex in question along the direction away from the opposing model;
2. Check the ray with the opposing model for intersections;
3. If the number of intersections is odd, the vertex in question is considered penetrating the opposing model.

Some examples are provided in Fig. 5 to illustrate the algorithm of this stage.

In the second stage, the system looks for a nearest vertex on the opposing model for each vertex that is not considered as a penetration. The distance to the nearest vertex is stored, and later used to generate the occlusion by assigning colors to different ranges of distances (Fig. 6). The occlusion on the models can be browsed and measured with the model, and the details and concealed areas of occlusion can be explored by zooming in and using clipping planes. The results provided are helpful for evaluating occlusal dysmorphology and designing dental prostheses with functional occlusal surfaces.

As an extensive application, the new algorithm can also give satisfactory results in TMJ analysis (Fig. 7). We use fossae and condyles models instead of teeth models, and assign colors to larger distance ranges to visualize interarticular space. Thus the users can analysis and diagnosis of the TMJ condition according to the relation between the joint space and the loading.

6 System Integration and User Interface

Fig. 8 shows the architecture of the system.

All the functions provided are satisfactorily integrated in the system, and can be smoothly switched using the toolbars around the view. Various visual cues are provided in order to guide users in 3D operations, such as trackball rotation[8], and further improve the performance of the user interface.

In addition, users are allowed to use 3D mice for navigation if available. Users can navigate with a 3D mouse in the non-dominant hand, while operating on the models with the usual pointing device (mouse, for example) in the dominant hand, according to the concept of two-handed interaction[9]. All these techniques are employed in order to build a more efficient, intuitive and natural user interface.

7 Results and Evaluations

Our system can be used to analyze mandibular movements, occlusal contacts and TMJ situation. All of these functions are helpful to dental education because many conceptions and parameters in dentistry become intuitionistic. In addition, many quantitative data are also provided in our system, which benefit dental research. As far as dental clinic is concerned, this system assists diagnoses and therapies, and monitors the develop of diseases. We invited ten specialists among the dentists to perform an evaluation. After a period of trial, the subjects were surveyed on the degrees of satisfaction of the functions and the interface of our system. The results are presented in Table 1. It is revealed that our system is sufficient for the education, the research, and the clinic in dentistry.

Table 1. Evaluation results

	The degrees of satisfaction				
	1	2	3	4	5
Movement Simulation					10
Specific Plane Visualization			2	6	2
Clipping Plane			4	6	
2D/3D Measurement			6	4	
Motion Paths				1	9
Occlusal Analysis				5	5
TMJ Analysis			2	7	1
Accuracy			1	7	2
System Integration and User Interface				8	2
Overall				7	3

8 Conclusion and Future Work

In this paper we present a new virtual dynamic dentomaxillofacial system. We integrate mandibular movement simulation, occlusal analysis and TMJ analysis into a single system, and design new algorithms to further improve the results, as shown in the figures. We also perform an evaluation to prove the functionality and usability of the system.

Our future work includes employing the Marker-controlled Perception-based Mesh Segmentation algorithm to enhance the speed and fineness of the occlusal analysis[10], and researching on utilizing this system to assist virtual implant.

Acknowledgements. This research is supported by No. 2002CB312103 from the National Grand Fundamental Research 973 Program of China, No. 60673188 and 60303019 from National Natural Science Foundation of China, and 863 Program under Grant No. 2006AA01Z328.

References

1. Hayashi, T., Saitoh, A., Ishioka, K., Miyakawa, M.: A Computerized System for Analyzing Occlusal Relations during Mandibular Movements. Int. J Prosthodont 7, 108–114 (1994)
2. Okamoto, A., Hayasaki, H., Nishijima, N., Iwase, Y., Yamasaki, Y., Nakata, M.: Occiusal Contacts during Lateral Excursions in Children with Primary Dentition. J Dent. Res. 79, 1890–1895 (2000)
3. Kordaß, B., Gar̈tner, C.: The Virtual Articulator—Concept and Development of VR-Tools to Analyse the Dysfunction of Dental Occlusion. International Congress Series 1230, 689–694 (2001)

4. Enciso, R., Memon, A., Fidaleo, D.A., Neumann, U., Mah, J.: The Virtual Craniofacial Patient: 3D Jaw Modeling and Animation. Stud Health Teachnol Inform 94, 65–71 (2003)
5. Gössi, D.B., Gallo, L.M., Bahr, E., Palla, S.: Dynamic Intra-Articular Space Variation in Clicking TMJs. J Dent Res 83, 480–484 (2004)
6. Pröschel, P., Morneburg, T., Hugger, A., Kordaß, B., Ottl, P., Niedermeier, W., Wichmann, M.: Articulator-Related Registration - a Simple Concept for Minimizing Eccentric Occlusal Errors in the Articulator. Int. J Prosthodont 15, 289–294 (2002)
7. Hayashi, T., Hashimoto, Y., Saitoh, A., Miyakawa, M.: A Method of Computing the Interface Distance between Opposing Teeth. Frontiers Med Biol Eng 5, 175–184 (1993)
8. Chen, M., Mountford, S.J., Sellen, A.: A Study in Interactive 3-D Rotation Using 2-D Control Devices. Computer Graphics 22, 121–129 (1988)
9. Buxton, W., Myers, B.A.: A Study in Two-Handed Input. In: Proceedings of CHI '86, pp. 321–326 (1986)
10. Zhang, C., Zhang, N., Li, C., Wang, G.: Marker-controlled Perception-based Mesh Segmentation. In: Proceedings of the Third International Conference on Image and Graphics (ICIG 2004), Hong Kong, China, pp. 390–393 (2004)

Mechanism of Bifurcation-Dependent Coherence Resonance of Excitable Neuron Model

Guang-Jun Zhang[1,2], Jue Wang[1], Jian-Xue Xu[3], Xiang-Bo Wang[2], and Hong Yao[2]

[1] School of Life Science and technology, Xi'an Jiaotong University, PRC
[2] School of Science, Air Force Engineering University, Xi'an, PRC
[3] School of Aerospace, Xi'an Jiaotong University, PRC
Zhanggj3@126.com, jxxu@mail.xjtu.edu.cn

Abstract. In contrast to the previous studies that have dealt with phenomenon of coherence resonance induced by external noise in excitable neuron model, in this paper the mechanism of bifurcation-dependent coherence resonance (CR) of excitable neuron model is investigated by researching the random transitions of system motion between attractors in the two sides of bifurcation point. The results of research show: For two excitable neuron model, FHN neuron model and HR neuron model, There exist different attractors in two sides of the two excitable neuron model Hopf bifurcation point, at the presence of internal or external noise the system motion may transit between attractors in two sides of bifurcation point; mechanism of bifurcation-dependent CR of excitable neuron model is related to this kind of random transitions, the frequency of transitions tend towards a certain frequency for a certain optimal noise intensity, and the signal-noise-ratio of system response evaluated at this certain frequency is maximal at the optimal noise intensity, CR occurs.

Keywords: coherence resonance, attractor, transition, bifurcation point.

1 Introduction

The phenomenon of stochastic resonance in nonlinear dynamical systems perturbed by noise has attracted considerable attention over the last two decades[1-3]. Conventional SR occurs in noisy nonlinear dynamical system when the nonlinear dynamical system are perturbed by a weak external signal[3]. The signal-noise-ratio (SNR) and other appropriate measures of signal coherence, pass through a maximum at optimal noise strength. It is well known that neurons work in a noisy environment, and it is therefore of great interest to study how information is encoded and transmitted when neurons work in such a nosy environment. Recent studies have shown that excitable sensory neurons can use the mechanism of SR to detect weak subthreshold periodic and aperiodic signals, and noise plays a constructive role in the transmission of information in neurons[4-11]. Other studies has also shown that, for an excitable neuron model with only noise input, noise activates the excitable system producing a sequence of pulses and the coherence of the noise-induced oscillations is shown maximal for a certain noise intensity. Thus coherence resonance (CR), which

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 757–766, 2007.

is an SR-like resonance phenomenon, has been observed in the noisy excitable neuron model[12-19]. As a noise-induced ordering phenomenon, CR is of particular importance in neurophysiology, where large ensembles of neurons at the absence of input signal become orderly due to interaction with a noisy environment[17，19]. The results of neurophysiology experiment also show that non-periodic firing neurons were more sensitive to stimulis than periodic firing neurons, which was termed as “sensitivity of non-periodic activity” of neuron[20-21]. The research of Tan and Xu shows the dynamic mechanism of “sensitivity of non-periodic activity” is mechanism of bifurcation[22]. To reveal the mechanism of neuronal information encoding and transmission, it is necessary for mechanism of CR to be researched.

In these previous studies about CR, most only observed the phenomenon of CR and application in the transmission of information of neurons. But for period-doubling bifurcation in the noisy celebrated Feigenbaum map and in Rossler system, The researches of Neiman A. etc. show the two nonlinear dynamical systems when excitable by noise give rise to precursors of instabilities of periodic motion, which are prominently close to points of bifurcation, manifest themselves as noise-induced peaks in the power spectrum of the system[16]. The previous researches show that when there exist coexisting attractors in nonlinear dynamical system[23] or the bifurcation parameter of system is near the Hopf bifurcation point or tangent bifurcation point or period-doubling bifurcation point etc[13-16, 23], CR of nonlinear dynamical system occurs. So the mechanism of CR of excitable neuron is probably related to bifurcation. The CR of nonlinear dynamical system when the bifurcation parameter is near the bifurcation point is named bifurcation-dependent CR in this paper. The researches of Zhang and Xu show that the mechanism of stochastic resonance of FHN neuron model with periodic signal at the presence of internal noise is related to random transitions of system motion between attractors in the two sides of bifurcation point[10-11]. The mechanism of CR of excitable neuron is probably related to random transitions of system motion between attractors in the two sides of bifurcation point. So different from the previous ways, in this paper, the mechanism of bifurcation-dependent CR excitable neuron model is researched by researching the random transitions of system motion between attractors in the two sides of bifurcation point.

2 The Bifurcation Characteristic of the Unforced Excitable Neuron Model

2.1 FitzHugh-Nagumo Neuron Model

We consider the unforced FitzHugh-Nagumo neuron model in the following form[4-6, 10,25]:

$$\begin{cases} \varepsilon \dfrac{dv}{dt} = v(v-a)(1-v) - w \\ \dfrac{dw}{dt} = v - dw - b \end{cases} \tag{1}$$

The variable v is the fast voltagelike variable and w is the slow recovery variable. Throughout the paper we fix the values of the constants to ε=0.005,d=1.0. A firing or spike is considered to occur if the variable v has a positive-going crossing of the firing threshold v_{th}, chosen as v_{th} =0.5. According to Ref.[5,12] one may see that When b<0.2623 there exists only one stable fixed point in FHN neuron model, when b=0.2623, a limit cycle with small amplitude begins to appear. The amplitude of the limit cycle suddenly jumps to a large value as the bifurcation parameter b approaches the critical point $b=b_c=0.26395$. For $b \geq b_c$ the motion of FHN neuron model is a limit cycle attractor of suprathreshold oscillation.

Because the bifurcation parameter b in Eq.(1) is independent of variables v and w, the internal noise added to bifurcation parameter b is equivalent to external noise added to the system. In order to add parameter noise to bifurcation parameter b in Eq.(1) the following transformation is made:

$v = v_1 + v_0, w = w_1 + w_0$

Then the equation (1) becomes:

$$\begin{cases} \varepsilon \dfrac{dv_1}{dt} = (v_1 + v_0)(v_1 + v_0 - a)(1 - (v_1 + v_0)) - (w_1 + w_0) \\ \dfrac{dw_1}{dt} = v_1 - dw_1 \end{cases} \tag{2}$$

where v_0 and w_0 are the coordinate of fixed point of Eq.(1) and are the function of bifurcation parameter b. Thus it can be seen that noise added to bifurcation parameter b in Eq.(2) is not equivalent to external noise added to the Eq.(2) any longer.

2.2 Hindmarsh-Rose Neuron Model

HR neuron model depicts the typical bursting spike behaviors of neural system. We consider the Hindmarsh-Rose neuron model in the following form[14,15,17]:

$$\begin{cases} \dot{x} = y - ax^3 + bx^2 - z + I \\ \dot{y} = c - dx^2 - y \\ \dot{z} = r[s(x - x^*) - z] \end{cases} \tag{3}$$

Where x is membrane electrical potential, y is recovery variable, z is adjusting current, I is synaptic stimulus of input. In the paper we fix the values of the other parameters to a=1, b=3, c=1, d=5, x*=-1.6 and s=4, only r and I are considered as control parameters.

With the varying of synaptic stimulus intensity I, for the membrane electrical potential two kinds of oscillation occurs, which refers to supercritical and subcritical oscillation. The membrane electrical potential displays intermittent bursting spike when it is supercritical state. And the number of spike in every bursting depends on I. Therefore, I is a bifurcation parameter of HR neuron model nonlinear dynamic behavior. If I is constant membrane electrical potential also displays intermittent bursting spike with r bifurcation parameter[14, 15].

According to [14, 15], system (3) has a unique fixed point at (x_s, y_s, z_s), which can be expressed as:

$$\begin{cases} x_s = -\frac{2}{3} + (-p+q)^{1/3} + (-p-q)^{1/3} \\ y_s = 1 - 5x_s^2 \\ z_s = 4x_s + 6.4 \end{cases}$$

Where $p = 449/270 - I/2$, $q = \sqrt{p^2 + (8/9)^3}$. For r>0 and 0<I<5.4, Hopf bifurcation occurs in: $r_c = [-1 - 6d - d^2 + \sqrt{f(d)}]/[2(5+d)]$,

Where $d = 3x_s^2 - 6x_s$, $f(d) = d^4 + 8d^3 + (14 - 40x_s)d^2 + (-8 - 240x_s)d + 1 - 200x_s$ 。

According to [14, 15], the critical point of system (3) is (r_c=0.001, I_c=1.26987), when $I>I_c$ ($r=r_c$) or $r>r_c$ ($I=I_c$) a Hopf bifurcation occurs. In the left side of bifurcation point there is a stable fixed point attractor. In the right side of bifurcation point there is a stable limit cycle attractor (bursting). There are some short time intervals in ISI of every limit cycle; there is a longer time interval in every two bursting.

Similar to FHN neuron model above, in order to add internal and external noise conveniently to HR neuron model the following transformation is made: $x=x+x_s$, $y=y+y_s$, $z=z+z_s$, Then the equation (3) becomes:

$$\begin{cases} \dot{x} = y - a(x^3 + 3x^2 x_s + 3xx_s^2) + b(x^2 + 2x_s x) - z \\ \dot{y} = d(x^2 + 2xx_s) - y \\ \dot{z} = r(sx - z) \end{cases} \tag{4}$$

where x_s, y_s and z_s are the coordinate of fixed point of Eq.(3) and are the function of bifurcation parameter I. Thus internal or external noise can be added conveniently to the equation (4).

3 The Mechanism of Bifurcation-Dependent CR of Excitable Neuron Model

3.1 The Mechanism of Bifurcation-Dependent CR of FHN Model

In this paper, the Gaussian distributed white noise $\xi(t)$ is used for perturbing the system. The mean and autocorrelation function are as follows respectively:

$$\begin{cases} <\xi(t)> = 0 \\ <\xi(t)\xi(s)> = \delta(t-s) \end{cases} \tag{5}$$

In this paper it will be denoted as D_i and D_e for the internal and external noise respectively.

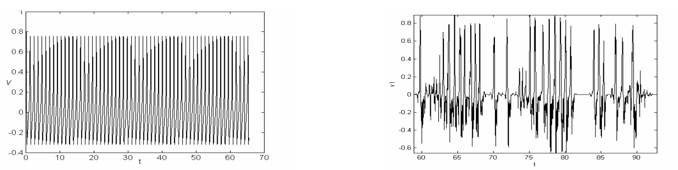
After the internal and external noises are added to equation (2), the equation (2) becomes:

$$\begin{cases} \varepsilon \dfrac{dv_1}{dt} = (v_1 + v_0)(v_1 + v_0 - a)(1 - (v_1 + v_0)) - (w_1 + w_0) \\ \dfrac{dw_1}{dt} = v_1 - dw_1 + D_e \xi(t) \end{cases} \tag{6}$$

where

$$\begin{cases} b = b_0 + D_i \xi(t) \\ v_0 = \sqrt[3]{-\dfrac{0.5-b}{2} + \sqrt{(\dfrac{0.5-b}{2})^2 + (\dfrac{0.75}{3})^3}} - \sqrt[3]{\dfrac{0.5-b}{2} + \sqrt{(\dfrac{0.5-b}{2})^2 + (\dfrac{0.75}{3})^3}} \\ w_0 = v_0 - b \end{cases}$$

The stochastic differential equation (6) is integrated using the fourth-order Runge-Kutta method with time step t=0.001 second. After the mean value of bifurcation parameters of system is fixed to b_0=0.264, the time history of the system response in two cases of noise is shown respectively in Fig.1. From the Fig.1 one may see that at the presence of internal noise the motion trajectory of system transits between attractors respectively in two sides of bifurcation point.



(a) (b)

Fig. 1. Time history of a realization of FHN neuron model response, b0=0.264; (a) Di=0.0; (b) Di=0.08

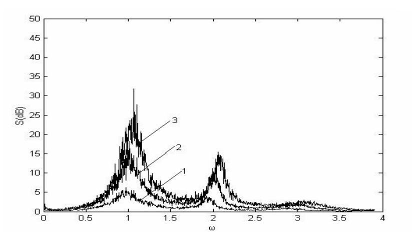


Fig. 2. Power spectrum of FHN neuron model without periodic stimulus for several values of internal noise D_i, b_0=0.264; 1-- D_i=0.023; 2-- D_i=0.040; 3-- D_i=0.060

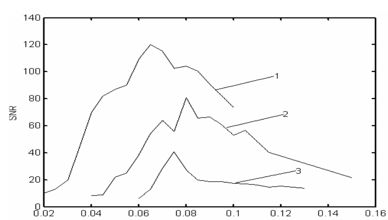


Fig. 3. The *SNR* vs noise intensity D_i at the presence of internal noise, 1---b_0=0.262; 2----b_0=0.260; 3--- b_0=0.258

According to Ref.[10,11], the bifurcation point of FHN neuron model will slightly shift to the left or to the right under the perturbation of Gaussian distributed external noise. Thus, the shift of bifurcation point induced by external noise may induce the motion trajectory of system to transit between attractors respectively in two sides of bifurcation point.

According to Ref.[10,11] noise could modulate the frequency of random transitions of FHN neuron model with periodic signal, and induce the frequency of random transitions in accordance with the frequency of input signal, then stochastic resonance occurs. In order to investigate the frequency of random transitions of FHN neuron model without periodic signal, the spectrum of response of system in this case is shown in Fig.2. As shown in Fig.2, one may see that under the perturbation of appropriate noise a peak of power spectrum of system response at a certain frequency will appear. That is to say the frequency of random transitions is spontaneously in accordance with a certain frequency. To measure the coherence of the system at the certain frequency, as in [16] we define the signal-noise-ratio as follow:

$$SNR = \frac{H}{w} \tag{7}$$

where $w=\Delta\omega/\omega_p$, ωp is the certain spontaneous frequency above, $\Delta\omega$ is width at the half-maximum height in the power spectrum depends on the noise intensity. H is the power of system response at the spontaneous frequency ω_p mentioned above. When the curve of the SNR defined above v_s noise intensity is a unimodal curve coherence resonance occurs[14-16].

Therefore, after the mean of bifurcation parameter being fixed to b_0=0.262 or b_0=0.260 or b_0=0.258 respectively, the signal-to-noise-ratio of FHN neuron model under the perturbation of internal or external noise are obtained respectively. The curves of SNR as a function of noise intensity D*i* or D_e are shown in Fig.3, 4 respectively.

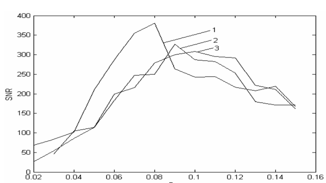


Fig. 4. The SNR vs noise intensity Di at the presence of external noise, 1--- b=0.262, 2--- b=0.260, 3 ---b=0.258

As shown in Fig.3, 4 in two cases of noise CR occurs, And the value of optimal noise intensity D_{opt} decreases as the bifurcation parameter b moves close to the critical firing onset value b_c. This is in accordance with the result of Gong and Xu in [12]. According to analysis above the conclusion can be drawn, that the mechanism of bifurcation-dependent CR of FHN neuron model at the presence of internal or external noise is related to the random transitions between attractors in the two sides of bifurcation point respectively.

3.2 The Mechanism of Bifurcation-Dependent CR of HR Neuron Model

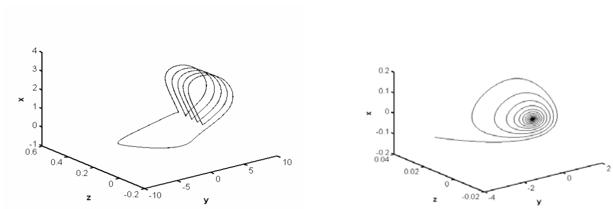
The internal and external noise are added to Eq.(4), and then the differential equation of HR neuron model is as follows:

$$\begin{cases} \dot{x} = y - a(x^3 + 3x^2 x_s + 3xx_s^2) + b(x^2 + \\ 2x_s x) - z + D_e \xi(t) \\ \dot{y} = d(x^2 + 2xx_s) - y \\ \dot{z} = r(sx - z) \end{cases} \tag{8}$$

Where
$$\begin{cases} I = I + D_{\mathrm{i}} \xi(t) \\ p = 449/270 - I/2 \\ q = \sqrt{p^2 + (8/9)^3} \\ x_s = -\dfrac{2}{3} + (-p+q)^{1/3} + (-p-q)^{1/3} \\ y_s = 1 - 5x_s^2 \\ z_s = 4x_s + 6.4 \end{cases}$$

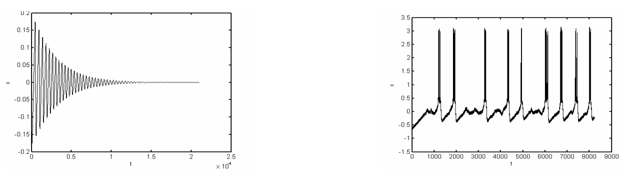
The stochastic differential equation (8) is integrated using the fourth-order Runge-Kutta method with time step t=0.01 second. To illustrate the stochastic transitions of HR neuron model motion between attractors in two sides of bifurcation point respectively under the perturbation of noise, the attractor of HR neuron model without noise is calculated respectively when I=1.27087 and I=1.25687. The results are shown as Fig.5. When the value of bifurcation parameters of system without noise is fixed to I=1.25987, a sample of the time history of the system response in two cases of noise intensity is shown respectively in Fig.6. From the Fig.6 one may see that at the presence of internal noise the motion trajectory of system transits between attractors respectively in two sides of bifurcation point. According to Ref.[10-11], the bifurcation point of Hopf bifurcation system will slightly shift to the left or to the right under the perturbation of Gaussian distributed external noise. Thus, the shift of bifurcation point induced by external noise may induce the motion trajectory of system to transit between attractors respectively in two sides of bifurcation point. According to above, for the FHN neuron model the transitions between attractors in two sides of Hopf bifurcation point induce CR to occur, however, in the present study, the question whether CR could be induced by this kind of transitions arises naturally in this context.

To investigate the frequency of random transitions mentioned above, the power spectra of HR neuron model response in three cases of internal noise intensity are calculated respectively. The results are shown as in Fig.7. As shown in Fig.7, one may see that under the perturbation of appropriate noise a peak of power spectrum of system response at a certain frequency will appear (the numerical result shows this certain frequency is only related to the inherent characteristics of system, but is not related to the noise intensity). That is to say that the frequency of random transitions spontaneously is in accordance with a certain frequency. Therefore, after the bifurcation parameter without internal noise being fixed to I=1.26887, I=1.26687 or I=1.26387, the signal-to-noise-ratio of HR neuron model under the perturbation of internal or external noise are obtained respectively. The curves of SNR as a function of noise intensity D_i or D_e are shown in Fig.8, 9 respectively. As shown in Fig.8, 9 in two cases of noise CR occurs, and the value of optimal noise intensity D_{opt} decreases as the bifurcation parameter I moves close to the critical firing onset value I_c. This is in accordance with the result of Gong and Xu in [12] and section 3.1 . According to analysis above the conclusion can also be drawn, that the mechanism of bifurcation-dependent CR of HR neuron model at the presence of internal or external noise is related to the random transitions between attractors in the two sides of bifurcation point respectively.



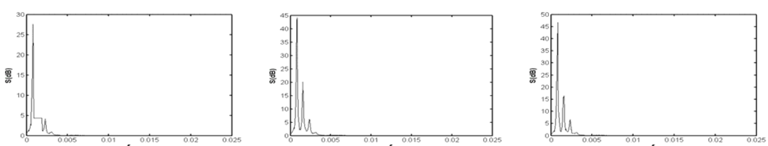
(a) (b)

Fig. 5. The phase figure of attractors of HR neuron model without noise in two cases of bifurcation parameter, (a) I=1.27087; (b) I=1.25987



(a) (b)

Fig. 6. A sample of the time history of the system response in two cases of noise intensity, I=1.25987; (a) Di=0.0, De=0.0; (b) Di=0.02, De=0.0



(a) (b) (c)

Fig. 7. The power spectra of HR neuron model response at the presence of internal noise, I=1.26687,De=0.0, (a) Di=0.06, (b) Di=0.08, (c) Di=0.10

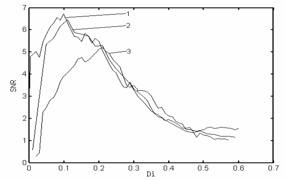


Fig. 8. The SNR of HR neuron model response versus model response versus internal noise intensity, D_e=0.0, 1—*I*=1.26887; 1—*I*=1.26887, 2—*I*=1.26687; 3— *I*=1.26387

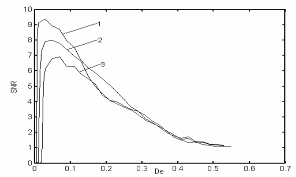


Fig. 9. The SNR of HR neuron external noise intensity, D_i=0.0, 2— *I*=1.26687, 3— *I*=1.26387

4 Conclusion

In summary, in this paper the mechanism of bifurcation-dependent CR of two excitable neuron models at the presence of internal or external noise have been investigated. For the two neuron model, FHN neuron model and HR neuron model, the same conclusions are drawn as follows: There exist respectively different attractors in two sides of Hopf bifurcation point of the two neuron models, The random transitions between attractors on two sides of bifurcation point respectively occur in the two neuron model at the presence of internal or external noise; The frequency of random transitions spontaneously is in accordance with a certain frequency under the perturbation of appropriate noise, and the SNR at the certain frequency as a function of noise intensity is a unimodal curve, Coherence resonance occurs. The mechanism of CR is related to the random transitions between attractors in the two sides of bifurcation point respectively.

Acknowledgement. The National Science Foundation under Grant 10432010, 30670660 supports the work reported in this paper.

References

1. Benzi, R., Sutera, A., Vulpiani, A.: The mechanism of stochastic resonance. J. Phys. A 14, L453–L457 (1981)
2. Benzi, R., Parisi, G., Sutera, A., Vulpiani, A.: Stochastic resonance in climatic change. Tellus 34, 10–16 (1982)
3. Gammaitoni, L., Hanggi, P., et al.: Stochastic resonance. Rev. Mod. Phys. 70, 223–287 (1998)
4. Longtin, A.: Stochastic resonance in neuron model. J.Stat Phys. 70, 309–327 (1993)
5. Gong, P.L., Xu, J.X.: Global dynamics and stochastic resonance of forced FitzHugh-Nagumo neuron model. Phys. Rev. E. 63, 31906 (2001)
6. Gong, Y.F, Xu, J.X.: Stochastic resonance: When does it not occur in neuronal models. Phy. Let.A 243, 351–359 (1998)
7. Wiesenfield, K., Moss, F.: Stochastic resonance and benefits of noise: from ice ages to crayfish and squids. Nature 373, 33–35 (1995)
8. Collins, J.J., Chow, C.C., Imhoff, T.: Aperiodic Stochastic resonance. Phys. Rev. E. 52, R3321–3324 (1995)

9. Collins, J.J., Chow, C.C., Imhoff, T.: Stochastic resonance without tuning. Nuture 376, 236–238 (1995)
10. Zhang, G.J., Xu, J.X.: Characteristic of nonlinear system stochastic resonance in the neighborhood of bifurcation point (in Chinese), Acta of Physica Sinica., 54(2), 557–564 (2005)
11. Zhang, G.J., Xu, J.X.: Stochastic Resonance induced by novel random transitions of motion of FitzHugh-Nagumo Neuron Model, Chaos,Solitons and Fractals, vol. 23(4), pp. 1439–1449 (2005)
12. Gong, P.L, Xu, J.X.: Coherence resonance in a noise-driven excitable neuron model. Chaos,Solitons & Fractals 13, 885–895 (2002)
13. Longtin, A.: Autonomous stochastic resonance in bursting neuron. Phys. Rev. E. 55, 868–876 (1997)
14. Wu, S., Ren, W., He, K., Huang, Z.: Burst and Coherence resonance in Rose-Hindmarsh model induced by additive noise. Phys. Lett. A. 279, 347–354 (2001)
15. Gu, H., Yang, M., Li, L., Liu, Z., Ren, W.: Dynamics of autonomous stochastic resonance in neural period adding bifurcation scenarios. Phys. Lett. A 319, 89–96 (2003)
16. Neiman, A., Saparin, P.I., Stone, L.: Coherence resonance at noisy precursors of bifurcation in nonlinear dynamical systems. Phys. Rev. E 56, 270–273 (1997)
17. Shi, X., Lu, Q.-S.: Coherence resonance and synchronization of Hindmarsh-Rose neuron model with noise. Chinese Physics 14(6), 1088–1094 (2005)
18. Namachchivaya, N.S., Ariaratnam, S.T: Stochastically perturbed Hopf bifurcation. International Journal of Nonlinear Mechanics 22(5), 363–372 (1987)
19. Manjarrez, E., Rojas-piloni, J.G., Mendez, I., Martinez, L., Velez, D., Vazquez, D., Flores, A.: Internal stochastic in the coherence resonance between spinal and cortical neuronal ensembles in the cat. Neuroscience Letters 326, 93–96 (2002)
20. Yang, H.J., Hu, S.J., Han, S., et al.: Relation between responsiveness to neurotransmitters and complexity of epileptiform activity in rat hippocampal CA1 neurons[J]. Epilepsia 43(11), 1330–1336 (2002)
21. Hu, S.J., Yang, H.J., Jian, Z., Long, K.P., Duan, Y.B., Wang, Y.H., Xing, J.L., Xu, H., Ju, G.: Adrenergic sensitivity of neurons with non-periodic firing activity in rat injures dorsal root ganglion. Neuroscience 101(3), 689–698 (2000)
22. Tan, N., Xu, J.X., et al.: The bifurcation mechanism arousing the phenomenon of ”sensitivity of non-periodic activity” in neurons. Acta. Biophysica Sinica 19(4), 395–400 (2003)
23. Hu, G.: Stochastic resonance without external periodic force. Phys. Rev. Lett. 71, 807–810 (1993)
24. Pikovsky, A.S., Kurths, J.: Coherence resonance in a noise- driven excitable system. Phys. Rev. Lett. 78, 775–778 (1997)
25. FitzHugh, R.: Biological Engineering, edited by H.P. Schwann. In: Nagumo, J., Arimoto, S., Yoshizawa, S. (eds.) Pro.ZRE, vol. 50, p. 2061. McGraw-Hill, New York (1962)

An Integrated Approach for Reconstructing Surface Models of the Proximal Femur from Sparse Input Data for Surgical Navigation

Guoyan Zheng and Miguel A. González Ballester

MEM Research Center, Institute for Surgical Technology and Biomechanics
University of Bern, Stauffacherstrasse 78, CH-3014, Bern, Switzerland
Guoyan.Zheng@MEMCenter.unibe.ch,
Miguel.Gonzalez@MEMCenter.unibe.ch

Abstract. A patient-specific surface model of the proximal femur plays an important role in planning and supporting various computer-assisted surgical procedures including total hip replacement, hip resurfacing, and osteotomy of the proximal femur. The common approach to derive 3D models of the proximal femur is to use imaging techniques such as computed tomography (CT) or magnetic resonance imaging (MRI). However, the high logistic effort, the extra radiation (CT-imaging), and the large quantity of data to be acquired and processed make them less functional. In this paper, we present an integrated approach using a multi-level point distribution model (ML-PDM) to reconstruct a patient-specific model of the proximal femur from intra-operatively available sparse data. Results of experiments performed on dry cadaveric bones using dozens of 3D points are presented, as well as experiments using a limited number of 2D X-ray images, which demonstrate promising accuracy of the present approach.

Keywords: surface reconstruction, surgical navigation, X-ray, point distribution model, statistical shape analysis.

1 Introduction

A patient-specific surface model of the proximal femur plays an important role in planning and supporting various computer-assisted surgical procedures including total hip replacement, hip resurfacing, and osteotomy of the proximal femur. The common approach to derive 3D models of the proximal femur is to use imaging technique such as computed tomography (CT) or magnetic resonance imaging (MRI). However, the high logistic effort, the extra radiation associated with the CT-imaging, and the large quantity of data to be acquired and processed make them less functional. The alternative presented here is to reconstruct the surface model using sparse input data consisting of dozens of surface points (e.g. 50 points) or a limited number of calibrated fluoroscopic images (e.g. 2 to 4 images).

Constructing an accurate 3D surface model from sparse input data is a challenging task. Additionally, inherent to the navigation application is the high accuracy and

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 767–775, 2007.

robustness requirements. When surface reconstruction is used for the purpose of surgical guidance, it requires that the algorithm satisfies the following criteria: (a) accurate geometrical information about the underlying anatomical structure can be derived from the reconstructed surface model; (b) target reconstruction error of the reconstructed surface model should be in the range of surgical usability, which is typically in the area of 1.5 mm average error (2 to 3 mm worst case) [1]; (c) 95% success rate is normally required, when an appropriate initialization is given [1]; (d) minimal user interaction during data acquisition and algorithm execution is highly appreciated for a sterilized environment; and (e) the algorithm should be robust to outlying data. In the present paper, we try to solve the problem in an accurate and robust way. At the heart of our approach lies the combination of sophisticated surface reconstruction techniques and a multi-level point distribution model (ML-PDM) of the target anatomical structure.

The paper is organized as follows. Section 1 reviews the related work. Section 2 presents the construction of the ML-PDM of the proximal femur. Section 3 describes the integrated approach combining our previous works on 3D-3D surface reconstruction [2, 3] and those on 2D-3D surface reconstruction [4, 5]. Experimental results using both 3D sparse point set as well as 2D X-ray images are described in Section 5, followed by conclusions in Section 6.

2 Related Works

Statistical shape analysis [6, 7, 8] is an important tool for understanding anatomical structures from medical images. A statistical model givens an effective parameterization of the shape variations found in a collection of sample models of a given population. Model based approaches [9, 10, 11] are popular due to their ability to robustly represent objects. Intraoperative reconstruction of a patient-specific model from sparse input data can be potentially achieved through the use of a statistical model. Statistical model building consists of establishing legal variations of shape from a training population. A patient-specific model is then instantiated through fitting the statistical model to intraoperatively acquired data. Thus, the aim of the statistical instantiation is to extrapolate from sparse input data a complete and accurate anatomical representation. This is particularly interesting for minimally invasive surgery (MIS), largely due to the operating theater setup.

Several research groups have explored the methods for reconstruction a patient-specific model from a statistical model and sparse input data such as digitized points [2, 3, 12, 13, 14, 15], a limited number of calibrated X-ray images [4, 5, 16, 17, 18, 19, 20], or tracked ultrasound [21, 22, 23, 24]. Except the method presented by Yao and Taylor [17], which depends on a deformable 2D/3D registration between an appearance based statistical model [25] and a limited number of X-ray images, all other methods have their reliance on a point distribution model (PDM) in common. In Fleute and Lavallée [12], a statistical shape model of the distal femur was fitted to sparse input points by simultaneously optimizing both shape and pose parameters. Their technology has been incorporated into a system for computer-assisted anterior cruciate ligament surgery and preliminary results were published in [13]. Chan et al. [21, 22] used a similar algorithm, but optimized the shape and pose parameters separately. Tracked ultrasound was used as the input in their work to instantiate 3D surface models of the complete femur and pelvis from their associated statistical shape models. Following the seminal work of Blanz and Vetter for the synthesis of 3D faces using a morphable model [26], Rajamani et al. [14, 15, 23] incorporated a Mahalanobis prior for a robust and stable surface model instantiation. In our recent work [2, 3], we proposed to use the dense surface point distribution model (DS-PDM) and a reconstruction scheme combining statistical instantiation and kernel-based deformation for an accurate and robust reconstruction of a patient-specific surface model of the proximal femur from dozens of points. This reconstruction scheme has also been combined with a novel 2D-3D correspondence establishing algorithm [27] for reconstructing surface model of the proximal femur from a limited number of calibrated X-ray images [4, 5].

3 Multi-level Point Distribution Model Construction

The ML-PDM used in this paper was constructed from a training database consisting of 30 proximal femoral surfaces from above the lesser trochanter. In the coarsest level, a sequence of correspondence establishing methods presented in [28] was employed to optimally align surface models segmented from CT volume. It started with a SPHARM-based parametric surface description [29] and then was optimized using Minimum Description Length (MDL) based principle as proposed in [30].

Following the alignment, the PDM in this level is constructed as follows. Let $\mathbf{x}_i = (p_0, p_1, \ldots\ldots, p_{N-1})$, $i = 0, 1, \ldots, m-1$, be m (here m = 30) members of the aligned training surfaces. Each member is described by a vectors $\mathbf{x}_i$ with N (N = 4098) vertices:

$$\mathbf{x}_i = \{x_0, y_0, z_0, x_1, y_1, z_1, \ldots, x_{N-1}, y_{N-1}, z_{N-1}\} \tag{1}$$

The PDM is obtained by applying principal component analysis on these surfaces.

$$\begin{aligned} D &= ((m-1)^{-1}) \cdot \sum\nolimits_{i=0}^{m-1} (\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})^T \\ P &= (\mathbf{p}_0, \mathbf{p}_1, \ldots); \qquad D \cdot \mathbf{p}_i = \sigma_i^2 \cdot \mathbf{p}_i \end{aligned} \tag{2}$$

Where $\overline{\mathbf{x}}$ and D represents the mean vector and the covariance matrix respectively.

Then, any one of the instance in this space can be expressed as:

$$\mathbf{x} = \overline{\mathbf{x}} + \sum\nolimits_{i=0}^{m-2} \alpha_i \mathbf{p}_i \tag{3}$$

And the estimated normal distribution of the coefficients $\{\alpha_i\}$ is:

$$p(\alpha_0, \alpha_1, \ldots, \alpha_{m-2}) = (2\pi)^{-\frac{m-1}{2}} \cdot e^{-\frac{1}{2}\sum_{i=0}^{m-2}(\alpha_i^2 / \sigma_i^2)} \tag{4}$$

Where $\sum\nolimits_{i=0}^{m-2} (\alpha_i^2 / \sigma_i^2)$ is the Mahalanobis distance defined on the distribution.

The vertices for constructing the denser point distribution model in a finer resolution are then obtained by iteratively subdividing the aligned surface model in the coarser resolution. The basic idea of subdivision is to provide a smooth limit surface model which approximates the input data. Starting from a mesh in a low resolution, the limit surface model is approached by recursively tessellating the mesh. The positions of vertices created by tessellation are computed using a weighted stencil of local vertices. The complexity of the subdivision surface model can be increased until it satisfies the user's requirement.

In this work, we use a simple subdivision scheme called *Loop scheme,* invented by Loop [31], which is based on a spline basis function, called the three-dimensional quartic box spline. The subdivision principle of this scheme is very simple. Three new vertices are inserted to divide a triangle in a coarse resolution to four smaller triangles in a fine resolution.

Loop subdivision does not change the positions of vertices on the input meshes. Furthermore, positions of the inserted vertices in a fine resolution are interpolated from the neighboring vertices in a coarse resolution. As the input surface models have been optimized for establishing correspondences, it is reasonable to conclude that the output models are also aligned. Principal component analysis can be applied on these dense surface models to establish a dense surface point distribution model (DS-PDM). In our previous work [2], we found that a single level subdivision is enough for our purpose, which results in 16386 vertices for each surface model. We thus created a two-level point distribution model (TL-PDM).

4 The Integrated Surface Model Reconstruction Approach

Based on the two-level point distribution model, we developed an integrated surface model reconstruction approach which can seamlessly handle both 3D sparse points and a limited number of X-ray images. When a set of 3D points are used, the fine level point distribution model (FL-PDM) will be chosen, which facilitates the point-to-surface correspondence establishment. But if the input is a limited number of calibrated X-ray images, we will use the coarse level point distribution model (CL-PDM) to speed up the computation. For completeness, we will briefly present these two methods below. Details can be found in our previously published works [2, 3, 4, 5].

3D-3D Reconstruction Method [2, 3]: The reconstruction problem is formulated as a three-stage optimal estimation process. The first stage, *affine registration,* is to iteratively estimate the scale and the 6 degree-of-freedom rigid transformation between the mean shape of the PDM and the sparse input data using a correspondence building algorithm and a variant of iterative closest point (ICP) algorithm [32]. The estimation results of the first stage are used to establish point correspondences for the second stage, *statistical instantiation*, which optimally and robustly instantiates a surface model from the PDM using a statistical approach [14]. The instantiated surface model is taken as the input for the third stage, *deformation*, where the input surface is further deformed by an approximating thin-plate spline (TPS) based vector transform [33] to refine the statistically instantiated surface model.

2D-3D Reconstruction Method [4, 5]: Our 2D-3D reconstruction approach combines statistical instantiation and regularized shape deformation as described above with an iterative image-to-model correspondence establishing algorithm [27]. The image-to-model correspondence is established using a non-rigid 2D point matching process, which iteratively uses a symmetric injective nearest-neighbor mapping operator and 2D thin-plate splines based deformation to find a fraction of best matched 2D point pairs between features detected from the fluoroscopic images and those extracted from the 3D model. The obtained 2D point pairs are then used to set up a set of 3D point pairs such that we turn a 2D-3D reconstruction problem to a 3D-3D one, which can be solved by the 3D-3D reconstruction approach as described above.

5 Experimental Results

We designed and conducted experiments on 18 cadaveric femurs (Note: none of them has been included for constructing the TL-PDM) with different shape to validate the present integrated approach. 3D point sets as well as calibrated X-ray images were used.

Reconstruction error measurement: To quantify the reconstruction error, Target Reconstruction Error (TRE) was used. The TRE is defined as the distance between the actual and the reconstructed position of selected target features, which can be landmark points or bone surfaces themselves.

Validation experiments: Two different types of sparse data were explored in two experiments: (1) using clinically relevant sparse points directly acquired from the surfaces of 7 cadaveric femurs; (2) using a limited number of calibrated fluoroscopic images of the other 11 cadaveric femurs. To evaluate the reconstruction error, we acquired another set of landmarks from each surface of the dry cadaveric femurs (Please note that these landmarks were not used in the reconstruction procedures but for the accuracy evaluation purpose). Then, the TRE was measured by calculating the distances between these landmarks and the associated reconstructed surface model.

Results of the first experiment: Total hip replacement and hip resurfacing procedures operated with posterior approach were identified as the potential clinical applications. At one stage of such surgeries, after internal rotation and posterior dislocation of the hip, most of the femoral head, neck, and some part of trochantric and intertrochantric (crest and line) regions are exposed [34]. Obtaining sparse surface points from these intraoperatively accessible regions and reconstructing a patient-specific 3D surface model of the proximal femur with reasonable accuracy will be useful for the above mentioned surgeries. In this experiment, one set of 50 points was used to reconstruct the surface model of each cadaveric bone and the other set consisted of 200 points was used to evaluate the reconstruction errors. The results of surface reconstruction using clinically relevant sparse points are presented in Fig. 1. For each case, the overall execution time was less than one minute.

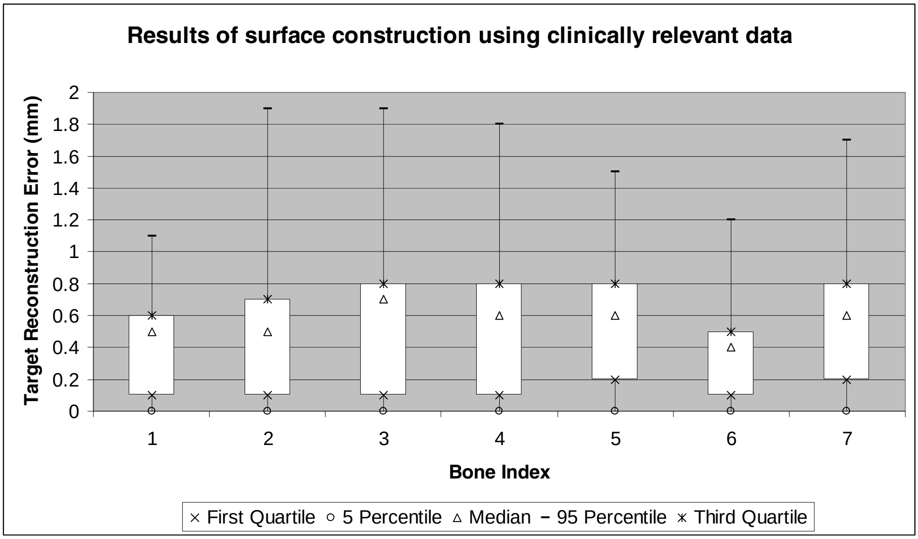


Fig. 1. Errors of reconstructing surface models of seven cadaver femurs using clinically relevant data

Results of the second experiment: In this experiment, two studies using different number of images were performed for each bone. In the first study two images acquired from anterior-posterior (AP) and lateral-medial (LM) directions were used to reconstruct the surface model of each cadaveric femur. In the second one, an image acquired from oblique direction was additionally used together with the above mentioned AP and LM images.

The reconstruction accuracies were evaluated by randomly digitizing 100 – 200 points from each surface of the cadaveric specimen and then computing the distance from those digitized points to the associated surface reconstructed from the images. The median and mean reconstruction errors of both experiments are presented in Table I. An average mean reconstruction error of 1.2 mm was found when only AP and LM images were used for each bone. It decreased to 1.0 mm when three images were used. Different stages of one reconstruction example are presented in Fig. 2.

Table 1. Reconstruction errors when different number of images were used

Reconstruction errors when only AP and LM images were used for each bone											
Bone Index	1	2	3	4	5	6	7	8	9	10	11
Median (mm)	1.3	0.8	1.5	1	1.3	1	1.1	1	0.8	1.1	1.2
Mean (mm)	1.5	0.8	1.4	1.3	1.4	1.2	1.2	1.2	1	1.1	1.6
Reconstruction errors when all three images were used for each bone											
Bone Index	1	2	3	4	5	6	7	8	9	10	11
Median (mm)	1.3	0.7	0.7	1.1	1	1.1	0.8	0.9	0.7	1	0.9
Mean (mm)	1.3	0.7	0.8	1.2	1.1	1.1	1.1	0.9	0.9	1.1	1.2

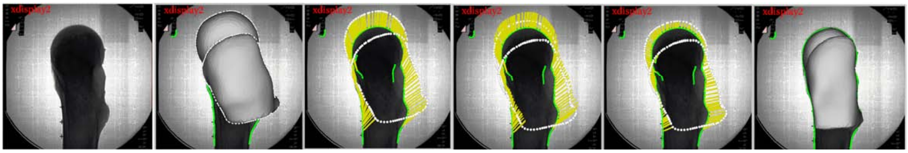


Fig. 2. Different stages of reconstruction. First: one of the acquired images. Second: the initialization of the mean model of the PDM. Third: after establishing image-to-model correspondence. Fourth: after 3D paired point matching. Fifth: after re-establishing correspondence; Sixth: the final reconstruction result after a series of computations.

6 Conclusions

We have presented an integrated approach using the ML-PDM for robust and accurate anatomical shape reconstruction from sparse input data. Based on the modalities of the input data, the appropriate level point distribution model was used. In this approach, the 3D-3D reconstruction problem is formulated as a three-stage optimal estimation process. In each stage, the best result is optimally estimated under the assumption for that stage, which guarantees a topologically preserved solution when only sparse points are available. The FL-PDM is employed in all stages to facilitate the correspondence establishment. When a limited number of calibrated X-ray images are used, the CL-PDM is employed to speed up the computation. A 2D-3D correspondence establishing algorithm based on a non-rigid 2D point matching process is applied to convert a 2D-3D problem to a 3D-3D one.

The proposed approach is generic and can be easily extended to other rigid anatomical structures, though in this paper we only demonstrate its application for reconstructing surface models of the proximal femur.

References

1. Livyatan, H., Yaniv, Z., Joskowicz, Z.: Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT. IEEE T Med Imaging 22(11), 1395–1406 (2003)
2. Zheng, G., Rajamani, K.T., Nolte, L.-P.: Use of a dense point distribution model in a three-stage anatomical shape reconstruction from sparse information for computer assisted orthopaedic surgery: a preliminary study. In: Narayanan, P.J., Nayar, S.K., Shum, H.-Y. (eds.) ACCV 2006. LNCS, vol. 3852, pp. 52–60. Springer, Heidelberg (2006)
3. Zheng, G., Dong, X., Nolte, L.-P.: Robust and accurate reconstruction of patient-specific 3D surface models from sparse point sets: a sequential three-stage trimmed optimization approach. In: Yang, G.-Z., Jiang, T., Shen, D., Gu, L., Yang, J. (eds.) Proceedings of the 3rd International Workshop on Medical Imaging and Augmented Reality (MIAR). LNCS, vol. 4091, pp. 68–75. Springer, Heidelberg (2006)
4. Zheng, G., Nolte, L.-P.: Surface reconstruction of bone from X-ray images and point distribution model incorporating a novel method for 2D-3D correspondence. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2006) vol. 2, pp. 2237–2244 (2006)

5. Zheng, G., Gonzalez Ballester, M.A., Styner, M., Nolte, L.-P.: Reconstruction of patient-specific 3D bone surface from 2D calibrated fluoroscopic images and point distribution model. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 25–32. Springer, Heidelberg (2006)
6. Dryden, I.L., Mardia, K.V.: Statistical Shape Analysis. John Wiley and Sons, New York (1998)
7. Kendall, D.: A survey of the statistical theory of shape. Statistical Science 4(2), 87–99 (1989)
8. Small, C.: The statistical Theory of Shape. Springer, Heidelberg (1996)
9. Turk, M., Pentland, A.: Eigenfaces for recognition. Journal of Cognitive Neuroscience 3(1), 71–86 (1991)
10. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models – their training and application. Computer Vision and Image Understanding 61(1), 38–59 (1995)
11. Corouge, I., Hellier, P., Gibaud, B., Barillot, C.: Interindividual functional mapping: a nonlinear local approach. Neuroimage 19, 1337–1348 (2003)
12. Fleute, M., Lavallée, S.: Building a complete surface model from sparse data using statistical shape models: application to computer assisted knee surgery system. In: Wells, W.M., Colchester, A.C.F., Delp, S.L. (eds.) MICCAI 1998. LNCS, vol. 1496, pp. 879–887. Springer, Berlin (1998)
13. Fleute, M., Lavallée, S., Julliard, R.: Incorporating a statistically based shape model into a system for computer-assisted anterior cruciate ligament surgery. Medical Image Analysis 3(3), 209–222 (1999)
14. Rajamani, K.T., Styner, M., Joshi, S.C.: Bone model morphing for enhanced surgical visualization. In: Proceedings of the 2004 IEEE International Symposium on Biomedical Imaging (ISBI): From Nano to Macro, pp. 1255–1258 (2004)
15. Rajamani, K.T., Gonzalez Ballester, M.A., Nolte, L.-P., Styner, M.: A novel and stable approach to anatomical structure morphing for enhanced intraoperative 3D visualization. In: Galloway Jr, R.L., Cleary, K.R. (eds): Proceedings of the SPIE Medical Imaging: Visualization, Image-guided Procedures, and Display. vol. 5744, pp. 718–725 (2005)
16. Fleute, M., Lavallée, S.: Nonrigid 3-D/2-D registration of images using statistical models. In: Taylor, C., Colchester, A.C.F. (eds.) Medical Image Computing and Computer-Assisted Intervention – MICCAI'99. LNCS, vol. 1679, pp. 138–147. Springer, Heidelberg (1999)
17. Yao, J., Taylor, R.H.: Assessing accuracy factors in deformable 2D/3D medical image registration using a statistical pelvis model. In: Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV 2003), vol. 2, pp. 1329–1334. IEEE Computer Society (2003)
18. Lamecker, H., Wenckebach, T.H., Hege, H.-C.: Atlas-based 3D-shape reconstruction from X-ray images. In: Proceedings of the 18th International Conference on Pattern Recognition (ICPR 2006), vol. 1, pp. 371–374 (2006)
19. Benameur, S., Mignotte, M., Parent, S., Labelle, H., Skalli, W., de Guise, J.A.: 3D/2D registration and segmentation of scoliotic vertebrae using statistical models. Computerized Medical Imaging and Graphics 27, 321–337 (2003)
20. Benameur, S., Mignotte, M., Labelle, H., de Guise, J.A.: A hierarchical statistical modeling approach for the unsupervised 3-D biplanar reconstruction of the scoliotic spine. IEEE Transactions on Biomedical Engineering 52(12), 2041–2057 (2005)
21. Chan, C.S.K., Edwards, P.J., Hawkes, D.J.: Integration of ultrasound-based registration with statistical shape models for computer-assisted orthopaedic surgery. In: Sonka, M., Fitzpatrick, J.M. (eds.): Proceedings of SPIE Medical Imaging 2003: Image Processing, vol. 5032, pp. 414–424 (2003)

22. Chan, C.S.K., Barratt, D.C., Edwards, P.J., Penney, G.P., Slomczykowski, M., Carter, T.J., Hawkes, D.J.: Cadaver validation of the use of ultrasound for 3D model instantiation of bony anatomy in image guided orthopaedic surgery. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) MICCAI 2004. Part II, LNCS, vol. 3217, pp. 397–404. Springer, Heidelberg (2004)
23. Talib, H., Rajamani, K.T., Kowal, J., Nolte, L.-P., Styner, M., Gonzalez Ballester, M.A.: A comparison study assessing the feasibility of ultrasound-initialized deformable bone models. Computer Aided Surgery 10(5/6), 293–299 (2005)
24. Talib, H., Zheng, G., Rajamani, K.T., Zhang, X., Styner, M., Gonzalez Ballester, M.A.: Feasibility of 3D ultrasound-initialized deformable bone-modeling. In: Langlotz, F., Ellis, R., Davis, B.L. (eds.) Proceedings of CAOS 2006, pp. 519–522 (2006)
25. Yao, J., Taylor, R.H.: Tetrahedral mesh modeling of density data for anatomical atlases and intensity-based registration. In: Delp, S.L., DiGoia, A.M., Jaramaz, B. (eds.) MICCAI 2000. LNCS, vol. 1935, pp. 531–540. Springer, Berlin (2000)
26. Blanz, V., Vetter, T.: A morphable model for the synthesis of 3D faces. In: Proceedings of the 26th Annual Conference on Computer Graphics, SIGGRAPH'99, pp. 187–194 (1999)
27. Zheng, G.: A novel 3D/2D correspondence building method for anatomy-based registration. In: Pluim, J.P., Likar, B., Gerritsen, F.A. (eds.) WBIR 2006. LNCS, vol. 4057, pp. 75–83. Springer, Heidelberg (2006)
28. Styner, M., Rajamani, K.T., Nolte, L.-P., Zsemlye, G., Székely, G., Taylor, C.J., Davis, R.H.: Evaluation of 3D correspondence methods for modeling building. In: Taylor, C.J., Noble, J.A. (eds.) IPMI 2003. LNCS, vol. 2732, pp. 63–75. Springer, Heidelberg (2003)
29. Brechbuehler, C., Gerig, G., Kuebler, O.: Parameterization of Closed Surfaces for 3D Shape Description. Comput Vision and Image Under 61, 154–170 (1995)
30. Davies, R.H., Twining, C.H., et al.: 3D statistical shape models using direct optimization of description length. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002. LNCS, vol. 2352, pp. 3–20. Springer, Heidelberg (2002)
31. Loop, C.T.: Smooth subdivision surfaces based on triangles. M.S. Thesis, Department of Mathematics, University of Utah (August 1987)
32. Besl, P., McKay, N.D.: A method for registration of 3D shapes. IEEE Transaction on Pattern Analysis and Machine Intelligence 14(2), 239–256 (1992)
33. Bookstein, F.: Principal warps: thin-plate splines and the decomposition of deformations. IEEE Transaction on Pattern Analysis and Machine Intelligence 11(6), 567–585 (1989)
34. Moreland, J.R.: Primary Total Hip Arthroplasty. In: Chapman, M.W. (ed.) Operative Orthopaedics, 1st edn., vol. 1, pp. 679–693. JB Lippincott, Philadelphia, PA (1988)

Part IV

Industrial and Ergonomic Applications

This page intentionally blank

Future Applications of DHM in Ergonomic Design

Heiner Bubb

Lehrstuhl für Ergonomie, Technische Universität München
bubb@lfe.mw.tum.de

Abstract. Until now DHMs are especially used to design the dimensions of products and production assembly according to anthropometric demands. Recently DHMs are additionally equipped with strengths simulation so that also the dimensioning of reaction forces is possible. First steps are done to describe and evaluate the contact between human body and environment. Some examples will be shown. However in this area further important steps are necessary. Especially the self paced calculation of posture depending on this contact is to be realized. Some proposals exist for the contact of seat and body. Also first basic research is done in order to simulate motion behavior. Especially the detection of "leading body elements" as basic idea for this simulation can be seen as an initial step to generate modeling of cognitive human properties. However, in order to realize it the simulation of the properties of sense organs is necessary. Certain properties of the eyes can be simulated rather simple. Meanwhile some experience exits to understand the glance behavior depending on specific tasks (e.g. car driving). That can serve as basic for input to cognitive models. The output of these can be the track in space of the leading body element. On the other hand sensor organs properties in the field of hearing and climate are possible. In both cases the more difficult problem is to simulate the properties of the environment. General application field of these future development is the computer aided ergonomic design of workplaces in production lines and of products especially vehicles already in the definition and development phase. In this connection is to be considered that in future especially the design of information flow in these areas becomes dominant. An example is the growing development of assistance systems in cars. The application of DHMs will allow achieving the connection between information design and the necessary geometric design of the equipment.

1 Introduction

Since the middle of the sixties digital human models have been developed in order to simplify the anthropometric lay-out of technical equipments. During this development we can observe two important lines:

- In the one line especially mechanical and physical properties are represented by computer programs. The original aim was to make predictions of the weightless human body in the space. Later on, this kind of simulation was used to calculate the behavior of dummies during a crash experiment in the car development. By this approach a part of the very expensive crash experiments should be substituted by

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 779–793, 2007.

the cheaper computer simulation. An example of industrial applied DHMs of this kind is the MADYDYMO, developed by TNO in the Netherlands.

- The other line is the development of anthropometric models that should represent the geometric and strength like human proportions. They are used to design the geometric layout of working places and especially narrow cabins as they are given in every kind of vehicle. The today most important models of this kind are JACK, developed in the USA, SAFEWORK, developed in Canada and nowadays incorporated in the CAD-tool CATIA V, and RAMSIS, a software tool developed by the enterprise Human Solutions and the Technical University Munich encouraged by the community of the German automotive industry. The last one is meanwhile world wide well established as tool for the packaging lay-out in the automotive industry. Further the new development SANTOS is to be mentioned. It is the project "Virtual Soldier" that is initiated and supported partly by the US Army TACOM project 'Digital Humans and Virtual Reality for Future Combat Systems (FCS)' and by the Caterpillar Inc. project 'Digital Human Modelling for Safety and Serviceability.' According to his authors (Abdel-Malek et.al, 2006) it promises to be the next generation of virtual humans.

The following paper will deal especially with the development of RAMSIS, as it generally represents the development of such software dummies. A further reason is that the author of this article knows and pushes this development as one of the researchers for this project. Of course, the trends demonstrated by this example can also be observed in connection with the other models, however in different expression.

2 Examples of Application

In order to understand the necessary future development some typical examples of today application of such DHMs are presented firstly in the following. In connection with working places in the production very often questions of reachability arise. The very good geometric representation of DHMs allows treating such problems with certain accuracy (See Figure 1) especially with view on the different anthropometries of the workers. However, the effort to create such pictures is rather height and until now no enough accurate posture and motion models exist to create a "natural" posture under all conditions. The good representation of the anthropometric variety allows designing even safety critical cabins. For example a special safety handle for roller coasters was designed that can be adapted to all appearance of body shape from small children up to fat big male people (see Figure 2). However, until now the most realistic representation of human behavior is given for the driver's posture in a passenger car. For this purpose e.g. in the case of RAMSIS extended research was carried out.

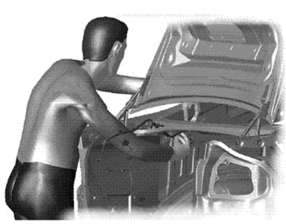


Fig. 1. Investigation of reachability with the help of DHMs in connection with questions of production

As reported by Bubb et al. (2006) instead of the usual inverse kinematic the most probable posture is

calculated by an optimalization algorithm that is based on experiments of the distribution of the angles in all joints of the human body during car using. By a regression appropriation even the discomfort caused by the necessary external conditions – the so-called restrictions - can be predicted. The correspondence between the calculated posture and real posture was evaluated by many studies (e.g. Kolling, 1997). This system can be used to adapt the so-called packaging of cars to
the variety of human's body size und to find out the optimal adjust equipment that is necessary to realize this aim (see Mergl et al., 2006).

During all applications the contact between the DHM and the environment is a big problem. Especially the pre-calculation of the contact in seats is of main interest. Different approaches to solve this problem are reported. In many cases the properties of the seat are described by the H-Point, measured by the H-point machine after SAE-J 826. In the case of RAMSIS actual seat experiments with subjects of different size are carried out in order to find out the offset between the H-Point that describes the properties of the seat and the hip-joint axis of the DHM RAMSIS. Ippili et al. (2002) describe an algorithm based on a simplified model, by which posture and force distribution can be calculated depending on the foam characteristics of the seat. A similar, however 3-D approach is reported by Schmale et al. (2002) that even provide a calculation of the

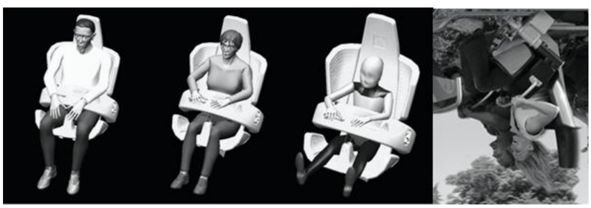


Fig. 2. Design of a rollercoaster seat by RAMSIS and the realization "Sky Wheel" by Maurer & Söhne

pressure distribution in the contact plane between the DHM COSYMAN and a parameterized seat. Mergl (2006) presents a FEM model of the deformability of the upper leg. It shows good agreement with observed pressure distributions of subjects. In connections with a posture prediction model so the pressure distribution in a virtual seat can be calculated. By Mergl et al. (2005) a discomfort model was developed by which this pressure distribution can be evaluated according to the expected discomfort. By all these measures the discomfort model based only on joint angels is enlarged and improved.

However, is this enough? All the described models cannot predict an arbitrary posture under every thinkable condition. It is only valid for sitting postures, especially driver sitting postures. As in future not only forces but also motion and recognition by the sense organs become important in connection with CAD applications, the modeling of human properties and behavior is to be started by deliberations that consider physiological and psychological knowledge from the beginning.

3 New DHM Developments

3.1 Primary Deliberations for New DHM Developments

A general human model that is very established in the area of ergonomic application divides between

- information reception, realized by the sense organs,
- information processing, realized by the central nerve system, especially the brain, and
- information transformation, i.e. motion by innervations of muscles.

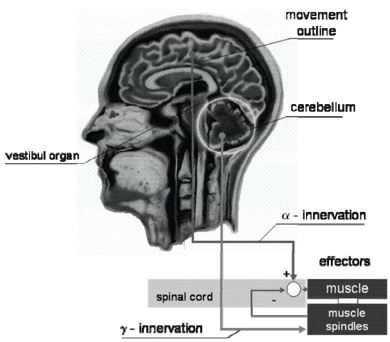


Fig. 3. The physiologic background of information reception, information processing, and motion innervations

Figure 3 shows the physiologic background of this categorization. By the receptor system external stimuli are transferred to nervous stimuli. The pattern of these is analysed by the complex and hyper-complex cells. The combined information of all sense organs stimulates memory contents. That means we combine information of the various sense organs to a singular recognition of the external world without direct conscious recognizing from what sense organ the information comes. The action that has been apparently successful in connection with this stimulus configuration is also a part of these. The action is realized by connecting the desired action to the effectors, i.e. the musculature. From the motor-sensory cortex this information runs via the so-called α-innervations to the spinal cord and excite there the so-called α-motor-neurons. These are directly connected to the corresponding muscles, the effectors. By the muscle spindles the expansion of the muscle is "measured" and fed back to the α-motor-neurons. By this closed loop system in the spinal cord a posture desired by the brain and signalled by the α-innervations is realized. In a special case this is also known as "knee-jerk". However, there exists a second way to initiate a motion: by the so-called γ-innervations the length of the muscle spindles can be influenced and so a motion can be produced. By these γ-innervations on the one side an adaptation to fine regulation is possible and on the other side the body equilibrium in the field of gravity is guaranteed. The reason is: γ-innervations are especially connected with the cerebellum and this has essential connections to the vestibule organ.

The task of human modelling is to rebuild these natural given conditions. In this connection two questions are to be answered in the first step and modelled in the second step:

- On the psychological level: "what is in a given situation the desired movement?"
- On the physiological level: "By what interaction of muscles is the motion realised?"

In the following first the physiological level is to be considered because it is the fundament of the psychological level that describes the intentional "willing" of the model.

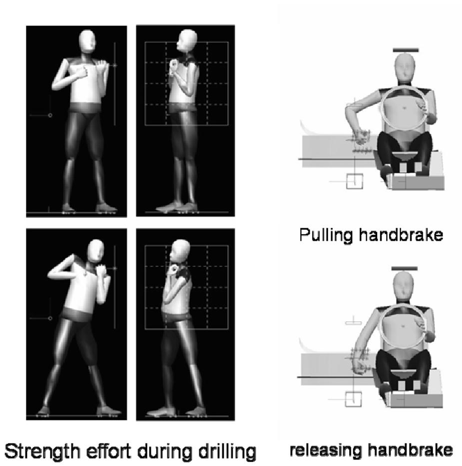


Fig. 4. Examples for the poster calculation by the FOCOPP-Model

By Seitz and Recluta (2005) the Force-Controlled Posture Prediction-Model FOCOPP was developed. Integral part of this approach is an accurate physical description of test subjects. The physical characterization includes anthropometry, masses of body parts, centre of gravity, angle dependent resistant torques and maximum torques for each joint of the body. This approach assumes that humans want to minimize the joint strain when taking a desired posture. This approach has been formulated within a MATHEMATICA and RAMSIS simulation and exemplary evaluated for different tasks (see Figure 4). By the experimental experience this model was confirmed. Furthermore, the individual load and discomfort based on the found correlation between load and discomfort assessment can be predicted (Zacher and Bubb, 2005). Of course, the gravity field is a further part of this calculation.

The generated approach for optimization considers the human biomechanical characteristics by connecting the effective moments inside a joint to relative strain. The posture optimization consists in the change of the degrees of freedom in different joints in that way that the strain is minimized. The modelled parameters of influence are maximal forces and passive resistance moments effected by the muscular system. In order to describe the passive resistance joint moments a measuring device was developed to analyze postures in a quasi-weightless state (Marach, 1999). Maximal forces were measured by Schaefer et al. (2000) and used for the model. The corresponding posture angles were measured by the contact-less stereo-photogram-metrical posture analyzing system PCMAN (Seitz, 1998).

3.3 Psychological Level

Arlt (1999) developed a new model that can acquire dynamic constraints for any human motion by external parameters. Than by applying the FOCOPP-Model the posture of the remaining body elements can be calculated. In order to generate dynamic constraints a theory of human motion from the field of neurophysiology was combined with psychological discoveries and adjusted to the problem of the use within the digital human model. The result is: for every movement exists a „leading body element (e.g. hand, finger, foot, buttocks etc.). The path of it defines additional “dynamic restrictions”. For this three control components exist:

- Transportations component: it defines the course in space and time between start-point and target-point. As could be shown in different experiments the moving path always is within a plane defined by the start and target point. The run of the path is like a parable and defined by the detaching and approaching vector and depends on the kind of grasping.
- Aligning component: the leading body element is to be aligned to the shape of start and target object. The detaching and approaching vector depends on the kind of grasping given by the form of the corresponding objects.
- Optical component: In order to adapt the motion to the external conditions information about spatial destination, position, and form of the aimed object must be obtained.

In order to model motion it is to be distinguished between:

1. “Conducted motions”: the corresponding leading body part (for example foot) is quasi fixed to an object, whose motion is controlled through technical parameters (for example pedal).
2. “Simple perfect motions”: the motions trajectories can be fully and autonomously found by the person. They are characterised by the detaching from an object and by the reaching of a target-object (for example from the steering wheel to the gear lever).
3. “Modified motions” are “perfect motions“, which are disturbed by the avoidance of an obstacle.
4. “Complex motions” are distinguished through the correct time-co-ordination of several simple perfect and modified motions (for example the process of ingress in a motor car).

Furthermore Arlt found that in the case of obstacles (“modified motion”) the human operator keeps a safety distance to the obstacle. This safety distance depends on the body element (it is very low for hand and foot and biggest for the head) and it depends on the moving speed. With increasing speed the distance becomes lower.

4 Modeling

4.1 Modeling and Evaluation of Motion

Cherednichenko et al. (2006) enlarged the model of Arlt by investigation of “complex motions” in the context of entering a car. By this he made an important step to a general modelling of motion. By experimental observation and by the attempt to model this experience he defined three main levels:

- Planning level: on this level the target of the motion is defined.
- Guiding level: on this level the motion more in detail is planned considering all conditions of environment and obstacles. An essential part of this level is the backward planning.

- Stabilization level: on this level the motion becomes reality. All autonomous reactions of the body, by which equilibrium and reduced effort is realized, becomes now important.

It is important to mention that these three levels seem to be of general importance for human activities. Also in the case of driving a motor car we know these three levels (see 4.2). Cherednichenko especially modelled the guidance level. An essential part of it is the definition of the leading body elements and the co-ordination of different body elements. As already found by Arlt also Cherednichenko could define experimentally the leading body element by the fact that the motion of it is always exactly in a plane. It is of interest that always only one leading body element exists. That seams to be attributed to the restriction of the active brain's capacity. However, in the case of a complex motion several leading body element are activated according to a motion plan.

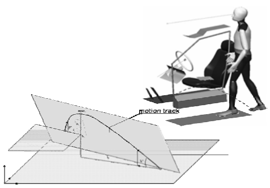


Fig. 5. Prediction of the motion of the leading body element in the case of entering a car

In the case of the ingress manoeuvre the leading body elements are

1. the right foot (in the case of steering wheel on the left side!),
2. the pelvis,
3. adjustment of the right foot,
4. adjustment of the left foot,
5. hands to the steering wheel.

Depending on conditions in the environment (e.g. position of the door sill, position of the A- and B-column etc.) he found rules, by which this plane is defined. Now, depending on the obstacles in the environment the exact motion track can be calculated (see Figure 5). On the stabilization level by use of the FOCOPP-Model the posture and motion of the rest of the body elements are calculated. It could be shown that by this procedure a good agreement between observed motion and calculated motion is achieved.

The next step is to evaluate such motion according to the expected discomfort. In combination with other discomfort factors, like body posture or seating pressure (see above), the overall discomfort may be predicted at any time instantly. This is an essential step introducing psychophysical relations to get on man's emotional world for further technical operations. The idea is to "explain" human behavior by some kind of "hedonic optimization", not in a physical but in a pure emotional world (Schaefer and Zacher, 2006). In the end, this approach may help to identify those postures and movements that are reducing perceived discomfort to a minimum, possibly ending up in some kind of autonomously moving manikins.

In order to draw a conclusion from the static discomfort model to a dynamic model it is important to measure movements and to analyze them in connection with the evaluation of discomfort. Therefore experiments were conducted with the aim to approve an approach for the dynamic modelling of discomfort. The subjects had to lift different weights on a rack positioned in various heights. After the movement the subjects were asked about their global discomfort feeling and the discomfort in defined body parts and joints (e.g. shoulder, elbow). The results assert that the maximum discomfort in one body part or joint is the same as the global discomfort. A movement analysis was conducted with PCMAN, so that the time histories of all joints were known.

Using the multi body system simulation software ALASKA with the human model DYNAMICUS the time dependant joint torques were calculated. If it is possible to adapt the static discomfort model to the dynamic requirements concerning the maximum force, it will be possible to calculate the development of the discomfort during a movement in every joint (dynamic model). Knowing the connection between the local and global discomfort one will also be able to state a value for the global discomfort for a complete movement or a sequence of movement. Thus, improved conditions for the analysis and assessment of discomfort in the area of ergonomics and design can be created. The background idea for this procedure is presented in more detail by Fritzsche (2007) in this proceeding. Figure 6 shows the idea of this evaluation approach in an overview.

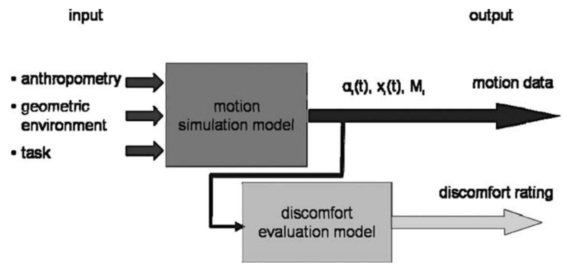


Fig. 6. General simulation model of motion and evaluation of motion

4.2 Modeling of Information Perception

The next important step to come to a complete human model is the modeling of the sense organs. Of course with view on application not the physiological properties of these sense organs are to be modeled but only their properties under the view of information transformation. The most properties of the sense organs are for a modeling rather well known enough. The much bigger problem is the modeling of the environment. For example for a pre-calculation of the acoustic field not primary the properties of ears are important but the exact pre-calculation of the sound field. The same calls for the thermal environment and for the influence of vibration (beside of the reaction of the human body on it, which is modeled rather well, e.g. by Verver, 2002 and Fritzsche, 2005). Under the view of human modeling especially the modeling of glance behavior and visibility is of importance. Although it is no problem to simulate the view from the two dummy eyes in the virtual world and thereby to evaluate the visibility of the environment e.g. depending on different anthropometries, it is a big problem to visualize the effect of ametropia and presbyopia. A further difficulty is to pre-calculate the effect of reflection and glare. In this case again a good modeling of the surrounding surfaces is necessary. By software dummies it is possible to calculate the sight conditions and hidden surfaces with respect to the variation of anthropometry. In this connection the challenge is to present the result of such calculations in understandable form. A main problem however is to calculate the glance behavior. In experiment together with DaimlerChrysler the viewing areas during driving were investigated. For this purpose a special experimental car was developed, the so-called glass dome car, by which under total unrestricted conditions the glance behavior in normal traffic situations could be investigated (see Figure 7). Beside the trivial result that, of course, the main glance direction is oriented in the area ahead the car, specific areas during curve driving could be detected. A further result was that the driver shows a strong inclination to keep the glance through the windscreen. As far as possible he avoids the glance through the side windows. That has impact to the design of position and form of the A-columns.

Fig. 7. The “glass-dome-car” of DC and a comparison between glance behavior and hidden viewing areas

The course of the road, vehicles and other traffic participants, as well as environmental conditions and weather determine the driving task. To accomplish this task the driver has to fulfil three hierarchically ordered tasks (Bernotat, 1970, Allen et. al., 1971 and many others): By the navigation task in the external loop he lays down generally, how the existing road network should be connected in order to reach the desired destination. This is the input for the task in the next lower loop, the guidance task. On this level the driver determines the desired course that decide in the immediate surrounding field of ca. 200 m how the vehicle should move in location and time. On the lowest level, the stabilization task, the driver detect the necessity information from the visual environment und transforms it in adequate handling of the control elements so that the actual lateral and longitudinal position of the car corresponds with a certain tolerance to the desired position, the so-called “result”. It should be mentioned that this process corresponds totally to the three levels of body motion initiating, described in 4.1. Altogether, the driving task can be reduced to the demand: every contact with standing or moving objects in the traffic field has to be avoided.

However, the description above does not mention all tasks that are to be accomplished during driving. According to a proposal of Geiser (1985) the total driving task can be divided in the following subtasks. The *primary driving task* is the task that is described already above. It is the actual driving process that aims to keep the car on the road.

Beside of that the driver has to accomplish additional tasks. Those arise in the framework of the primary driving task depending on traffic and environmental conditions as well as those that are necessary in order to inform other traffic participants about intended manoeuvres. For example the operation of the flasher, the wiper, the lamp switch, the horn, and in the case of a hand switched car the use of gear lever and clutch pedal belong to these tasks. Also assistance systems like Tempomat or ACC-system have to be operated on this level. All these control elements have to be used always depending directly or indirectly on the driving task. Therefore they are assigned as *secondary tasks*. In a further breakdown they can be divided in *reactive tasks* and *active tasks*. The reactions on a change of external conditions like to dip the headlight in the case of oncoming traffic or to switch the wipers in the case of rainfall are reactive tasks. Principally they can be automated. By active tasks the driver shows his intention (e.g. by using the horn).

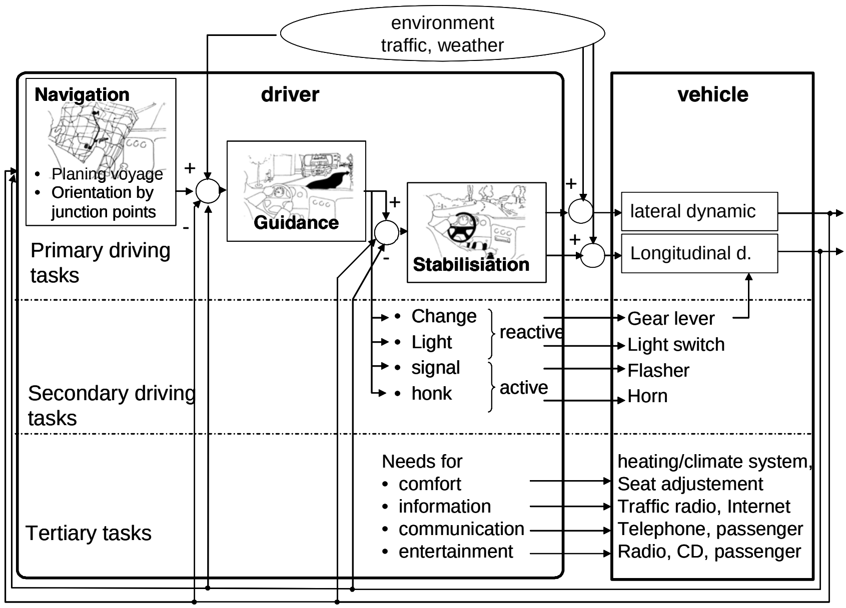


Fig. 8. Closed Loop Driver - Car

Tasks that have nothing to do with the driving tasks but aim to improve the needs of comfort, entertainment, and information are called *tertiary tasks*. The use of the heating/climate system, the radio, the telephone, and in further future additional equipments like internet and communication with office and home technologies belong to these tasks. Also in this case it can be divided between active and reactive tasks. Figure 8 shows a structure image of the driver's tasks.

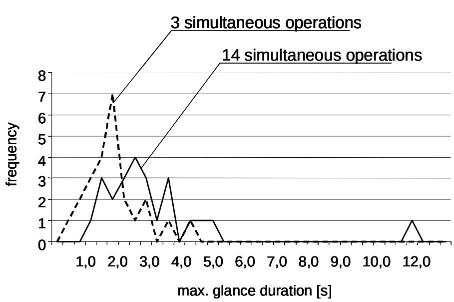


Fig. 9. Glance attention in connection with tertiary tasks

Rassl (2005) investigated the influence of tertiary task and different layout of such tasks on the glance behaviour. Beside others his subjects had to choose one of 3, 5, 8, and 14 options by a central control element (similar to the BMW i-drive controller) during driving. It is of interest that the selection needs in average about 1.2 seconds and there is no significant difference between these deviation times depending on the number of presented options. However, when we look at the maximum duration of distracted glances the case of 14 options is significantly different from the others. The maximum distraction time in this case is on average 2.2 seconds. As Figure 9 shows, in this case one time even a distraction time of 12 seconds was observed. That is not a singular event! In the research of Rassl in another part of the experiment even 16 seconds distraction was observed and in further not jet published experiments with the operation of the climate system we also found distraction times of 12 seconds. By other experiments Treugut (2007) could show that the glance duration to a tertiary task depends on the position of the indicator or the screen in the environment of the driver. When this position is closer to the view direction on the road, the driver keeps the glance longer on the display because he has the impression to keep the road in view in the periphery. It is an actual question, where is the optimum between a certain peripheral perception, however inexact observing of the main task - it is the driving task - and the more urgent impulse to turn the glance back to the main task after a duration of normally 2 seconds.

The experience of such experiments can be modelled and in not too far future it will be a further simulation part of DHMs.

4.3 Modeling of Cognitive Behavior

Whereas presently - caused by the technical development – the load of information processing by secondary and especially tertiary tasks and the design of the equipment for it is object of actual research, the investigation and design of the primary task is the core of the design that promise a reduction of accident frequency and an improved handling of the car. It makes necessary to model the core of cognitive behaviour. We have to answer the question: how does the human information processor work in connections with driving tasks? Answering this question also other interaction between the human operator and the technical equipment can be predict.

In general glance behaviour is one key to consider how information gets into the brain. As by the fovea centralis of the eyes only a small angle sphere of about 2° - 3° can be resolved sharply, the eye must scan sequentially the scene, in order to receive the situation. Therefore, the strategy of this scan behaviour depends on the attractiveness of the objects in the environment. So, the glance behaviour can be seen as a peer scope into the internal information processing behaviour. However, any changing stimulus that reaches the peripheral sphere of the eyes effects immediately to direct the glance at the moving object.

When we observe glance behaviour by an eye tracking system during driving we can distinguish between “scanning” and “processing” (Cohen, 1976) Scanning glances have a rather short duration of averagely 400 ms. By scanning especially the edges of the road, other traffic participants and traffic signs are received. Processing means that special - so-called - areas of interest (AO’s) are fixated. That can be the instruments, the mirror, and the display of the car but also objects in the environment, which are not relevant for the traffic. The processing glances need on average twice of the time of scanning glances (see Figure 10).

By scanning internal models are stimulated that are stored as a general concept in form of a structural engram in the long term memory. By this stimulation the

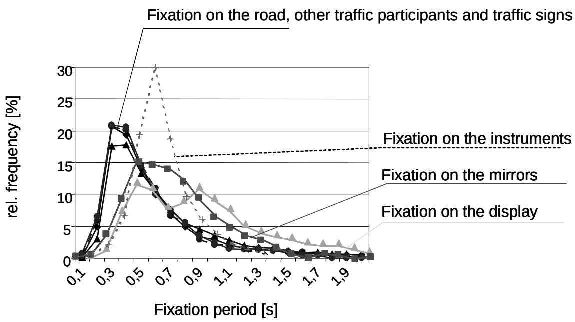


Fig. 10. Experimental results of eye tracing researches (Schweigert, 2003)

corresponding internal model is "waked up" and becomes in this way a part of the working memory. This process shall be understood more exactly by the example of driving on a road bended to the left. The information received by the scanning stimulates a general internal model of a concept of left bended road. By adapting this concept to the real stimuli in the working memory the driver recognizes the real width and the real curve of this road. Further scanning stimuli give him information about traffic participants on this road. The scanned information of these objects stimulates also internal models of their behaviour. By this way only a short glance is enough to recognize the speed and the course of an oncoming car or the expected behaviour of a foot passenger. As the glance can only scan the scene in a sequence, it is possible that relevant objects are not observed. The combination of all this information gives the driver a feeling of the presence of the situation. He thinks only this internal image as reality.

Felt present has duration of 2 to 3 seconds (Pöppel, 2000). This has big importance for the glance behaviour. All experiments with glance behaviour in car driving show that normally only in a distance of 1 to 1.5 seconds ahead information is scanned (distance = speed x preview time; Donges,1978; Yuhara et al., 1999; Guan et al., 2000, Schweigert, 2003). And there are a lot of experimental results that show that the driver accepts normally to take the view from the road for up to two seconds (e.g. Zwahlen et al. 1988, Gengenbach, 1997, Schweigert, 2003).

This felt present encourages the driver to keep not always the glance on the road. Schweigert (2003) carried out glance research in enlarged experiments. He found out that the glance was directed also to non traffic relevant objects in 89% of all observed traffic situations under normal non stressing conditions. When the traffic situation becomes more difficult, the driver reduce first the glance on these non traffic relevant objects, than to specific areas of interest, than he reduce the attention to traffic signs, tachometer and mirrors. When the situation becomes very complex it can be observed that the driver in 41% of all cases counts on other traffic participant's behaviour according to the rules. When the situation becomes still more complex in 7 % of the situations the driver omits even the glance to primary necessary information.

All these experiments allow step by step a modeling of human behavior. The control of glance behavior by internal models enables modeling more deeply the information receiving process. However, the same models are connected with the information realization by innervations of muscles. This process can be modeled by the idea of the leading body element. Some ideas from the modeling of human behavior with methods of control theory, as it was done in the research field round McRuer in the 60^{th} and 70^{th}, can help to predict the dynamic properties more exactly. So, knowing the technical design of given equipment represented in CAD, it will be possible to judge what information the operator can take in view and what is probably neglected. In connection with existing methods in this area the human error probability may be estimated. All taken together a prediction of the usability already on the basis of CAD becomes attainable. Figure 11 shows a survey on such a future model for that actually research is done.

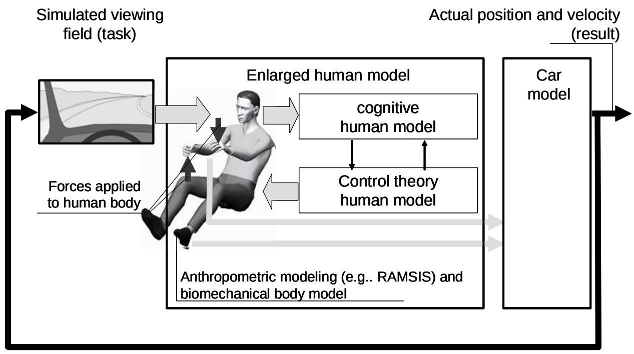


Fig. 11. Illustration of the combination of cognitive and control theory based human model

5 Conclusion

It was shown that the actual application of DHM is characterized by gravity in the area of anthropometry. Immediate research is done for applying forces and predicting simple motion. More complex motions give already the connection to modeling cognitive behavior. An essential concept in this connection is the idea of the leading body element. This represents so to say the connection between information processing and information transformation. Research of glance behavior is concerned with the connection between information reception and information processing. In the course of more and more possibilities to model this behavior the total process between information input and information output of the human operator can be represented by computer based models. That can serve as basis for “CAD experiments”, by which the probable human behavior in context of a virtual task and a virtual environment can be calculated. So the influence of changing task and/or environment on the reliability of the human operator could be tested in a very early and not jet cost intensive phase of development. However, it is to be stated that such virtual experiments can done only in connection with few well known and well investigated tasks. An example is car driving. Presently in the context of the excellence research project “CoTeSys” of the Technical University Munich beside this mentioned example also for some application in assembly line and in the production industry such a modeling of human behavior is set in realizing.

References

1. Abdel-Malek, K., Yang, J., Marler, T., Beck, S., Mathai, A., Zhou, X., Patrick, A., Arora, J.: Toward a new generation of virtual humans. International Journal of Human Factors Modeling and Simulation 1(1), 2–39 (2006)
2. Allen, R.W., Jex, H.R., Magdaleno, R.E.:: Manual Control Performance and Dynamic Response during Sinusoidal Vibration. Wright-Patterson AFB, AMRL-TR-73-78 (1973)
3. Arlt, F.: Untersuchung zielgerichteter Bewegungen zur Simulation mit einem CAD-Menschmodell, Dissertation am Lehrstuhl für Ergonomie der Technischen Universität München, 1999 (1999)
4. Bernotat, R.: Anthropotechnik in der Fahrzeugführung. Ergonomics 13(3) (1970)
5. Bubb, H., Engstler, F., Fitzsche, F., Mergl, C., Sabbah, O., Schaefer, P., Zacher, I.: The development of RAMSIS in past and future as an example for the cooperation between industry and university. International Journal of Human Factors Modeling and Simulation 1(1), 140–157 (2006)
6. Cherednichenko, A., Assmann, E., Bubb, H.: Experimental Study of Human Ingress Movements for Optimization of Vehicle Accessibility. In: Proceedings IEA Congress Maastricht 2006 (2006)
7. Cohen, A.S.: Augenbewegungen des Autofahrers bei Vorbeifahren an unvorhersehbaren Hindernissen und auf freier Strecke. Zeitschrift für Verkehrssicherheit 22 (1976)
8. Donges, E.: Ein regelungstechnisches Zwei-Ebenen-Modell des menschlichen Lenkverhaltens im Kraftfahrzeug. Zeitschrift für Verkehrssicherheit 24, S. 98–112 (1978)
9. Fritzsche, F.: Prediction of Discomfort during Arm Movements. In: Proceeding of the HCI-Congress, 2007, Beijing (2007)
10. Fritzsche, F.: Software-Schwingungs-Dummy. In: FAT-Bericht 189: Darstellung des Schwingungsverhaltens von Fahrzeug-Insassen. Symbiose aus Experiment und Simulation. Frankfurt (2005)
11. Geiser, G.: Mensch-Maschine-Kommunikation im Kraftfahrzeug. ATZ 87, 77-84 (1985)
12. Gengenbach, R.: Fahrerverhalten im Pkw mit Head-Up-Display. Gewöhnung und visuelle Aufmerksamkeit. VDI Fortschrittsbereichte Reihe 12: Verkehrstechnik/Fahrzeugtechnik. Nr. 330 (1997)
13. Guan H., Gao, Z., Guo, K., et al.: An Optimal Preview Acceleration Model for Velocity Control in Driver-Vehicle-Environment System. In: Proceeding of AVEC 2000, pp. S. 525–530, Ann Arbor (2000)
14. Ippili, R.K., Davies, P., Bajaj, A.K.: Prediction of Static Settling Point in a Seat Occupant System. VDI Bericht Nr. 1675. In: Digital Human Modeling Conference, Munich, 2002, pp. 377–404 (2002)
15. Jürgensohn, T.: Bedienermodellierung. In: Timpe, K.-P., Jürgensohn, T., Kolrep, H. (Hrsg.). Mensch-Maschine-Systemtechnik - Konzepte, Modellierung, Gestaltung, Evaluation, 2000, Symposion Publishing, Düsseldorf (2000)
16. Kolling, J.: Validierung und Weiterentwicklung eines CAD-Menschmodells für die Fahrzeuggestaltung, Dissertation am Lehrstuhl für Ergonomie der Technischen Universität München (1997)
17. Lindsay, P.H., Norman, D.H.: Human Information Processing. In: An Introduction to Psychology, Academic Press, New York, London (1972)
18. Marach, A.: CAD-fähiges Prognosemodell für statische Körperhaltungen des Menschen, Dissertation am Lehrstuhl für Ergonomie der technischen Universität München (1999)
19. Mergl, C., Bubb, H., Vogt, C., Kress, H.: Ergonomic Layout Process for a Driver Working Place in Cars. In: DHMC 2006 (2006)

20. Mergl, C., Klendauer, M., Mangen, C., Bubb, H.: Predicting long term riding comfort in cars by contact forces between human and seat. In: DHMC 2005 (2005)
21. Mergl, C.: Entwicklung eines Verfahrens zur Optimierung des Sitzkomforts auf Automobilsitzen, Dissertation am Lehrstuhl für Ergonomie der Technischen Universität München (2006)
22. Pöppel.: Grenzen des Bewusstseins - Wie kommen wir zur Zeit und wie entsteht Wirklichkeit? Insel Taschenbuch, Frankfurt a. M. und Leipzig (2000)
23. Rasmussen, J.: The definition of human error and a taxonomy for technical system design. In: Rasmussen, J., et al. (Hrsg.): New technology and human error, Wiley & Sons Ltd, New York (1987)
24. Rassl, R.: Ablenkungswirkung tertiärer Aufgaben im Pkw - Systemergonomische Analyse und Prognose. Dissertation an der Technischen Universität München (2005)
25. Schaefer, P., Rudolph, H., Schwarz, W.: Digital man models and physical strength - a new approach in strength simulation. In: SAE Proceedings: digital human modelling for design and engineering, Dearborn (2000)
26. Schaefer, P., Rudolph, H., Schwarz, W.: Variable Force Limits for Optional Target Populations - a New Approach realized in CEN-Standardization. In: Proceedings of the 13th IEA, Tampere, vol. 4, S. 533–535 (1997)
27. Schaefer, P., Zacher, I.: On the way to autonomously moving manikins - empowered by discomfort feelings. In: submitted to the proceedings of the 16th IEA, Maastricht (2006)
28. Schmale, G., Stelzle, W., Kreienfeld, T., Wolf, C.-D., Härtel, T., Jödicke, R.: COSYMAN, A Simulation Tool for Optimization of Seating Comfort in Cars. VDI Bericht Nr 1675, In: Digital Human Modeling Conference, Munich 2002, pp. 301–311(2002)
29. Schweigert, M.: Fahrerblickverhalten und Nebenaufgaben Dissertation an der Technischen Universität München (2003)
30. Seitz.: The optical measurement system PCMAN. In: Proceedings of the workshop on 3D Anthropometry and industrial product design, Université Paris V (1998)
31. Seitz, T., Recluta, D., Zimmermann, D.: An Approach for a Human Posture Prediction Model using Internal/external Forces and Discomfort. In: Proceedings of the SAE Conference, 2005 (2005)
32. Shannon, C.E., Weaver, W.: The Mathematical Theory of Communication. University of Illinois Press, Urbana (1949)
33. Steward, T.B., Ferrel, W.R.: Human Perception of Angular Acceleration and Implications in Motion Simulator. J. Aircraft 8, Mo. 4 (1971)
34. Treugut, F.: Untersuchung des Einflusses visueller Aufmerksamkeitsverlagerung auf die Sicherheit der Fahrzeugführung; Diplomarbeit an der Rhodes University, Grahamstown, Südafrika (2007)
35. Verver, M.M., van Hoof, J.: Vibration Analysis with MADYMO Human Models. VDI Bericht Nr. 1675. In: Digital Human Modeling Conference, Munich, 2002, pp. 447–456 (2002)
36. Yuhara, N., Horiuchi, S., Asanuma, N., et al.: Improvement of Vehicle Handling Quality Through Active Control of Steering Reaction Torque. In: Proceeding of AVEC 1992, Jokohama, S. 407–412 (1992)
37. Zacher, I., Bubb, H.: Ansatz zur Entwicklung eines Diskomfortmodells für Bewegungen, In: Proceedings of the Spring University, 2005, Balatonfüred, Hungary (2005)
38. Zwahlen, H.T., Adams Jr, C.C., DeBald, D.P.: Safety aspects of CRT touch panel controls in automobiles. In: Gale, A.G. (Hrsg.) Vision in Vehicles II. S.335/344, Elsevier North Holland Press, Amsterdam (1988)

The Design and Exploitation of Visual Feedback System for Rowing

Chunmei Cao[1], Chuncai Wang[1], Linhong Ji[1], Zixi Wang[1], and Xiaoping Chen[2]

[1] Department of Precision Instruments and Mechanology, Tsinghua University
[2] Institute of Physical Education, Tsinghua University, 100084 Beijing, P.R. China

Abstract. Based on Neural Mechanism theory of short-term memory, a predigested model was established. It demonstrated learning, establishment and strengthening process concerning with the motor pattern domination of high-level nerve center system. According to the model three feedback loops and their dialectic correlation were synthetically analyzed during rowing athletes studied the technique in training. The analyzed results provided the technical requests and measuring parameters of the system, then, a visual feedback system for rowing was designed and developed. This system has been primarily used in the national rowing team.

Keywords: visual feedback, Neural Mechanism, rowing.

1 Introduction

Rowing is one of the most important items in athletic sports. Its characteristic is that the athlete impels the boat in virtue of oars. The environment of training and competing for rowing is in largo fields out of door and the athletes scatter everywhere, which makes the coaches' exact observation of the action of athlete difficult. Also, there is a limitation to use different instruments on the largo fields. All of above increase the difficulty of training for coaches and athletes. How could we provide more effective evaluating parameters for coaches, which made them give more veracious guidance? Many researchers investigated it in theory and experiment in- and outland. Yoshiga [1] and Desgorces [2] evaluated the rowing performance of female and male rowers with regard to their body size and observed the influence of anthropometrics characteristics on rowing performance at national and international level. Galand [3] determined the pacing strategies adopted by elite rowers in championship 2000-m races. Hill [4] evaluated different critical velocity models applied to rowing ergometry and to investigate prediction of performance time in a 2000-m race based on results from shorter trials. Baudouin and Hawkins [5] investigated the influence of biomechanical factors on rowing performance. Zheng Weitao [6], [7], [8] et al tested the hydrodynamic performance of typical scull under different oblique angles, depths and angles of attack in circle water channel. In the training aiding aspect, Shao Guihua [9] constituted a technique expert system of rowing whose main function was to answer the technique questions on rowing from abecedarian. Zheng Weitao [10] at al developed a Visualization Testing and Training

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 794–802, 2007.

System of rowing which could be not only used as specific physical training but also tested and diagnosed the technique of rowing. Ou Yangbo [11] at al manufactured a Specific Strength Training System on Land Rowing which could simulate the motor pattern on water.

Although the investigations above could provide some help for the training of rowing, there are a few shortages in which the athletes could not see their training state real-timely.

2 System Designing

2.1 Theory

The paddling technique of athlete in each period in the process of rowing had its own neural control which lay in alliums and cerebellum. According to Neural Mechanism theory of short-term memory: for the great mass of nerve center system except for the motor neuron in spinal cord, if these nerve cells were stimulated for several seconds, they would appear irritability buildup from several second to several hours; if these nerve cells were stimulated during the period of irritability buildup once again, they would appear more forceful irritability buildup than the last time. Apparently this kind of remembrance rests with the change of excitability in interrelated nerve cells, and we could describe a predigested model which was about how the motor pattern dominated by high-level nerve was established, learned and strengthened, shown in figure 1. This model concluded two loops. The one was described in real line which showed the process how the controlling pattern of nerve center system was founded: when the consciousness dominated the movement, cerebra calculated the behavior signal that arrived at the motor aim through the motor aim and the state at present. The behavior signals were translated into nerve impulsion by behavior encoder, and then the nerve impulsion were transmitted to motor system by sequence. Executive results were fed back to nerve center system by different feels which were used to adjust the process of behavior. The other one was described in broken line which showed the process how the controlling pattern of nerve center system was learned and strengthened: at the end of each action evaluation organ evaluated all the behavior and gave out the executive results according as motor aim. The nerve center system decided the next feedback was positive or negative on the base of executive results. When the executive result was bad the nerve center system would choose negative feedback and restrain the behavior signals. Because the nerve center system didn't know which kind of pattern was the best behavior signal, it would attempt to choose another kind of pattern. When the executive result was good the nerve center system would choose positive feedback and enhance the behavior signals. The nerve center system would strengthen nerve impulsion concerning this kind of behavior signal in the next period. The process of how the movement control was learned and strengthened was the process of that the nerve center system enhanced and restrained the activity of some kind of nerve cell. After one behavior occurred, the contact of nerve cells would be reserved for some time. If in this period of time the contact of nerve cells was strengthened continually, the pattern of motor which was controlled by this contact would become main motor pattern.

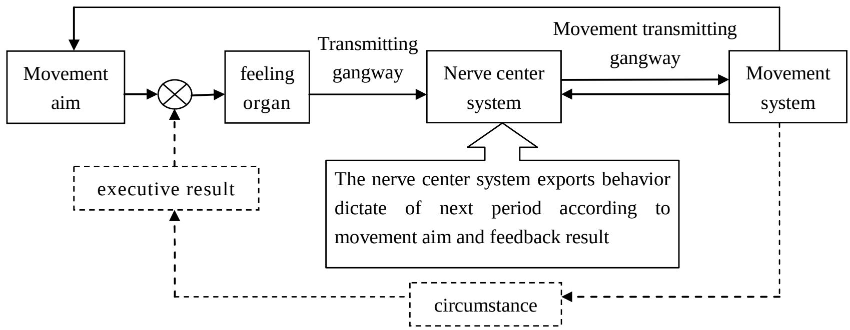


Fig. 1. A model of the motor pattern dominated by high-level nerve established, learned and strengthened

Analyzing the pattern of nerve center system dealing with information through the process of the athlete's training, we could find there were two biofeedback loops, shown in figure 2. The one was composed of the athlete's noumenal feeling and nerve center system: nerve center system of the athlete evaluated his behavior himself according to the motor result after the motion, estimated the dispersion between the motor aim and the motor result, and then gave motor orders adjust motor pattern to minish that dispersion. This loop was in dominant station if the coaches did not participate in the process. The other loop was made up of the guidance of coach and the athlete's nerve center system: after the athlete finished the motion, the coach evaluated the dispersion between the motor aim and the motor result, guesstimated the reason of the dispersion, and then told the athlete how to ameliorate. Receiving the information from coach the nerve center system of the athlete gave out new motor orders in term of the information. This loop usually was in dominant station in the process of coach giving guidance, because the athlete believed that the evaluation of coach was more veracious than the one of himself.

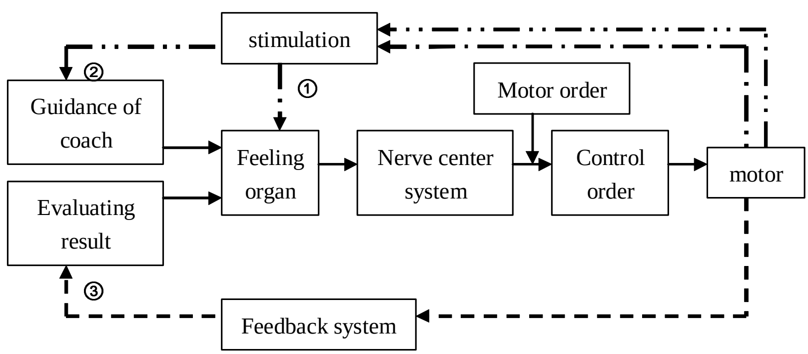


Fig. 2. Three feedback loops in rowing training

The two biofeedback loops above could only offer the qualitative result for the athlete because of the limit of the evaluation method. All results were blurry in most time, so athlete could not determinately know to strengthen or to restrain current motor pattern. The use of visual feedback system for rowing increased the third real-time feedback loop. This visual feedback system must have some technical qualifications if it could be brought into play.

2.2 Design Demand of System

Through the academic analysis above we could conclude that the visual feedback system for rowing must have the technical qualifications as follows so that it could be brought into play in used.

The feedback information must be reliable. Human nerve biology system had the characteristic of plasticity and the ability of adaptability to condition stimulation. Whether the result fed back to the athlete could be effective during the process that the motor pattern came into being, it greatly depended on the degree of the athlete's credit. The athlete chose to use the visual feedback system in daily training only when the information of the system fed back to athlete was more veracious than the result that the athlete evaluated himself. Therefore this characteristic of the biofeedback system requested the parameter measuring subsystem of the visual feedback system should work steadily at the same time could survey the parameters well and truly.

Evaluation result must be real-time. Short-timely after the athlete finished motion, the visual feedback system brought the evaluation result to the athlete. At this time the memory trail of that motion still existed in the nerve center system, so it was easy to strengthen or restrain the motor signal. This characteristic of the biofeedback system requested the parameter evaluation subsystem of the visual feedback system could measure and evaluate real-timely and we should shorten the time lapse to the best of our abilities.

The format of feedback must be compendious and intuitionistic. This would make athlete expediently and duly get the feedback result in any condition at any time. Further more because the best form of feedback was the firsthand physical stimulation the information had better been transmitted to athlete directly, that was we should not use many medial configuration. This characteristic of the biofeedback system requested the feedback subsystem of the visual feedback system could feed back information directly to athlete and not been obstructed.

The process can not affect the training pattern of athletes. The intention of biofeedback was to improve the training efficiency of athlete, so when the athlete made use of biofeedback system he should not change his daily training pattern and the technique motion. This characteristic of the biofeedback system requested the parameter evaluation subsystem and the feedback subsystem of the visual feedback system could not affect the training pattern of athletes and not change the technique motion.

The process can not aggrandize burden in athletes' mentality. Because the biofeedback did the deed through the biologic process of human nerve center system, the burden in athletes' mentality would sometimes affect the efficiency of biofeedback. This characteristic of the biofeedback system requested the feedback stimulation of the visual feedback system could not affect the training pattern of athletes and make the athlete think the stimulation in the nature of things.

2.3 Choosing Feedback Parameters

According the design demand of biofeedback system, parameter choosing and measuring well and truly were the kernel variable in the visual feedback system. We should think over conditions as follows in designing parameters.

1. **The speed of boat:** If the athlete wanted to get good performance, he must get across the stated distance in shortest time, which was the row must be the highest speed during the grand distance. In term of nowadays training condition in china, coach gave guidance that instruct athlete's technique motion each period mainly according to the judgments about the technique parameters such as speed of row from eyes. There was not a real-time apparatus which could help coach watch and control the speed of row on water training. In the process of daily training athlete estimated the speed of row each period by his feeling. Because the information from human feelings was subjective and qualitative it was difficulty for athlete to give veracious and impersonal estimation in a period (1.5 s). In the process of judgments the experience of coach took the main effect. Therefore veracious data of the speed of boat was necessary, and the speed of boat became the main parameter in parameter measuring subsystem beyond all doubt.
2. **The postural parameters of boat:** Resistance from water was another important factor that affected the speed of boat. During the process of movement there were resistances from air, friction, mucosity and wave. In most time the three fore resistances were aptotic. If the athlete wanted to enhance the speed of boat he must ameliorate and optimize his technique of paddling in order to minish the resistance of wave. Therefore the postural parameters (pitching angle and obliquity) of boat also were important parameters for the training.
3. **The frequency of paddling:** The frequency of paddling meant the times of paddling in one minute. In biomechanics, the speed of boat was decided by the frequency of paddling and the distance of one paddling, so if the athlete wanted to enhance the speed of boat he had two ways, one was to increase the frequency of paddling and the other one was to prolong the distance of one paddling. In certain apparatus, these two factors were mainly influenced by the technique of paddling. The times of paddling in each period was not boundless, furthermore the athlete could not depress the effect of paddling in order to increasing the frequency of paddling. It was necessary for athlete to watch and control the frequency of paddling in games. In the process of daily training monitoring the frequency of paddling was also important because different frequency of paddling inflected the different intensity in training.

From the analysis above we concluded three main parameters such as the speed of boat, the postural of boat and the frequency of paddling.

2.4 The Project of Design

According the investigation on theory of biofeedback above, the project of the visual feedback training system for rowing was composed of three sections, shown in Figure 3. The first section was pivotal parameter measuring subsystem, whose main function were measuring, analyzing, evaluating and storing the main parameters data of boat in the process of training, at the same time transmitting respectively the evaluation results to the feedback subsystem for athlete and the monitoring subsystem for coach. The second section was feedback subsystem, whose main function was displaying the evaluation results from pivotal parameter measuring subsystem to athlete. The third section was monitoring subsystem, whose main function was display the evaluation results from pivotal parameter measuring subsystem to coach, so that the coach gave guidance to athlete according to the evaluation results.

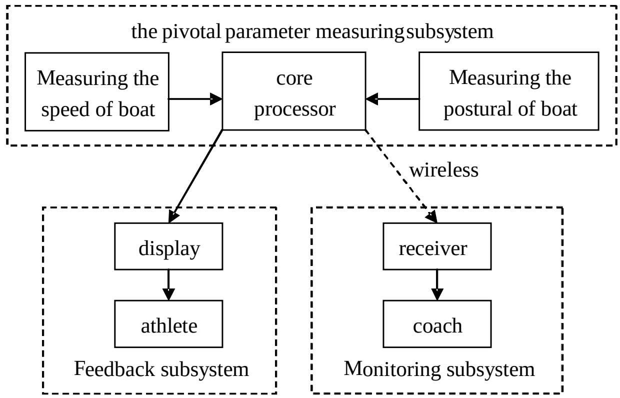


Fig. 3. The project of visual feedback system

2.5 Checking Up Reliability of the System

The real-time speed of boat was the core parameter in the content of the visual feedback training system. On the analysis of the characteristics of the circumstance and motion during rowing training, the sensor used in the system must possess the

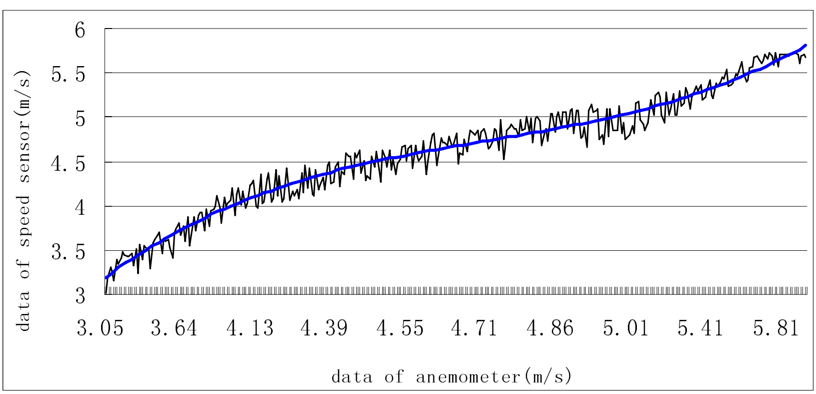


Fig. 4. Checking up reliability of the speed sensor

characteristics of the high precision (0.1m/s), light weight(<50g) and big moving space(500m~8000m). In surveying all of the speed sensors in market, we could not find any one fit these characteristics. So we devised a new kind of speed sensor on the theory of press in hydrodynamics. After the design finished we used anemometer to checkout and test the precision of the speed sensor, shown in Figure 4. The x-axis denoted the data of anemometer and the y-axis denoted the data of speed sensor we machininged. From Figure 4 we could get that the error of the speed sensor did not exceed 5%, which accorded with the demand of design.

2.6 The Real-Time Design of the System

The evaluation results were displayed on the screen at the beginning of pulling oar, so that athlete could see the results last period during the process of pulling oar. During the pressing oar and drawing oar, the sensorium and motor-center of athlete focused on the force between water and oars, however the process of lifting oar was short when the distance between athlete and screen was long, so at this time the athlete could not see the results syllabify. Only in the process of pulling oar, the body and nerve center system of athlete was correspondingly relaxative, and it was the optimal time for athlete to see the evaluation results.

3 System Application

When the design and exploitation of the visual feedback training system finished and the precision of speed sensor had been checked up, the visual feedback training system were brought to national team. Some elite rowers used this system in daily training.

3.1 Experimental Object

In November 2006 the system was tested in Water Training Base in Qiandao Lake in Hangzhou City. Seven elite rowers used the system in their daily training. The basic information of these rowers was as follows, shown in Table 1:

Table 1. The basic information and testing times of rowers

athlete	sex	stature(cm)	weight(kg)	age	level	testing times
Athlete 1	man	196	93	21	master sportsman	4
Athlete 2	man	193	96	22	master sportsman	1
Athlete 3	woman	175	58	28	master sportsman	1
Athlete 4	woman	177	59	20	master sportsman	1
Athlete 5	woman	180	78	21	master sportsman	1
Athlete 6	woman	181	74	26	master sportsman	1
Athlete 7	man	200	95	24	master sportsman	1

3.2 Methods

1. Checking up the practicability of the system. This test had two purposes, one was checking up the influence of feedback content and interface of the system on athlete, the other one was testing whether the use of the system affected the motion

pattern and mentality of the athlete or not. The test carried through during the daily training. After the system was installed on the boat the athlete trained in term of his training plan.

2. Before testing, we requested athlete to choose the feedback parameters which he thought it important.
3. During the process of testing, there were the parameters which athlete chose and a line of the real-time speed of the boat. Before the testing, we introduced the meaning of each parameter and line on the screen in detail, which made the athlete understand the relationship between the parameters and training.

3.3 Results

1. Through the results obtained during the test carried through in national team, we could concluded that the visual feedback training system satisfied the demands that the system could not affect the training pattern of athletes and not aggrandize burden in athletes' mentality. In the testing athlete finished 8km long distance using the visual feedback training system, and the purpose of this kind of training was to experience the technique of paddling in low frequency so that to cultivate the feelings of paddling.

   In the beginning of the test all of the athletes felt the system uncomfortable, but after several paddle the uncomfortableness disappeared. This was most because in the process of drawing oar new content was appended into the vision system of athlete, which changed the content structure of vision that athlete had been used to. Now the nerve center system must judge the new central vision according to the content structure in vision system which made for the uncomfortableness of the athlete. After several periods the vision system and the nerve center system had adapted to the new content structure in vision, and the uncomfortableness disappeared.
2. In the aspect of what parameter was chosen, the athlete inclined to the ones which were used by coach frequently or the ones which they thought important. In the test we found that the coaches and the athletes manifested nicer consistency in the parameters chosen, and 7 athletes and 5 coaches all chose the average speed of boat and the frequency of paddling. It was in the nature of things that the athletes and coaches focused on the average speed of boat because it was the purpose of athletes. Although the frequency of paddling had some relationship with the speed of boat, it was only one factor that influenced the speed of boat, and in daily training it was only used as a parameter estimating the intensity of training. The athlete and coaches chose it as the feedback content, which illuminated that the athlete inclined to the ones which were used by coach frequently.
3. The elite athlete could choose to use the feedback information or not according to the precision of the feedback content. Some investigation showed that in studying motor technique the abecedarian would unconditionally use the feedback information if the information was brought to them no matter the information was right or not. This was mainly because the abecedarian did not know how to use the information from other feeling organs, and they thought the feedback information as pivotal information which helped them confirm the direction of adjusts. However the elite athletes had trained for many years, and had constituted a much

accurate feedback system in their bodies. When the feedback information was not more accurate than the information from bodies, the athletes chose to not use the feedback information.

4 Conclusion

The visual feedback training system for rowing which was designed based on the theory of Neural Mechanism theory of short-term memory. It has been used in real training by elite athletes, and the primary effect has been observed. In the aspect of which parameters to be chosen, the athlete inclined to the ones which were used by coach frequently or the ones they thought them important. The elite athlete could decided whether to use the feedback information or not according to the precision of the feedback content.

References

[1] Yoshiga, C.C., Higuchi, M.: Rowing performance of female and male rowers. Scand J Med Sci Sports 2003 13, 317–321 (2003)

[2] Desgorces, F.D., Chennaoui, M., Guezennec, C.Y.: Influence of anthropometrics parameters on rowing performance at national level. Science & Sports 2004 19, 327–329 (2004)

[3] Galand, S.W.: An analysis of the pacing strategy adopted by elite competitors in 2000 m rowing Br J Sports Med 2005 39, 39–42 (2005)

[4] Hill, D.W., Alain, C., Kennedy, D.M.: Modeling the relationship between velocity and time to fatigue in rowing. Med. Sci. Sports Ecerc. 2003 35(12), 2098–2105 (2003)

[5] Baudouin, A., Hawkins, D.: Investigation of biomechanical factors affecting rowing performance. Journal of Boimechanics 37, 969–976 (2004)

[6] Weitao, Z., Jiurui, H., Xinfa, G., et al.: The Performance Effect Of Different Oblique Angle Of Scull. Jounal Of Wuhan Institute Of Physical Education, 1998.2 125, 51–54

[7] Weitao, Z., Jiurui, H., Qian, H., et al.: Reseach On Hydrodynamic Test Of Flexible Oar. Journal of hydrodynam ICS, 2000.6 15, 163–168

[8] Weitao, Z., Jiurui, H.: The Hydronamic Performance Reseach Of Equipments Of Rowing, Canoeing And Kayak. Jounal Of Wuhan Institute Of Physical Education, 129, 91–95 (1999)

[9] Guihua, S., Qingzhu, S., Jianghong, M.: Research On The Technique Expert System Of Rowing. Jounal Of Systems Engineering, 81, 48–51 (1997)

[10] Weitao, Z., Jiurui, H., Xinfa, G., et al.: A Study Of The Visualizationtesting And Training System Of Rowing. Journal Of Hubei Sports Science, 65–69 (1998)

[11] Yangbo, O., Peiyong, W., Du, S., et al.: Design Of Specific Strength Training System On Land Rowing. Jounal Of Wuhan Institute Of Physical Education 39(2), 32–36 (2005)

A Method for Garment Pattern Generation by Flattening 3D Body Scan Data

Young Lim Choi[1], Yunja Nam[1,*], Kueng Mi Choi[2], and Ming Hai Cui[1]

[1] Department of Clothing and Textiles, Seoul National University, Seoul, Korea
[2] Dept. of Fashion Design, Dong Seoul College, Seongnam-si, Korea
orangebk@empal.com, yunja@snu.ac.kr, orogi1961@hanmail.net, mh-cui@hanmail.net

Abstract. The purposes of this study are to manufacture a flat pattern according to silhouettes through the Surface Flattening process of a triangle by the grid method, and to verify the result to present the method for manufacturing patterns according to designs. Based on the data measured by Size Korea, this study modeled a representative figure ranging from 18 year-old to 24 year-old female, modified the modeled body shapes to a semi-fit silhouette and loose-fit silhouette and did the surface flattening development. We manufactured patterns according to silhouettes and analyzed the area differences between 3D-piece and flat patterns and the deviation between a muslin scan data manufactured by a flat pattern and modeled data. The result of the surface flattening development by the method can be utilized for manufacturing garment patterns by converting a 3D scan data into 2D pattern and proved to be excellent at piece re-creating because the result did not show a big difference from a modeled silhouette.

Keywords: 3D scan data, apparel pattern, flattening, grid method, silhouette.

1 Introduction

A size designation represents a set of garment sizes in a sizing system designed to reflect the body sizes of most individuals in a population. Since the anthropometric data on which the Ready-to-Wear (RTW) sizing system is outdated, RTW does not properly fit the size of current population(s). Previous research showed that consumer dissatisfaction with garment fit range from 62% among men to 50% among women (DesMarteau, 2000; LaBat & DeLong, 1990). For the enhancement of customer satisfaction, there have been a number of studies in the areas of MTM, E-tailoring and mass customization which aims to maximize customer satisfaction and to minimize inventory costs. Mass customization became an effective way for the apparel industry to succeed in the competitive marketplace.

New technologies under development could make custom-fitted clothing affordable by automating some of the traditionally labor-intensive processes. For

* Corresponding author.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 803–812, 2007.

instance, the process of body measurements using 3D body scanners takes about 12 seconds. Some apparel companies have already adopted and utilized automated patternmaking processes (Lands End, 2006) and body scanning (Brooks Brothers, 2004) in order to create custom clothing. However, it is still difficult to convert the 3D human figures to the 2D flat patterns since the human bodies are composed of the various compound curves. In fact, soft wares, which provide the mass customization by using 3D scan data, adopt the methods drafting by sampling 2D measures rather than 3D ones.

A little number of research groups is exploring ways to create apparel patterns directly from an individual's 3D body scan data. Many of the proposed methodologies utilize a few linear dimensions derived from the scan and used to alter an existing pattern. A similar approach developed by McCartney and Hinds (1992) utilized a combination of curve-fitting technologies and linear measurements of selected points from a 3D scan. Conversely, some researchers have investigated ways of wrapping and fitting existing 2D patterns onto the 3D form (Okabe et al., 1992; Zhong & Xu, 2004). Their process is currently being used in the apparel industry to develop the first sample garments with the commercial software, Browzwear (Browzwear, 2006). In addition, Textile and Clothing Technology Corporation (TC2), the apparel research group, is integrating its light scanning system with an automated pattern-altering system using a wrapped pattern (Giannovario, 1998). Research is now underway in Europe to develop intelligent pattern alteration software using wrapping technologies on a 3D scan (e-tailor, 2002; Fontana, 2005).

In reality, in many industrial processes a two dimensional raw material needs to be taken and made to assume a 3D surface. Unless the 3D surface is a surface that could be develop, the material will have to distort in some particular way so that a satisfactory fit can be achieved. If the 3D surface deviates from the developable property severely or if the material is relatively inelastic, then incisions have to be made in the 2D pattern so as to enable material to be removed or additional material to be inserted (McCartney et al., 1999).

The Grid method, which flattens 3D data creating curves those are usually different from remeshing, has a merit when it has less curves as well as the regular interval among curves. The method is suitable to the Surface Flattening because it flattens the body figure by equally subdivided compound curves. In fact, three-dimensional body scans provide an enormous amount of information about the body of the person being scanned while the majority of cited research only uses a fraction of these data. The specific objective of this study is to propose a method to flatten 3D body figure and to develop various apparel patterns suitable for mass customization.

2 Methods

2.1 Modeling of the Representative Body Figure and Silhouettes

This study selected a young adult group ranging from 18 years old to 24 years old based on the measured data of Size Korea in 2004, modeled the representative figure of a female fallen under the KS designation of 82-66-91(165) and named it as YA82 and made it as a standard for the development of a body surface (Table.1). It modified the modeling of the body figure into a semi-fit silhouette and a loose-fit silhouette and adopted them as the object of the surface flattening development. A semi-fit silhouette in this study means that it offers the least margin as if one has a basic block on and is a simplified figure connecting concave parts of a body such as the beneath part of the busts, the interval of the busts and the center of the back in a straight line. A loose-fit silhouette is a figure connecting a semi-fit silhouette projected on a transverse plane to the outmost line.

2.2 Shell Partitions of the Representative Body Figure

The modeling of the human body was partitioned horizontally based on the line of the waist measurement and vertically based on the midsagittal plane and the under line of the posterior and anterior axillary fold.

The under line of the posterior axillary fold consists of the central point of the waist curve, a point of the posterior axillary fold and a plane passing a point projected to the waist line. The under line of the posterior axillary fold was decided by the plane creating a curve onto the body and the same method was applied to the under line of the anterior axillary fold (Fig 4). This study partitioned the upper body into three shells by the method.

Table 1. Size of representative body figure (unit;mm)

Part	Bust circumference	Waist circumference	Hip circumference	Waist back length	Shoulder slope
YA82 model	819.93	653.48	903.43	378.01	26.49°

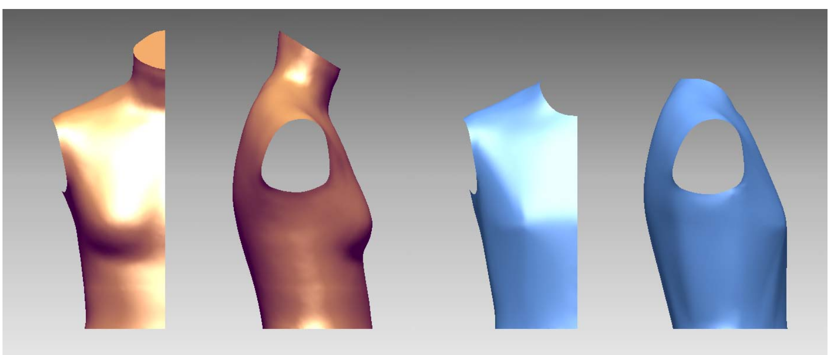


Fig. 1. Representative Body figure **Fig. 2.** Semi-fit silhouette

2.3 Block Subdivision of the Upper Shells

The human body consists of various curved surfaces. It needs to be subdivided according to the features of the body figure in order to develop a shell. Upper front

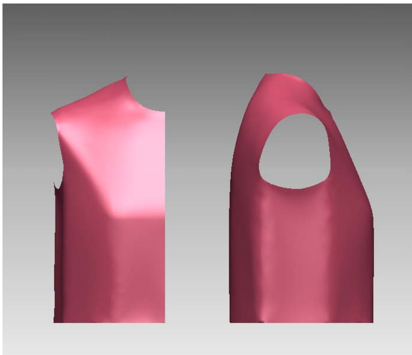


Fig. 3. Loose-fit silhouette

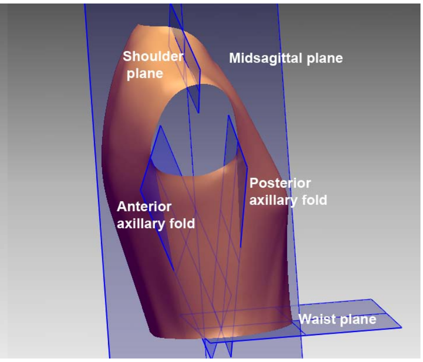


Fig. 4. Shell Partition of Body figure

shell was divided based on the line of the bust circumference and the princess line and was partitioned into four blocks according to the features of the body. Upper side shell was partitioned into four blocks based on the sideline and the line of the bust circumference. Upper back shell was partitioned into eight blocks based on the point of inflection of the armhole, the connecting line of the back neck point, the back yoke line, the bust circumference and the princess line (Fig.5).

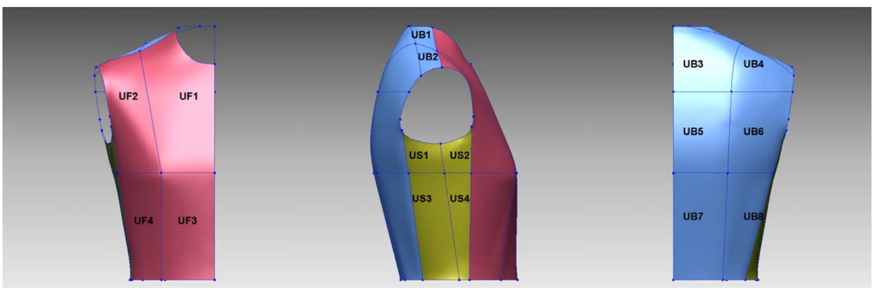


Fig. 5. Block partition of upper front shell, upper side shell, and upper back shell

2.4 Surface Flattening by the Grid Method

Curves can be created by the grid method at each block of a semi-fit silhouette and loose-fit silhouette. The grid method is the way to divide the width and height of a block by the number of a row and column and to arrange the curves on the piece. Since flattening is made by using the length of a segment of a triangle, curves were created at each grid diagonally. Using 3D curves, we manufactured 2D triangles we also made grids by connecting two triangles, arranged each grid to keep the length of the outer curve and calculated the surface flattening piece of each block. This surface flattening piece is called an apparel pattern consisting of each silhouette (Fig. 6).

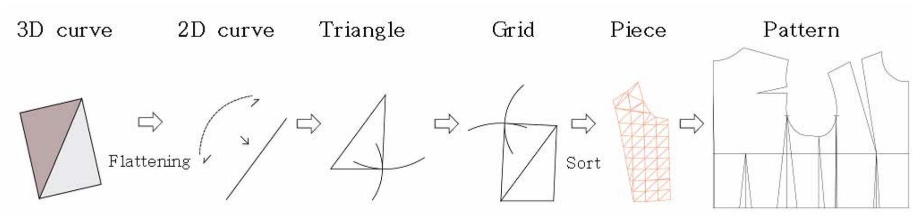


Fig. 6. Surface Flattening process

2.5 Comparison of the Area Differences

We calculated the area of a piece, which is an arranged state of a whole grid included in one block, measured the difference ratio of the areas of the block and piece, compared the difference and confirmed an error of the surface flattening process by the grid method. Rapidform2006 (INUS Technology, Inc, Korea) was employed for partitioning development parts and creating curves by a block. AutoCAD2005 (AUTODESK, Inc) was utilized for the surface flattening of a 3D scan data.

2.6 Verification by a Piece Recreating

Previous progressed experiment was executed by a representative figure modeled by the 3D scan data of a specific size and age group. Accordingly, the result can be regarded as a representative feature. However verification is needed whether the derived apparel pattern reflects the modeled silhouette correctly. Consequently we recreated the 3D silhouette piece by making muslin of the apparel pattern given by the surface flattening development of a semi-fit silhouette and loose-fit silhouette, dressed the representative figure RP (Rapid prototype) in the muslin and scanned it with a 3D body scanner. We compared the 3D scan data with the modeled data and analyzed the differences. WB4 (Cyberware, USA) was utilized for the 3D scan of the muslin recreating silhouettes and Rapidform 2006 (INUS Technology, Inc, Korea) was used for analysis.

3 Results

3.1 Analysis of the Area Differences

With dividing four blocks of Upper Front shell, four blocks of Upper Side shell and eight blocks of Upper Back shell of YA82 model, we created curves by the grid method, flattened the curves and examined the area differences. We showed the difference ratio of the area of the block and the surface flattening piece and observed the area differences according to the surface flattening process at Table 2 and 3. The increase of the area in the surface flattening process was revealed and the tendency of decreasing area at UB 4 and 5 blocks was shown. Since the human body was flattened and partitioned according to various compound curved surfaces, it is thought as the result of the curves of each block itself. In case of UB 4 and 5 blocks, because

Table 2. Comparison of the 3D piece of semi-fit silhouette and the area difference of the surface flattening development

Shell	Block	Modeled(㎟)	Flattened(㎟)	Difference(㎟)	Diff. Ratio %)
Upper front shell	UF1	20642.8011	20787.5099	144.7088	0.7010
	UF2	14419.2863	14491.8001	72.5138	0.5029
	UF3	12898.1249	13031.7871	133.6622	1.0363
	UF4	12538.8129	12584.3027	45.4898	0.3628
Upper side shell	US1	2646.2457	2707.1779	60.9322	2.3026
	US2	2261.4195	2286.9602	25.5407	1.1294
	US3	8973.1570	9098.9805	125.8235	1.4022
	US4	4408.6248	4409.7854	1.1606	0.0263
Upper back shell	UB1	2119.9539	2148.4742	28.5203	1.3453
	UB2	2507.8956	2526.5343	18.6387	0.7432
	UB3	9852.4161	9921.1846	68.7685	0.6980
	UB4	6787.5990	6636.6648	-150.9342	-2.2237
	UB5	10164.5455	10110.5301	-54.0154	-0.5314
	UB6	11441.9505	11472.4686	30.5181	0.2667
	UB7	12743.4787	12750.9183	7.4396	0.0584
	UB8	10330.6507	10349.9348	19.2841	0.1867
Total		144736.9622	145315.0135	578.0513	0.3994

the feature of convex curved surfaces, during the surface flattening process, keeps the length of the outer curve, convex parts can be overlapped and showed the decrease of the flattened surface area (Fig. 7).

3.2 Analysis of the Area Differences

We manufactured a muslin garment, dressed it on RP(rapid prototype) and measured deviation by the 3D scanner in order to examine the 3D piece recreating of a flat-developed apparel pattern. Table 4 and 5 are the deviation measurement pictures of the 3D piece modeled and 3D data scanning, manufacturing the muslin according to the flattened patterns of silhouettes. As shown at Table 4, the average deviation of the semi-fit silhouette is 2.4 mm. The upper part of the bust circumference of Front and Side and the upper part of the lower projection of the scapulas in case of Back are similar to the modeled data. However, in the lower part, the front and back came to loose and the sides adhered more closely to the body. In the case of the loose-fit silhouette (Table 5), the average deviation is 2.7 mm. The similarity between the degree of close adhesion of the muslin scan data and modeled data and the tendency of the front, side, back of the semi-fit silhouette was confirmed. The reason of this tendency is that the upper part of the bust circumference at the front side is the part where a garment falls down and gets fixed. In case of the back side, a garment is fixed at the most projecting part of the back from the lower projection part of the scapulas so that the upper part shows a similar result to most modeled pieces. It is considered

Table 3. Comparison of the 3D piece of loose-fit silhouette and the area difference of the surface flattening development

Shell	Block	Modeled(㎟)	Flattened(㎟)	Difference(㎟)	Diff. Ratio(%)
Upper front shell	UF1	20395.6760	20745.9550	350.2790	1.7174
	UF2	14785.0264	14832.2626	47.2362	0.3195
	UF3	12596.8411	12575.2007	-21.6404	-0.1718
	UF4	16090.5261	16082.4835	-8.0426	-0.0500
Upper side shell	US1	2902.2173	2956.2184	54.0011	1.8607
	US2	2229.4586	2260.8449	31.3863	1.4078
	US3	11935.9802	11947.4505	11.4703	0.0961
	US4	4782.4589	4791.8581	9.3992	0.1965
Upper back shell	UB1	2125.7860	2151.5842	25.7982	1.2136
	UB2	2506.2808	2532.0512	25.7704	1.0282
	UB3	9793.6377	9902.1203	108.4826	1.1077
	UB4	6788.1879	6638.4100	-149.7779	-2.2064
	UB5	10125.2090	10114.2274	-10.9816	-0.1085
	UB6	11750.0470	11772.9310	22.8840	0.1948
	UB7	12279.6520	12284.1403	4.4883	0.0366
	UB8	16265.2999	16276.8413	11.5414	0.0710
Total		157352.2849	157864.5794	512.2945	0.3256

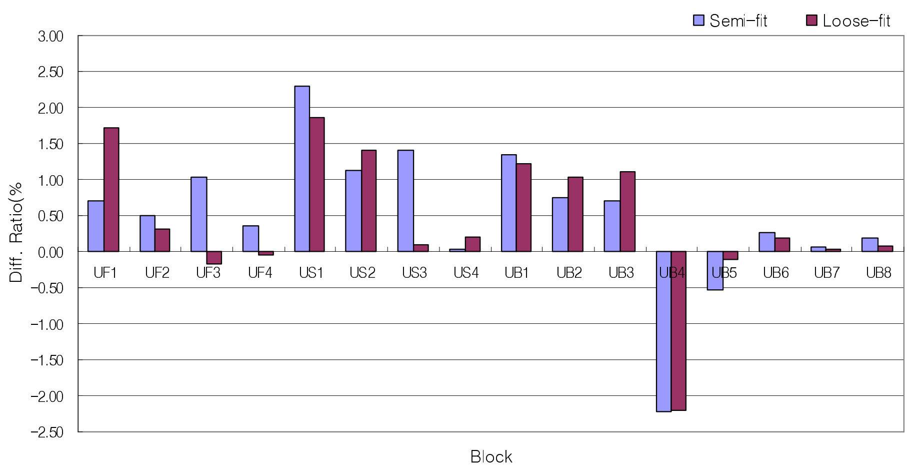


Fig. 7. The area differences of the 3D piece and the surface flattening development of semi-fit silhouette and loose-fit silhouette

that the bust circumferences line and the lower part of the projection under the scapulas are the parts where a margin is included according to the feature of a garment and where there are no fixed parts and, as a result, the similarity to the modeled result falls down.

Table 4. Analysis of the deviation of muslin scan data and modeled data of semi-fit silhouette

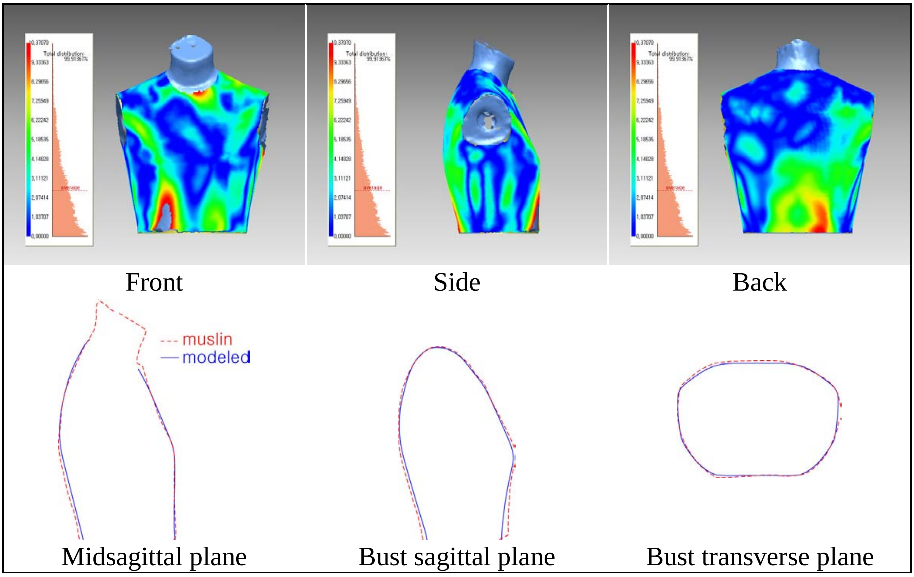
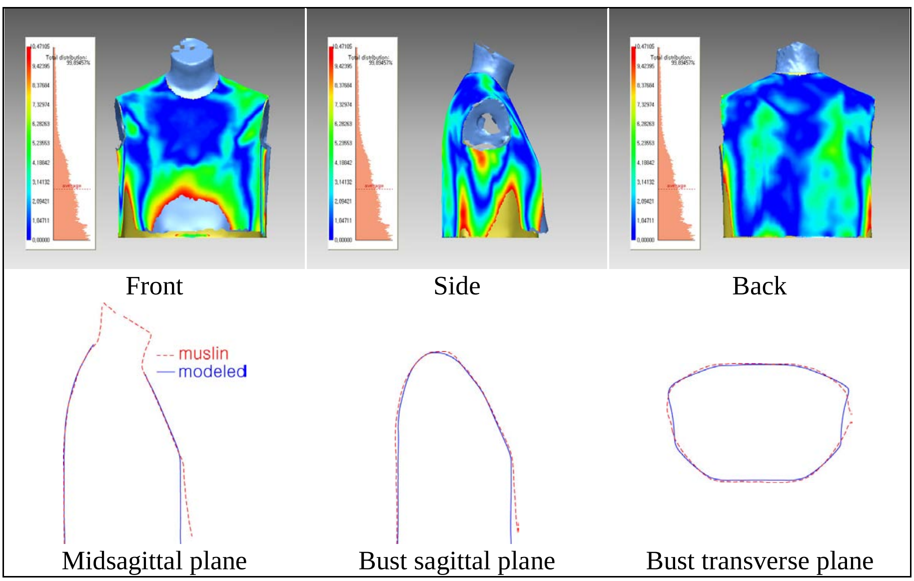


Table 5. Analysis of the deviation of muslin scan data and modeled data of loose-fit silhouette



4 Conclusion

This study targeted the representative figure model ranging from 18 years old to 24 years old female, modeled a state in a semi-fit silhouette and loose-fit silhouette garment, flattening-developed a 3D piece data according to silhouettes by using the grid method and manufactured an apparel pattern. It analyzed the area difference between the 3D piece and the surface flattening pattern, the deviation between the muslin scan data and modeled data manufactured by the flattened pattern, and verified the accuracy of the flattening method using the grid method. As a result of analyzing the area difference between the 3D piece and the flattened pattern, the difference ratio of the semi-fit silhouette and the loose-fit silhouette is 0.4% and 0.33% respectively. It is shown that the area differences barely occurred. However, with respect to the area decrease of the lower part of the back of the scapulas, more detailed researches on the curve ratio of the human body are needed. The deviation between the muslin scan data and modeled data could not control the piece modification by the properties of matter of fabric so that it showed the differences of the average 2.4 mm in the semi-fit silhouette and 2.7mm in the loose-fit silhouette. The result of the surface flattening development by this method can be used for manufacturing a garment pattern by converting 3D piece to 2D and is evaluated as an excellent at recreating a piece because it did not show a big difference from the modeled silhouette. Finally, in this study, the surface flattening development by the grid method, which makes it possible to remodel between the 3D data and the 2D pattern and to convert the bidirectional data of the pattern development, can be presented as a method of automated patternmaking to manufacture garment patterns according to various designs.

Acknowledgements. This work was supported partially by “The Second Stage of BK21” and by the Technology Infrastructure Foundation Program funded by the Ministry of Commerce, Industry and Energy, South Korea.

References

1. Brooks Brothers Digital Tailor.: (Retrieved November 13, 2006) http://www. brooksbrothers;. com/digitaltailor.tem
2. Browzwear.: AIE Apparel Takes a Leap Towards Virtual Sampling with the VStitcher? (November 13, 2006) http://www.browzwear.com/news.asp
3. DesMarteau, K.: CAD: Let the fit revolution begin. Bobbin, pp. 42–56 (2000)
4. E-tailor. (n.d.).: Integration of 3D body measurement, advanced CAD, and e-commerce technologies in the European fashion industry (Retrieved November 13, 2006) http://www.atc.gr/e-tailor/ETAILOR_BROCHURE.pdf
5. Fontana, M., Rizzi, C., Cugini, U.: 3D virtual apparel design for industrial applications. Computer-Aided Design 37(6), 609–622 (2005)
6. Giannovario, B.: Implementation of 3-dimensional body scanning for apparel mass customization. 1998 AAMA Scholarship Winning Paper (1998)
7. Korean Agency for Technology and Standards. 5th Size Korea Technical Report (2004)
8. LaBat, K.L., Delong, M.R.: Body cathexis and satisfaction with fit of apparel. Clothing and Textiles Research Journal 8(2), 43–48 (1990)
9. Lands End. (n.d.): (Retrieved November 13, 2006) http://www.landsend.com
10. McCartney, J., Hinds, B.: Computer aided design of garments using digitized three-dimensional surfaces. Journal of Engineering Manufacture 206(B3), 199–206 (1992)

11. McCartney, J., Hinds, B.K., Seow, B.L.: The flattening of triangulated surfaces incorporating darts and gussets. Computer-Aided Design 31(4), 249–260 (1999)
12. Okabe, H., Imaoka, H., Tomiha, T., Niwaya, H.: Three dimensional apparel CAD system. Computer Graphics 26(2), 105–110 (1992)
13. Textile and Clothing Technology Corporation. (n.d.): (Retrieved November 13, 2006) http://www.tc2.com
14. Zhong, Y., Xu, B.: 3D garment wrapping using adaptive force fields. AATCC Review 4(5), 17–20 (2004)

Human Body Modeling for Riding Comfort Simulation

Hyung Yun Choi[1], Kyung Min Kim[1], Jiwon Han[1], Sungjin Sah[1], Seok-Hwan Kim[2], Su-Hwan Hwang[2], Kwang No Lee[2], Jong-Kweon Pyun[2], Nicole Montmayeur[3], Christian Marca[3], Eberhard Haug[3], and Inhyeok Lee[3]

[1] Hongik University, ME Dept. Seoul, Korea
hychoi@hongik.ac.kr
[2] Hyundai-Kia Motor Company, Korea
[3] ESI Group, France

Abstract. In order to assess the seating and riding comfort design of a vehicle seat in an objective manner, finite element occupant models with anatomically precise features have been developed. The human body models are incorporated into Pam-Comfort, a tool for automotive seat design and analysis, as parts of occupant surrogates that also include HPM I and II. In this paper, the detailed process of FE human body modeling including an effort on the implementation of new anthropometry will be introduced. The validation of some features of human body models in seating and riding comfort simulation against human experiments will be also presented.

Keywords: Seating and riding comfort, Pam-Comfort, Finite element human body model, Anthropometry.

1 Introduction

The design of a comfort seat which earns a good market evaluation needs to be based on the recent anthropometric information of the population so that it can minimize the discomfort feeling possibly caused by failing to fit the diverse sizes and shapes of current occupants. The human body models in PAM-Comfort, a simulation tool for virtual seat design and testing, are in the process to be equipped with new anthropometry obtained from SizeUSA, a nationwide survey in the US [1]. There are three targets in size, two from the edges in "normal" distribution curve that are the 5th percentile female and the 95th percentile male and one from the center which is the 50th percentile male. Using the extreme size such as 5th percentile female and the 95th percentile male in a design process is to minimize the discomfort, e.g. from too wide or too deep a seat pan, while the medium size is testing maximum comfort features. Only caucasian between 36 and 45 years old were retained for the further considerations from the SizeUSA database for narrowing down the analysis to the majority in the USA.

The acquisition of the skin geometry for each model was done by taking the three-dimensional whole body laser scans of volunteers in driving position (the 5th percentile female model is not completed at present). For the recruitment of volunteers with representative body size for each target group, a statistical factor

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 813–823, 2007.

analysis with the SizeUSA database was performed to determine several primary dimensions. The primary dimensions are classified into three orthogonal factors which can also provide overall shape information of each target group.

In order to build the finite element models with anatomically detailed skeletal structure, medical images such as X-ray pictures and ultrasonic scan images were taken from volunteers to gauge the flesh margins. Body pressure distribution on urethane seat foam was also measured for the model validation.

2 Anthropometry

2.1 SizeUSA Survey Data

The overall size and shape of the people change with time. They are taller and larger than 30 years ago. The current human body model [2-4], a 50^{th} percentile adult male in Pam-Comfort is based on the anthropometry of a motor vehicle occupant developed in the early 80's [5]. Therefore, a recent US national size survey, SizeUSA [1] has been adopted in order to update the human body model in Pam-Comfort. In the SizeUSA survey performed from 2000 to 2003, the anthropometries of 10,800 subjects (18-65 years old) were directly obtained by taking 3D laser scans of individuals in a normal standing position. Body measurement extraction software was then used to produce a list of body measurements such as length, width, and girth of body segments from the 3D body data.

There are 3 targets in size, two from the edges in the “normal” distribution curve which are the 5^{th} percentile female and the 95^{th} percentile male and one from the center which is the 50^{th} percentile male. Caucasian between 36 and 45 years old were selected from the SizeUSA database for narrowing down the analysis to the majority in the country. Among 10,800 subjects in total, 687 and 354 were used for female and male objects, respectively. Fig. 1 shows whole body scans for a typical small (5^{th} percentile) female, medium size (50^{th} percentile) male and large size (95^{th} percentile) male from the SizeUSA survey. The corresponding height and weight of each size are also listed in Fig. 1. From the statistical factor analysis, 6 dimensions, i.e., height, weight, hip height, hip girth, bust, and back waist length were selected as primary dimensions among 18 segmental measurements. The 6 primary dimensions are classified into 3 factors which exhibit high orthogonality to each other and thus can provide overall shape information of the population. The results of factor analysis for both male and female are listed in Table 1.

2.2 Volunteer Recruitment and Selection

The specific ranges for three target sizes of each primary dimension were assigned as the selection criteria of volunteers for data collection of body shapes. There was a two-stage recruitment process of volunteers. Since people are well aware of their height and weight but not the other four dimensions, volunteers were invited on the condition of specified range of those two dimensions. The ranges of height and weight were set to $\pm\sigma/4$ (one quarter of standard deviations for small female and

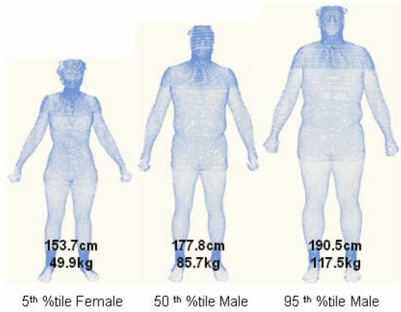


Fig. 1. Typical whole body scans of three target sizes and their standard sizes

Table 1. Results of factor analysis using SizeUSA

	Male			Female		
	F1	F2	F3	F1	F2	F3
Height	0.915	-0.265	0.191	0.884	-0.354	0.212
Back Neck Height	0.951	-0.234	0.147	0.925	-0.314	0.159
Long Shoulder Height	0.937	-0.250	0.132	0.915	-0.322	0.130
Hip Height	0.925	-0.256	-0.094	0.920	-0.236	-0.147
Top Hip Height	0.924	-0.298	-0.212	0.944	-0.231	-0.202
Upper Hip Height	0.907	-0.312	-0.263	0.941	-0.225	-0.225
Waist Height	0.882	-0.321	-0.308	0.931	-0.218	-0.244
Arm Length Straight	0.847	-0.266	0.095	0.791	-0.239	0.052
Knee Height	0.828	-0.423	0.013	0.777	-0.403	-0.079
Crotch Height	0.772	-0.546	-0.007	0.762	-0.597	-0.066
Hip Girth	0.532	0.784	-0.123	0.539	0.815	-0.007
Bust	0.474	0.777	-0.038	0.443	0.793	-0.029
Thigh_Max	0.540	0.759	-0.037	0.535	0.786	0.033
Thigh_Mid1	0.479	0.751	-0.016	0.479	0.790	0.032
Weight	0.662	0.717	0.019	0.662	0.717	0.019
Back Shoulder Width	0.482	0.691	0.036	0.458	0.684	0.065
Hip Width	0.574	0.666	-0.077	0.587	0.690	0.034
Back Waist Length	0.447	0.064	0.883	0.320	-0.219	0.913

medium male) and $\pm\sigma/2$ (half of standard deviations for large male) from average values and used in local advertisement to recruit Americans staying in Korea. Table 2 lists the ranges of height and weight used in the volunteer recruitment. The visiting volunteers were further screened to be selected with the range of four other primary dimensions shown in Table 3. The upper and lower limits of four primary dimensions represent the respective maximum and minimum dimensions obtained from the data of the subjects in SizeUSA those who satisfied both height and weight. Two Caucasian volunteers of 44 and 43 years old were finally selected for 50^{th} and 95^{th} percentile objects, respectively. Both of them meet the required ranges for the segmental dimensions in Table 3.

Table 2. Height and weight ranges for volunteer recruitments

	No. of subjects*	Height (cm)		Weight (kg)	
		Range	S.D. (σ)	Range	S.D.
Small female**	19	152.1 - 155.3	6.4	45.5 - 54.3	17.4
Medium male**	17	175.9 - 179.9	7.7	81.5 - 89.9	16.8
Large male***	6	186.7 - 194.3		109.8 - 125.9	

* Number of subjects who satisfied both height and weight ranges.

** avg.-σ/4 ~ avg.+σ/4, ***avg.-σ/2 ~ avg.+σ/2.

***σ/2 was used since no subjects were found with σ/4 among 354 male subjects in SizeUSA.

Table 3. Ranges for other primary dimensions for volunteer selection

Dimension\Size	5^{th} %tile female	50^{th} %tile male	95^{th} %tile male
Hip Height	73.2 – 79.7	87.0 - 92.5	93.9 - 103.5
Back Waist Length	36.3 – 44.0	47.6 - 54.8	44.8 - 58.3
Bust Girth	83.1 – 99.6	105.2 - 117.0	111.9 - 128.5
Hip Girth	87.4 – 103.0	100.0 – 108.0	110.4 - 125.7

3 Three-Dimensional Scanning and CAD Modeling

The external skin and internal skeletal geometries of two volunteers were obtained by using three-dimensional scanning devices. Merging these external and internal images of each volunteer, geometrically precise CAD models were constructed.

3.1 Three-Dimensional Laser Scanning

The external geometry of selected volunteers in driving posture was measured by a whole body 3D laser scanner (Model: Cyberware WB4). Undeformed shape of buttock and thigh was separately scanned from the volunteers propping on elbows while maintaining the knee angles the same as in driving posture. The synthesis between patches of undeformed parts and the whole body scan image was carried out by using a 3D scan modeling software, RapidForm [www.rapidform.com]. Fig. 2 shows some pictures of volunteer scanning process.

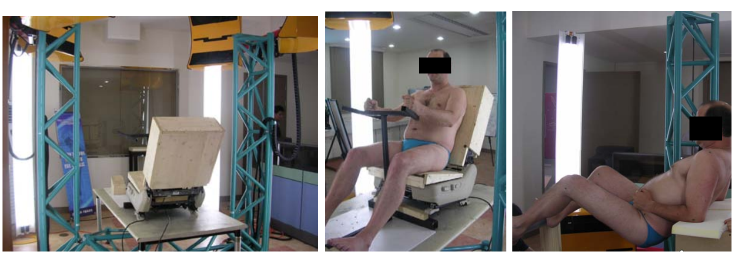
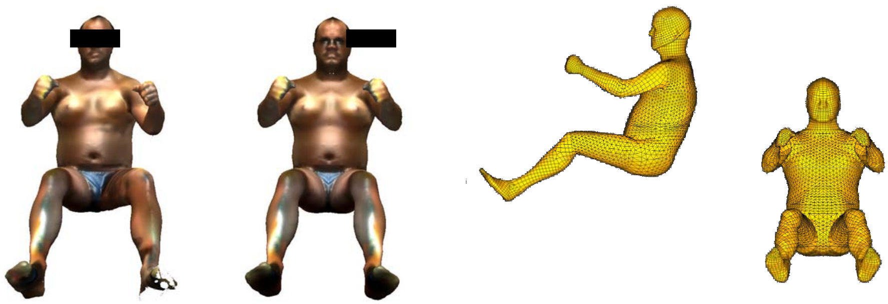


Fig. 2. Whole body 3D scanning of a volunteer in a driving posture

3.2 CAD Modeling

The scanned surface model was reflected along the sagittal plane to make it symmetric. The image with two right sides showed better match ratio with the original scanned image than the one with two left sides. Smoothing the local details, the reflected image was simplified and converted into a surface mesh model. Fig. 3 shows both the original scanned and the reflected images of the 50%tile percentile male volunteer. The surface mesh model is also depicted in the Fig. 3.



Original image Reflected image Surface mesh model

Fig. 3. Scanned images and surface mesh model of 50th percentile male volunteer

The precise location of the skeleton to the skin surface in the model is important for the accurate prediction of sitting pressure distributions in the finite element analysis. In order to gauge the flesh margin, an ultrasonic scanning as shown in Fig. 4 was employed. Ultrasonic scanning was performed over the trunk, hip, and thigh regions of the volunteers in a driving position. Both compressed and uncompressed bone depths were measured in particular for the ischial tuberosity where the peak sitting pressure occurs and thus could be an adequate point for a validation of the finite element model and its prediction. X-ray radiography was also taken but used as supportive information. CT scan was not an effective measure since it was limited to scan the body in a supine position.

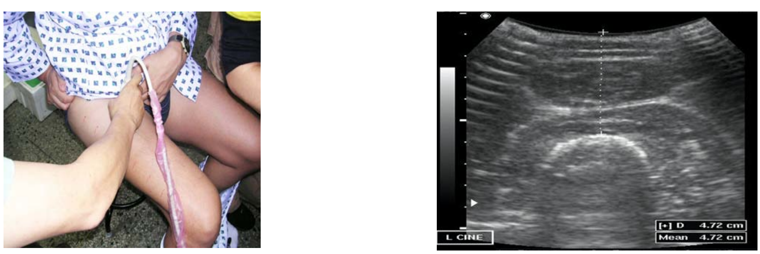


Fig. 4. Ultrasonic scanning for gauging a flesh margin

The locations of joints relative to the surface landmarks were calculated using simple linear scaling relationships [6-8]. Fig. 5 shows schematic drawings of the

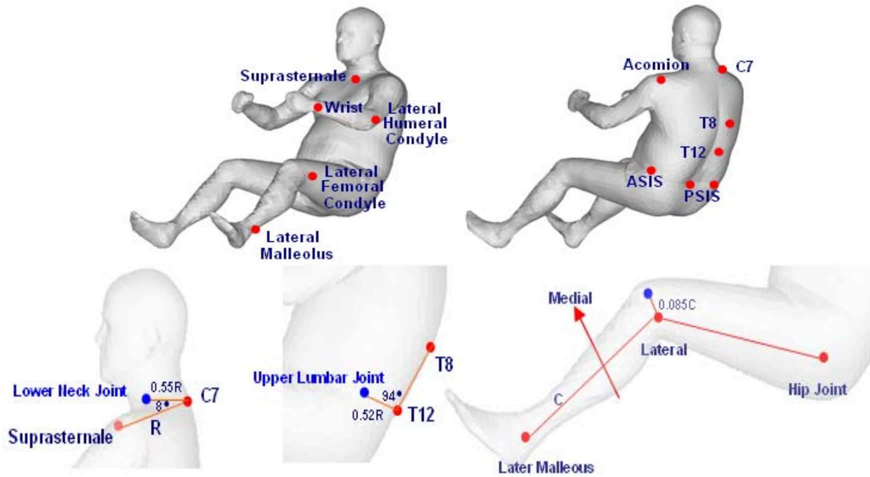


Fig. 5. Schematic drawing of the procedure for calculating various joint locations from surface markers

Table 4. Landmarks for the calculation of joint locations

Joints	Landmarks
Upper Neck (atlanto-occipital)	Infraorbitale., Tragion
Lower Neck (C7/T1)	C7, Suprasternale
Upper Lumbar (T12/L1)	T8, T12, C7, Suprasternale
Lower Lumbar (L5/S1)	ASIS, PS, PSIS
Hip	ASIS, PS, PSIS
Shoulder	Acromion, C7, Suprasternale
Knee	Lateral Femoral Condyle, Hip joint, lateral malleolus
Elbow	Lateral humeral Condyle, shoulder joint, wrist landmark
Wrist	Wrist
Ankle	Lateral malleolus, lateral femoral condyle, hip joint

Fig. 6. Skin with skeleton for 50th percentile male model

procedure for calculating various joint locations from surface markers. Table 4 lists the joint locations of the model associated with landmarks used in the calculation.

A generic skeletal structure, which was used for H-model [9] was incorporated into the skin model by scaling it to match the flesh margins as well as the joint locations computed as in Fig 5. The skin model with skeleton inside for the 50th percentile male volunteer is shown in Fig. 6.

4 Construction of Comfortable Driving Posture

A driving posture is to be determined according to the design of the interior package layout and the operating pattern of a driver. A certain level of fatigue might be produced as the consequence of an uncomfortable driving posture. Therefore, both the package layout of vehicle interior and the anthropometry of a driver need to be taken into account simultaneously to set up the standard seating posture in the design process. Reed [10] proposed the Cascade Prediction Model (CPM), which consists of linear regression equations and inverse kinematics to build a comfortable driving posture. The linear regression equation based on the major factors of package layout and anthropometric data provides eye and hip locations of a given driver. Segmental angles in a comfortable posture can then be optimally obtained from an inverse kinematics computation. The CPM was constructed from the volunteer tests with 68 subjects taking 916 numbers of comfortable voluntary postures. Seat height, steering wheel position, and cushion angles were selected as independent variables affecting the driving posture. In order to make the model posture changeable, the skin model was segmented into 15 parts, i.e. head, neck, upper torso, abdomen, hip, two upper arms, two forearms (including hands), two upper legs, and two lower legs, and two feet, which are articulated at the joints. This kinematic linkage model was visualized by adding a number of stick elements between each joint location as shown in Fig. 7.

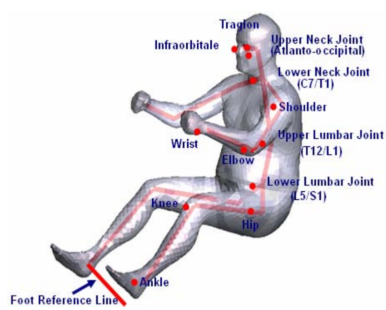


Fig. 7. Kinematic linkage model for 50th percentile male

Comfortable driving postures of 50th and 95th percentile male models were constructed against the package layout of a medium size sedan. As a first step, the eye point and the hip point were calculated from the regression equation in CPM by applying the given package information such as ‘Seat Height’, ‘SW to BOFx’ (horizontal distance between center of steering wheel and ball of foot), and 'Cushion Angle' together with anthropometry data of 50th percentile male model.

The following experimental coordinate equations proposed by Reed [10] were utilized to determine the torso segmental articulation angles within the fixed eye point and hip point.

$$HN(X) = \sum_{i=1}^{4} L_i Sin(\theta_i + \gamma_i \alpha + \beta) \quad (1)$$

$$HN(Z) = \sum_{i=1}^{4} L_i Cos(\theta_i + \gamma_i \alpha + \beta) \quad (2)$$

where L_i is the length of each of four segments between hip and head/neck, θ_i is the starting segment orientation from overall segment angles which was obtained by taking the average from the test measurement. γ_i means the sensitivity on the segment

Table 5. Articulation angles between eye and hip points of 50th percentile male model

Segments	Reed (1998)		50%-ile model	Calculated value		
	Theta (°)	Gamma	segmental length (mm)	Alpha	Beta	Segmental angle (°)
Neck	2	0.399	128.14	-1.47	1.00	2.42
Torso	-3.6	0.617	297.66	-1.47	1.00	-3.50
Lumbar	32.9	1.199	172.99	-1.47	1.00	30.97
Pelvic	63.4	1	111.94	-1.47	1.00	62.93

Err: $Eye_{prediction} - Eye_{calculation} = 1.32$(mm)

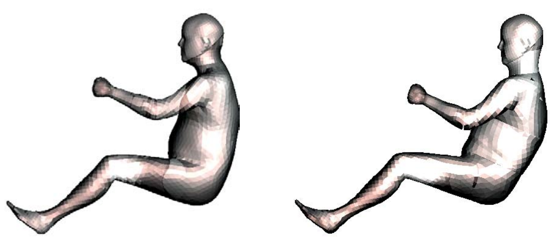


Fig. 8. Scanned and computed comfortable driving postures of 50th percentile male models

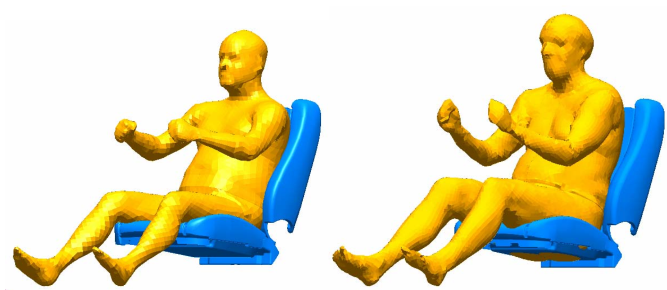


Fig. 9. Computed comfortable driving postures of 50th and 95th percentile male models against a medium size sedan seat

orientation depending on posture changes and was obtained by dividing each slope value in slope estimate table.

The simulated annealing method was used to compute α and β in the above coordinate equations by minimizing the error of eye position relative to the hip point. The optimized segmental angles in the torso region for 50%tile male model are listed in Table 5.

The process of computing the segmental articulation angles for upper and lower limbs were followed by torso fitting. The pedal angle and restraints of hand position on steering wheel (2 and 10 o'clock) were used to decide the posture of legs and arms for a comfortable driving posture. The difference in driving posture of 50^{th} percentile male model between the scanned and the computed are comparatively displayed in Fig. 8. Fig. 9 shows the constructed comfortable postures of 50^{th} and 95^{th} percentile models against a package layout of a medium size sedan.

5 Finite Element Modeling

In order to compute the seating kinetics such as body pressure distribution over the seat pad foam, finite element models of 50^{th} and 95^{th} percentile models were developed. A fine mesh with tetragonal solid elements was applied to thigh, buttock and trunk back, the regions in contact with a seat for the body weight support in seating positions and thus the mesh quality of those parts was closely related to the prediction of seating pressure distribution. Skeletons such as femur, pelvis, spine, and rib cage were modeled as rigid bodies and articulated with kinematic joint elements. In particular, the relative placements of femur, pelvic bone, lumbar spine and scapular to skin surface, and visco-elastic deformation of thigh, hip and torso flesh, and the motions of hip joint and lumbar articulation are all taken into account in the modeling process. The basic properties of the lumbar spinal articulation were based of the literature data [11-14]. The mass and inertia properties of body segments were obtained from GEBOD (http://biodyn1.wpafb.af.mil/GebodIntro.aspx) with the user-supplied body dimension option. The other detail FE modeling process of human body for the seating comfort simulation can be found in author's prior publications [2-4, 9].

Fig. 10 shows a typical setup for the measurement of body pressure distribution on a raw foam pad (hexahedron in 400 x 400 x 100 (30) mm for the cushion(back)) of the 50^{th} percentile subject in the package layout of a medium size sedan. The subject

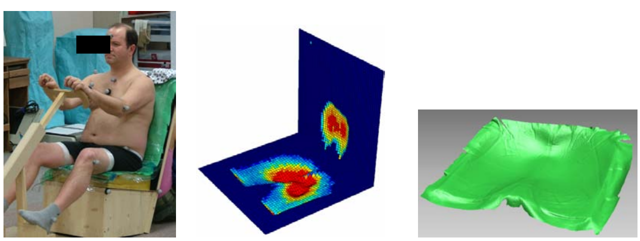


Fig. 10. Test setup for a measurement of body pressure distribution, seating posture, and deformed shape of buttock

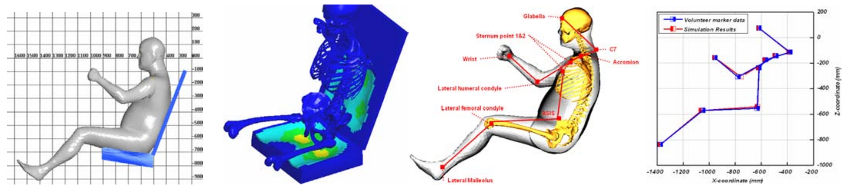


Fig. 11. Simulated seating process with FE human body model

was asked to take a comfortable posture with the given handle and pedal positions. The sitting pressure distributions on both cushion and back were measured by using a pressure sensor system (X-sensor, 43x43 cells). A thin sheet layer of cast (3M's Scotchcast) was also placed between the buttock and raw foam pad for a measurement of deformed buttock shape. The pattern of body pressure distribution in a seating position is generally known to closely depend on the posture of a subject such as torso and thigh angles. Therefore, a precise position of the subjects was measured by an optical sensor system (MotionAnalysis, Model: Eagle 5).

Applying a gravity load into the FE human body model, a simulation of seating process was performed. The foam block in Fig. 11 was modeled through a highly compressible, non-linear elastic material. The material characteristics were determined through preliminary compression tests on a foam sample. A non-linear contact algorithm with friction (Coulomb's law) was used at the interface with the human model. The horizontal (Ux) and vertical (Uz) translations of the wrist and the ankle markers of human body model were fixed to meet the geometric packaging requirement of the vehicle layout. The head was constrained only for the horizontal translation and the pitching (Ry) but free to move in other directions to maintain the natural eye position. The seating posture of the FE human body model in Fig. 11 indicates a good correlation between the experiment and simulation results. A further quantitative analysis of body pressure distribution is now in progress to assess a seating comfort.

References

1. http://www.tc2.com/what/sizeusa/index.html
2. Montmayeur, N., Haug, E., Marca, C., Choi, H.Y., Sah, S.: Numerical and Experimental Analyses of Seating and Riding Comfort. In: DHMC (2004)
3. Montmayeur, N., Marca, C., Choi, H.Y., and Sah, S.: Experimental and Numerical Analyses of Seating Pressure Distribution Patterns. In: DHMC (2005)
4. Choi, H.Y., Sah, S., Na, S., Montmayeur, N., Marca, C.: Human Body Modeling for Virtual Seat Comfort Testing. In: DHMC (2006)
5. Robbins, D.H.: Anthropometry of Motor Vehicle Occupants, vol. 2: Mid-sized Male, vol. 3: Small Female and Large Male (UMTRI-83-53-2, University of Michigan) (1983)
6. Snyder, R.G., Chaffin, D.B., Schutz, R.: Link system of the human torso. Report no. AMRL-TR-7 1-88. Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory (1972)

7. Seidl, G.K., Marchinda, D.M., Dijkers, M., Soutas-Little, R.W.: Hip joint center location from palpable bony landmarks–A cadaver study. Journal of Biomechanics 28(8), 995–998 (1995)
8. Bell, A.L., Pedersen, D.R., Brand, R.: A comparison of the accuracy of several hip center location prediction method. Journal of Biomechanics 23(6), 617–621 (1990)
9. Handbook of Numerical Analysis, Computational Models for the Human Body – Human Models for Crash and Impact Simulation, vol. XII, pp. 297-361, Elsevier, Amsterdam (2004)
10. Reed, M.P.: Statistical and biomechanical prediction of automobile driving, PhD thesis, University of Michigan (1998)
11. Markolf, K.L.: Deformation of the thoracolumbar intervertebral joints in response to external loads. A biomechanical study using autopsy material. J Bone. Jt. Surg. 54-A, 511–533 (1971)
12. Adams, M.A., Dolan, P.: A technique for quantifying the bending moment acting on the lumbar spine in vivo. J. biomechanics 26, 513–522 (1991)
13. Miller, J.A.A., Schultz, A.B., Warwick, D.N., Spencer, D.L.: Mechanical properties for lumbar spine motion segments under large loads. J. Biomechanics 19(1), 79–84 (1986)
14. Panjabi, M.M., et al.: Mechanical Behavior of the Human Lumar and Lumbosacral Spine as shown by Three-Dimensional Load-Displacement Curves. J Bone Jt Surg. 76(A3), 413–424 (1994)

Applications of Digital Human Modeling in Industry

H. Onan Demirel[1] and Vincent G. Duffy[1,2]

[1] School of Industrial Engineering
[2] School of Agriculture and Biological Engineering
Purdue University
47907 West Lafayette, IN, USA
hdemirel@purdue.edu, duffy@purdue.edu

Abstract. This paper represents current and probable applications of Digital Human Modeling (DHM) in different industries. Three real-world challenges, Cummins Inc. (New Product Launches), NASA (Ground Operations), Purdue University (DHM in Curriculum), are investigated for current/probable benefits of DHM. Direct contacts with company representatives and academic faculty were established to understand the current challenges and probable demands of tasks/operations in industry where DHM tools can be utilized. Dassault Systemes' CATIA V5 PLM solution package and UGS Tecnomatix JACK DHM software was utilized to offer a resolution for industrial (Cummins Inc. and NASA) challenges and a software manual (JACK) is developed to enhance the engineering curriculum at Purdue University. The results indicated that DHM tools have potential to improve the product development challenges and provides control of the entire process of designing and analyzing a product before it's ever launched. DHM in engineering curriculum would also increase the integration/adaptation of DHM tools in industry.

Keywords: Digital Human Modeling (DHM), Digital Manufacturing, Computer Aided Design (CAD), Product Development, Human Factors and Ergonomics, Product Lifecycle Management, Engineering Design, Motion Capture.

1 Introduction

Many industries have the same challenges where the human element is not being effectively considered in the early stages of the design, assembly, and maintenance of products. However the use of DHM tools can increase the quality by minimizing the redundant changes and improve safety of products by eliminating ergonomics related problems [1].

A possible software solution available in the industry to address the challenges of improving product development and controlling process of product design and analysis is to utilize the integration between DHM and PLM solution packages. DHM uses digital humans as representations of the workers inserted into a simulation or virtual environment to facilitate the prediction of performance and/or safety. Also, it includes visualizations of the human with the math and science in the background [2].

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 824–832, 2007.

DHM helps organizations design safer and efficient products while optimizing the productivity and cost. Applications that include DHM have the potential to enable engineers to incorporate ergonomics and human factors engineering principles earlier in the design process [3]. An example at Ford Motor Co. involved the UGS Tecnomatix Jack human simulation solution to evaluate ergonomics and worker safety in installing a new satellite digital antenna radio system. This analysis occurred early, at the product design review stage. This reduced the late design modification efforts and helped assess performance and ergonomics based problems prior to prototyping [1].

UGS Tecnomatix JACK [4], one specific piece of the PLM software, is a part of a UGS Product Life-cycle Management (PLM) product solution package, which has been used to prevent human factors and ergonomics related deficiencies in product development. This provides benefits such as a shorter design time, reduction in redundant design changes, lower development costs, improved quality, increased productivity, enhanced safety, and heightened morale. JACK focuses on improving the design of a particular product in the ergonomics and human factors aspects by using Digital Human Modeling [4].

2 Studied Applications of Digital Human Modeling

This paper investigates different application areas of DHM in industry. Three real-world challenges, Cummins Inc. - New Product Launches (IE431 Senior Design Project), NASA - Ground Operations (IE486 Semester Project), Purdue University - DHM in Curriculum (IE499 Honors Program Research), are investigated for current/probable benefits of DHM. Direct contacts with company representatives and academic faculty were established to understand the current challenges and probable demands of tasks/operations in industry where DHM tools can be utilized. Dassault Systemes' CATIA V5 PLM [5] solution package and UGS Tecnomatix JACK DHM software was probed to offer an inclusive resolution for industrial (Cummins Inc. and NASA) challenges and a software manual (JACK) is developed to enhance the engineering curriculum at Purdue University.

2.1 Cummins Inc. – New Product Launches

Cummins Inc., headquartered in Columbus, Indiana, is a corporation of complementary business units that design, manufacture, distribute and service engines and related technologies, including fuel systems, controls, air handling, filtration, emission solutions and electrical power generation systems. Cummins serves customers in more than 160 countries and reported net income of $550 million on sales of $9.9 billion in 2005 [6].

Approximately every two years the design of the engine is modified to comply with EPA emission regulations as well as other issues relating to customer requirements. These modifications, although usually minor with few visible differences, sometimes require drastic changes in the engine assembly process which poses the problem of retraining more than 600 people. Training this large of a population is challenging due to the constant production requirements which demand that lines do not shut down for training activities. Historically, training may take as long as 6 months, which inhibits the Cummins Inc.'s ability to deliver final products to their customers. Cummins Inc. has requested assistance to review their current training standards and procedures to help to improve the transition from engineering design to production and reduce the cycle time [7].

In order to find the reasons of long training period, direct contact with the training manager, engineers and assembly line workers was established. Also, current training methods were evaluated and compared to other organizations [8,9]. It was found that the current training system is appropriate for the organization. However the real problem lies with the constant design changes resulting primarily from customer feedback. The rerouting of tubing and too much vibration in the test trucks that have new engines in them are examples of required changes due to customer demands. When problems arise there is an immediate effort to fix the problems before the engine is launched which sends large amounts of changes to the assembly line after the line has already been set up and workers trained on the existing engines. For each design change, the training cycle must be redone and many workers who were already trained on the previous process must be retrained on a different process. The only shift to see these multiple design revisions is the first shift since training on the other shifts only happens once the first shift is fully trained. As a result, on the most recent new engine transition the first shift took six noncontiguous months to train compared with two weeks for the second and third shifts. This disparity is due in large part from the first shift having to experience multiple assembly task revisions. Better integration is needed between design, manufacturing and training departments. The team also takes into account worker interactions on process steps by using their EASE database during assembly which is a single source of standard times to perform particular tasks. Although relatively accurate, EASE task times sometimes do not match up with reality, a problem that may not arise until engines are being produced on the assembly line while training is taking place. In this regard, a digital work simulation with an up-to-date ergonomics data base is needed to improve the assembly line training. [7] (see Fig.1.)

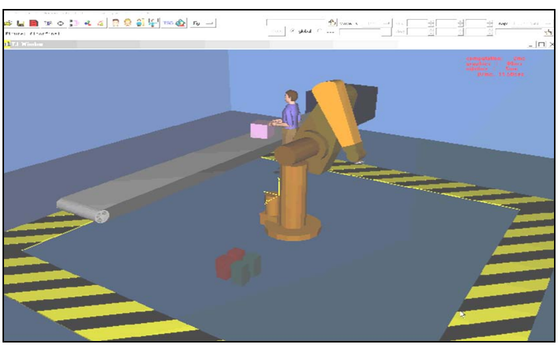


Fig. 1. Assembly Line and virtual worker

2.2 NASA – Ground Operations

In September 2005, NASA [10,11] unveiled its new plan, which includes the NASA's next spaceship, Crew Exploration Vehicle (CEV), Orion [12]. Orion is a manned space craft, which is currently under development, being designed to carry astronauts to the moon and back as well as having future goals to reach Mars [13,14].

One of the biggest challenges in Orion program faces is the ability to carry and shift current operations, skills and expertise to a new technology system. Furthermore, given the value of the expertise that veteran technicians have, many of which will be retiring within a few years time, it is of critical importance to catch this knowledge to pass it along the newer generation of NASA technicians [15].

The proposed approach in this project was to divide each critical knowledge area into two categories depending on the demands of each sub-task: knowledge and methodology. "Knowledge" category was defined to be all of the concepts and background information that technicians need to know in order to successfully perform each task. The other category, labeled as "methodology", includes the entire practical and hands-on experience that is needed to perform a specific task within mission preparation and pre-launch operations. In other words, "knowledge" describes the abstract concepts while "methodology" focuses on the specific practical applications of these concepts [16].

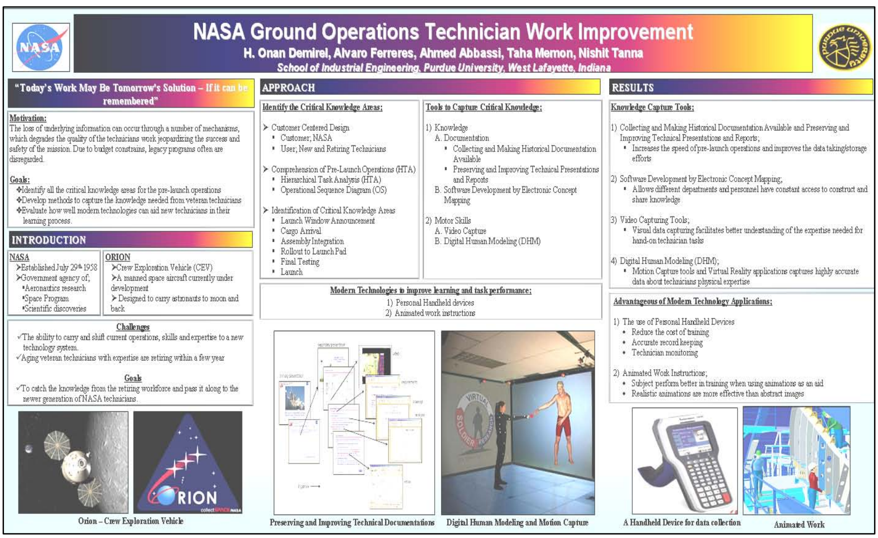


Fig. 2. NASA – Ground Operations and DHM

After capturing all necessary knowledge to cover the technician critical areas, the next step is to capture all the hands-on expertise that is used during the pre-launch operations of any given NASA mission [17,18]. To do so, a process consisting of two parts was developed: The first part consists of capturing simple tasks by means of video [19,20], while the second part concentrates on very critical aspects of the pre launch operations. DHM was regarded as a way to capture all of the detailed motions as well as to correct any bad practices during the process so they are not passed along to the next generation [16]. Motion Capture tools and digital simulation of workers are proposed as possible solutions for both knowledge capture and further training of new workers. (see Fig.2.) It is unlikely that all of the information exposed to the new technicians through the formal class room as well as the case studies will be remembered in full [15,21]. This is one of the reasons why DHM is proposed as one of the new/alternative knowledge capture and training technologies in this project.

2.3 Purdue University – DHM Curriculum

The growing interest in human-centered design practices brought the importance of human safety, usability and performance. This motivates engineers to understand the

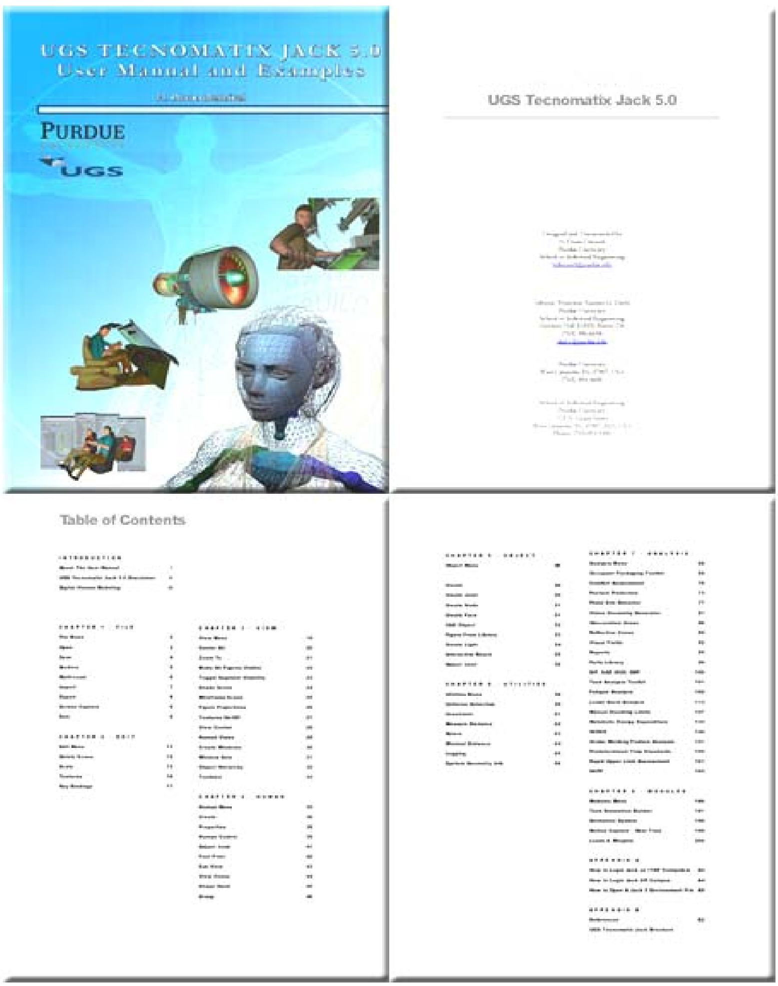


Fig. 3. Jack Software Manual developed at Purdue

human values and needs to allow them to design products, services and experiences that people truly value as individuals and as a culture [22]. Also, engineering problem-solving is not only a technical challenge; it also lies on societal and human side of engineering [23]. Industrial leaders are seeking students with research based teaching and learning experience, as well as strong communication and team based working skills. These challenges in engineering practice require sculpting the engineering curriculum with up-to-date technological tools and knowledge to enhance the technical, social and intellectual capabilities of students to be more creative to answer complex real-life engineering problems. Also continuing to update the technical skills, expertise and knowledge with latest scientific and technological developments will ensure the future competitiveness and success of engineers [23]. In this regard, comprehensive JACK software manual (see Fig.3.) is developed as a research work in Purdue University to offer students a background support in area of DHM, and applied ergonomics. The purpose of this manual is to be a reference to someone using DHM software and to help the individuals' ongoing professional development. At Purdue University three courses (IE 590D Applied Ergonomics, IE 656 DHM Research Seminar, and IE 486 Work Analysis & Design II) are available to provide students an up-to-date engineering curriculum with applied DHM. In addition, Purdue Envision Center and Purdue Discovery Park Center of Advanced Manufacturing have additional software and hardware available for further interest in DHM research and development.

3 Results

After three different real-world challenges (Cummins Inc., NASA, and Purdue University) are analyzed, a common answer for both challenges is highlighted. DHM and its integration with PLM have potential to improve the product development challenges and provide control of the entire process of designing and analyzing a product before it's ever launched. Proposing a DHM curriculum with a ready to use software package would also improve/update the engineering quality and provide a unique educational experience for students.

In the Cummins Inc. study, a PLM-based JACK software could be a good solution for industries where constant design changes are slowing the expected training time and speed of production due to customer and ergonomics based modification needs. Before launching a new product, human factors and ergonomics components of JACK software can be utilized in the design department by considering the needs of workers and customers. Moreover, simulation toolkits of JACK software may help the training teams to visually train their workers on the job site. Also, the PLM integration will keep all departments (design, manufacturing, training, management, etc.) in contact. This could improve the quality of the final product and customer satisfaction. In Cummins' specific training improvement challenge, training leaders can see any changes done by the design department before having their training plans. This integration will benefit the company by lowering training costs and the training time.

In order to capture specific physical task attributes of a technician in NASA's Prelaunch Operations, DHM tools can be implemented as the last step to capture and correct any awkward physical work habits (wrong postures, excessive lifting…).The technician's physical expertise can be captured in detailed through motion capture sensors. Furthermore, the usage of certain body functions (hand dexterity, reach and move motions, etc.) can be recorded and stored to be later corrected before they are passed along. Ergonomics analyses of new working environment, tools and transportation vehicles within NASA prelaunch facilities can be accomplished as well.

In order to meet the objectives of the new DHM curriculum, software manual for JACK is developed and proposed for students in IE486 Work Analysis and Design II [24]. The software manual is intended for use in introducing students to the operation and applications of DHM and Computer-Aided Design (CAD) systems. The purpose of this manual is to be a reference to someone using the JACK software. The user manual covers JACK's tools and menu commands. It explains each tool and menu command in details with pictorial examples. It also covers frequently used industrial applications and demonstrates those in details through step by step method.

4 Conclusion

Use of DHM applications and integration of those with PLM solutions have noticeable advantage in product development. Number of design changes and cost of rapid prototyping can be optimized by DHM and PLM integration. Another advantage of UGS Jack software is its capability of integration with UGS PLM packages, which allows engineers to have control of the entire process of designing and analyzing a product before it's ever launched. It also provides real cost savings to the industry by integrating different departments (design, management, marketing, manufacturing, training, etc.) within a company. This cross-departmental integration increases the quality and the speed of information flow. A leading PLM consultancy, CIMdata, concluded: On average, organizations using digital manufacturing technologies can reduce lead time to market by 30%, the number of design changes by 65% and time spent in the manufacturing planning process by 40%. Production throughput can be increased by 15% and overall production costs can be cut by 13%. [25].

The educational goal of this work is to develop an enhanced engineering curriculum, building on human factors and ergonomics science fundamentals, in which students benefit from the latest technology tools and knowledge. This collaborative research & development teaching and learning approach would increase the engineering problem solving capabilities of the students and will also allow students to interact in interdisciplinary engineering industrial research projects. It is authors' opinion that an engineering curriculum supported with DHM education would increase the integration and adaptation DHM in industry.

5 Future Work

Simulation tools help engineers and scientists to replicate the operation of complex systems by using computer software. Advanced computer integration in engineering is necessary for understanding the complex systems ranging from building jet engines to simulation of earthquake shocks. DHM software facilitates this integration possible and provides solution for complex engineering design and analysis problems. Advancements in computer technology will increase the engineering design and analysis capabilities, improve the product ergonomics and enable engineers to control of human-machine virtual reality applications.

Despite its advantageous, further research and development is necessary to develop graphically advanced, virtually interactive DHM models which include cognitive and physiological aspects of human and working environment. Graphically advanced (more realistic) DHM models will allow engineers to work with virtual reality environments to gain control in product design in earlier stages of product development. Also advanced computational processing capabilities will make it possible to develop advanced mathematical models for DHM applications. A future software development which includes full-body finite human models may take the place of crash test bodies in transportation design and in military applications. This will enable engineers to incorporate advanced human-object and object-object simulations before physical prototyping. Also, advance simulation capabilities with easy micro and macro motion control capacity is required to allow engineers to test robotic devices which are controlled by a human operator.

Acknowledgements. The Cummins Inc. – New Product Launches and NASA – Ground Operations studies were part of the IE431 and IE486 semester projects for Fall 2006 semester. The authors would like to thank students for their valuable contributions who involved in referenced projects.

References

1. Tersko, John.: Process Planning: Human Simulation. Industry Week (1 November 2006) http://www.industryweek.com/ReadArticle.aspx?ArticleID=12822
2. Sundin, A., Ortengren, R.: Digital human modeling for CAE applications. In: Salvandy, G. (ed.) Handbook of Human Factors and Ergonomics, pp. 1053–1078. John Wiley & Sons, New York (2006)
3. Chaffin, D.B., Faraway, J., Reed, M.B., Woolley, C., Rabourn, R.: From modeling to improved ergonomics in designed systems, June, 2005, University of Michigan HUMOSIM Laboratory, p. 27 (2005)
4. UGS Tecnomatix Jack. UGS.: (November 2006) http://www.ugs.com/ products/ tecnomatix/human_performance/jack/
5. Catia V5 Dassault Systemes: (October 2006) www.3ds.com
6. Cummins Inc. About Cummins.: (2006) http://www.cummins.com
7. Demirel, H.O., Bachle, C., Rodriguez, A.M., Siegler, J.: Cummins MidRange Engine Plant - Defining New Training Standards for New Product Launches, IE431 Senior Design Project Purdue University (2006)
8. Pennathur, A., Mital, A., Rajan, V., Kaber, D., Ray, P., Huston, R., Thompson, D., Markle, G., Pitman, M.A., Bishu, R., Crumpton, L., Deivanayagam, S., Fernandez, J., Mcelwee, M., Mcmulkin, M., Sule, D.: A framework for training workers in contemporarymanufacturing environments. International Journal of Computer IntegratedManufacturing 12(4), 291–310 (1999)
9. Sisson, G.R.: Development of training for a new manufacturing process. Training and Development Journal, pp. 22–31 (1972)

10. NASA, National Aeronautics and Space Administration. NASA (November 2006) http://www.nasa.gov/
11. About NASA.: (17 September 2005) (NASA. November 2006) http://www.nasa.gov/ about/ highlights/ what_does_nasa_do.htmln
12. Roach, J.: NASA Aims for Moon by 2018, Unveils New Ship. National Geographic News (September 19, 2005)
13. NASA - Constellation Program. Lockheed to Build Orion Crew Vehicle, NASA (31 August 2006) http://www.nasa.gov/mission_pages/constellation/main/index.html
14. The Vision for Space NASA Names New Crew Exploration Vehicle Orion. Exploration. 22 Aug. 2006. NASA. (November 2006) http://www.nasa.gov/mission_pages/ exploration/ mmb/orion_announcement.html
15. Caldwell, Barrett (ed.): Lecture Notes, IE Work Analysis and Design II. Purdue University (November 2006) http://web.ics.purdue.edu
16. Demirel, H.O., Ferreres, A.M., Abbasi, A., Memon, T., Tanna, N.: NASA Ground Operations Technician Work Improvements, Purdue University IE486 Semester Project (2006)
17. NASA Launch and Flight Operations. NTST 1988 News Reference Manual http://science.ksc.nasa.gov/shuttle/technology/sts-newsref/stsover-launch.html
18. NASA People. Wokforce Data. NASA (November 2006) http://nasapeople.nasa.gov/ Workforce/data/page17.htm
19. Rosenschein, S.: New Techniques for Knowledge Capture. Knowledge Management Comes of Age, TTI/Vanguard (September 23, 2003) (November 2006) http://www. quindi.com/Vanguard_paper.pdf
20. Sobanek, G., Phair, D.: Meeting Knowledge Capture Tools. MITRE. MITRE, p. 9. (2005) http://www.mitre.org/news/events/tech05/briefings/TS376.pdf
21. Gonzalez, Cleotilde.: Does Animation in User Interfaces Improve Decision Making? Universidad De Las Americas, 27–34 (1996)
22. Stanford Mechanical Engineering. Human Centered Design. Stanford University (January 2007) http://me.stanford.edu/research/design.html
23. Grimson, J.: Re-Engineering the Curriculum for the 21st Century. European Journal of Engineering Education 27, 31–37 (2002)
24. Demirel, Onan, H.: UGS Tecnomatix Jack 5.0 User Manual and Examples, 1–202 (2006)
25. Teresko, J.: Digital Manufacturing: Not Only Manufacturing. Industry Week (November 2006) http://www.industryweek.com/ ReadArticle.aspx?ArticleID=12821

A Computer-Aided Ergonomic Assessment and Product Design System Using Digital Hands

Yui Endo[1], Satoshi Kanai[1], Takeshi Kishinami[1],
Natsuki Miyata[2], Makiko Kouchi[2], and Masaaki Mochimaru[2]

[1] Hokkaido University, Japan
yendo@minf.coin.eng.hokudai.ac.jp
[2] Digital Human Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Japan

Abstract. The purpose of this research is to develop a system for computer-aided ergonomic assessment of products without real subjects and physical mockups by integrating a digital hand model with a product model. In previous work, we developed functions of a semi-automatic grasp planning for the digital hand and of quantitatively evaluating the grasp stability of the product based on the force-closure and the grasp quality in our system. We also confirmed the validity of the results of these functions by comparing them with the real grasp postures. In this paper, we propose a new function of evaluating "ease of grasping (EOG)" for the grasp posture based on EOG-map constructed from principal component analysis for finger joint angles in real subjects' grasps.

Keywords: Digital Human, Digital Hand, Computer Mannequin, Ease of Grasp, Principal Component Analysis.

1 Introduction

Recently, handheld information appliances such as mobile phones, handheld terminals and digital cameras etc. have widely spread to general users. Especially some of these appliances for professional use have to be operated with one grasp posture for long hours, so manufactures of these appliances have to take into consideration the ergonomic design. However, the ergonomic assessments of these appliances need the physical mockups and subjects for testing, and therefore take large time and cost. So there are an increasing need to do the ergonomic assessments in virtual environment using digital mockups of the appliances and digital human models especially including digital hand models.

Some simulation software using digital human models have been commercialized and are being used in the design of automobiles and airplanes [1]. However, the digital hand models included in the digital human models of such software do not necessarily satisfy desired accuracy and size variation of human hands when operating the handheld appliances. Moreover, the software does not have the functions of automatically generating the grasp postures, of evaluating the ease of grasp and of evaluating the ease of manipulations of the user-interface by fingers.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 833–842, 2007.

Therefore, our research purpose is to develop a computer-aided ergonomic assessment and product redesign system for handheld information appliances using the digital hand model and the 3-dimesional product model of the appliance.

In our system, we realize the following feature functions for ergonomic assessment to satisfy our purpose:

1. Generation of kinematically and geometrically accurate 3-dimensional digital hand models with rich dimensional variation: We apply a digital hand model called "Dhaiba-Hand" [2] to the ergonomic assessment in our system.
2. Automatic generation and evaluation of the grasp posture: By only inputting a few user-interactions, our system automatically generates one of the possible grasp postures determined by the product shape and the digital hand geometry. It also quantitatively evaluates two indices of the grasp: i) *grasp stability* for the product geometry and ii) *ease of grasping* from aspect of finger joint angle configuration.
3. Automatic evaluation of easiness of the finger motions in operating the user interface: The system automatically moves fingers of the digital hand by following an operation task model of the user-interface (e.g. which button has to be pushed). It also automatically evaluates easiness of finger motions during operation of the user interface based on the flexion joint angles of fingers.

In this paper, we mainly discuss technical background of above second function especially of evaluating the "ease of grasping". This computer-aided ergonomic assessment system is being developed as a part of our government-funded project called "Sapporo IT Carrozzeria" [3] whose mission is developing rapid prototyping technology of information appliances.

2 Previous Works

2.1 Related Works

Some researches (for example, [4,5,6,7]) have proposed the generation and evaluation methods of the grasp postures of the digital hand for the objects. The methods of generating the grasp posture are roughly classified into two types: a variant method [5] and a generative method [4,6,7].

In the variant method, real grasp postures of many subjects for sample objects have been measured in advance using dataglove to build a grasp posture database. Then, in generation step, one grasp posture where his/her grasping object shape is most similar to the given object model shape is chosen from the grasp posture database. The selected posture is then modified to fit to the product model shape. If a very similar product shape can be found in the database, a nearly appropriate grasp posture for the given product can be obtained after this modification process.

On the other hand, in the generative method, the grasp posture is generated by full-/semi-automatic grasping algorithm. This method does not need any database of the real grasp postures. For any unknown shape to be grasped, this method can generate a grasp posture which satisfying geometric boundary conditions when the hand is contacting with the product surface.

However, the both of these researches did not discuss whether the obtained grasp postures are truly appropriate and possible one or not.

2.2 Digital Hand Model and Grasp Stability Evaluation [6]

2.2.1 Digital Hand Model

To perform effective digital ergonomic assessment, it is insufficient to use only one digital hand model with a fixed dimension because the physical dimensions of appliance users differ from person to person. Therefore, we needed to generate a digital hand model with possible anthropometric variation. In this purpose, we used a digital hand model based on the Dhaiba-Hand [2]. The digital hand model used in our system consists of the following four parts and relation among these parts is shown in Figure 1(b):

1. Link structure model: A link structure model approximates to the rotational motion of bones in the hand. The model was constructed from the measurement by MRI and the motion capture [8,9].
2. Surface skin model: A surface skin model is a 3-dimensional polygonal mesh for the hand surface generated from CT images, as shown in Figure 1(a). The geometry of the skin model is defined at only one opened posture.
3. Surface skin deformation algorithm: This algorithm defines the deformed geometry of the surface skin model when the posture of the link structure model is changed, as shown in Figure 1(b).
4. Finger closing motion sequencer: The finger closing motion sequencer is a function to automatically and naturally generate a finger-closing motion path of the hand model from a fully opened state to a clenched one, as shown in Figure 1(b). This motion reflects the joint angle constraints of the link structure model.

A link structure and a surface skin model are generated by inputting the 82 dimensional parameters of a specified subject's hand into the *generic hand model*

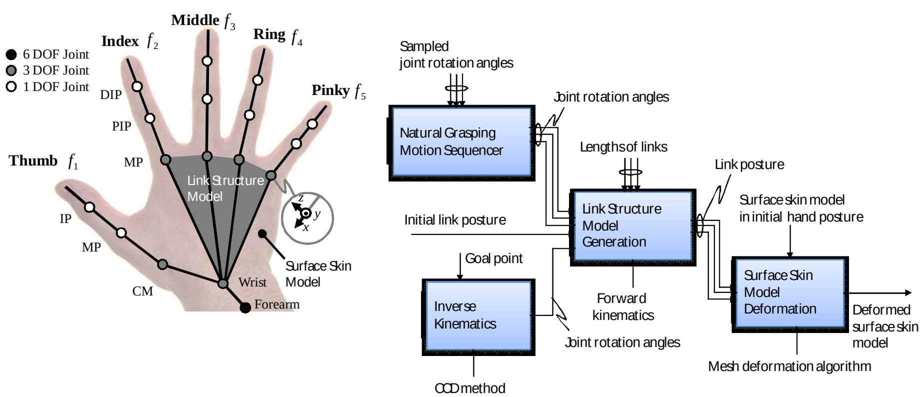


Fig. 1. The digital hand model of our system, (a) The link structure model, (b) The flow of the deformation of the surface skin model

which are implemented in the Dhaiba-Hand [2]. On the other hand, a surface skin deformation algorithm and a finger closing motion sequencer were originally developed by us [6].

2.2.2 Grasp Posture Generation and Evaluation of the Grasp Stability [6]

The generation of the grasp posture of the digital hand model for the product shape model is the first step in our system. The process consists of four phases: 1) selection of the contact point candidates, 2) generation of the rough grasp posture, 3) optional correction of the contact points and 4) maximization of the number of the contact point.

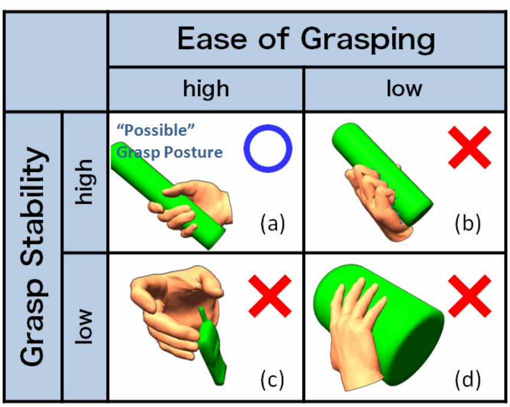
After generating the grasp posture, the system automatically evaluates the grasp stability for the product in this estimated posture. We introduced the *force-closure* and the *grasp quality* into the evaluation of the grasp stability [10, 11].

We also described the verification results of our system by comparing the estimated grasp postures given from the system with the ones from experiments, and also by comparing the grasp stability evaluated by our system with the ones felt by real subjects.

2.2.3 The Problems on Generating Grasp Postures and Our Solution

As described in the previous section, our generation and evaluation approach for the grasp posture based only on grasp stability does not necessarily ensure that the obtained grasp posture is truly “possible” one. There is a possibility that only using the grasp stability causes the wrong evaluation result. For example, the grasp posture in Table 1(b) satisfies the force-closure condition, and has high grasp quality value. Therefore, this posture has high grasp stability. But this grasp has the impossible grasp posture where the configuration of finger joint angles is quite different from the one we usually take in holding objects. Therefore, we need another index which can evaluate the validity of the configuration of finger joint angles of the grasp posture. We call this second index “ease of grasping (EOG)”. Of course, only using this index

Table 1. The relation between grasp stability and ease of grasping



may sometimes cause the wrong evaluation results (Table 1(c)). Therefore, we define the truly "possible" grasp posture as the one which has high index value of the grasp stability and the ease of grasping, as shown in Table 1(a).

To evaluate the EOG, we have to calculate one index value at a certain configuration of finger joint angles. However, total degrees of freedom of fingers in a hand ranges from 25 to 32 [12], and it becomes very difficult to calculate an index value from the combination of these many finger angle values.

In neuroscience field, some research described that the finger joint angles of human hand are strongly interrelated with each other during the grasp motion and at holding state [13]. So the "possible" grasp postures are represented by less variables than the degree of freedom of the hand. Therefore, in this paper, we propose to apply this interrelation of the finger joint angles to decreasing the degrees of freedom of the hand and to constructing a new index for the grasp posture evaluation.

3 The Proposed System for Generation and Evaluation of Grasp Posture

In this section, we describe the method of how to evaluate the ease of grasping the object, which is the second evaluation index of the grasp. Figure 2 shows an overview of our evaluation method for the ease of grasping an object in our computer-aided ergonomic assessment system. The system consists of the following three steps:

1. As preprocess of the method shown in Figure 2(A1-A3), we construct an "ease of grasping evaluation map (EOG-map)" defined in an *M*-dimensional space, where a large number of actual hand postures from the opened state to the grasping state are plotted.
2. An initial estimated grasp posture is generated for a product model to be grasped (Figure 2(B1)), as described in our previous work [6].
3. The ease of grasping for the optimized grasp posture is evaluated (Figure 2(B2)).

We describe the above 1 and 3 in the following section.

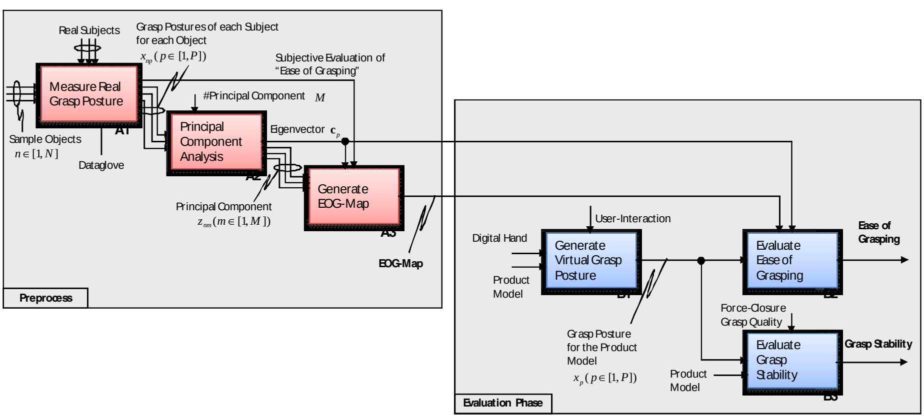


Fig. 2. The proposed system for evaluating ease of grasping

3.1 Building the EOG-Map

The EOG-map consists of a set of plots in multi-dimensional space, and each of plot corresponds to a represent to finger joint angle configuration measured from real grasping examples. This plot is generated by the following process:

1. For a subject to grasp and an object to be grasped, a sequence of finger joint angles of the subject's hand from the opened state to the grasping state are measured.
2. A subject is asked to hold a set of objects including primitive shapes and real daily products.
3. A number of subjects are required to carry out the above experimental process of 1 and 2.
4. All recorded sets of finger joint angles are processed by PCA and the results are plotted as points on the EOG-map.

In this map, one posture example is indicated as a set of scores of a principal component analysis (PCA) for finger joint angles of some "real" human hands, measured by a dataglove. We can estimate the ease of grasping which is generated from a product model and a digital hand by plotting the principal component score for this posture on the map.

3.2 Principal Component Analysis for Finger Joint Angles of Hand Postures in Grasp Motion

Let's define a set of P joint angles of a hand posture as $\{x_p \mid p = 1,2,...,P\}$. P indicates the degree of freedom of the hand model and is generally defined as more than 30. However, it has been known that finger joint angles are strongly interrelated with each other during the grasp motion [13]. Therefore, we reduce the degree of freedom of the hand posture to less than P by PCA.

Suppose joint angles of the hand postures in grasping for N objects are recorded. We define a set of the standardized measurements of these angles as $\{x_{np} \mid n = 1,2,...,N;\ p = 1,2,...,P\}$, and define a N x P matrix as $X = [x_{np}]$ $(x_{np} \in \Re; n \in [1, N]; p \in [1, P])$. Then we obtain a variance-covariance matrix V and the matrix $W = [w_{ij}]$ $(w_{ij} \in \Re;\ i, j \in \{1,2,...,P\})$ as

$$V = \frac{1}{N-1} X^T X \quad , \tag{1}$$

$$W = [w_{ij}] = [\mathbf{c}_1\ \mathbf{c}_2\ \cdots\ \mathbf{c}_P], \tag{2}$$

where $\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_P$ are eigenvectors of the matrix V which are sorted in decreasing order of eigenvalue.

Suppose the P finger joint angles can be approximated by M ($M \le P$) principal components, then we approximately describe the hand posture as the first M principal component scores $\mathbf{z}_n = [z_{n1}\ z_{n2}\ ...\ z_{nM}]$ as follows:

$$z_{nm} = \sum_{p=1}^{P} w_{pm} x_{np} \qquad (m = 1,2,...,M,\ n = 1,2,...,N). \tag{3}$$

3.3 Ease of Grasping Evaluation Map (EOG Map)

3.3.1 Method for Generating the EOG Map

Using the above eq. 3, we can evaluate the sequence of the principal component scores z_{nm} of the hand postures from the opened to grasping states for each sample object with "real" subjects. By plotting the first M principal component scores of these postures in a M -dimensional space, we can generate the "ease of grasping evaluation map (EOG-map)". At the same time, we also record the two-level subjective evaluation for the grasp postures (i.e. "easy to grasp / not easy to grasp"). By partitioning the EOG-map into the equally spaced M -dimensional voxel, an EOG value $\text{eog\_vox}(\gamma)$ is attached to each voxel γ. The range of this EOG value $\text{eog\_vox}(\gamma)$ can be one of the following three values, which are:

1. *easy_to_grasp* — a voxel which includes grasp postures whose subjective evaluations are "easy to grasp"
2. *able_to_grasp* — a voxel which includes hand postures during the grasping sequence but does not include final grasp postures
3. *unable_to_grasp* — a voxel which includes neither hand postures nor grasp postures

Therefore, the function $\text{eog\_vox}(\gamma)$ is defined as

$$\text{eog\_vox}\,(g) : \Gamma \rightarrow \{\ \textit{easy\_to\_grasp},\ \ \textit{able\_to\_grasp},\ \ \textit{unable\_to\_grasp}\ \} \quad (4)$$

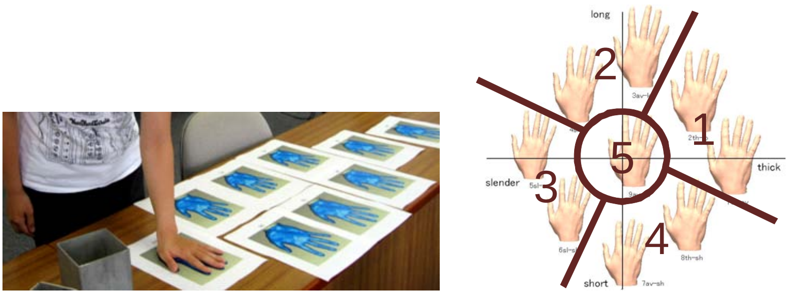
where, Γ is a set of all voxel.

3.3.2 Generation of the Multiple EOG-Maps Classified by the Hand Dimension

There is a possibility that a large variance in the hand sizes of the subjects causes different results in the EOG evaluation. To consider this difference, we classify subjects into five groups with respect to their hand dimension and generate five different EOG-maps for the five groups. In the measurement of the preprocess, we prepare nine sheets of the paper where the nine representative Japanese hands shown in Figure 3(a) are printed. By putting the subject's hand on these sheets, each subject chooses one representative hand that has the most equivalent hand dimension. Based on the class of their chosen representative hand, the dimensions of subject's hand are classified into five groups, as shown in Figure 3(b).

3.3.3 Results of Generating the EOG Map

Figure 4(a) shows one of the EOG evaluation maps generated by the above method (the 5^{th} hand group in Figure 3(b)). The blue/red circular points show the grasp postures that have the subjective evaluation of "easy to grasp / not easy to grasp". The small gray points show the instantaneous hand postures in grasping sequence. The EOG evaluation maps are generated by 8 subjects and 70 sample grasped objects, as shown in Figure 4(b). The finger joint angles of this hand postures were measured by the "CyberGlove", which has 19 sensors. We determined that the number M of the



(a) (b)

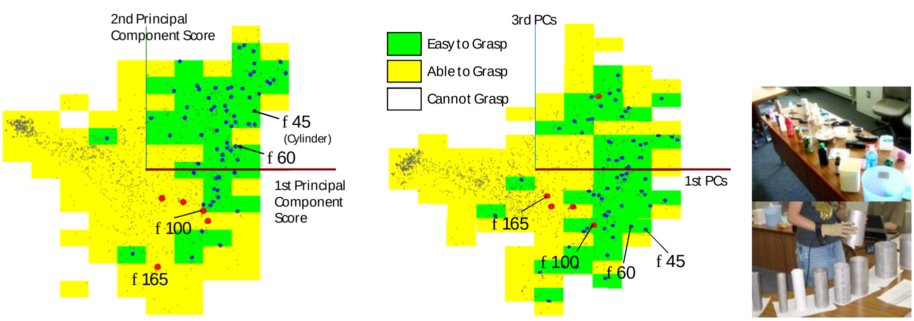
Fig. 3. The generation of multiple EOG-map classified by the hand dimension

principal components should be three, because the cumulative proportion of the first three principal components was more than 80% at every EOG evaluation map of each hand group.

3.4 The Evaluation and Verification for Ease of Grasping

In the evaluation process in Figure 2(B2), the optimized grasp posture i of the digital hand is generated, and the vector of the principal component scores $\mathbf{z}_i$ for this posture i is calculated from eq. 3. Then, the *ease of grasping* for this grasp posture i eog(i) can be defined as $\mathrm{eog}(i) = \mathrm{eog\_vox}(\gamma^i)$ where γ^i is a voxel including a point $\mathbf{z}_i$, and the function $\mathrm{eog\_vox}(\gamma)$ returns one of the EOG values of the voxel γ as defined in the previous section.

We verified the ease of grasping evaluation by plotting some grasp postures of the digital hand on the EOG-map, as shown in Figure 5. These grasp postures are generated from the digital hand and the product models which have the same geometry as the real shape of the product used in generating the EOG-map. The blue/red rectangle points show the principal component scores of the grasp postures



(a) (b)

Fig. 4. (a) The EOG map and (b) sample objects used for generating the EOG-map

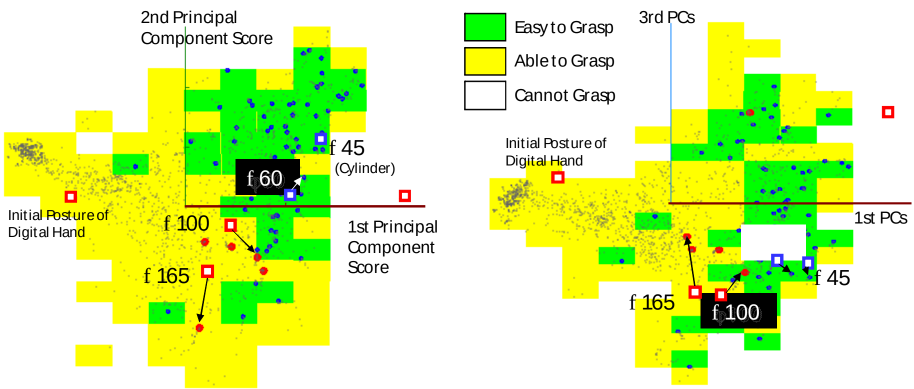


Fig. 5. The evaluation and verification results of the ease of grasping

that were estimated as “stable/unstable” grasp posture by force-closure. The two-point pair connected with an arrow shows that these grasp postures are for the same grasped object. The EOG evaluation results by taking a digital hand and by a real subject’s hand for the same object fell into the same category for almost all cases. Therefore, the proposed EOG-map-based method is effective for evaluating the ease of grasping a product model with a digital hand.

4 Conclusions

The conclusions of our research summarizes as follows:

1. We proposed a system of automatic ergonomic assessment for handheld information appliances by integrating the digital hand model with the 3-dimensional product model of the appliances.
2. The ease of grasping evaluation (EOG) map was introduced by measuring the grasp postures from real subjects. Based on this EOG map, we proposed an evaluation method of ease of grasping from the product model and the digital hand model. The EOG values by taking a digital hand and by a real subject’s hand for the same object fell into the same classification of the easiness. PCA for the finger joint angles enables to calculate the posture similarity with much less variables than the DOF of the human hand.

In our future research, we will develop a new function to evaluate the easiness of finger operation for the use-interface of the products, and to aid the designers in redesigning the housing shapes and the user-interfaces in the product model.

Acknowledgements. This work was financially supported by a grant-in-aid of Intelligent Cluster Project (Sapporo IT Carrozzeria) funded by MEXT.

References

1. For example, Jack.: http://www.ugs.com
2. Kouchi, M. et al.: An Analysis of Hand Measurements for Obtaining Representative Japanese Hand Models. In: Proceedings of the 8th Annual Digital Human Modeling for Design and Engineering Symposium, 2005-01-2734 (2005)
3. Carozzeria, S.T.: http://www.it-cluster.jp/english/
4. Miller, et al.: GraspIt!: A Versatile Simulator for Robotic Grasping. IEEE Robotics and Automation Magazine, 110–122 (2004)
5. Kyota, F., et al.: Detection and Evaluation of Grasping Positions. In: Proceedings of the 2005 ACM SIGGRAPH, sketches (2005)
6. Endo, Y., et al.: An Application of a Digital Hand to Ergonomic Assessment of Handheld Information Appliances. In: Proceedings of SAE 2006 Digital Human Modeling for Design and Engineering Conference, 2006-01-2325 (2006)
7. Yang, J., et al.: Posture Prediction and Force/Torque Analysis for Human Hands. In: Proceedings of the SAE 2006 Digital Human Modeling for Design and Engineering Conference, 2006-01-2326 (2006)
8. Miyata, N., et al.: Modeling of Human Hand Link Structure from Optical Motion Capture Data. IEEE/RSJ Int. Conf. Intelligent Robots and Systems, pp. 2129–2135 (2004)
9. Miyata, N., et al.: Method to Derive Hand Link Model with MRI images. In: Proceedings of the 21st Annual Conference of the Robotics Society of Japan, 1A11 (2003)
10. Nguyen, V.: Constructing force-closure grasps. The International Journal of Robotics Research 7(3), 3–16 (1988)
11. Pollard, N.S.: Closure and Quality Equivalence for Efficient Synthesis of Grasps from Examples. The International Journal of Robotics Research 23(6), 595–613 (2004)
12. Lee, J., et al.: Model-Based Analysis of Hand Posture. IEEE Computer Graphics and Applications 15(5), 77–86 (1995)
13. Santello, M., et al.: Postural Hand Synergies for Tool Use. The Journal of Neuroscience 18(23), 10105–10115 (1998)

Human Age and Vehicle Speeds Affect on Vehicle Ingress Motion Pattern

Hanson, L.[1], Yong, L.[2], and Falkmer, T.[3, 4]

[1] Ergonomics, Department of Design Sciences, Lund University, Sweden
[2] Zhongyan Institute of Technology,
[3] Rehabilitation Medicine, Department of Neuroscience and Locomotion, INR, Faculty of Health Sciences, Linköping University, Sweden.
[4]School of Health Sciences, HHJ, Jönköping University, Sweden

The vehicle ingress and egress are important issues for the automotive industry, both for minimizing assembly work load and for maximizing end-users' comfort. Digital human modeling tools are used for evaluating and visualizing these issues. The assembler and end-user are more or less performing the very same entering task if the vehicles have identical geometrical prerequisites. The major difference is the vehicle speed; an assembler is entering a vehicle slightly moving forward on the assembly line with a speed of 5 meter/minute whereas the end user's vehicle is standing still. The human motion when entering a car is a complex biomechanical process, which affects many different body parts. Car ingress techniques, such as flopper, swiveler, and glider vary among humans; for which humans' agility may be one affecting factor. Agility is affected by joint diseases, which is more frequent among older people. There are several studies regarding ingress motion patterns[1,2], but studies on the differences in car ingress motion between car assemblers and end-users, or older and younger people are rare. Thus the purpose of the present study was to compare the ingress motion between younger versus older persons, and assemblers versus end-users.

A Saab 900 s and a treadmill were used in the study. The treadmill was used to simulate the forward moving vehicle. The car was elevated to not change treadmill-vehicle ingress height from normal ground vehicle ingress height. The motion pattern of the subjects was recorded with a ShapeWrapII motion capturing system[3,4]. In total, 40 subjects participated. The subjects, all holding a valid driving license and healthy with respect to joint disease diagnose, were divided according to their age into two groups: 20 – 25 years old and 60 – 65 years old; each group containing 20 persons. Each group performed ingresses twice, both as end-users, from a stable ground, and as car assemblers, from a moving treadmill.

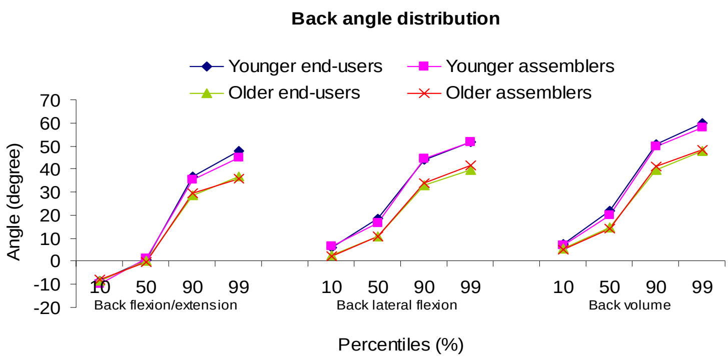
The procedures of the test were divided in five steps: initial part, mounting the ShapeWrapII system, warming and calibrating ShapeWrapII, performing the ingress, closure part. In the initial part, the subjects were introduced to the procedure and asked to answer a questionnaire about age, driving experience and joint disease history. The closure part included a questionnaire, in which the subject described and rated his/her experience of ingress. In the closure part, subjects were rewarded for their participation (as fig.1).

From the ShapeWrapII motion capture system, joint angles and angles velocities were calculated and plotted (graph.2 -graph.5). Joint angles were also used in the ergonomic evaluation tool REBA. The total time to perform ingress motion was also

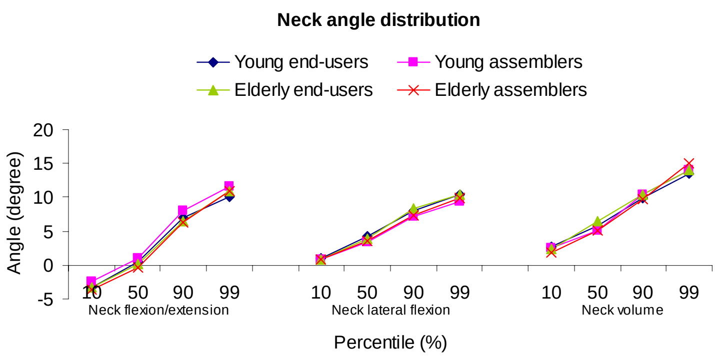
V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 843–846, 2007.

Fig. 1. Start and ending position

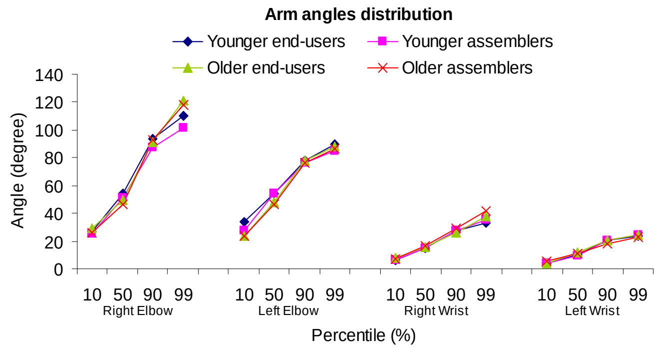
calculated from motion capture data. Descriptive statistics were used to summaries the ingress motion pattern and Friedman's test to investigate differences between young and old as well as assembler and end user.



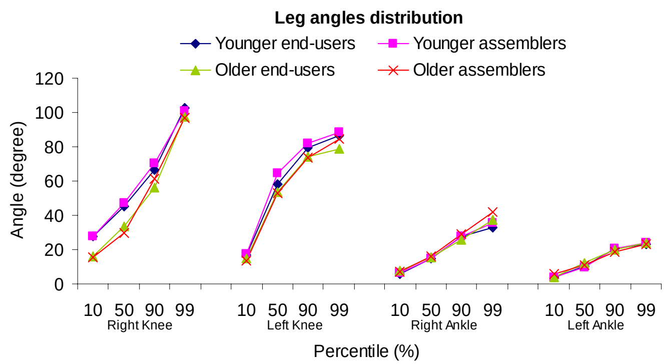
Graph 2: Back angle distribution



Graph 3: Neck angle distribution



Graph 4: Arm angles distribution



Graph 5: Leg angles distribution

All the above experimentation, calculations and data analysis gave the following results: there was a significant difference in the car ingress motion between older and younger people. The joint in which the differences were most significant was the back. The younger people’s ingress motion was more risky according to REBA evaluation. Furthermore, the older needed a significantly longer time to perform the ingress motion. No significant difference in motion pattern was found between assemblers and end-users, and the slow forward motion of the car on assembly line had no effect on the ingress pattern.

These differences and similarities found in this study lead to the following conclusions:

1. There is a need to introduce age as a parameter in digital human modeling tools, due to its affect on motion behavior. With age as a parameter, similar to nationality, anthropometrics motion pattern predicted, reach envelopes generated and vision cones generated can represent also an older human.

2. Within the automotive industry, responsibility for assembler and end user is normally divided into two departments. The communication and sharing of information between these are normally limited. However this result shows that ingress information from digital human modeling analysis could be shared between them and facilitates an effective and efficient design process.

References

1. Menceur, A.E., M. O., Pudlo, P., Découfour, N., et al.: An experimental protocol to study the car ingress/egress movement for elderly and pathological population. AM'06: In: European Annual Conference on Human Decision-Making and Manual Control, alenciennes France (2006)
2. Kazmierczak, K., Mathiassen, S.E., Forsman, M., Winkle, J.: An integrated analysis of ergonomics and time consumption in Swedish 'craft-typ' car disassembly. Applied Ergonomics 36, 263–273 (2005)
3. Hignett, S., McAtamney, L.: Rapid Entire Body Assessment (REBA). Applied Ergonomics 31, 201–205 (2000)
4. Measurand_Inc_2.: ShapetapeTM Manual (2004) (Retrieved August 22, 2006) ftp://ftp.measurand.com/shapeware/ShapeWare_Release_II/ShapeWare_II_manual_only/help_shapeware%20II.pdf

Digital Human Model Based Participatory Design Method to Improve Work Tasks and Workplaces

Kaj Helin[1], Juhani Viitaniemi[1], Jari Montonen[2], Susanna Aromaa[1], and Timo Määttä[1]

[1] VTT, P.O. Box 1300, FIN-33101 TAMPERE, Finland
[2] VTT, P.O. Box 1000, FIN-02044 VTT, HELSINKI, Finland
firstname.lastname@vtt.fi

Abstract. The objective of this research project was to improve manual work tasks and workplace design with a new digital human model based design method. The idea of the method was to make the design and analyze of work and workplaces easy for floor level development case. It also should to be exploitable in the context of participatory design approach. The developed method was implemented on a production design simulation platform. It was designed to be used in design of human factors, performance and functionality of a production concurrently. The implemented tool includes basic human motions which exploit real human motion data, effective work design features to easily generate variational solutions, embedded ergonomic analyses and checklists to help analyzing different work environment solutions, and to document the design outcome. Four industrial case studies were completed with the tool. The results show that the tool is feasible for individual and group design work, and has positive impacts on the design process and on the way how individuals can influence on her or his future work in production system.

Keywords: Computer aided ergonomics, workplace design, ergonomic analyses, participatory design, posture, motion analysis, digital human model.

1 Introduction

The design process is complex combination of different information and it is often difficult to handle and use the information effectively. There are plenty of assessment methods to help principal design process, but they rarely take acknowledgement to human safety and ergonomics. Ergonomics and human work tasks are not systematically considered since several hindering design aspects such as technical orientation or fragmentation of design [6]. Production design is based on to take account well technical and cost-effective limitations, and it is common that priority of human factors come after the other designs problems. However, at the machinery typical human work tasks are control tasks, loading and removing tasks. Therefore, safety and ergonomics are important part of the design and they should be acknowledged in the design process concurrently with or even before applying the other aspects [9], [10], [11] to ensure good usability and overall performance, and to avoid extra costs of re-design and corrective actions.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 847–855, 2007.

To make design and especially work place design more effective the special computer aided methods and tools such as digital human models have been developed during past few decades. Most of the digital human models aren't available to ordinary designers because of their poor availability in context of common design tools, high costs, or software compatibility limitations, and beside the fact that also special knowledge may be needed to use digital human models in accurate way. Moreover few of the tools actually include necessary versatile assessment methods.

There are various human models available e.g. biomechanical [4], 2D versions or such models that are used in 3D environment. Despite of their complexity and accuracy they seldom have same varieties of motions as real human being does being a question of realism versus algorithmic efficiency. Especially in case of exploiting virtual environments also motion capture systems are typically applied because they can help to obtain and record real human motions as they are performed in activities related to use cases. However, exploitation of motion capture and motion analysis would be even more active to comprehend human motion prediction in design of work, workplaces, and especially concurrently with the corresponding machinery and production functionalities design already in early design phases.

There is strong argument for a participative approach to ergonomics at an individual and organizational level in workplaces [13]. Commonly participative development groups consist of workers, experts and decision-makers. They have been increasingly used in industrial cases to develop work in context to the production. Also professional teams who are responsible of their own work development have been utilized. According to the trend also the workers on the floor level will be more responsible for the development of their own work and working environment and conditions. When workers responsibility and possibility to influence increases, then the workers will be more committed to their work. As a result to this scene there is a demand for a computer based digital human model tool which includes ordinary basic postures and motions with novel features and ways of exploit them in context of common design. However, the necessity is that such a tool can be used easily and concurrently to design overall environment variations for both existing systems and systems being at the design and development stage.

2 Objectives

The objective of this study was to develop a new digital human model based participatory design method to improve manual work task and workplace design, and to be used concurrently with production design. It was emphasized that by using this method also an ordinary worker would identify quickly workplace and work task problems and spots that require improvement. Moreover, the objective was to create a tool with which one can also produce various illustrative 3D models as solution proposals to control physical load of a human and increase productivity.

3 Material and Methods

The background for this study was composed of the state of the art in this field (questionnaires, literature search). According to this background investigation the specified goals for the method development was defined. Furthermore the requirements for the platform, implementation of the method and development of the tool were specified.

There were various options available for the platform. One was to develop even an individual design tool. However, the comparison of possible outcomes and resources needed to develop the tool, gave a solution which also the case companies supported. The decision was to choose a platform which has been developed especially for of production design purposes and which was implemented already in earlier cases.

The software platform of Visual Components Ltd was based on so called component modeling which means that each component model includes most of the simulation codes. The idea in this software was simple enough for a desired tool - just collecting the components of a work environment from existing libraries which are made available by the manufactures of equipment and machinery. Also the designers of a company could easily make and code their own company specific models to be attached to the component library. Also an important aspect was the easiness of an ordinary worker to compose his or her own workplace just by collecting the components from the libraries in house and in internet. The components could be connected by dragging and dropping them into the layout. The component was able to be placed only in a correct way to make production flow alive.

The new method has been implemented as a tool, the OSKU-tool. The tool was tested in a laboratory case. The development work was expanding the basic human motion database and implementing sufficient analysis, and ensuring the integration of added features. The tool has been implemented also in use cases which correspond to real work, such as manual handling task of a company. Also training for the use of the tool was provided for the case companies.

4 OSKU – Digital Human Model and Analysis Toolkit

4.1 Digital Human Model and Motion Database

The digital human model was one of the primary component models to be developed to follow the methodology of digital human models and its exploitation by using participatory approach. The method is based on exploitation of a kinematical digital human model with pre-analyzed basic human motions composed in an accessible database. The method in detail describes features that provide means to model, analyze, simulate and evaluate ordinary industrial production tasks, and to compare properties between variational production systems and their related tasks considering production key figures and human factors. To improve the easiness of human action modeling it was decided to obtain and couple a basic human motion library to the usage of a digital human mode to obtain predicted required human motions. Library creation was to be as easy and effective as possible. A sufficient way was exploiting the real human motion recording which is partially similar type task that is performed e.g. with virtual reality applications. Also an easy modification of the library data content was a necessity. The library was later implemented as database, a part of the tool implementation.

For achieving the necessary accuracy of digital human model the motion capture system was used. The data suit MotionStar Wireless® by Ascension Inc. and data gloves 5th Glove ’95 by 5DT were used (Fig. 1). The data suit was equipped with 8 tracking sensors for obtaining position, orientation and to perform calibration. The data gloves were attached to wireless data transmission system unit of the data suit. The position of the head was obtained from a sensor on top of the head. Other sensors were attached to the arms and feet, the lower and upper back.

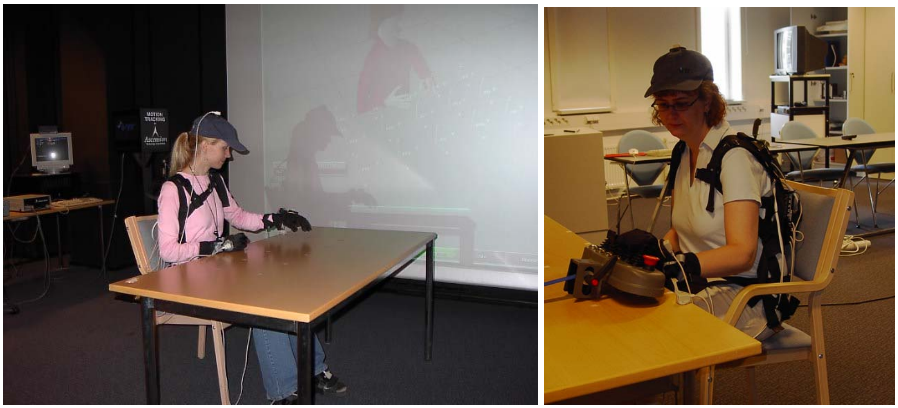


Fig. 1. Recording movements by a motion capture system.

A new digital human model was created with three different size model, namely P05, P50 and P95 (5 -percentile, 50 -percentile and 95 -percentile respectively)(Fig. 2).

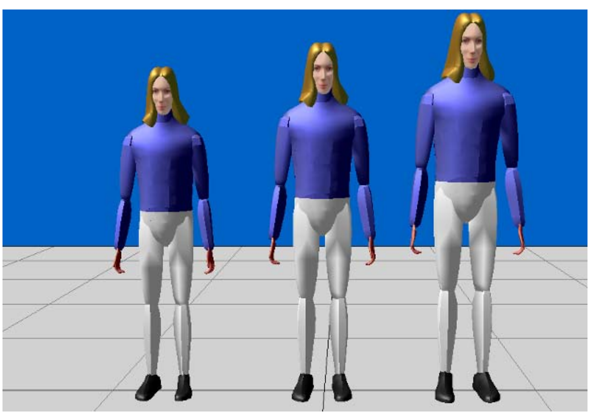


Fig. 2. Osku – female digital human models left to right: P05, P50, P95

4.2 Controlling the Digital Human Model

The developed digital human model - Osku – had the features for controlling its motions and producing work description in three different ways:

1. Direct use of recorded basic human motions from the database
2. Pointing objects in the working area for obtaining sufficient motions; software calculating and selecting the nearest basic human motion from the database.
3. Utilization of pre-defined sub work task, which are based on corresponding recorded motions from database or created interactively by integrating the basic human motions from the database.

User interface has been divided in three main areas (see Fig. 3).

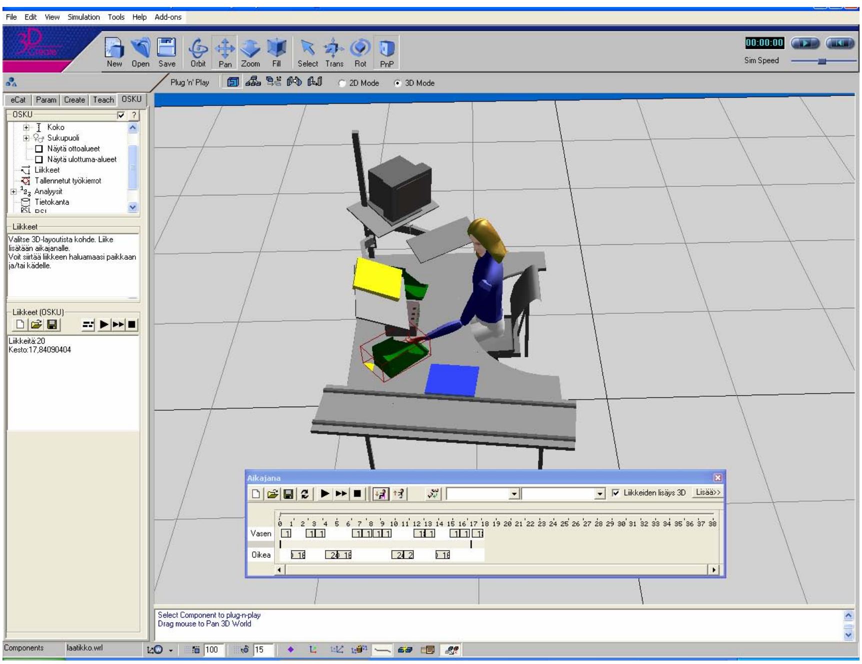


Fig. 3. Osku – digital human model and user interface. Also timeline dialog box with left and right hand movements.

Main use of upper area is focused on creation and modification of digital human model and also analysis of designed work tasks. Middle area consist list of recorded movements. Lover area consist lists of selected movements for designed works tasks.

4.3 Analysis Toolkit

The developed method was based on easy exploitation of recorded and pre-analyzed motions to predict the human motions of work in tight context to production and machinery. Therefore the implemented tool included various safety and ergonomic analysis, and if possible, embedded partly into motion data. The use of standards, posture analyses and general ergonomic analyses were taken account in the development of the new digital human model (Fig. 4). Common standards for anthropometrics (e.g. EN ISO 14738, 2002 and EN ISO 7250:1997), [1], [2], [3] and other workplace dimensions were used in the development as well.

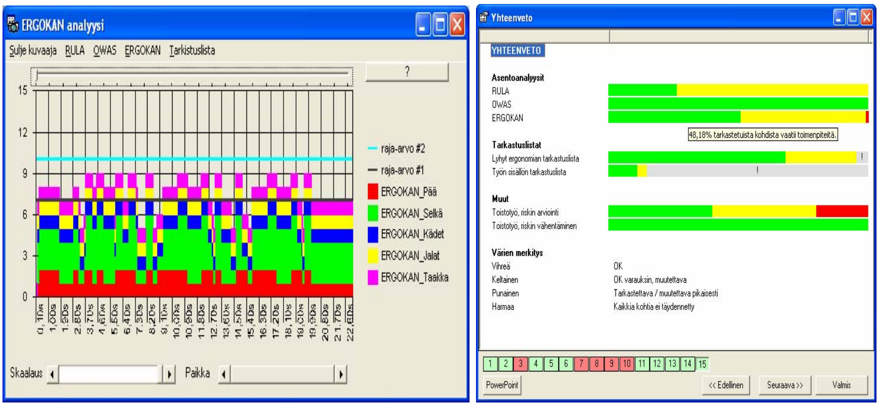


Fig. 4. Left: ERGOKAN – analysis result windows. Right: Conclusion windows of analysis, which based on ergonomic analysis and work analysis checklists.

For analyzing ergonomic postures RULA [7], Owas [5] and Ergokan [8] were implemented based on motion capture data and connected to the motions (Fig. 4). To find out sufficiency of forces tables like SNOOK [12] and standard based calculation were used. For general ergonomics and work environment analyze checklists were implemented. Checklists were also created to elicit the issues and aspects of cognitive ergonomics [2].

5 Results

The novelty of the new method developed was expected to be in the capability to fulfill the demands for modeling the human actions of a work and work tasks tightly connected to the design of the production and the machinery. The development of this new method adding the required efficiency for participatory design, and applying the digital human models succeeded well. The implementation of this method as a tool, the OSKU-tool was successful, too. The method and the OSKU-tool combine the commonly different design areas to be able to carry out concurrently and exploiting participatory design approach. Following an ordinary human natural performance to work and to interact with machinery can be done by using the OSKU-tool. Thus a designer or even a member of involved employees in production, can in a comprehend way ensure about the application of human based limitation and possibilities for technical design of workplace and machinery, and also influences to the production. Thus a designer or end-user can concurrently develop variational solutions to be visualized and assessed by many ways as described, and also exploit participatory approach benefits.

Main feedback from the companies' end-user was that the OSKU-tool can well be applied for participatory design of work and workplaces. Also the digital human model is easy to use and the interface has several user-preferable ways to choose and to carry out the design tasks. This is important for the easy use of the tool to be able apply each user's own habits to design the details.

There were unexpected problems in installation of the software platform itself in some companies because of the licensing process and non-floating usage. The basic interface of the software platform itself is simple but it includes plenty of features to be learned and trained, and also a requirement to be used frequently to be smoothly adopted in common practice of work and workplace design.

The installation of OSKU-tool features was easily done which is an important characteristic for a tool. The OSKU-tool was so far implemented using the most powerful version of the software platform which had strong influences to feedback, too. The modeling capabilities of human work were easy to use for those users experiencing other 3D software applications. The adoption of the OSKU-tool features were smooth to get easily started applying also OSKU in design besides the other software platform design and simulation properties. The OSKU-tool interface is added as a part of the software platform interface.

However the feedback was for ordinary users that some of the features included in the OSKU-tool at the moment still need too much additional expertise. Therefore an ordinary worker can't use tool as effectively as it was expected. Because of still the early development phase of the OSKU-tool, some of the dialog boxes are unlike and have still lacks of compatibility and visual features compared to those of the platform itself. There was a notice for some human motion prediction creation problems resulting to unnatural or too rough motion predictions for design. Another notice was for supporting weakly the design purposes, because of including too much of simplicity in obtaining the basic human motion data for database, or producing additional data to support design and dimensioning process yielding to difficulties comprehend the content.

6 Summary

The results show that the tool is feasible for individual and group design work, and has positive impacts on the design process and on the way how individuals can influence on her or his future work in production system. This tool can be used by most of people in every workplace or workgroup for work development illustrations and simulations. Moreover it can be enlarged and updated with model libraries and analyze packages for company's specific work and work environment. The functionality properties and data may be added or modified later on taking account more precisely the characteristics and requirements of use cases.

This kind of easiness will build a very strong level of confidence and that is nowadays the key issue in the fast paced industry. Thus this kind of tool will enable in a company that different people can work together to arrive at the best possible solution with the greatest level of confidence and in the shortest possible period of time.

The users of these kind of tools have noticed that the focus is on the creation of a solution, a concept and the speed of understanding it by all the involved. It is known that use of an interactive visual representation will remove ambiguity and make the involved even more innovative and keen to continue further. The production workers can be more responsible for the development of his or hers own work and working environment, and thus they may be more committed to their work.

In addition the data generated during the development stage is able to be utilized again in all areas of product marketing, documentation and training material, too. Thus the digital 3D models are used to create a common understanding of the unique operational issues of production and benefits throughout companies own organization as well as their potential sub suppliers and partners. These kinds of tools have also great potentials in enhancing communication internally in a company between design, manufacturing, sales, applications, operation and service.

The advanced tool of this work is still under development and needs still further feedback and further research to enhance the capabilities utilizing the database and in analyzing and designing different kind of production systems and human activities in the system. However the experiences of different ways to usage in different design areas the OSKU-tool are enough encouraging continuing the development and the studies complementary to each other to get a better design method and tool.

Acknowledgments. The study was funded by TSR (The Finnish Work Environment Fund) carried out in OSKU project and by TEKES (Finnish Funding Agency for Technology and Innovation) carried out in KONEMASINA project.

References

1. EN ISO 7250:1997 Basic human body measurements for technological design (ISO 7250:1996)
2. EN ISO 10075-1 - 3: Ergonomic principles related to mental work-load. Part 1: General terms and definitions (ISO 10075:1991) (2000) Part 2: Design principles (ISO 10075-2:1996) (2000) Part 3: Principles and requirements concerning methods for measuring and assessing mental workload (ISO 10075-3:2004) (2004)
3. EN ISO 14738:2002 Safety of machinery - Anthropometric requirements for the design of workstations at machinery (ISO 14738:2002)
4. Chaffin, D.B., Andersson, G.: Occupational biomechanics, 3rd edn. John Wiley & Sons, Inc, New York (1999)
5. Karhu, O., Kansi, P., Kuorinka, I.: Correcting working postures in industry: a practical method for analysis. Applied Ergonomics 8(4), 199–201 (1977)
6. Launis, M., Kiviluoma, M.: Considerations of ergonomics in designing manufacturing machinery. In: Proceedings of 33rd Annual Congress of the Nordic Ergonomics Society, Publications 7, University of Tampere, pp. 70–73. School of Public Health, Tampere (2001)
7. McAtamney, L., Corlett, E.: RULA: a survey method for the investigation of work related upper limb disorders. Applied Ergonomics 24(2), 91–99 (1993)
8. Määttä, T.: Analysing working postures using ”ERGOKAN” -method - An example of packing work at a bakery. In: Proceedings of the 12th Triennial Congress of International Ergonomics Association, August 15-19, Toronto, Canada (1994)

9. SFS-EN ISO 6385. Ergonomic principles in the design of work systems (ISO 6385:2004) 1st edn. (2004)
10. SFS-EN 614-1. Safety of machinery. Ergonomic design principles. Part 1: Terminology and general principles. 1st edn. (1995)
11. SFS-EN 614-2. Safety of machinery. Ergonomic design principles. Part 2: Interactions between the design of machinery and work task. 1st edn. (2001)
12. Snook, S.H., Ciriello, V.M.: The design of manual handling tasks: revised tables of maximum acceptable weights and forces. Ergonomics 34(9), s. 1197–1213 (1991)
13. Wilson, J., Corlett, E.: Evaluation of human work, A practical ergonomics methodology, 2nd edn., p. 1134. Taylor & Francis Ltd, London (1995)

Investigation on Ergonomics Characteristics of Protective Clothing Based on Capture of Three-Dimensional Body Movements

Huimin Hu[1], Li Ding[2], Chunxin Yang[1], and Xiugan Yuan[1]

[1] School of Aeronautic Science and Technology, Beihang University, Beijing 100083, China
[2] Dept. of Biological Engineering, Beihang University, Beijing 100083, China
{Huimin,Hu}hhm312@163.com

Abstract. The objective of this investigation is to measure and document any significant difference in motion performance between the partial pressure suit (PPS) and the common clothing (CC), and therefore to explore the ergonomics evaluation indexes and appraisal methods of protective clothing (PC). This paper presents the results of three human mechanics tests and two operational performance tests conducted on nine male subjects which made possible the comparative evaluation under these two different situations. The human mechanics tests included a range of motion (ROM) test, a grip strength test, and a torque test while the operational performance tests included a target-aiming test and a screwing nuts test. Significant difference was found between the two suits in the ROM test, the grip strength test as well as the target-aiming test. This study demonstrates that motion performances vary significantly between the two suits and ROM together with grip strength can serve as ergonomics evaluation indexes of PC.[1]

Keywords: motion performance, ergonomics evaluation, range of motion (ROM), partial pressure suit (PPS), protective clothing (PC).

1 Introduction

Almost the design of any equipment could not be realized without the ergonomics evaluation which provide evidence for its feasibility and reliability [1~5]. At present, the capability of life guarding provided by protective clothing including spacesuit, anti-G suit, partial pressure suit (PPS) and other safety clothing has been developed fairly well. Therefore, the emphasis on research at home and abroad has been attached to the conversion from the life protection to the work efficiency. Ergonomics of protective clothing demands that PC should meet the requirements of human mechanics, namely the mechanical characteristics of clothing joints should accord with the motion DOF (degree of freedom), range of motion (ROM), force and moment of human joints[6~14]. For example, the joint motion angle of spacesuit should meet the need of space

[1] Sponsored Project: National Youth Nature Science Fund (50406008).

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 856–864, 2007.

manipulation task under the two conditions with pressure and without pressure. According to the requests of ergonomics, the integral joint moment of spacesuit couldn't exceed 20 percent of the corresponding maximal moment of human joints mobility [5]. Hence the designers of clothing should scheme out motion structure of clothing joint with less moment, lager range and lighter weight to the best of their abilities [2].

Dressed in PC, human will undergo restricts from clothing pressure, which directly confine human motions and touching distance, so as to affect the workspace envelope [3, 15, 16]. The reason that the dressed person would always keep moving when he completes his task drives us to investigate the motion performances of PC and find out ergonomics evaluation indexes and appraisal methods for the objective of analyzing work efficiency of PC. During this process, how to take ergonomics parameters into account during the primary stage of designing work becomes the most desirable issue to be overcome.

Vykukal, Newman, et al applied direct measuring methods to analyze the motion performance of spacesuit [17], [18]. And there were also some researchers who proposed to apply robots to build ergonomics measure system [19~21]. White, et al used subjective methods to investigate work efficiency of anti-G suit [22]. Kozycki put forward a modeling and simulation approach to examine the encumbrance of helicopter aircrew clothing and equipments [23]. O'hearn used a video-based motion analysis system and a force plate to capture the maximum ROM in various planes of the body and the analysis of gait kinematics and kinetics [24]. There were also some subjective methods, directive measuring methods, and performance comparisons methods of operational items while evaluating the ergonomics of other PC [25~30]. Nevertheless, problems still exist with regard to system ergonomics evaluation indexes, real wearing tests, advanced measuring methods, the relationship between ergonomics indexes and operational performance, and standard appraisal system, etc.

This paper applies three-dimensional motion capture system and some measuring equipments of operational performance to gain basic human mechanics datum with nine healthy male undergraduates under two kinds of wearing state. Experimental items include the tests of range of motion (ROM), strength, torque and the accomplishing time of operational task. All the human movements used to measure joint motions were designed according to the theories of physiology, anatomy, kinematics, and ergonomics. Comparative results of two suits and the relationship between ergonomics indexes are presented in this investigation. Conclusions of this research could provide experimental basis for ergonomics evaluation of PC, and help to develop simulation analysis and ergonomics appraisal software of PC.

2 Materials and Methods

2.1 Experiment Equipments

2.1.1 Three-Dimensional Body Motion Capture System

VICON460 system was used to measure ROM of joints in this study. This equipment consists of the hardware (Workstation, cameras etc.) and the software applications systems for the complete recording and analysis of body motion. The principal parts are

VICON Workstation and VICON Datastation. The theory of motion capture is actually the recording of body movements by an array of infrared video cameras in order to represent them in a digital environment.

2.1.2 Hand Dynamometer

It is an instrument for testing strength which is composed of a sensor and a digital display which can display the moment strength. The minimum graduation of dynamometer is 0.1kg and the measuring range is from 0 to 100kg.

2.1.3 Torque Instrument

It is a screwing device with a 30-*mm* diameter knurled head. Its minimum graduation is 0.01*N.m*, and the measuring range is from 0 to 5*N.m*.

2.1.4 Target Board

It is a rectangle flat board with five equidistant buttons which are fixed on top left corner of the panel. The gradient of board is adjustable according to requirements.

2.1.5 Nuts-Assemble-Panel

This is an iron board with three bolts and suited nuts. The diameters of three bolts are respectively 16*mm*, 12*mm* and 8*mm* which are fixed on the board.

2.2 Experiment Methods

2.2.1 Range of Motion (ROM)

Before the experiment is carried out, 39 reflective markers are attached to subjects' body to acquire the body parameters for the experiment. Then the subjects stand at the center of the defined positions of the VICON capture volume and fulfill all the movements of Upper limbs and Lower limbs as required by recording design ahead. Then the subjects put on PPS and repeat all the above-mentioned actions.

2.2.2 Grip Strength

The subject sits down right before the test board and makes his trunk erect. All subjects must outstretch their right arms as long as possible and grip the Hand Dynamometer to the best of their abilities. The maximal force of Grip Strength is recorded.

2.2.3 Torque

The subject keeps the same posture as 2.2.2 and twist Torque Instrument clockwise. The maximal torque is recorded.

2.2.4 Target-Aiming

Subjects sit down right in front of Target-Aiming-Board with their right arms put on the appointed place. The Target-Aiming-Board is placed within the motion envelope of subjects. Keep the same sitting state as Grip Strength. While receiving the signal, the subject outstretches the right arm to touch and press down the buttons one by one as quickly as possible. The experiment regulates that the subjects should not only point to the correct place but also press down the buttons (This task is designed based on the real task properties of panel operations).

2.2.5 Screwing Nuts

The subject takes up quickly the Nuts-Assemble-Panel placed on the ground and set on the experimental table. And then they must screw the three nuts onto the suited bolts. Take record of the total accomplish time.

2.3 Experimental Contents and Design

2.3.1 Experiment Contents

Table 1. Summary of experiment items and tested joints in CC and PPS

Sorts	Items	Contents	Equipments
Human Mechanics	ROM	Shoulder Joint Elbow Joint Knee Joint	VICON460
	Grip Strength	Hand	Hand Dynamometer
	Torque	Hand	Torque Instrument
Operational Performance	Target Aiming Time	Right Arm	Target Board
	Screwing Nuts Time	Right Hand	Nuts-Assemble-Panel

2.3.2 Tested Movements of ROM

Table 2. Summary of tested joints and status of movements

Tested Joints	Tested Movements
Shoulder Joint	The Upper arm Abduction/Adduction/Swing
Elbow Joint	Forearm Flex/Extension
Knee Joint	Calf Flex/Extension

2.4 Subjects

A total of nine healthy male subjects give their informed consent to participate in this study. Anthropometric dimensions, height and weight are also taken before the experiment is carried out. They are all undergraduates and the ages range from 22 to 25, height 165*cm*~176*cm*, weight 55*kg*~67*kg*, all right-handed and voluntary to take part in this research. Before the experiment they attended training and master the experiment essentials very well.

2.5 Data Analysis

The raw video data are collected and processed by VICON Workstation. A three-dimensional digital motion model is reconstructed according to the anthropometric parameters with the help of Lower Body; the reconstructed double-line model can be analyzed by Polygon in terms of Kinematics and Kinetics; then the program of C-language is applied to collect the data which reflect human mechanics parameters; When actions are done more than three times, we adopt the mean; finally analysis of variance and Paired-Samples T Test are completed by SPSS12.0.

3 Results

3.1 Range of Motion (ROM)

Table 3. Comparison between the Joint motion degree in two wearing states

Tested States	Tested Actions	Tested Joints	ROM(deg)
CC◆	Up/Down Swing of Right Arm	Shoulder	125.0±22.4
PPS			95.8±13.0★★★
CC◆	Forth/Back Swing of Right Arm		121.3±8.7
PPS			107.6±8.7★★★
CC◆	Adduction/Abduction of Right Arm		136.4±17.7
PPS			114.9±17.5★★★
CC◆	Flex/Extension of Forearm	Elbow	129.9±6.4
PPS			113.7±8.7★★
CC◆	Flex/Extension of Calf	Knee	122.2±14.1
PPS			111.8±12.6★

Compared with ◆the control group, ★P<0.05, ★★P<0.01, ★★★P<0.001

3.1.1 ROM of Shoulder Joint

Seen from fig.1, 2 and 3, the three maximal ROM of shoulder joint of PPS (round points) are all less than that of CC (rectangle points) with total subjects. Difference showed in table 3 is very significant (P<0.001). Therefore, based on this experimental result, we could believe that PPS heavily affect the motion performance of shoulder joint, and the ergonomics indexes of ROM are very sensitive to PPS. By regressive analysis we cannot find effective functional relationship among the three indexes of ROM.

3.1.2 ROM of Elbow Joint

ROM of elbow joint of CC as showed in fig.4 is all greater than that of PPS with nine subjects. The result of variance testing demonstrates that the difference is also very significant (P<0.01, showed in table 3). Experimental result shows that ROM of elbow joint could act as an ergonomics evaluation index of PPS.

3.1.3 ROM of Knee Joint

We can see from the experimental results that the maximal ROM of Knee joint of CC as exhibited by nearly all the subjects (with one exception) are greater than that of PPS. The difference demonstrated in table 3 is significant (P<0.05), and the index of ROM of Knee can be used to evaluate motion performance of PPS.

3.2 Operational Performance Items (TAT and SNT), Strength and Torque

Table 4. Experimental results of operational performancej tems, strength and torque

Tested States	TAT(s)	SNT(s)	GS(Kg)	Torque(N.m)
CC◆	7.3 ± 0.9	20.6 ± 2.7	45.5 ± 6.9	2.4 ± 0.8
PPS	8.2 ± 0.8★★	21.2 ± 2.7	41.8 ± 8.4★	2.5 ± 0.7

TAT: Target Aiming Time; SNT: Screw Nut Time; GS: Grip Strength

3.2.1 Target Aiming and Grip Strength

Seen from Table 4, PPS significantly affects TAT ($P<0.01$) and GS ($P<0.05$). These experimental results accord with the conclusion that PPS significantly weaken ROM of shoulder-elbow joints and Grip Strength. Because when the subjects conduct the action of Target Aiming, the human mechanics characteristics of upper limb directly affect the accomplishing time of Target Aiming. When the ROM and Grip Strength are larger, the TAT will be shorter and vice verse. However, further research should be performed on this issue.

3.2.2 Screwing Nuts and Torque

There isn't significant difference between SNT and Torque in these two wearing states ($P>0.05$, showed in table 4). From the perspective of testing actions, these two items mostly require the participation of wrist. As a result, we can conclude from this experiment that PPS doesn't interfere with the motion of wrist.

3.3 Functional Relationship Among the Ergonomics Indexes

3.3.1 Functional Relationships Between TAT and ROM of Shoulder

Given that we name the shoulder's ROM of Up/Down Swing, Forth/Back Swing, and Adduction/Abduction respectively as ROM^1, ROM^2, and ROM^3. As the work efficiency of Target-Aiming is opposite to TAT, we suppose ROM^1, ROM^2, and ROM^3 are respectively x_1, x_2, and x_3, 1000*(1/TAT) is y. By linear regression analysis and variance test, we find out an effective functional relationship between the two indexes of y and x_2.

$$\hat{y} = -4.64 + 1.14x_2, R = 0.76, P = 0.03 < 0.05 \tag{1}$$

Formula (1) shows ROM of Forth/Back Swing of shoulder is linear relative to TAT.

In this research, we didn't find any effective functional relationship between x_1 and y, x3 and y.

3.3.2 Functional Relationships Between TAT and ROM of Elbow

Suppose ROM of elbow (ROM^4) is x_4, y is the same to 3.3.1; by regression analysis we find out that there isn't any effective functional relationship between y and x_1.

3.3.3 Functional Relationships Between TAT and GS

If GS is named as x_5, y is still TAT. By regression analysis we establish the regression equation as follow.

$$\hat{y} = 3.79 + 3.19x_5, R = 0.70, P < 0.05 \tag{2}$$

Equation (2) demonstrates that TAT is linearly relative to GS.

3.3.4 Functional Relationships Among TAT, ROM^2, ROM^4 and GS

There isn't effective functional relationship among four indexes.

4 Conclusions

PPS significantly affects ROM of shoulder joint, elbow joint, and knee joint, of which the shoulder joint is the most prominent and hence we could conclude that the human mechanics of shoulder is the most sensitive to PPS; ROM of the three joints can serve as evaluation indexes of motion performance of PPS; There aren't effective functional relationships between the ROM of three shoulder movements. Therefore, when we appraise the motion performance of PPS, these three indexes must be all taken into consideration.

Grip Strength is also sensitive to PPS and reflects the human mechanics of upper limbs, so this strength can be regarded as an index to evaluate the work efficiency of PPS. However, in order to keep consistence with the real operation task of Upper Limb, the experimental conditions must be strictly controlled.

PPS doesn't significantly interfere with motion performance of wrist, and in later tests hand Torque shouldn't be adopted.

Target Aiming item could comprehensively reflect the ROM and strength of human body movements and it is very sensitive to PPS. We can take this item as standardized ergonomics item for appraisal since by regression analysis we find out effective functional relationship between TAT, ROM_2, and GS. From this study we conclude that TAT can act as an integrative ergonomics item of PC. Under restricted experimental conditions, we can choose TAT to replace the other two indexes. Nevertheless, further research is required to be developed for shedding light on this relationship.

References

1. EJ. MS.: Human Factors Engineering Design. National Defense Industry Press (1992)
2. Wang, X., Wu, J.: Ergonomics of Applied Design. Press of China Textile University (2001)
3. Xu, J., Tao, K.: Human Engineering. China Textile Press (2002)
4. Lv, Z.: China Machine Press, Ergonomics (2006)
5. Chen, J.: Spacesuit Engineering. National Defense Industry Press (2004)
6. Siguang, J.: Hypoxia and Oxygen of Aeronautics and Astronautics-Physiology and Protective Equipment. People's Military Medical Press, Beijing (1987)
7. Jia, S.: The application of human mechanical. China Safety Science Journal (1995)
8. Li, L., Guo, X.: Some studied problems of aviation ergonomics. Ergonomics (1996)
9. Teng, Y., Yuan, C.: Study and development of foreign spacesuit. Technical report (1997)
10. Ji, B., Yuan, X.: Studies on astronaut intravehicular activity aboard. China J Aerospace Med (1998)

11. Wang, H., Wu, G.: Development of aviation ergonomics research. Chin J Aerospace Med (1998)
12. Jia, S., Chen, J.: Physiological basis of human mechanics and its application in the design of pressure suit. Space Medicine and Medical Engineering (1999)
13. Zhou, Q., Wei, Z.: Ergonomics design and study progress of human machine interface about EVA spaceship. Aerospace China (2005)
14. Zhou, Q.: The work efficiency of spacesuit for EVA and its research way. Aerospace Shanghai (2005)
15. Alexander, M., Laubach, L.: The effects of personal protective equipment upon the arm reach capability of USAF pilots. N76-10738 (AH) (1973)
16. Vykukal, H.C: Spacesuit Mobility Joints. NTIS (1978)
17. Vykukal, H.C., Webbon, B.W.: Pressure Suit Joint Analyzer. NASACASEARC113141 (1982)
18. Newman D.J., Schmidt P.B.: Modeling the extravehicular mobility utility (EMU) space suit: physiological implications for extravehicular activity (EVA). In: 30th international conference on environmental system (2000)
19. Nowrowzi: A overview of robot measurement technology. The Industrial Robot (1998)
20. Du, L., Gao, X., Liu, H., Li, T.: Arobot measurement system for spacesuit joint torque. Space Medicine & Medical Engineering (2003)
21. Wang, H., Jin, M., Gao, X.: Modeling of spacesuit's arm based on kinematics and measurement of mechanical characteristics of its joint. Robot (2004)
22. White, D.J., Utley, R.D., Bridges, C.W.: Integration testing of full coverage anti-g suit technology. In: SAFE Association, Annual Symposium, 31st, Las Vegas, NV (1993)
23. Kozycki, R.: Developing a modeling and simulation paradigm for assessing the encumbrance of helicopter aircrew clothing and equipment. In: SAFE Association Annual Symposium (1998)
24. O'Hearn, B.E., Bensel, C.K., Polcyn, A.F.: Biomechanical Analyses of Body Movement and Locomotion as Affected by Clothing and Footwear for Cold Weather Climates. ADA432258 (2005)
25. Schafer, L.E., et al.: Comparison of two shuttle launch and entry suits: research envelop, isokinetic strength, and treadmill tests. SAE (1992)
26. Adams, P.S., et al.: Three methods for measuring range of motion while wearing protective clothing. International Journal of Industrial ergonomics (1993)
27. Adams, P.S., et al.: Methods for assessing protective clothing effects on worker mobility. ASTM (1994)
28. Bishop, P., Ray, P., Reneau, P.: A review of the ergonomics of work in the US military chemical protective clothing. Industrial Ergonomics (1995)
29. Malapane, N.C., Shaba, M.N.: Comparison of functional performance of a solider in full chemical and biology protection versus battle dress. Industrial Ergonomics (2001)
30. Havenith, G., Heus, R.: A test battery related to ergonomics of protective clothing. Applied Ergonomics (2004)

Appendix

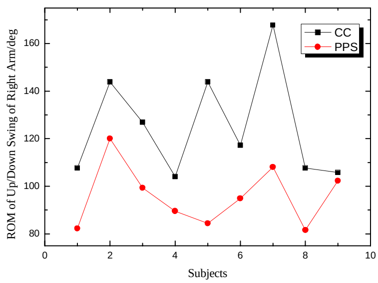


Fig. 1. ROM of Up/Down Swing of Right Arm

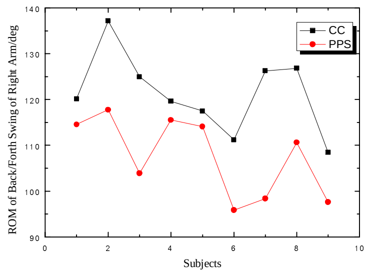


Fig. 2. ROM of Forth/Back Swing of Right Arm

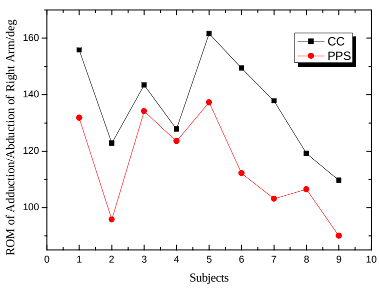


Fig. 3. ROM of Adduction/Abduction of Right Arm

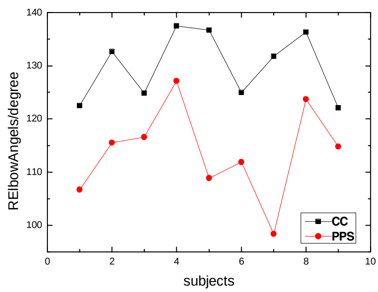


Fig. 4. ROM of Right Elbow

Strategy to Operate Cylindrical Interface-Operation Difference According to the Dimension of the Cylinder and That of the Hand

Ohki Kanezashi[1], Natsuki Miyata[2], and Jun Ota[1]

[1] The University of Tokyo
Hongo 7-3-1, Bunkyo-ku, Tokyo, Japan
{kanezashi,ota}@robot.t.u-tokyo.ac.jp
[2] National Institute of Advanced Industrial Science and Technology(AIST)
Aomi 2-41-6, Koto-ku, Tokyo, Japan
n.miyata@aist.go.jp

Abstract. In this article, we study how users determine the way to operate cylindrical interfaces (lever, switch, etc.) according to their dimensions, shape, and the user's hand dimensions. With a target of avoiding control failure, it is important to estimate user's intuitive operation of a given interface from design. Eight subjects were observed operating different cylinders with various height and diameter. The boundary between the tilting operation and other operations was found to be related to an aspect ratio; the ratio of the diameter to the height. The boundary between the sliding operation and rests was determined by a comparison of hand width with cylinder height. The boundary between the pushing and rotating operations was thought to be related to the length of the distal phalanxes of the fingers.

Keywords: Interface-design, Human-Model, Affordance, Ergonomics.

1 Introduction

Various types of cylindrical interfaces (lever, switch, button, etc.) are available to realize the same control function, such as switching a discrete status or adjusting a continuous value. When people are in a hurry, they are apt to operate the given shape of an interface as induced by the interface itself. Therefore control failure occurs if the interface requires a different method of operation from what the interface induces the user to do, which can result in serious accident. Interface operation depends on a user's physical description and interface factors such as dimension, shape, position, posture and texture. In order to avoid control failure, it is important to estimate a user's intuitive operation of the given interface in the design stage.

Research on the usability of interfaces is carried out in the field of ergonomics.[1)] This research creates guidelines for human-machine interfaces assuming that users know how to operate interfaces. However, the guidelines cannot be applied to situations where users do not know how to operate interfaces.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 865–873, 2007.

In the field of psychology, D.A. Norman suggests that "affordance", which was advocated by J.J. Gibson,[2] contributes toward informing users of how to operate interfaces.[3] However, it is unclear how users operate interface according to their physical description and interface factors.

In this paper, we focused on the dimensions of a cylinder and that of the hand. Our goal is to observe and explain how users determine interface operation in relation to the dimensions of the cylinder and the hand.

2 Classification of How to Operate Cylinders as an Interface

The way of operating a cylindrical interface is classified according to the direction in which the cylinder is moved. Considering interfaces used in our daily life, we have five kinds of operation when using cylindrical interfaces (Fig.1.).

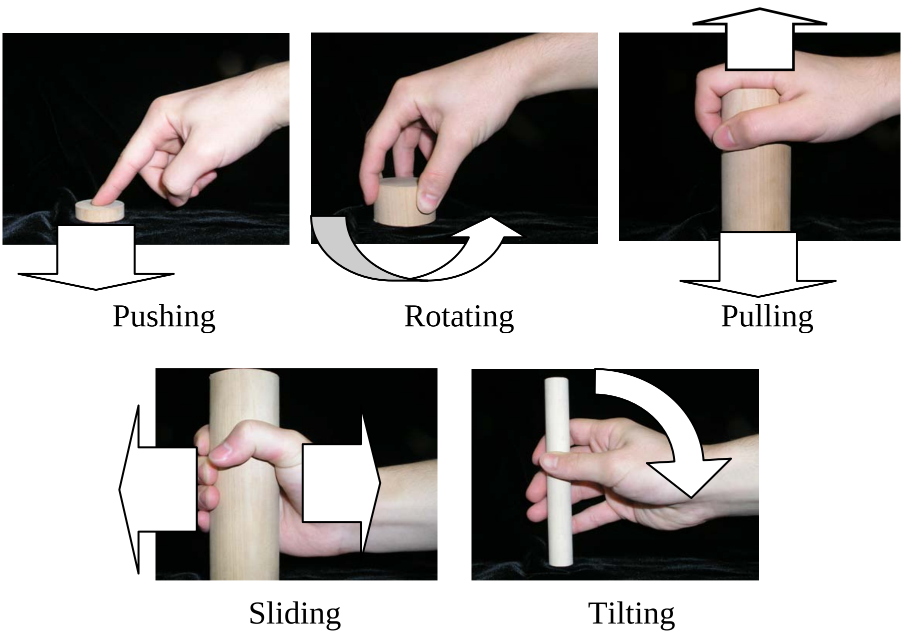


Fig. 1. Various ways to operate a cylindrical interface

- Pushing: the cylinder is pushed in the vertical direction, such as a button.
- Rotating: the cylinder is rotated around the vertical axis, like a rotary switch.
- Sliding: the cylinder is slid in the horizontal direction, like a sliding switch.
- Tilting: the cylinder is tilted around the horizontal axis, such as a lever.
- Pulling: the cylinder is pushed and pulled in the vertical direction, like a push-pull switch.

3 Experimental

Eight subjects were observed operating various cylinders, with different height and diameter, which were positioned on a desk. The interfaces were assumed to be interfaces to "switch from one discrete status to another" or to "adjust a continuous quantity".

3.1 Method

Experimental Conditions. Fig. 2 shows the experimental setup. The eight subjects were male students in their twenties. The cylinders had eight different heights (5, 10, 20, 30, 50, 100, and 200 mm) and seven different diameters (5, 10, 20, 30, 40, 50, and 60 mm).

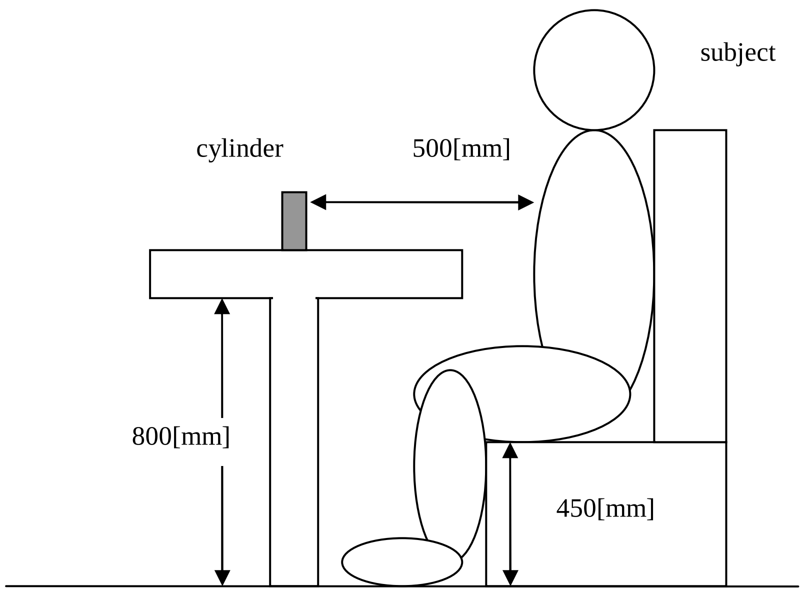


Fig. 2. The experimental setup

Experimental Procedure. Subjects were asked to operate 56 cylinders in turn as interfaces to "switch from one discrete status to another" and then to operate them as interfaces to "adjust a continuous quantity". The cylinders were presented at random. For each cylinder, the subjects began the trial with their eyes closed, and then operated the cylinder that was presented, as soon as possible after opening their eyes once given a cue.

3.2 Results

Figs. 3 and 4 show the observed operation when the cylinders were assumed to be a switching-control interface (Fig. 3) and a continuous-control interface (Fig. 4.) The background of each square in the tables was colored according to the type of operation. The color distribution shows that favorable operations exist according to the cylinder dimension and the tendency for favorability is different for the different control functions.

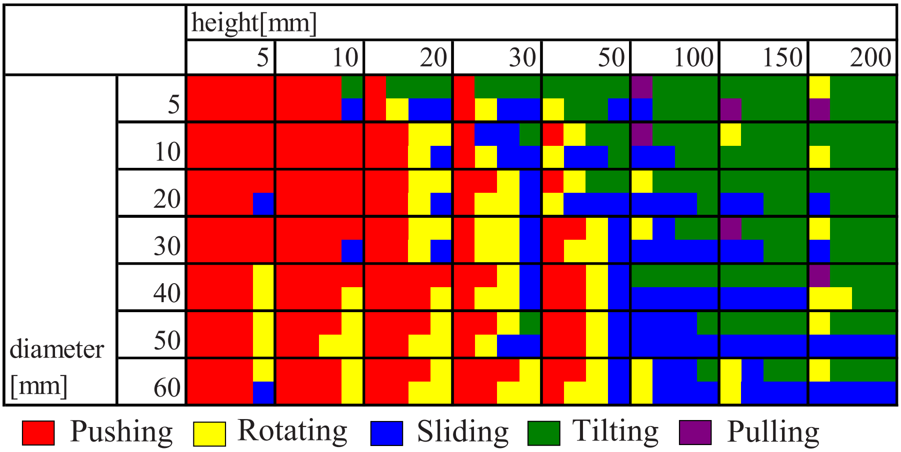


Fig. 3. Ways to operate a *switching-control interface* according to the cylinder dimensions

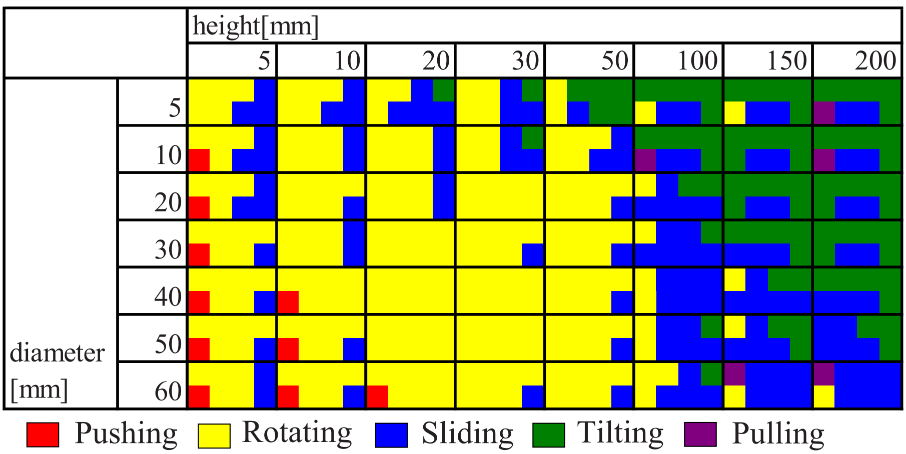


Fig. 4. Ways to operate a *continuous-control interface* according to the cylinder dimensions

4 Explanation for the Boundary Between Each Region

Based on the results shown in section 3, the cylinders assumed as switching-control interfaces can be divided into four groups. Firstly, by extracting the tilting operation region, and secondly, the sliding operation region, and then the pushing or rotating operation region (Fig. 5.) About half of the subjects pushed cylinders in the pushing or rotating operation region, and the remaining subjects rotated them. In the pushing or rotating operation region, some subjects selected only the pushing operation. Similarly, the cylinders assumed as continuous-control interfaces can be divided into three groups by first extracting the tilting operation region and then the sliding operation region from the rotating operation (Fig. 6.).

In order to estimate how users operate interfaces of arbitrary dimension, it is necessary to understand what creates the boundary between each region.

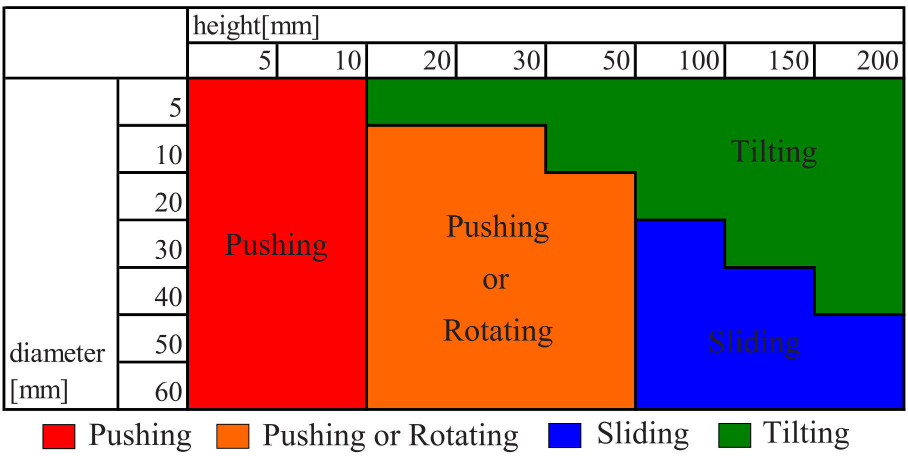


Fig. 5. Regions of operation for a *switching-control interface*, according to the cylinder dimensions

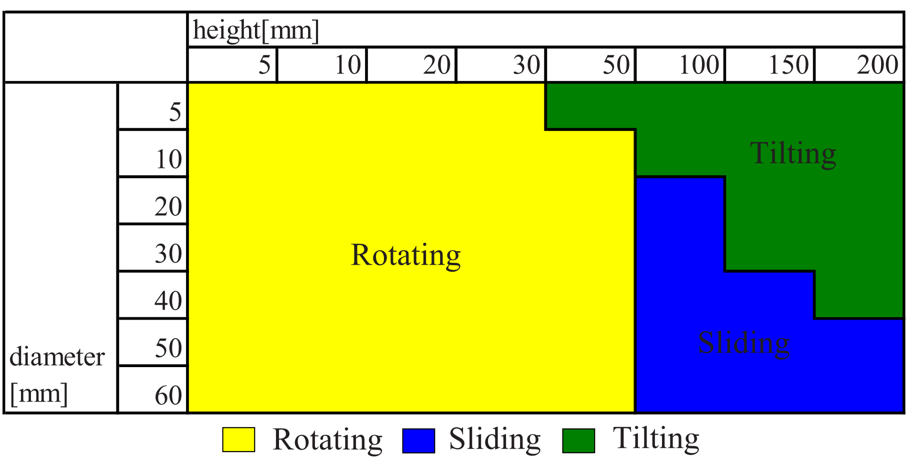


Fig. 6. Regions of operation for a *continuous-control interface*, according to the cylinder dimensions

4.1 The Boundary Between the Tilting Operation and Other Operations

Let us focus on the boundary between the tilting operation and other operations.

The tall and thin cylinders belong to the tilting operation region. The diagonal boundary indicates that the boundary is determined equally by its height and diameter.

Operation of each cylinder was estimated as tilting or otherwise. A successful estimation rate was defined as the separation ratio to find appropriate indices for estimation. We compared various indices that were expressed as a function of the cylinder dimensions.

The inertia moment multiplied by the height or diameter to the N_th power was used as the indices, because the inertia moment generally corresponds to the difficulty in tilting a cylinder. The separation ratio of the tilting operation was the highest when N was equal to five (Fig. 7.).

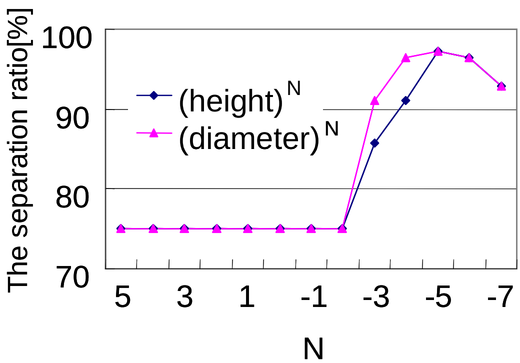


Fig. 7. The separation ratio of the tilting operation from other operations

Equations (1) and (2) show the inertia moment multiplied by the height and diameter to the fifth power. I represents the inertia moment when cylinders are tilted, D represents the diameter of the cylinder, H represents the height of the cylinder and ρ represents the density of the cylinder.

$$I/H^5 = \rho\pi\left(\frac{1}{64}\left(D/H\right)^4 + \frac{1}{12}\left(D/H\right)^2\right) \tag{1}$$

$$I/D^5 = \rho\pi\left(\frac{1}{64}\left(D/H\right)^{-1} + \frac{1}{12}\left(D/H\right)^{-3}\right) \tag{2}$$

As shown above, the index can be converted into an expression that includes only the proportion of the diameter to the height (the aspect ratio: D/H) as a variable. This means that the boundary between a tilting operation and other operations is related to the aspect ratio.

4.2 The Boundary Between the Sliding Operation and Other Operations

In the previous section, the tilting operation region was extracted. Next, the boundary between the sliding operation and other operations is explained.

Subjects tended to slide those cylinders that were higher than 100 mm, and to push or rotate cylinders that were lower than 50 mm. A change in operation was associated with the change of the relative disposition of the hand. Subjects grasped the side of cylinders to slide them, and the top of cylinders to rotate them, and touched the top of cylinders to push them. The fact that the subjects' hand width was from 65 to 95 mm indicated a dependency of the boundary height of the cylinders on the hand width (Fig. 8.).

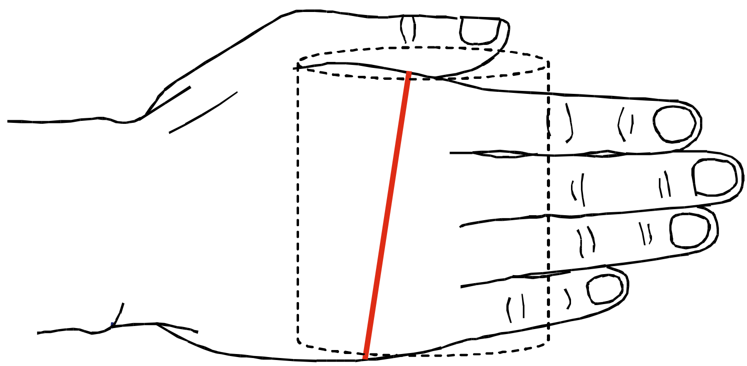


Fig. 8. Hand width and the height of the cylinder

To more precisely investigate the relationship between hand width and the boundary height of cylinders, additional experiments were conducted using cylinders with the same diameter (50 mm) and seven different heights (65, 70, 75, 80, 85, 90, and 95 mm). Subjects operated the seven cylinders ten times at random, assuming them as switching-control and continuous-control interfaces. The experimental setup was the same as the setup described in chapter 3.

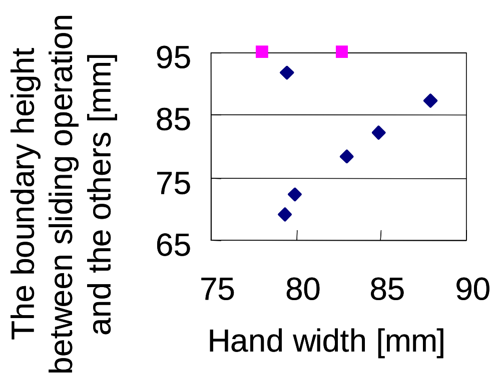


Fig. 9. The boundary height of a cylinder between the sliding operation and other operations, according to hand width

The boundary height was calculated by logistic regression.[4] The boundary heights of two subjects were over 95 mm, which was outside of the measurement range. Fig. 9 shows the boundary height according to the hand width. The hand width was measured by the method given by Kouchi.[5] The boundary height was proportional to the hand width in five subjects. The index of correlation was 0.99, which showed that the boundary height was determined according to the hand width.

4.3 The Boundary Between Pushing or Rotating Operations and the Pushing Operation

Finally, we take up the boundary between the “pushing or rotating” operation and the pushing operation when a cylinder is assumed to be a switching-control interface.

Five of eight subjects tended to rotate cylinders that were higher than 20 mm, and to push cylinders that were lower than 10 mm. They pinched the side of cylinders to rotate them, and touched them from the top to push them. The lengths of their phalanxes were from 20 to 35 mm. These results indicate that cylinders, which have a height apparently lower than the length of the user’s phalanx, are pushed, because of the difficulty in pinching the cylinder (Fig.10.).

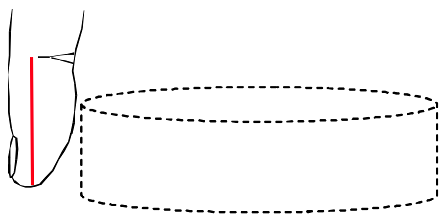


Fig. 10. The length of the phalanxes of the fingers and the height of cylinder

5 Conclusion

We observed how users determine interface operation in terms of the dimensions of the cylinder and the hand. The boundary between the tilting operation and other operations was found to be related to an aspect ratio. The boundary between the sliding operation and other operations was found for most of people to be determined by comparing the hand width with the cylinder height. The boundary between the rotating and pushing operations depends on a comparison of the length of the phalanxes of the fingers with cylinder height.

References

1. Grandjean, E.: Fitting the task to the man, Taylor &Francis (1990)
2. Gibson, J.J.: The Perception of the Visual World, Greenwood Pub Group (1974)
3. Norman, D.A., Draper, S.W.: User Centered System Design: New Perspectives on Human-Computer Interaction. Lawrence Erlbaum Assoc Inc, Mahwah (1986)

4. Walker, S.H., et al.: Estimation of the probability of an event as a function of several independent variables. Biometrika 54, 167–179 (1967)
5. Kouchi, M., Miyata, N., Mochimaru, M.: An analysis of hand measurements for obtaining representative Japanese hand models, SAE 2005 Transactions Journal of Engineser Cars - Electronic and Electrical Systems (2005)

Modeling of Human Head for Custom Wig Production

Youngjun Kim, Jungbum Cho, Bohyun Kim, and Kunwoo Lee

Mechanical Engineering and Aerospace Department, Seoul National Univ.,
Shillim-dong, Kwanak-ku, Seoul 151-744, South Korea
yjkim@cad.snu.ac.kr

Abstract. We propose an innovative 3D wig designing system which would replace the traditional manual method. The customer's head is scanned by a 3D head scanner and the whole design process is computerized. After sending the design data to the wig-making factory through the internet, the data are reproduced by an NC milling machine to make a wig pattern. To realize our integrated 3D wig designing system, various CAD technologies are implemented: 3D scanning, triangular mesh generation, texture mapping, mesh smoothing, mesh splitting, mesh offsetting, 3D sketch using NURBS curve, tool-path generation, etc.

1 Introduction

With the rapid development of computer graphics and computer vision, many applications of 3D human modeling are being introduced and widely used these days. The field of human modeling and its application is attracting great attention in recent years. Many 3D scanners for the human head have been developed and their scanned head data have been used to design helmets, face masks, glasses, etc. Although an ample number of these topics have been researched and applied to the practical field, few trials for the custom wig industry have been done. The wig manufacturing industry is still out-of-date and depends on manual labor. In this paper we propose a novel and practical CAD system for the custom wig industry. Our scanning method accurately reconstructs the customer's head data. Our scanning operation for head modeling is very simple and all designing procedures of the scanned data are computerized. The designed data file can be transferred to a wig factory through the internet, instead of sending the physical pattern. In the factory, the duplication of the customer's head is machined by an NC milling machine, then the duplicated polyurethane foam is used for wig manufacturing.

The remainder of the paper is organized as follows. Section 2 is devoted to head modeling and Section 3 explains the procedures for handling the head data for designing a custom wig. The results are shown and concluded in Section 4.

2 Shape and Color Reconstruction of Human Head

2.1 Hardware Configuration

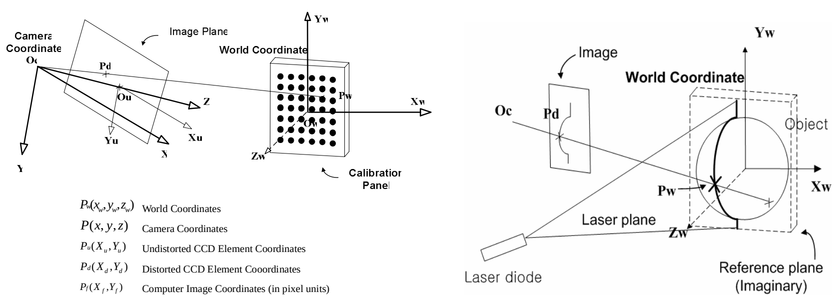
For the 3D head scanning, the cylindrical scanning type is suitable so that we do not need to perform further operations such as registration or merging. Laser triangulation

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 874–883, 2007.

is an appropriate method for rotational scanning. Our scanning unit, which rotates around human head, consists of two CCD cameras and a laser slit sensor. We used a B/W CCD camera for shape reconstruction and a color CCD camera for texture mapping. The laser slit sensor should not be harmful to a human's naked eye, so we adopted the class 1 laser sensor for safety. Our laser sensor's wavelength is 650nm and a bandwidth filter is attached to reduce the noise points for the B/W CCD camera, which obtains the shape of head.

The images for both the shape and color reconstruction are captured simultaneously while the scanning unit turns around the head. These captured images are for calculating points-cloud with geometric and texture coordinates. It takes about 10 seconds for one scan, which is not a short time for a person to sit with no swaying motion. So we devised a head support whose height and directions can be adjusted in order to set the subject's head at the proper position and help him not to sway. Fig. 1 shows the appearance of the head scanner and head support.

Fig. 1. Photos of our 3D head scanner (left) and head support (right)



(a) (b)

Fig. 2. (a) Camera geometry associated with calibration, (b) Calculation of 3D points by laser triangulation

2.2 Calibration

The aim of camera calibration is to find the relation between world coordinates and computer image coordinates. We assumed a pin-hole camera whose intrinsic and extrinsic parameters are computed by Tsai's coplanar camera calibration method[2]. The intrinsic parameters are focal length, optical center and radial distortion factor. The extrinsic parameters are the transform matrix between the world-coordinates and the camera-coordinates. A calibration panel that has a sufficient number of calibration points is fabricated precisely and attached to the scanner's rotational axis.

First we assign the pairs of world coordinates P_w and computer image coordinates P_f in the image of the calibration panel. The calibration parameters are computed by optimization of the input pairs of P_w and P_f. The following equations show the detail relations between P_w and P_f.

- Relation of world coord. & camera coord.

$$\begin{bmatrix} Xw \\ Yw \\ Zw \\ 1 \end{bmatrix} = \begin{bmatrix} & & & Px \\ \mathrm{Rot}(\alpha,\beta,\gamma) & & & Py \\ & & & Pz \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} Xc \\ Yc \\ Zc \\ 1 \end{bmatrix} \tag{1}$$

where $\mathrm{Rot}(\alpha,\beta,\gamma)$: 3*3 rotation matrix
(P_x, P_y, P_z) : Translation vector

- Relation of camera coord. & undistorted CCD Element coord.

$$x_u = f\frac{x_c}{z_c} \qquad y_u = f\frac{y_c}{z_c} \tag{2}$$

where f: focal length of camera

- Relation of undistorted CCD element coord. & distorted CCD element coord.

$$x_d = (1+\kappa_c r^2)^{-1} x_i \qquad y_d = (1+\kappa_c r^2)^{-1} y_i \tag{3}$$

where $r = \sqrt{x_d^{\,2} + y_d^{\,2}}$
κ_c: radial distortion coefficient of lens

- Relation of distorted CCD element coord. & computer image coord.

$$x_f = s_x d'_x x_d + C_x \qquad y_f = s_y d'_y y_d + C_y \tag{4}$$

where s_x, s_y: uncertainty image scale factor in x, y direction
d'_x, d'_y : effective x, y dimension of pixel in computer image plane
C_x, C_y : row, column numbers of the computer image plane center

We compute all the calibration parameters except for the uncertainty image scale factor (S_x, S_y) by using Eq. (1)-(4). Finally, we obtain all the calibration parameters including S_x and S_y by an optimization method that minimizes the distance between the world-coordinates of calibration points P_w and those resulting from the computed calibration parameters as initial values.

2.3 Calculation of 3D Points

During the process of scanning the laser slit beam is projected onto the subject's head. The each image of the reflected laser line captured by B/W CCD camera is converted via image processing of binarizing, closing and thining operation. Each final pixel of the processed image is converted to a 3D point using the camera calibration information. The computed world point, P_w in Fig. 2-(b), is the intersection point of the laser plane and the line which goes from the image pixel, P_d in Fig. 2-(b), to the relevant position on the virtual calibration target plane.

Line: $$\vec{x} = \overrightarrow{p_0} + t\vec{d} \quad (5)$$

Laser plane: $$\hat{n} \cdot (\vec{x} - \vec{x}_p) = 0 \quad (6)$$

Substitute (5) to (6): $$t = \frac{\hat{n}(\vec{x}_p - \vec{p}_0)}{\hat{n} \cdot \vec{d}} \quad (7)$$

where $\vec{x}_p$: point on plane, $\vec{d}$: directional vector of line, $\hat{n}$: normal vector of laser plane.

The colligation of the sections becomes the whole points-cloud. Each section data's coordinates are rotated with the rotation angle at the moment of capture. This enables us to get all the points-cloud data with no additional processes, such as registration or merge.

2.4 Mesh Generation

The aim of our triangulation algorithm is to be fast enough to generate a triangular mesh within about five seconds for a cloud with 10,000 to 20,000 data points. Triangles having all acute angles are most desired. The algorithm does not have to consider all data points for triangulation if we can accelerate the calculation.

Various triangulation algorithms have already been proposed. Approximation approaches are suggested by Hoppe et al [4] and Curless et al [5]. Interpolation algorithms, which use mainly alpha shapes, voronoi diagrams and delaunay triangulation, have been proposed [6], [7], [8] and [9]. As far as the final results are concerned, these approaches suit our needs, but these algorithms are inherently slow and cannot match the desired speed.

Another approach is the 'intuitive' category. The algorithm suggested by Linsen et al. [10] creates facets very fast. The algorithm can mainly be used for visualization because it does not create one single mesh out of cloud of points. And Bernardini et al. [11] have suggested a ball pivoting algorithm. In this algorithm, a ball is rolled over the mesh and those three points on which the ball is resting are connected to form the triangle. This approach is most suited to our requirement but it has following drawbacks. The critical factors for this algorithm are selection of ball radius and uniform density of cloud of points. Bigger ball would loose features and smaller ball may fall through the cloud of points during the pivoting operation.

Considering the suggested issues, we devised an intuitive approach satisfying the speed criteria. Our approach starts with a seed point and tries to grow the mesh in different directions. The design issues are described in the subsequent section. Fig. 3 shows an example to explain our approach step by step.

- Step 1: Getting the points around a seed point

First we find all points within a certain distance from a seed point p_0. This certain distance is a function of the density of points-cloud. With these points we find the least square plane and project all the points onto it.

- Step 2: Getting the first edge

The first edge is taken by joining the seed point and the point closest to it. Fig. 3-(b) shows the result.

- Step 3: Getting the triangle

We find out the third point to complete the first triangle at this step. As triangles with all acute angles are desired, we choose the nearest point from the first edge which can make an acute triangle. If we do not get any triangle meeting the criteria, then we select a triangle with an angle closer to 90 degrees. Fig. 3-(c) shows the result after this step.

- Step 4: Closing the triangles around the seed point

We select any edge of the triangle starting from the seed point. Let the edge be e_1. The line passing through edge e1 divides the plane into two half spaces. This line is shown as the dotted line in Fig. 3-(d). We create the facet, using Step 3, around the edge e_1 in the other half space of point p_1. In our example of Fig. 3, point p_2 is rejected during the process because it forms a facet with one angle greater than 90 degrees. By repeating the previous step around edge e_2 we get the next facets. We repeat this step until all facets around the seed point have been created.

- Step 5: Creating the triangles around the next point

The next seed point p_1 is the closest point to the first seed point p_0. By repeating Step 1 using the new seed point, we get the triangles. The computation of the new least square plane around the new seed point would be a costly operation. If the variations in the normal vectors of the generated triangles are within a certain acceptable value, the previously computed plane may be treated as the same least square plane for this seed point. We obtain the first triangle by using edge e_4, and we obtain the second triangle by using edge e_5 as shown in Fig. 3-(h). We proceed in

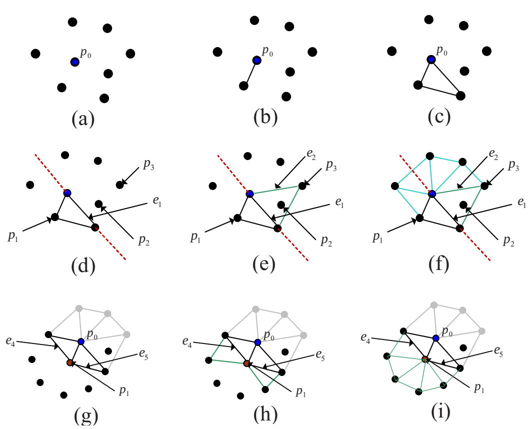


Fig. 3. Mesh generation algorithm (example)

this same way and create all triangles around the new seed point. The angle around the seed point may be measured to avoid folding. Fig. 3-(i) shows the result.

- Step 6: Repeating for all the points on the free boundary

We obtain the triangles by repeating the previous step for all the points on the mesh's free boundary.

As our algorithm is specifically tailored to human head triangulation, it would work well near flatter regions but may have difficulty near sharp edges. But our algorithm is computationally efficient and yields good quality mesh data. It takes about three seconds to generate a mesh from about 20,000 data points.

After the mesh has been generated, it is refined with a smoothing process for the visualization and the successive operations. We referred to Ohtake's algorithm [12] for mesh-smoothing. We applied a coefficient of 0.7 with 50 iterations for the most part, and a coefficient of 0.01 with 1000 iterations for the top area around the scanning axis, in his algorithm.

2.5 Texture Mapping

For color reconstruction, texture images are mapped on the mesh. Forty texture images are taken, while gray images are taken in the scanning process. For this operation we should select which images to use and find the pixel coordinates of those texture images. The selected region of texture is mapped on each facet of the mesh. It is decided by back-projection from the vertices of each facet onto the camera's origin using camera calibration information. The texture coordinates of the texture image represent the intersection point of this project line and the image plane.

We obtain texture images while gray images are being taken, so we select the region of texture except for the 0-15 degrees from the scanning line to avoid the red laser line. Of course the vertices of one facet must be associated with the same texture image(s). And when a facet has more than two texture images, those textures are blended by the OpenGL blend function to get a smooth effect on the image-shifting region. Fig. 7-(b) shows a result of texture mapping.

3 Designing and Machining

In order to manufacture a custom wig, the reconstructed head model data are processed in several steps. Our software provides some functions for designing a wig pattern. The procedures are as follows: 3D sketching, mesh splitting, mesh expansion and tool-path generation.

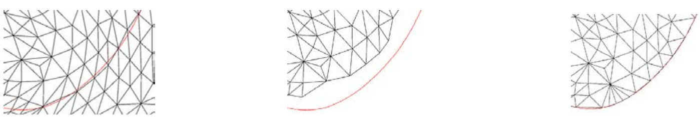
3.1 3D Sketch on Mesh

Sketching on three-dimensional polygonal mesh is a necessary process for communication between the wig designer in the shop and the wig manufacturer in the factory. The exterior boundary curve should be defined for splitting out the useless outer region. Our software provides various types of drawing and editing tools.

For 3D sketching, a 2D sketch is drawn first on the screen by a mouse or a pen tablet device. Then the sketched curve is projected onto the mesh with the viewing direction. A 2D sketch is created using NURBS curve interpolation with given control points. Then, a 2D sketch curve is sampled by appropriate intervals and these sampling points are projected onto the mesh. In order to accelerate finding the projected facets, all the facets are grouped by longitude and latitude. Searching is done in the 25 groups (5 longitude * 5 latitude direction), and then the projected facet is finally found in that group. Fig. 7-(c) shows an example of our 3D sketching result.

3.2 Mesh Splitting

The mesh splitting process is done to cut out the useless outer region and get only the concerned region of mesh. Using the 3D sketch tool, an exterior boundary curve is defined. The boundary curve should be a closed curve. In the first stage of our development we tried to cut the facets that were intersected by the boundary, but we changed the algorithm of mesh splitting to a mesh expanding problem by deleting facets which are on or out of the boundary curve and generating facets between the inner facets and the boundary. This is a much simpler way than cutting all the facets on the boundary curve. The inner facets are selected recursively. The search is started with a user-input facet by clicking. Then the new facets are created using the boundary curve and the inner facets' vertices. Fig. 4-(b) shows the inner facets and boundary, and Fig. 4-(c) illustrates the new facets created. The algorithm for creating the new facets is described in detail in Section 3.4.



(a) (b) (c)

Fig. 4. Mesh split procedures; (a) Original mesh before splitting (b) Deleting facets on/outside of boundary curve (c) Making new boundary facets

3.3 Mesh Expansion

It is necessary to generate extra space by expanding the boundary for manufacturing convenience. Because the boundary curve is a NURBS curve, we can easily obtain the offset curve. Considering the facets' size, we restricted the offset value with a moderate value in one step and expanded it gradually in several steps. The algorithm is explained below with pseudo code.

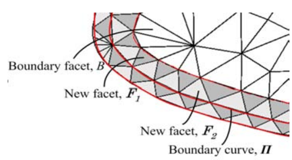
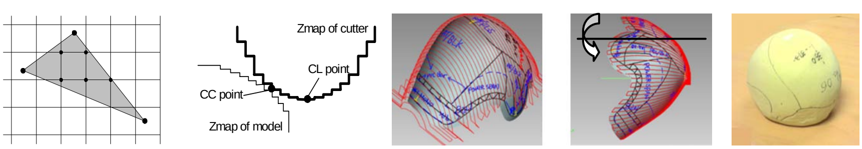


Fig. 5. Mesh expanding algorithm

For each expansion step:
  Set boundary facet list, B
  Find offset curve Π considering the tangential and normal directions of B
  For $^{\forall}B$,
    If the facet has one null edge (which has no neighbor facet):
      Find the min. distance position on Π from the mid position of the null edge
      Create a new vertex V and a new facet F_1 (only if the new vertex is not close to the previously added vertex)
  For $^{\forall}V$,
    If Π's curvature changes rapidly between the successive two V s, add some new vertices according the distance
    Connect all V s and make new facets F_2

3.4 Tool-Path Generation

Tool-path generation uses a Z-map algorithm. A Z-map is a 2D array which stores the Z-value of the model at the regular grid points. In order to make the Z-map model, as shown in Fig. 6-(a), intersecting points are sampled on each lattice of the X-Y plane for every facet. From the coordinates of three vertices, the intersecting points are calculated and stored. The Z-value of a cutter (ball endmill) is also calculated. Then, as shown in Fig. 6-(b), the cutter is located at the lattices and is lowered in the z-direction. When the cutter touches the model's surface, the cutter's contact point (CC point) can be obtained, and it is used to determine the cutter's location point (CL point). The tool-path is generated by arranging the CL points. Because the human head has a convex hemispherical shape, an X-Y directional tool-path is produced for simple and fast machining. For small wigs, 3-Axis machines are used. To machine larger molds, X, Z and rotary axis (C-axis) are used. While rotating the model from 1 to 360° in the C-axis, the cutter's location is calculated in the X-Z plane. Fig. 6-(c),(d) show examples of the generated tool-path. Depending on the materials' size, the tool-path for rough cut can be included. Because sketch lines include important information for making wigs, such as the wig's border, hair direction, material type, customer's name, etc., the sketch lines should be displayed at exact positions on the mold's surface. An additional pen device is attached to the NC-machine to write the sketch lines. After machining the molds, the pen device draws lines on the mold's surface. The tool-path for pen drawing is also generated from the 3D sketch points on the model.



(a) (b) (c) (d) (e)

Fig. 6. (a)&(b) Tool-path generation algorithm using Z-map (c) Generated tool-paths for 3-axis machining (d) Generated tool-paths for 4-axis machining (e) A duplicated polyurethane foam after machining

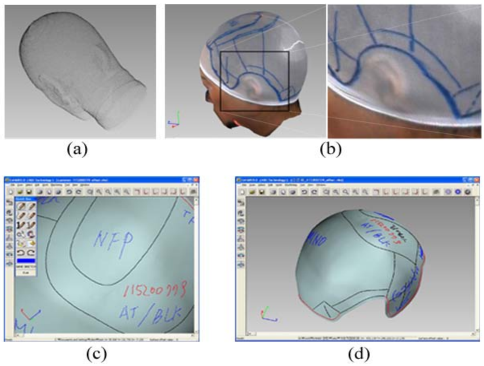


Fig. 7. (a) A screen shot of computed points-cloud, (b) texture-mapping result (left) and its detail (right), (c) sketch toolbox and sketched 3D curves on mesh, (d) designed custom wig-pattern data

4 Conclusion

The whole process of human head modeling and its use in the wig industry are described in this paper. We achieved a fully customized and optimized system for the custom wig industry. Our system perfectly replaces the wig industry's traditional, manual method. The processing time from scanning to final data design is only about 10 minutes. The processed data of each stage are automatically saved in the specified folder to avoid the problems of operator mistakes. The final data are sent through the internet, making it possible to make a custom wig much faster and reduce the transportation cost by not transferring the physical pattern. And because customers' data are stored in a database, any customer who has scanned before can reorder his wig in a different style or hair density without visiting the shop again. Currently our patented system has been applied in over hundred of wig hair shops, and its robustness and practical usefulness have been proven.

Acknowledgements. We would like to thank K&I Technology and Hi-MO for this project. This work is financially supported by the Ministry of Education and Human Resources Development(MOE), the Ministry of Commerce, Industry and Energy(MOCIE) and the Ministry of Labor(MOLAB) through the fostering project of the Lab of Excellency. This research was also supported in part by the Institute of Advanced Machinery and Design at Seoul National University.

References

1. Levoy, M.: The Digital Michelangelo Project: 3D Scanning of Large Statues. In: Proc. Second Int. Conf. on 3-D Digital Imaging and Modeling, 1999, Ottawa Canada, pp. 2–13 (1999)
2. Tsai, R.Y.: A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lens. IEEE Journal of Robotics and Automation RA-3(4), 323–344 (1987)

3. Agrawala, M., Beers, A.C., Levoy, M.: 3D painting on scanned surfaces. In: Proc. of the 1995 symposium on Interactive 3D graphics, ACM Press, New York (1995)
4. Hoppe, H., DeRose, T., Duchamp, T., McDonald, J., Stuetzle, W.: Surface Reconstruction from Unorganized Points, Doctoral Thesis, University of Washington (1995)
5. Curless, B., Levoy, M.: A Volumetric Method for Building Complex Models from Range Images. In: SIGGRAPH 96, Computer Graphics Proceedings (1996)
6. Bajaj, C., Bernardini, F., Xu, G.: Automatic Reconstruction of Surfaces and Scalar Fields from 3D Scans. In: Int. Conf. on Computer Graphics and Interactive Techniques (1995)
7. Edelsbrunner, H., Mucke, E.: Three-Dimensional Alpha Shapes. ACM Transactions on Graphics 13, 43–72 (1994)
8. Amenta, N., Bern, M., Kamvyselis, M.: A new Voronoi-Based Surface Reconstruction Algorithm. In: SIGGRAPH 1998, pp. 415–421 (1998)
9. Amenta, N., Choi, S., Kolluri, R.: The power crust, Symposium on Solid Modeling and Applications 2001, pp. 249–266 (2001)
10. Linsen, L., Prautzsch, H.: Local Versus Global Triangulations. In: Eurographics 2001, vol. 20(3) (2001)
11. Bernardini, F., Mittleman, J., Rushmeier, H., Silva, C., Taubin, G.: The Ball-Pivoting Algorithm for Surface Reconstruction. Transactions on Visualization and Computer Graphics 5(4) (1999)
12. Ohtake, Y., Belyaev, A.G., Bogaevski, I.: Mesh regularization and adaptive smoothing. Computer-Aided Design 33(11), 789–800 (2001)
13. Matsushita, K., Kaneko, T.: An Efficient Image Mapping Method for 3D Surface Texture. Systems and Computers in Japan 32(4) (2001)
14. Lee, K.W.: CAD System for Human-Centered Design, Computer-Aided Design and Applications 3(5) (2006)
15. Marco, T., Hitoshi, Y., Jörg, H., Hans-Peter, S.: Texturing Faces. In: Proc. Graphics Interface 2002, pp. 89–98 (2002)
16. http://www.knitech.co.kr/
17. http://www.himo.co.kr/

Evaluation of Navy Shipboard Habitability for a Warship Design Using Human Model

Hongtae Kim, Jin H. Park, Hojin Hwang, and Chang-Min Lee

Maritime and Ocean Engineering Research Instititue/KORDI, Daejeon, Korea
P.O. BOX 23, Yusong, Daejeon, 305-600, Republic of Korea
{kht,jhpark,hjhwang,cmlee}@moeri.re.kr

Abstract. Recently, ergonomic design is a key issue in the warship design. Ergonomic ship design aims to improve the safety and convenience of crew and passengers. In this paper, as a basis of analyzing and evaluating ergonomic design criteria and evaluation methods for a warship, existing military design guidelines and navy criteria in the fields of ergonomics were surveyed. Also, the methodology including an ergonomic human model is introduced. Finally, it is shown that the ergonomics evaluation of a design can be performed at the early stage of warship design using Navy Ergonomic Human Model based on Delmia's ERGO.

Keywords: Shipboard habitability, Human model, Warship design.

1 Introduction

Recently, ergonomic design criteria has become an essential issue in a warship since it can improve safety and reliability as well as simplify maintenance, operation and education references [7], [8]. Ergonomic warship design criteria can be defined as a standard for design, selection, arrangement and operation of equipments, facilities and systems considering Korean body type so as to increase comfort, efficiency and safety of navy crew while performing their mission.

Ergonomic warship design is an important part in the design, manufacturing, safe navigation, accident prevention, and operation efficiency. There are many cases from other countries that are not only about design process - prevention of design error, reduction of design period - but also about introducing ergonomic idea - efficient arrangement of equipment, most suitable working environment for navy crew - to decrease uncertainties and risks of warship.

In this paper, technical trends about standard and criteria in relation to ergonomic human model for warship design are surveyed, and an evaluation method of ergonomic considerations in weapon systems including warship itself is introduced with examples using 3D simulation.

2 Naval Weapons System and Ergonomics

It is very important to consider ergonomic factors in weapon systems such as warship, fighter and tank. The life cycle of a weapon system is reducing while importance of

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 884–893, 2007.

system flexibility is increasing. Therefore, optimization of a system in the design stage is becoming more important than gradual improvement of the system after development.

In general, the development process of a weapon system, similar to conventional product development process, can be divided to general idea formation according to requirement, experimental prototype/design, system development, manufacturing, operation and maintenance. Ergonomic considerations are especially important in general idea forming and design stage, and should also be considered in other stages.

2.1 Ergonomic Consideration for Warship Design

Ergonomic considerations are essential in a warship since, compared to a commercial ship, the number crew is larger and the length of operation at sea is longer. The purpose of introducing ergonomic notion in warship design is to increase operation efficiency and fighting power through crew centered design, reduces unnecessary cost caused by design error, reduce design period by using standard, and reflect the necessity of reestablishing design guidelines considering crew body type.

The U.S. military recognized the necessity of ergonomic design early and tried to apply an ergonomic approach and reformed organizations for this purpose in 1983 [8]. The items to apply ergonomic technology in warship design and operation part are information input/output, control device, car control device, working area, space arrangement, environment, equipment, and software interface, etc.

Within application areas of ergonomic technology for operating and design of warship, the most important part is shipboard habitability. Shipboard habitability is influence on crew's health and spirit since they spend a lot of time in the warship. Eventually, it will be directly connected with fighting power of the warship. Most of the lists considering shipboard habitability in standard are connected with combinations between man and machine or man and environment. Meeting space, sanitary space, sleeping space, eating space, bed, exercise space, leisure space, reading space, lounge, religion activity space come under this. In addition, environment factors such as heating, ventilation, cooling, drainage, lighting, and coloring should also be considered.

Diverse shipboard habitability studies are under progress in the navy of U.S. and United Kingdom [1], [7], [9], [10], and most of them are to improve shipboard habitability through analysis of habitability equipment and space.

The following things should be done to develop an ergonomic warship design standard with warship original mission and crew relaxation. In addition, suitable design items for the Korea Navy should be identified, and detailed lists in each item should be analyzed and stored in database. Some items require Korean anthropometrical data, while other items can be constructed by modifying previous materials. In order to identify item type, information gathering through surveys and interviews with navy personnel and shipyard designers should be done first.

Lastly, anthropometrical database of the ROK (Republic of Korea) Navy should be constructed from existing materials and actual measurement, which can be the basis for building a Navy Ergonomic Human Model. Using this ergonomic analysis human model, ergonomic evaluation is possible by simulation in virtual space such as a living area, transferring area, and control area.

2.2 Related Regulations and Standards

A few of representative materials in regulations and standards related to ergonomics are ISO/TC 159 of ISO (International Organization for Standard), MIL-STD-1472 and MIL-HDBK-759B(Military Handbook of Human Engineering Deign for Army Materiel) of America national defense, ASTM(American Society For Testing and Materials), and U.S Navy Shipboard Habitability Manual. Among these, special features of MIL-STD-1472, ASTM guidelines and U.S. Navy Shipboard Habitability Manual, which is directly connected with warship design, are as follows.

MIL-STD-1472 (Design Criteria Standard Human Engineering) of USA is the most significant material in the field of national defense. It has under gone many revisions and the current version is MIL-STD-1472F. It includes general ergonomic design standard for military systems, equipments and facilities. The purpose of MIL-STD-1472 is to achieve various missions successfully through effective integration of systems, equipments and facilities with human.

ASTM F1166-95A establishes general human engineering design criteria for marine vessels, and systems, subsystems, and equipment contained therein. The purpose of this practice is to present ergonomic design criteria, principles, and practices to achieve mission success through integration of the human into the vessel system, subsystem, and equipment with the goals of effectiveness, simplicity, efficiency, reliability, and safety for operation, training, and maintenance. Also, ASTM F1337-91 establishes and defines the requirements for applying ergonomics to the development and acquisition of ships and marine systems, equipment, and facilities based on the ASTM F1166-95A. This practice includes about 37 sub lists regarding accessibility, anthropometrics, and auditory indicator.

Naval sea systems command of US Navy utilize shipboard habitability manual to improve living or working condition to guarantee navy crew's safety and health. Shipboard habitability manual states habitability as one of the most important factors in mission preparation considerations, and stresses about the importance of shipboard quality of work life. Shipboard habitability manual consists of shipboard habitability design criteria manual that describes general specification, and shipboard habitability design practices manual describing actual practice.

3 Survey of Shipboard Habitability

The basic survey for criteria of ergonomic ship design suitable for the ROK Navy was carried out. It questioned a shipboard habitability with a target of officer, senior rate, and junior rate. The questionnaire was filled out by 208 person in the ROK Navy, which were organized by crews in MHC(Mine Hunter Coastal), PCC(Patrol Combat Corvette), and MLS(Mine Laying Ship).

The specificity of ships surveyed in this research would be summarized as follows:

- MHC (Mine Hunter Coastal) : a warship(small size) for hunting mines. It speeds 30 knots and displaces 500 tons.
- PCC (Patrol Combat Corvette) : a warship(small size) for local convoy and antisubmarine watch and attack. It speeds 30 knots and displaces 500 tons.

- MLS (Mine Laying Ship) : a warship(big size) for laying mines. It speeds 15 knots and displaces 3,000 tons.

We visit the ROK Navy and explain the purpose and contents of this questionnaire. Respondents filled out it immediately. The questionnaire consists of 12 questions and following contents was included:

- General information about respondents (Unit, Position, Age and Career)
- Evaluation of current warship
- Personal preference of shipboard facilities & improvement needs for new ships

3.1 General Information About Respondents

The ranks occupy over 90 percent in the survey of 208 navy crews. Fig. 1 shows a distribution of rank, age, and sea time for participations.

3.2 Satisfaction of Shipboard Habitability

Fig. 2 shows analysis results of total satisfaction of shipboard habitability classified by rank, ship type, sea time, and age. The total satisfaction was reported under 3 points (Normal) as shown in Fig. 2. It indicates most of navy crews have a dissatisfaction of shipboard habitability. We classified habitability with following 6 items to analyze detail satisfaction of shipboard habitability and questioned satisfactions of these items.

- Sleeping Accommodation - Bathroom - Messroom
- Cooking - Laundry & Service - Recreation & Study

According to the survey, navy crews are more satisfied with Sleeping Accommodation, and they satisfied with in order of Bathroom, Messroom, Cooking, Laundry & Service, and Recreation & Study.

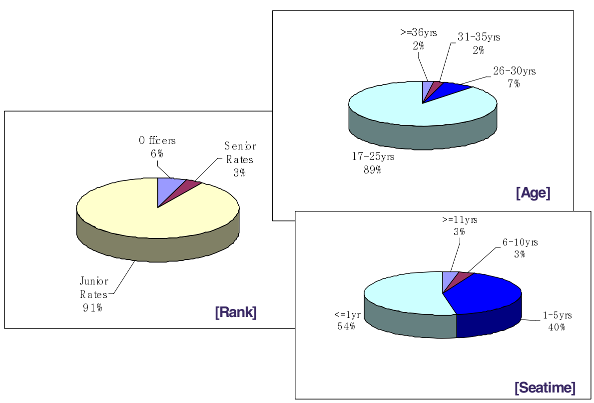


Fig. 1. General information about respondents

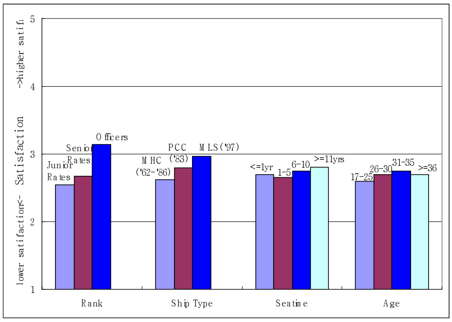


Fig. 2. Analysis results of total satisfaction of shipboard habitability

Table 1. Satisfaction and preference of habitability

Items	Satisfaction (1=lowest, 5=highest)	Preference (Rank)
Sleeping Accommodation	2.91	1
Bathroom	2.78	2
Messroom	2.71	3
Cooking	2.56	4
Laundry & Service	2.42	5
Recreation & Study	2.15	6

Fig. 3 shows difference of satisfaction by ship types for habitability items. As shown in this figure, the highest satisfaction was founded at MLS items except Recreation Area. Sleeping Accommodation and Bathroom & Laundry have relatively high satisfaction in MHC and Service Area and Recreation Area have relatively high satisfaction in PCC.

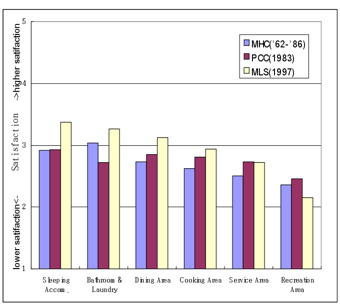


Fig. 3. Satisfaction of shipboard habitability by ship types

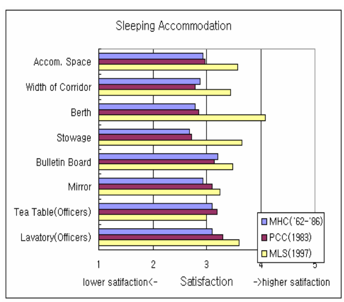


Fig. 4. Satisfaction of Sleeping accommodation by ship types

Additional questions for detailed shipboard habitability were conducted. MLS got similarly the highest satisfaction at all items for analysis results of detailed habitability. Fig. 4 represents analysis results of 3 type ships for Sleeping Accommodation.

3.3 Facility Preferences and Improvement Needs for New Ships

The preference of personally important facilities in ship life was reported in Fig. 5. As shown in the figure, navy crews preferred a sleeping facility. They also preferred privacy/personal space, eating facilities, and hygiene/washing facilities/laundry.

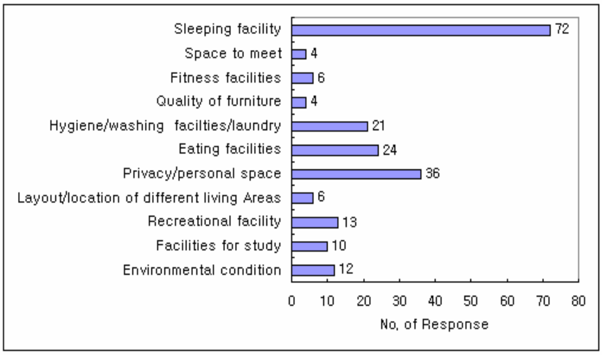


Fig. 5. Facility preferences of navy crews

Fig. 6 shows analysis results of improvement needs of new ships. Navy crews had more needs for facilities to customize living/personal space and personnel accommodated in cabins. They also had needs for recreation areas separate from cabins, separate spaces/facilities for study, and dedicated areas for fitness activities.

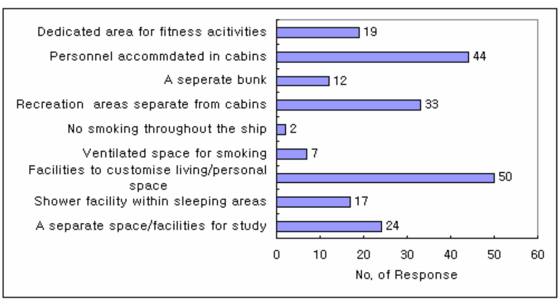


Fig. 6. Improvement needs for new ships

3.4 Comparison with the Royal Navy

Comparison for shipboard habitability of the ROK Navy and the Royal Navy [9] was carried out. As shown in Fig. 7, the overall satisfaction o the Royal Navy was higher than the ROK Navy. Sleeping Accommodation has relatively high satisfaction in the ROK Navy, while Dining Area, Service Area and Recreation Area have relatively high satisfaction in the Royal Navy.

Result of two surveys it will not be able to compare evenly because has a difference from investigation item, warships type and the no. of respondents. Both navy crews preferred a sleeping facility in the preference of personally important facilities in ship life. They also had more improvement needs of new ship for personnel accommodated in cabins.

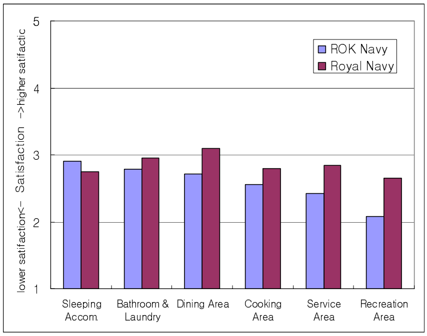


Fig. 7. Comparison for shipboard habitability of the ROK Navy and the Royal Navy

4 Ergonomic Evaluation Using Human Model

4.1 Ergonomic Human Model

For the ergonomic evaluations of warship design, diverse basic data is required, but to get the data from statistical research or real human body experiment is very costly and sometimes hazardous. Therefore, computer simulation is required. An ergonomic human model is a computerized visualizing human to be used in CAD systems for product design [2].

Using an ergonomic human model enables evaluation of clearance, reach and visibility without the use of physical mock-ups in conventional design process. Also, it does not require several mock-ups because it can easily be modified to expresses various types of users for different ergonomic evaluations. A few of general purpose systems using an ergonomic human model are ERGO from Delmia, SafeWork and Virtual Man from Softwork Inc., and ERGOPlan from EAI-Delta Inc. [4]. These general- purpose systems offer an advanced human body model with ergonomic evaluation capabilities.

In this study, Delmia's ERGO is used as ergonomic human model. ERGO can evaluate human resources using human model integrated with design environment for ergonomic analysis, and can apply ergonomic considerations such as Design for Assembly (DFA) at the early stage of design. Namely, rapid prototyping and analysis of human activity such as reach ability, field of view, posture analysis, lifting guidance and energy expenditure is. The ROK Navy crew/officer model is constructed by modifying human data in ERGO. This model can be applied for building guidelines of warship design, and additionally, can be used for analysis such as human walkthrough and interaction/collision detection with virtual objects.

4.2 Development of the ROK Navy's Standard Crew Model and Its Applications

To develop of a ROK Navy Ergonomic Human Model, crew and officer model database is constructed based on the "National Anthropometric Survey" [3]. Among 120 measurements in National Anthropometric Survey, total of 25 measurements that have direct relation with warship design are used. The measurements are composed of 7 height, 1 length, 1 width, 7 thickness and 1 circumference measurements in standing position, and 10 measurements in sitting position. To make detailed anthropometrical database approximately 100 measurements are required, but only above 25 measurements are used construct the database. Among gender and age groups of "National Anthropometric Survey", only two groups (18-24 male and 25-39 male) are used. Based on this database, commercial human modeling tool, ERGO is supplemented to build a ROK Navy Ergonomic Human Model is completed.

To revise ERGO model, each physical parameter is modified both as a ratio between original model's height and a ROK Navy model's height, and by direct modification of each parameter. Among the parameters in ERGO model, only those that affect 25 selected parameters from "National Anthropometric Survey" are modified. Also a ROK Navy model is divided to crew model and officer model, and each model is further divided 5, 50, and 95-percentile model.

A ROK Navy Ergonomic Human Model can be modified later when actual measurement of the ROK Navy personnel is performed, and female crew and officer model will be developed as well.

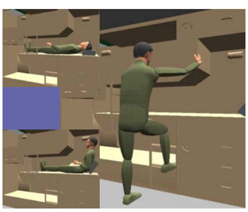
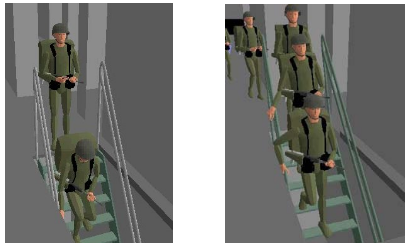


Fig. 8. Usage of the sit-up berth

Fig. 9. Analysis of the width of the corridor in LPX ship

We conducted ergonomics analyses for an amphibious assault/landing platform dock ship (LPX) using a ROK Navy Ergonomic Human Model. Fig. 8, Fig. 9 and Fig 10 show the results of analyses. Figure 8 shows usage of the sit-up berth – a three-tiered bunk bed system. It provides provide 40 percent more storage space and still provide enough headroom to sit in an upright position. And Fig. 9 shows the example of analyses for the width of the corridor. We investigated the width of the corridor (1.2 m, 0.95 m, 0.7 m), in case 2 persons cross the corridor with full military equipment. In this case, 2 persons could cross the corridor in 1.2 m wide.



a) before (55 degrees) b) after (45 degrees)

Fig. 10. Analyses of gradient of inclined ladder

As shown in Fig. 10, we improved gradient of inclined ladder from 55 degrees to 45 degrees in the accommodation space considering a comfortable position. The case of 45 degrees is more comfortable posture than the case of 55 degrees.

5 Conclusions

This paper proposes ergonomic evaluation criteria required for evaluation of a warship design and operation. Basic research about the usage of human engineering as a basis of developing ergonomic design criteria is done to this purpose. And the basic survey for criteria of ergonomic ship design was carried out.

Also, a Korean Navy anthropometric database is constructed to use for making ergonomic guidelines for warship design by simulation in virtual environment, and a ROK Navy Ergonomic Human Model is developed using commercial human modeling system. Developed model is used in analysis of transferring area and accommodation area. Ergonomic considerations in warship design can increase mission performance by increasing comfort of crew in the sense of design for quality of work life. In addition, in the viewpoint of safety, it can reduce work fatigue and human error. In the development of next generation warship and submarines, researches about ergonomic considerations in warship design should be continued.

Acknowledgements. The contents of this paper are the results of the Inherent Research Project of MOERI/KORDI, “Development of smart operation technologies for exploration fleet based on ubiquitous concept”. And this research was partially supported by the ROK (Republic of Korea) Navy.

References

1. Castle, J.E.: US Navy Shipboard Habitability Design Standards and Trends, Naval Engineers Journal, pp. 66–75 (1984)
2. Feyen, R., Liu, Y., Chaffin, D., Jimmerson, G., Joseph, B.: Computer-aided Ergonomics: a Case Study of Incorporating Ergonomics Analyses into Workspace Design. Applied Ergonomics 31, 291–300 (2000)
3. Kim, D.J., et al.: National Anthropometric Survey of Korea, Project Report, Korea Research Institute of Standards and Science(in Korean) (1997)

4. Kim, H., Lee, J.K., Lee, D., Park, J.H.: State of the Art of Human Factors Technologies for Ships and Ocean Engineering. Journal of the Ergonomics Society of Korea (in Korean) 20(2), 99–111 (2001)
5. Kuo, C.: The Role of Human Factors in the Safety of Marine System. In: Proceedings of the 3rd International Offshore & Polar Eng. Conference, Singapore (1993)
6. Rothblum, A.M., Carvalhais, A.B.: Marine Applications of Human Factors of Human Factors Test & Evaluation. In: O'Brien, T. G., et. al. (ed.): Handbook of Human Factors Testing & Evaluation (1996)
7. Saklem, A.A.: US Navy Habitability Control in Relation to Human Factors. In: Proceedings of the SNAME Spring Meeting (1985)
8. Stein, N.I.: An Overview of the Role of the NAVSEA Human Factors Engineering Program in Ship Design, Naval Engineers Journal, 139–152 (1983)
9. Strong, R.: RN Habitability Survey -Ship Design Implications. In: Proceedings of Human Factors in Ship Design & Operation, London (2000)
10. Ware, H.D.: Habitability in Surface Warships, The Royal Institution of Naval Architects, 41–58 (1986)

Two Vibration Modes of a Human Body Sitting on a Car Seat-
The Relationship Between Riding Discomfort Affected by the Material Properties of the Seat Cushion and the Two Vibration Modes

Mitsunori Kubo, Fumio Terauchi, and Hiroyuki Aoki

Department of Design and Architecture, Faculty of Engineering,
Chiba University, 1-33, Yayoi-cho, Inage-ku, Chiba-shi,
Chiba 263-8522, Japan
mkubo@faculty.chiba-u.jp

Abstract. The relationship between the properties of flexible polyurethane foam composing seat cushions for standard cars and the riding discomfort evaluated by subjects who sat on the seat were investigated. The properties of five kinds of polyurethane foam were measured using the DMA to define dynamical properties. Riding discomfort was subjectively evaluated by 17 subjects at four frequencies, and the correlation coefficients between the material properties and riding discomfort were analyzed. The results suggested that there were two relationships between material properties and riding discomfort, and that these relationships strongly relied on frequency. Furthermore, a digital human model was created to confirm the influence of frequency on these relationships, which suggested that the relationships affected sensitivity by the change in the vibration mode of the human body-seat vibration system.

Keywords: Riding Discomfort, Vibration Mode, Human Model.

1 Introduction

There are only a few reports relating to the riding discomfort of car seats. In these reports, subjective responses are used as a reliable measure of discomfort. Such experiments, which try to directly treat riding discomfort, generally tend to include too much uncertainty. Furthermore, the physical responses of subjects who subjectively evaluate riding discomfort tend to be difficult to reproduce, as a precise measure of dynamic human behaviour has not yet been developed. Recent research relating to riding discomfort seem to assume that riding discomfort would be improved if vibration is reduced, with few proper treatments for measuring subjective response or riding discomfort. In recent studies concerning riding discomfort and human models, the aims of the studies have been varied, in one case to improve the

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 894–903, 2007.

impact absorptive properties of military vehicles [1], another proposed active damping using an air cushion system for improving the ship siding discomfort using a mathematical model [2], and a further study aimed to identify suspension parameters in order to prevent spinal injury [3]. These studies were performed from an engineering perspective, and seem to be indirectly related to riding discomfort, even though the research objectives purportedly aimed to improve riding discomfort. In reality, it is difficult to directly and realistically evaluate riding discomfort by subjective methods. It can be said, then, that the objectively measurable and subjectively recognizable definition of this discomfort is yet to be identified. Although we have performed some experiments in order to clarify the discomfort [4-7], we have not been able to discover how to reduce the discomfort, nor why the subjects sense the discomfort in the first place. However, we have observed a trend in the relationship between riding discomfort and the physical responses of the subjects sitting on experimental car seats, which has been observed in all experiments we have implemented to date. This trend reflects the existence of a transition point, at around 5 Hz or 6 Hz of vertical sinusoidal vibration, in the subjective descriptions of subjects exposed to the vibration whilst sitting. This transition point represents a change in the relationship between riding discomfort and physical responses. There are two distinctive relationships, one before and after the frequency transition point.

This report was written based on results obtained from our previous experiments, in order to confirm the existence of the transition point, which has presented repeatedly under various experimental conditions. In this report, we would like to consider the reason for the transition point, how the existence of the transition point affects riding discomfort.

2 Experiments

The experiment was divided into two parts. The first part was a laboratory experiment to measure the subjective response regarding riding discomfort and the physical responses of the subjects exposed to vibration whilst seated. The second experiment was a numerical simulation using an FEM human model in order to elucidate the reason for any relationships. In this study, relationships were investigated with reference to riding comfort and the material properties of polyurethane foam. The following descriptions of the experiments have been extracted from a previous report [4].

2.1 Evaluation of Riding Discomfort Affected by the Properties of Flexible Polyurethane Foam

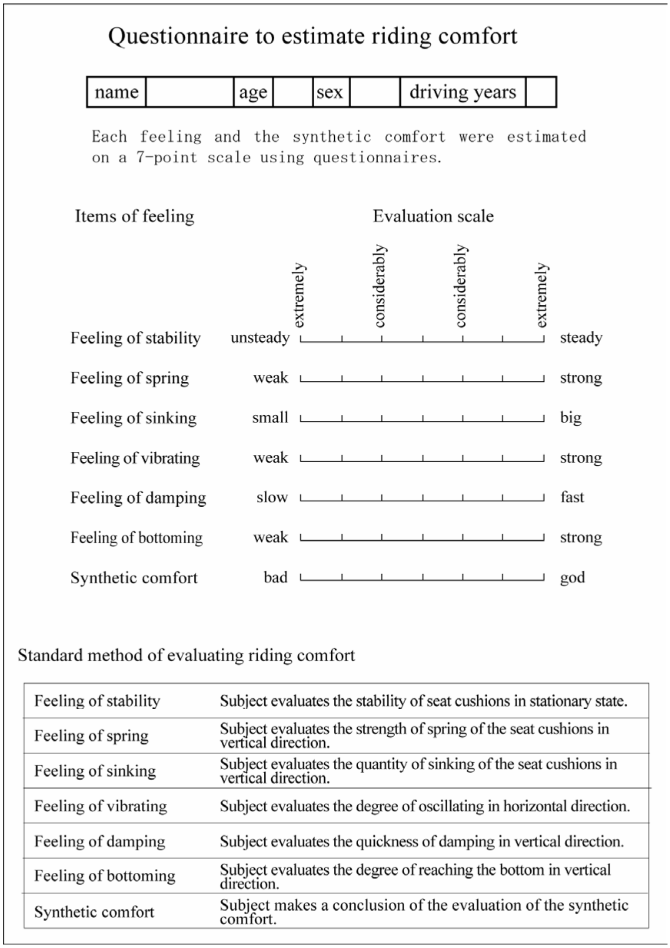
The material properties of the five experimental seat cushions made of five different flexible polyurethane foams are shown in Tables 1. These five samples were designated A through E. Subjects were 17 regular car drivers (15 male; 2 female). The subjects evaluated riding discomfort during exposure to a vibration according to SD-method. The evaluation comprised the following factors: feeling of stability, feeling of spring, feeling of sinking, feeling of vibration, feeling of damping, feeling of bottoming, and synthetic discomfort. Each factor was evaluated using a 7-point scale using questionnaires shown in Figure 1. Synthetic discomfort was the name applied to riding discomfort. Figure 2 is the experimental devices and the subject who is evaluating the feelings.

2.2 A Human Model Reproducing the Behaviour of Human Body Sitting on the Seat by Using FEM

All finite elements composing the human model were defined within the same solid cubic elements, where each element had differential properties and sizes. The outline of the human body was measured with a 3-D measurement device (FAROARMS, FARO Technology). This measurement data was translated to a DXF-file, and then the DXF-file data was imposed into a CAD-application (CADKEY2000, CADKEY SYSTEMS). A solid human model, which was to be applied to a FEM-application (ANSYS Ver.5.2, ANSYS), was created in the CAD-application. The seat model was constructed as a simple shaped seat with the properties of polyurethane foams. The seat model was formed according to the representative characteristics of an actual seat: the position of the seat cushion and the seat back. Finally, the human model was applied to the seat model in the same way as a real human would sit in such a seat. Then the human-seat model data was translated into an SAT-file to be imposed onto ANSYS. All parts were meshed with a solid cubic element: SOLID45, which had 8

Table 1. Material properties of the polyurethane form

Properties ＼ Sample		A	B	C	D	E
E' at 2Hz (real part)	Pa	7.96E+04	8.47E+04	9.30E+04	1.21E+05	1.28E+05
E' at 4Hz	Pa	8.06E+04	8.58E+04	9.42E+04	1.25E+05	1.31E+05
E' at 6Hz	Pa	8.12E+04	8.62E+04	9.48E+04	1.27E+05	1.33E+05
E' at 8Hz	Pa	8.13E+04	8.64E+04	9.51E+04	1.28E+05	1.34E+05
E" at 2Hz (imarginary part)	Pa	3.58E+03	3.81E+03	4.26E+03	1.19E+04	1.13E+04
E" at 4Hz	Pa	3.41E+03	3.66E+03	4.08E+03	1.11E+04	1.10E+04
E" at 6Hz	Pa	3.27E+03	3.44E+03	3.98E+03	1.09E+04	1.05E+04
E" at 8Hz	Pa	3.08E+03	3.24E+03	3.65E+03	1.02E+04	9.82E+03
Compression-tan δ at 2Hz		3.87E-02	3.83E-02	3.92E-02	8.42E-02	7.69E-02
Compression-tan δ at 4Hz		4.06E-02	4.01E-02	4.23E-02	8.74E-02	8.01E-02
Compression-tan δ at 6Hz		4.21E-02	4.24E-02	4.30E-02	8.74E-02	8.22E-02
Compression-tan δ at 8Hz		4.40E-02	4.41E-02	4.49E-02	9.26E-02	8.42E-02
G' at 2Hz (real part)	Pa	9.73E+03	9.53E+03	1.04E+04	2.60E+04	1.95E+04
G' at 4Hz	Pa	1.02E+04	9.88E+03	1.08E+04	2.75E+04	2.05E+04
G' at 6Hz	Pa	1.03E+04	9.91E+03	1.09E+04	2.83E+04	2.11E+04
G' at 8Hz	Pa	1.02E+04	9.78E+03	1.08E+04	2.87E+04	2.13E+04
G" at 2Hz (imarginary part)	Pa	1.45E+03	1.29E+03	1.36E+03	3.91E+03	2.96E+03
G" at 4Hz	Pa	1.58E+03	1.34E+03	1.43E+03	4.17E+03	3.18E+03
G" at 6Hz	Pa	1.63E+03	1.39E+03	1.46E+03	4.37E+03	3.32E+03
G" at 8Hz	Pa	1.63E+03	1.43E+03	1.65E+03	4.53E+03	3.42E+03
Shearing-tan δ at 2Hz		1.49E-01	1.35E-01	1.31E-01	1.50E-01	1.52E-01
Shearing-tan δ at 4Hz		1.55E-01	1.35E-01	1.32E-01	1.52E-01	1.55E-01
Shearing-tan δ at 6Hz		1.58E-01	1.40E-01	1.34E-01	1.54E-01	1.58E-01
Shearing-tan δ at 8Hz		1.59E-01	1.46E-01	1.53E-01	1.58E-01	1.60E-01


Questionnaire to estimate riding comfort

name		age		sex		driving years	

Each feeling and the synthetic comfort were estimated on a 7-point scale using questionnaires.

Items of feeling — Evaluation scale (extremely, considerably, considerably, extremely)

Items of feeling			
Feeling of stability	unsteady		steady
Feeling of spring	weak		strong
Feeling of sinking	small		big
Feeling of vibrating	weak		strong
Feeling of damping	slow		fast
Feeling of bottoming	weak		strong
Synthetic comfort	bad		god

Standard method of evaluating riding comfort

Feeling of stability	Subject evaluates the stability of seat cushions in stationary state.
Feeling of spring	Subject evaluates the strength of spring of the seat cushions in vertical direction.
Feeling of sinking	Subject evaluates the quantity of sinking of the seat cushions in vertical direction.
Feeling of vibrating	Subject evaluates the degree of oscillating in horizontal direction.
Feeling of damping	Subject evaluates the quickness of damping in vertical direction.
Feeling of bottoming	Subject evaluates the degree of reaching the bottom in vertical direction.
Synthetic comfort	Subject makes a conclusion of the evaluation of the synthetic comfort.

Fig. 1. Questionnaire to estimate the riding discomfort

nodes and was a normal 3-D element in ANSYS. The total element number was 5567. The primary properties for each element were defined with the pre-processor of the ANSYS. In essence, all properties of the elements had been modified to be almost synchronized with experimental results through several simulations on harmonic response analysis in ANSYS.

3 Results

Correlation coefficients between the subjective evaluations of riding discomfort and material properties reveal the existence of two distinct relationships, which were also confirmed through numerical simulation using the human model.

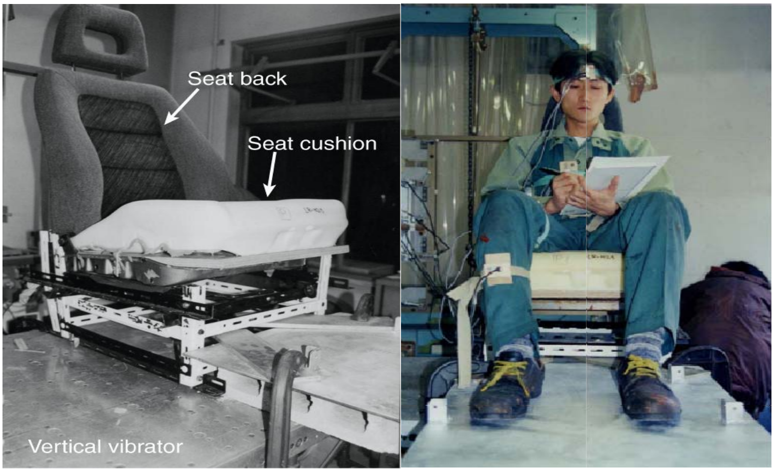


Fig. 2. The real car seat, with only the seat cushion of the seat pan removed was fixed on the vibrator. Each experimental seat cushion was installed in the car seat, to comparatively evaluate the sensory difference between the experimental seat cushions. And, the seat cushions were installed without seat fabric to avoid influences of tension induced by the fabric.

3.1 Correlation Between the Riding Discomfort and the Properties of Flexible Polyurethane Form

Table 2 shows correlation coefficients between riding discomfort and the properties of flexible polyurethane foam used in the seat cushions. Boldfaced letters in Table 2 are correlation coefficients with significance levels larger than 5%. It was thus demonstrated that, at 2 and 4 Hz, riding discomfort was affected by impact resilience, stress relaxation, compressibility and compression-tanδ. The compression-tanδ is an index representing the vertical vibration damping. On the other hand, it was demonstrated that riding discomfort at 6 and 8Hz was affected only by shearing-tanδ. The shearing-tanδ is also an index representing the horizontal vibration damping. It was thus possible to summarize that riding discomfort at 2 and 4 Hz relates to the properties of the cushions affecting the behaviour of the vertical vibration, while riding discomfort at 6 and 8 Hz relates to the properties affecting the horizontal vibration behaviour.

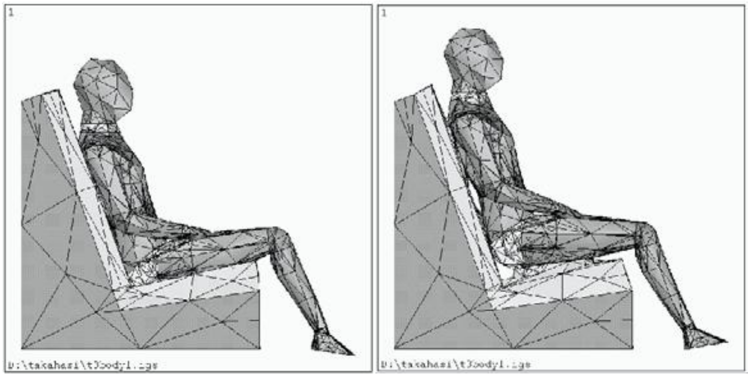
3.2 The Eigen Vibration Modes Simulated by the Human Model

From the modal analysis of the human-seat vibration system, it was demonstrated that the lowest characteristic frequency was 4.78 Hz, and the eigen vibration mode, or eigenvector, relating to this frequency showed strong vertical vibration (Fig.3). This suggested that the real human-seat vibration system would show clear vertical vibration at around 4 to 5 Hz. This result almost agrees with the experimental one, in which the resonance frequency induced by the external vertical sinusoidal vibration

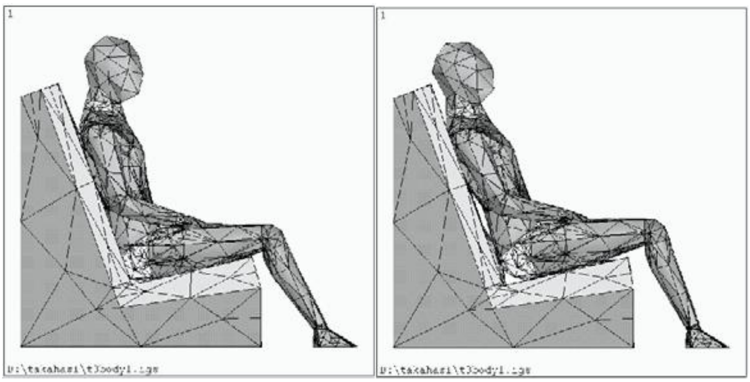
Table 2. Coefficients of correlation between riding comfort and the properties

	Seating	2Hz	4Hz	6Hz	8Hz
Over-all Density	**-0.928**	-0.841	**-0.896**	-0.268	-0.390
Impact resilience	**-0.912**	**-0.886**	**-0.918**	-0.280	-0.456
Air flow ability	-0.250	-0.457	-0.460	0.167	-0.194
Hardness:25%ILD	0.486	0.530	0.378	0.739	0.674
Stress relaxation	**0.923**	**0.892**	**0.889**	0.253	0.417
Ratio of hysteresis loss	0.820	0.747	0.800	0.059	0.217
Quantity of elastic hysteresis loss	**0.933**	0.857	0.873	0.251	0.373
Deflection - 500N	**-0.926**	-0.874	-0.843	-0.743	-0.746
Deflection - 40N	-0.802	-0.859	-0.700	-0.610	-0.682
Spring constant - 500N	-0.815	-0.823	-0.670	-0.535	-0.573
Spring constant - 200N	**0.895**	0.759	0.732	0.599	0.532
Spring constant - 40N	0.773	0.817	0.655	0.617	0.662
Ratio of spring constant	**0.886**	0.876	0.765	0.591	0.635
Compressibility	**-0.922**	**-0.916**	-0.834	-0.645	-0.703
Resonance frequency	0.533	0.553	0.612	-0.201	0.054
Transmissibility at peak	-0.047	-0.220	-0.307	0.179	-0.124
Transmissibility at 6Hz	0.338	0.491	0.528	-0.153	0.181
E' - 2Hz	0.826	0.776	0.851	0.139	0.316
E' - 4Hz	0.833	0.781	0.855	0.146	0.321
E' - 6Hz	0.838	0.786	0.858	0.151	0.326
E' - 8Hz	0.839	0.788	0.859	0.153	0.327
E" - 2Hz	**0.915**	0.857	**0.890**	0.269	0.413
E" - 4Hz	**0.911**	0.844	**0.889**	0.266	0.417
E" - 6Hz	**0.913**	0.855	**0.892**	0.260	0.399
E" - 8Hz	**0.915**	0.855	**0.892**	0.262	0.408
Compression-tan δ - 2Hz	**0.934**	**0.881**	**0.894**	0.238	0.387
Compression-tan δ - 4Hz	**0.930**	**0.883**	**0.896**	0.235	0.380
Compression-tan δ - 6Hz	**0.930**	0.867	**0.892**	0.239	0.390
Compression-tan δ - 8Hz	**0.931**	**0.881**	**0.889**	0.242	0.389
G' - 2Hz	**0.911**	**0.922**	0.859	0.255	0.437
G' - 4Hz	**0.912**	**0.923**	0.860	0.257	0.439
G' - 6Hz	**0.915**	**0.924**	0.862	0.264	0.443
G' - 8Hz	**0.916**	**0.924**	0.864	0.265	0.444
G" - 2Hz	**0.936**	**0.927**	**0.879**	0.302	0.463
G" - 4Hz	**0.943**	**0.931**	**0.887**	0.322	0.478
G" - 6Hz	**0.943**	**0.929**	**0.885**	0.321	0.474
G" - 8Hz	**0.929**	**0.933**	**0.883**	0.298	0.472
Shearing-tan δ - 2Hz	**0.906**	0.707	0.826	0.759	0.646
Shearing-tan δ - 4Hz	0.826	0.637	0.770	0.859	0.715
Shearing-tan δ - 6Hz	0.815	0.595	0.730	0.814	0.642
Shearing-tan δ - 8Hz	0.781	0.744	0.877	**0.905**	**0.917**

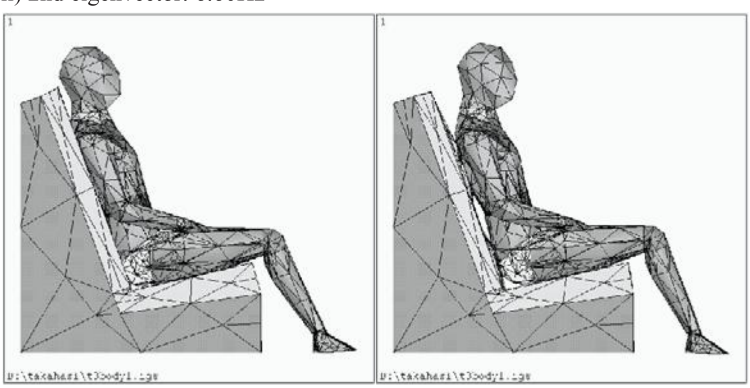
was at around 4 Hz. Also, the second and third characteristic frequencies were 6.60 Hz and 7.76 Hz, respectively. The natural vibration modes relating to both frequencies show strong horizontal vibration as compared with the eigenvector of the lowest characteristic frequency: 4.78 Hz (Fig.3). This suggests that the real human body vibrates easily in horizontal direction at around 6 to 8 Hz. This result is consistent with the experimental results at 6 and 8 Hz.



i) 1st eigenvector: 4.78Hz



ii) 2nd eigenvector: 6.60Hz



iii) 3rd eigenvector: 7.76Hz

Fig. 3. The eigen vibration modes or eigenvectors at around 5Hz were obtained by using the FEM human model. The above figures describe the minimum displacement (left) and the maximum one (right) at each eigen frequency: 4.78Hz, 6.60Hz, and 7.76Hz. The minimum displacement means that the human body sinks down most deeply in the seat.

4 Discussion

Two distinctive tendencies were observed in the relationship between the material properties of a seat cushion made of flexible polyurethane foam and riding discomfort subjectively evaluated by 17 subjects who sat on the car seat and were exposed to vertical sinusoidal vibration. The results confirmed that there was a transition point at around 5 Hz. The tendency of the relationship at below 5 Hz was that riding discomfort was affected by the dynamic property in the vertical vibration because the compression-tanδ, which governs the vibration dumping factor in the vertical direction, correlated clearly with riding discomfort. On the other hand, the tendency at above 5 Hz was that shearing-tanδ, which governs the horizontal vibration dumping factor of the seat cushion, correlated with discomfort, and the correlation between compression-tanδ and riding discomfort was completely absent. In other words, the relationship between riding discomfort and frequency may be described using the dynamic material properties compression-tanδ and shearing-tanδ, as defined by the DMA.

Although we could verify the transition point and the change in the relationship between riding discomfort and the material properties in relation to frequency, the whole vibration mode of the human body on the seat, and the reason why the subjects unconsciously felt the change were not investigated at this time. As such, we tried to implement a numerical simulation with the FEM human model, in order to grasp the eigen vibration mode of the whole human body, and to elucidate the reason why the subjects evaluated their riding discomfort that way.

The results of the FEM human model also reproduced a distinctive change in vibration modes at around the 5 Hz. The changes observed in the simulation seem to sufficiently explain the subjectively described transition from the relationship between the compression-tanδ and riding discomfort to the relationship between the shearing-tanδ and the discomfort. The human model suggested that the human body tended to vibrate vertically at below 5 Hz, while it tended to vibrate horizontally at above 5 Hz. In this way, riding discomfort was affected by compression-tanδ of the seat cushion at 2 or 4 Hz, while it was affected by shearing-tanδ at 6 or 8 Hz. These results seem to suggest the possibility that the relationship between subjective human sensation and material properties may describe an accurate human function of evaluation using subjective measuring methods. Finally, the human model also suggests the way in which material properties of a seat cushion affect riding discomfort. Therefore, we may be able to apply this human model from the perspective of discomfort when designing car seats, if precision regarding anthropometry and proof of the reliability of the simulation can be sufficiently obtained.

5 Conclusions

The correlation coefficients between riding discomfort evaluated by 17 subjects who sat on a standard seat made of flexible polyurethane foam and the 41 material properties describing the static and dynamic characteristics of the seat cushion material revealed two distinct relationships. These relationships strongly relied on the frequency of the vertical sinusoidal vibration that the subjects were exposed to. Interestingly, the turning point for these relationships was distinguished at around 5 Hz. Notably, compression-tanδ, which governs the vertical vibration dumping factor of the seat cushion, affected riding discomfort at 2 and 4 Hz, while shearing-tanδ, which is predominant in the horizontal dynamic characteristic of the cushion, affected discomfort at 6 and 8 Hz.

Furthermore, a human model which was composed of simple 3D finite solid elements was proposed in order to confirm the above-mentioned relationships, or to discover the reason why the subjects unconsciously felt the effects of compression-tanδ and shearing-tanδ in evaluating their riding discomfort. The vibration system consisting of the human model and the simplified car seat using FEM was solved as an eigen value problem, and a few interesting eigen vibration modes of the human model were extracted. The finite element parameters controlling the characteristics of both the human model and the seat were aligned with the experimental data in order to conform to the lowest frequency vibration model of the real human-seat vibration system. The results confirm that the first eigen vibration mode at lower than about 5 Hz shows that the human vibrates mainly in the vertical direction (up-and-down), while the second and third modes at above 5 Hz show that the human is governed by fore-and-aft movement. In other words, at a lower frequency than about 5 Hz, the human tends to vibrate vertically and consequently, riding discomfort will be affected by the compression-tanδ governing the vertical characteristics of the seat; while at higher than about 5 Hz the human tends to vibrate in the horizontal direction and shearing-tanδ will affect riding discomfort.

In conclusion, the human model seems to indicate that subjective estimation of the relationship between riding discomfort and seat material properties is feasible. In addition, it may be possible to design seats using a digital human model to evaluate the relationship between the material properties and riding discomfort.

References

1. Chung, J., Jung, T.-C.: Optimization of an air cushion vehicle bag and finger skirt using genetic algorithms. Aerospace Science and Technology 8, 219–229 (2004)
2. Sørensen, A.J., Egeland, O.: Design of ride control system for surface effect ships using dissipative control. Automatica 31, 183–199 (1995)
3. Man-tractor system dynamics: Towards a better suspension system for human ride comfort. Journal of Biomechanics 11, 397–406 (1978)
4. Kubo, M., Terauchi, F., Aoki, H., Isobe, M., Okubo, K.: Riding comfort affected by the properties of flexible polyurethane form. In: Proceeding of 35th UK Group Meeting on Human Responses to Vibration, pp. 29–38 (2000)
5. Kubo, M., Terauchi, F., Aoki, H., Matsuoka, Y.: An investigation into a synthetic vibration model for humans – An investigation into a mechanical vibration human model. International Journal of Industrial Ergonomics, (27) (2001)

6. Tamemura, R., Terauchi, F., Kubo, M., Aoki, H.: A study on riding and sitting comforts of polyurethane car seats under real driving situation. In: Proceedings of 36th UK Group Conference on Human Response to Vibration, pp. 97–105 (2001)
7. Chiew, L.T., Kubo, M., Terauchi, F., Aoki, H.: Inter-subject variability of vibration characteristics of human seated in car seat. Bulletin of Japanese Society for the Science of Design, vol. 48 (2002)

The Effects of Human Interaction on Biometric System Performance

Eric P. Kukula[1], Stephen J. Elliott[1], and Vincent G. Duffy[2]

[1] Purdue University, Department of Industrial Technology, 401 N. Grant Street,
[2] Purdue University, Department of Industrial Engineering, 315 N. Grant Street,
West Lafayette, IN 47906 USA
{kukula,elliott,duffy}@purdue.edu

Abstract. This paper discusses the impact of human interaction with biometric devices and its relationship to biometric performance. The authors propose a model outlining the Human-Biometric Sensor Interaction and discuss its necessity through case studies in fingerprint recognition, hand geometry, and dynamic signature verification to further understand the human-sensor interaction issues and underlying problems that they present to the biometric system. Human factors, human-computer interaction and digital human modeling are considered in the context of current and future biometric research and development.

Keywords: human-computer interaction (HCI), biometrics, human-biometric sensor interaction (HBSI), digital human modeling (DHM).

1 Introduction

Biometric technologies are predominantly used to authenticate individuals in applications such as access control – to enter a facility or sign-on to a personal computer, identity management – search photographs for multiple aliases (Visa applications), or time and attendance. Biometrics is defined as the automated recognition of behavioral and physiological characteristics of an individual [1]. Biometric systems, by definition, require individuals to provide their personal characteristics to a sensor. Moreover, this interaction is critical, as many times, how to use a sensor might not be intuitive to the inexperienced user. The Human-Biometric Sensor Interaction (HBSI) is a conceptual model that is centered on the relationship between the human providing the biometric characteristic, such as a fingerprint to the biometric sensor; a relationship that becomes more critical to understand and research as biometric sensors become pervasive in society. The successful deployment of biometric systems, regardless of application, needs to take into consideration how individuals interact with the device. Failure to do so will cause a degradation of the optimal performance of the biometric sensor, causing problems such as: failure to acquire, failure to enroll, and impacts in the false reject rate. And if the individual cannot successfully interact with the biometric device then although the device has been implemented, there is a potential for a failure to use. So even though a device such as a laptop or a personal digital assistant (PDA) has a biometric sensor

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 904–914, 2007.

integrated, it does not necessarily mean that the individual will use it, especially if they have had previous negative experiences with the biometric sensor. The use of biometrics will likely be dependent on individuals' ability to not only use it more effectively, but also find it more useful than the technology that it replaces (such as the username / password combination in a computer sign-on application). Not only do designers need to consider these factors, those deploying biometric systems need to examine the procedure and methodology sections of articles and test reports to establish the type of population that the device was tested on. Gray and Salzman discussed the importance of experimental design in their review of five of the more popular (in terms of citations) experiments that compared usability evaluation methods (UEMs) for practitioners in the field of Human-Computer Interaction (HCI) [2]. The authors discussed that small problems in design called into serious question the recommendations and advice given from the results of those experiments due to experimental design, statistical power, and other validity issues [2]. This paper provides a discussion on problems documented in the literature, as well as those researched by the authors in the area of the Human-Biometric Sensor Interaction that may be caused by environmental conditions (illumination, temperature, or installation height), demographic impacts (age, gender, or medical/physical impairments), as well as device attributes (for example digitizer feedback for dynamic signature verification), and possibly the design or shape of a device. As noted in this paper, all modalities are affected by this human-biometric sensor interaction. However, those modalities that require either alignment with the sensor (iris), or contact (fingerprint and hand geometry) are probably more susceptible to human-biometric interaction issues than those sensors that do not require physical interaction. Therefore biometric practitioners and system designers should take the concepts of HBSI into consideration when designing the sensor and its associated form factor, in order to increase performance of the system. What may seem desirable from a marketing or artistic sense, may reduce user efficiency, effectiveness (system performance), and/or user satisfaction. This paper will investigate such topics further.

2 Biometrics

Human-Biometric Sensor Interaction is the examination of how individuals interact with a biometric sensor. Biometrics are concerned with the statistical analysis of human characteristics, such as the fingerprint, hand shape, face, iris, and signature, typically for security applications. The General Biometric Model as originally presented in [3] and updated in [4] covers the generic association of data collection, signal processing, matching, and decision-making. The Human Biometric Sensor Interaction model is concerned with the data collection portion of this model, which is vital component as subsequent modules rely on this portion to perform. The data collection component consists of the individual (human) presenting their biometric characteristic(s) to a sensor. The specific trait being collected and subsequently measured (fingerprint, hand samples) for example, needs to be presented in a consistent manner, providing unique samples for downstream modules to process.

While much research has concentrated on the improvement of algorithms and image processing to increase performance, how the target population interacts with the sensor is also of importance, but has not been thoroughly researched.

Understanding the design of the system and sensor is critical, as it must accommodate as many of the intended users as possible, as well as work in the targeted environment. Once these factors have been taken into consideration, the last step in system evaluation is to evaluate whether expected performance of the biometric system will ultimately satisfy the intended purpose. According to Jain, et al., the complexity of designing a biometric system is based on three main attributes – accuracy, scale (size of the database), and usability [5]. As utilization of biometric technology becomes more pervasive in society; understanding the interaction between the human, the environment, and the biometric sensor becomes increasingly imperative. There are a number of common biometric modalities (hand geometry, fingerprints, face, iris, signature, and voice recognition), and the majority of these follow the general biometric model discussed in [3]. For the purposes of this paper, we discuss hand geometry, fingerprint, and dynamic signature verification.

2.1 Common Biometric Modalities and the Impacts on the HBSI

Hand Geometry. Hand geometry has a well documented history, coming of age in the 1980's for applications such as time and attendance, and access control [6, 7]. When discussing the issues of the Human-Biometric Sensor Interaction with regard to hand geometry, those implementing such a system need to take into consideration several factors. One factor is the universality of the hand (or lack thereof) - an individual's hand may be significantly different from the “average” expected hand. For instance, an individual might be missing a digit, which could cause an acquisition problem, known as a failure to acquire (FTA). A hand geometry system requires the user to place their hand on the platen so that all of their digits touch all of the guide pins simultaneously for a set period. If some of digits do not touch all of the guide pins, then an individual with small hands or missing digits must compensate by rotating the hand, causing either ulnar or radial deviation of the wrist.

Fingerprint Recognition. Fingerprint recognition is the most widely used biometric technology with nearly 60% of the market [8]. Fingerprint recognition has an extensive history, with awareness of fingerprint individuality dating back to Ancient Mesopotamia circa 3000 B.C. [9, 10] Although it would be ideal for a biometric to be stable throughout the life of the individual, this may be difficult to achieve. Over time, general wear on the body will change the characteristics. A person can also change his or her biometric characteristics by growing a beard, losing weight, intentionally altering a signature (including changing his or her name), or injuring a specific feature (e.g., losing a digit, cutting and scarring a finger). Unintentional changes resulting from occupational wear and tear, as well as normal aging, can also cause scarring or wrinkling of fingerprints, further challenging the repeatability of these biometric characteristics. Each of these conditions or alterations can affect image quality, and thus the performance of the biometric recognition algorithm. Several papers discuss how physical changes such as scarring, wrinkles, contact with the sensor (force and speed), as well as age can impact repeatability of image capture or image quality that can impact fingerprint algorithm performance [11-14].

Dynamic Signature Verification. Dynamic signature verification has long been used to authenticate individuals based on their signing characteristics. Applications such as document authenticity, financial transactions, and paper-based transactions have all, at one time or another, used the signature to convey the intent to complete a transaction [15, 16]. The signature is primarily a behavioral biometric. As such, it includes challenges that need not be factored into some of the other modalities; examples include the fact that a signature is learned, contains measures that can be altered depending on the sensor, can be changed by the owner, and can be one of many versions, all depending on the transaction or intent of the signer. Current dynamic signature verification applications are normally employed in lieu of the traditional paper/ink combination (e.g., at a retail point-of-sale); as such, the signature might not be verified at the specific moment of the transaction — it might not be challenged until some later date. The signature has a limited number of extractable features, including the x, y coordinates, time, and depending on the digitizer, pressure [17]. These input variables are used to create the global and local features described in various accounts in the literature [18, 19]. Nalwa observed that the temporal "characteristics of the production of an on-line signature are the key to the signature's verification" [20]. Typical signature functions include pressure, horizontal and vertical components of position, velocity, acceleration, and force, all compared to time. Another way of characterizing a signature is through the analysis of the "stroke" — that is, the pen-down, pen-up movement of the pen on the digitizer. All of these various dynamic traits collected during the act of signing are said to make an impostor signature easier to detect.

2.2 Biometric Properties and Ergonomic Implications

In addition to classifying biometric modalities as either physiological or behavioral, biometrics can also be classified by five desirable properties, outlined in [3, 21], and amended by numerous others. Desirable properties of biometric characteristics are that they offer: universality – available on all people; invariance – features extracted are non-changing; high intra-class variability – features extracted are distinct from each other user; acceptability – characteristic is suitable for use by everyone; and extractability – a sensor can extract the features presented in a repeatable manner. Although commonly described in the literature as the ideal characteristics of the biometric, each must overcome challenges. Herein lies one of the challenges associated with large-scale deployment of biometrics and the purpose behind research in this area – the majority of biometrics are challenged to satisfy all of these five categories.

3 Influence of the Scientific Discipline of Human Factors and Human-Computer Interaction on Biometrics

One can argue that one of the most influential inventions of the twentieth century with regards to information technology was the computer mouse. The invention of the mouse in 1963 changed the way users interacted with computer interfaces; decreasing the level of complexity required to interact and manipulate items on a computer visual display terminal. While other devices were crafted, comparative testing undertaken by NASA in the 1960s between light pens, cursor keys, joysticks, trackballs, revealed that mouse designed by Engelbart and English resulted in their device outperforming the other devices due to its ease of use [22]. **Fig. 1** shows early designs of the mouse. Stu Card, one of the pioneers of the mouse, recalls that Fitts's Law was used to test the movement of the hand with and without the mouse device [22]. Observations and tests showed that the slope of that curve was about 10 bits per second, which was very similar to the measurements for just moving the hand alone without a device, revealing the limitation was not the mouse device, rather was the hand-eye coordination system of the user. Since the time to move the hand with and without the device was very similar English and Card concluded the design was nearly optimal.

Fig. 1. First mouse design 1963-64 (left), first production mouse (3rd picture from left). Xerox mouse designs during the 1980s (3 images to the right in order of increasing radical design) [22]

Those working with examining the human-biometric sensor interaction can learn much from the early pioneers in HCI, especially from the development of the mouse, as it was a vital component of humans interacting with computers. Historically the biometrics community has performed limited work in the area of human-computer interaction and related fields of ergonomics and usability. In 2003, leading biometric experts met under the sponsorship of the National Science Foundation (Grant EIA-0240689) to "develop a rational and practical description of crucial scholarly research to support the development of biometric systems" [5]. Within the proposed research agenda was an item on ergonomic design of the capture system and usability studies to evaluate the effect on system performance [5]. At the time of writing, literature on the subject of ergonomics, perceptions of biometrics, and usability has been limited, but is gaining interest. Papers discussing ergonomic and usability studies regarding biological biometrics can be found in [23-26], while [27-29] discuss perceived acceptability of behavioral biometrics.

Recent work conducted by the National Institute of Standards and Technology (NIST) examined the effect of the height of a fingerprint sensor on the image quality and the time required to collect fingerprints to answer three questions: Does work surface height affect: the time required to capture fingerprint images or the quality of captured images And if users prefer a particular work surface height [26]. Results from the study consisting of 75 NIST employees revealed a counter height of 36 inches (914 mm) gives fastest performance, while the 26 inch (660 mm) counter height gave the highest quality fingerprint images, while counter heights of 32 or 36 inches (813 or 914 mm) were the most comfortable for users. Similar work has been conducted by Kukula, Elliott, et. al with hand geometry to investigate the effect that different working heights have on the performance of a hand geometry device [30]. Four hand geometry devices were used in this evaluation with the heights at which devices were placed at the three recommended height ranges provided by Grandjean for the different types of work: precision, light, and heavy [31]. The fourth device was mounted at a similar height as the hand geometry device that was found on the Purdue BSPA laboratory door. The results of the 32 participant evaluation revealed that there were no differences among the median value of match scores across the different four heights (30, 35, 40, and 45 inches), thus allowing the biometric device installer some flexibility when implementing hand geometry devices. Users were also surveyed to establish which device satisfied them. The results showed that 63% of the users preferred the hand geometry device mounted at 40 inches. In addition, the study revealed a correlation of preferred hand geometry height and that of the user, therefore a practitioner should take into consideration the height of the intended user population before installing a hand geometry device.

Up to this point, the paper has discussed biometrics and the need for designers to include usability metrics, as well as ergonomic principles during the design stage of biometrics devices. More studies are needed in the area of the HBSI to inform designers and practitioners of potential problems with the design of a biometric form factor, device, or system implementation. Two case studies (using fingerprint as the biometric modality) that illustrate these issues discussed in the beginning part of this paper highlight these concepts.

The swipe fingerprint sensor experiment examined a capacitance swipe sensor embedded in a commercially available laptop computer. The purpose of the evaluation was to examine issues related to fingerprint acquisition of all ten digits to evaluate the biometric-user interface in terms of biometric performance metrics, error rates, in particular the failure to acquire (FTA) and failure to enroll (FTE) rates. The study involved 88 participants, most of which were between the ages of 18 and 30. The research identified multiple factors warranting consideration when designing a biometric system in order to better understand why failures occur and how individuals react to systems. The results of this experiment revealed that the medial digits (thumb, pointer/index, and middle) had fewer acquisition problems than the lateral digits (ring and little fingers) of both hands, which can be attributed to a decrease in finger dexterity as one moves from the index to the little finger. In addition, if you examine the left image in **Fig. 2**, you can easily see the ulnar deviation required for a user's thumb to interact with the sensor. To read more about this experiment, please refer to [30].

Fig. 2. The HBSI between a user's thumb and swipe fingerprint sensor producing ulnar deviation (left) and an optical fingerprint sensor mounted on a force sensor (right)

The optical fingerprint sensor experiment used the optical fingerprint sensor and housing that is used in the United States Visitor and Immigrant Status Indicator Technology (US-VISIT) program, to see if image quality is correlated with the amount of force an individual applies on the sensor. Four force levels were evaluated: 3, 9, 15, and 21 newtons. In addition to image quality, user comfort level was evaluated to determine what level of force was acceptable to users. The results of the 29-person experiment revealed that there was a statistically significant difference between image quality and applied force $F_{(.95, 3, 344)} = 22.56$, $p = 0.000$. Investigating the quality scores further, the scores significantly increased between 3N and 9N, but there was minimal benefit of applying more than 9N of force, as the quality scores increased minimally. The subjective comfort level question also revealed that the more pressure a user applied, the more uncomfortable it was for subjects to interact with the device. The image to the right in **Fig. 2** shows the experimental setup for the experiment.

3.1 The Human-Biometric Sensor Interaction

Biometric systems are largely dependent upon the sensor to acquire the sample and the segmentation and feature extraction modules to manage data from the collected sample so that the matcher can determine the correct response. By observing how users interact with biometric sensors, several design issues could be resolved by integrating knowledge of industrial and interactive design, ergonomics, and usability. Rubin discusses five reasons why products and systems are difficult to use. The main problem is that the emphasis and focus has been on the machine/system and not on the end user during development [32]. These factors are true within the context of the biometric sensor. Humans will adapt to the sensor, as discussed in [30] where users will rotate their hands if they have missing digits or spread their fingers if they have small hands in order to satisfy the requirements of a hand geometry device. Many times, sensors and their form factors are not tested on sufficiently large numbers of the general populations, namely due to the cost of doing so. The biometric community may test the algorithms exhaustively off-line, using pre-collected images. In addition to the physical interaction, the biometric-sensor interaction also poses cognitive interaction problems, that is, problems that arise due to the burden of conscious thought needed to operate the system, which can be addressed by work related to cognitive theory of explanatory learning (CE+) as discussed in [33].

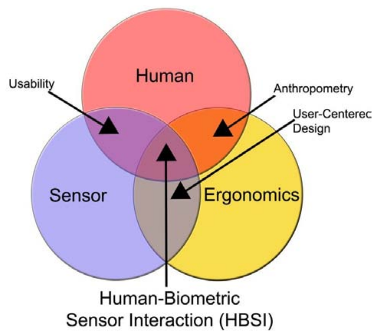


Fig. 1. Interaction of the human, biometric sensor / device, ergonomics, and related fields forms the HBSI

To address the issues discussed throughout this paper, the authors have proposed a conceptual model called the HBSI, which is a combination of various methodologies: anthropometry, ergonomics, user-centered design, and usability (see). The HBSI applies generic ergonomic principles outlined in the approaches of [32, 34-40], among others, to the design of a biometric device or system. A result of this model will hopefully provide a framework for biometric device designers to follow so that stressors on the human body are excluded or at least minimized, as well as reducing the loading on the user's mental model, making it easier for one to become habituated, or perform an action using subconscious thought, to a device more quickly. The model will also aid designers to help place the body in a repeatable, stress-free location so that the biometric can be collected under optimal conditions. The authors are currently undertaking research that validates this model through an increase in overall system performance, improved through a higher rate of sample repeatability, a more intuitive user interface, or through further examination of interaction producing better training methods for biometrics.

4 Discussion

The intent of this paper was to discuss the effect human interaction has on biometric system performance to begin dialogue with Human Factors, usability, and Human-Computer Interaction experts. We have outlined work in the fields of ergonomics, usability, and HCI, and discussed how they not only relate to biometrics, but can be integrated into the research to create better and more usable biometric devices that users are satisfied with, which will lead to lower failure to acquire, failure to enroll, and false rejection rates. The authors are not alone in their thoughts and opinions that this dialogue needs to begin. Smith stated that some members of the HCI community believe that interfaces of security systems do not reflect good thinking in terms of creating a system that is easy to use, while maintaining an acceptable level of security [41]. Moreover Adams and Sasse discussed the fact that security systems are one of the last areas to embrace user-centered design and training as essential [42]. Lastly, Maple and Norrington, noted three observations that align with the authors advocating for continued research in the HBSI:

– People with different physical characteristics interact differently with equipment,
– People have different sensory abilities to perceive the test area and equipment, and
– People have different cognitive abilities [25].

5 The Future of Biometrics in Digital Human Modeling

An understanding of the mismatch between predicted and actual performance, along with the need for further development of devices incorporating biometrics, provides an opportunity for considering the human aspects from the outset. Digital human modeling (DHM) can facilitate these efforts through further cross-disciplinary research. In this case, electrical engineering and computer science practitioners, their users, and their designs could be the beneficiaries of past research in human factors and ergonomics. DHM uses humans as representations of the user inserted into a simulation or virtual environment to facilitate the prediction of performance and/or safety. Traditional DHM has focused mostly on industrial operators or product users. DHM includes not only the visualization, but the math and science in the background. Hand geometry, fingerprint recognition, dynamic signature verification, and face recognition, as well as the other biometric modalities provide opportunities for developing predictive performance models for designers that utilize computer intensive resources along with information available in current biometric databases.

Improvements in biometric device design can be realized from current human factors and ergonomics principles. Biometrics may be further benefited from a strategic plan developed in cooperation with computer-aided engineering related research and DHM.

References

1. International Organization for Standardization, ISO/IEC JTC1/SC37 Standing Document 2 - Harmonized Biometric Vocabulary, ANSI. p. 66 (2007)
2. Gray, W., Salzman, M.: Damaged merchandise? A review of experiments that compare usability evaluation methods. Human-Computer Interaction 13(3), 203–261 (1998)
3. Mansfield, A., Wayman, J.: Best Practices in Testing and Reporting Performances of Biometric Devices, Biometric Working Group. pp. 1–32 (2002)
4. International Organization for Standardization, ISO/IEC 19795-1: Information technology - Biometrics performance testing and reporting - Part 1: Principles and framework, ANSI. p. 64 (2006)
5. Rood, E., Jain, A.: Biometric Research Agenda: Report of the NSF Workshop, 2003, National Science Foundation: Morgantown (2003)
6. Ashbourn, J. (ed.): Biometrics: Advanced Identity Verification: The Complete Guide, vol. 200. Springer, New York (2000)
7. Miller, R.: Finger Dimension Comparison Identification System. United States (1971)
8. IBG. Biometrics Market and Industry Report 2006-2010 (2006) (cited 2006 September 12) Available from: http://www.biometricgroup.com/reports/public/market_report.html
9. Berry, J., Stoney, D.: History & Development of Fingerprinting. In: Lee, H.C., Gaensslen, R.E. (eds.) Advances in Fingerprint Technology, pp. 1–40. CRC Press, Stoney (2001)
10. Burke, J.: AuthenTec Ships Record 10 Million Biometric Fingerprint Sensors (2006) (cited 2006 August 10) Available from: http://www.authentec.com/news.cfm?newsID=217
11. Elliott, S., Kukula, E., Sickler, N.: The challenges of environment and the human biometric device interaction on biometric system performance. In: International Workshop on Biometric Technologies - Special forum on Modeling and Simulation in Biometric Technology. 2004. Calgary, Alberta, Canada (2004)
12. Kang, K., et al.: A Study on Performance Evaluation of Fingerprint Sensors in Audio- and Video-Based Biometric Person Authentication. In: Goos, G., Hartmanis, J., van Leeuwen, J. (eds.): Springer, Berlin / Heidelberg. pp. 574–583 (2003)
13. Modi, S., Elliott, S.: Impact of Image Quality on Performance: Comparison of Young and Elderly Fingerprints. In: 6th International Conference on Recent Advances in Soft Computing (RASC 2006). Canterbury, UK (2006)
14. Sickler, N., Elliott, S.: Evaluation of fingerprint image quality across an elderly population vis-a-vis an 18-25 year old population. In: 39th Annual International Carnahan Conference on Security Technology (ICCST). 2005. Las Palmas de Gran Canaria, IEEE, Spain (2005)
15. Nelson, W., Kishon, E.: Use of dynamic features for signature verification. In: 1991 IEEE International Conference on Systems, Man, and Cybernetics, Decision Aiding for Complex Systems 1991, IEEE, Charlottesvill, VA (1991)
16. Hamilton, D.J., et al.: Low cost dynamic signature verification system. In: 1995 European Convention on Security and Detection, 1995, IEEE, Brighton, UK (1995)
17. Vielhauer, C., Steinmetz, R., Mayerhofer, A.: Biometric hash based on statistical features of online signatures. In: 16th International Conference on Pattern Recognition, 2002, IEEE (2002)

18. Leclerc, F., Plamondon, R.: Automatic Signature Verification: The State of the Art - 1989-1993. World Scientific Publishing Co, Singapore (1994)
19. Brault, J.-J., Plamondon, R.: A Complexity Measure of Handwritten Curves: Modeling of Dynamic Signature Forgery. IEEE Transactions on Systems, Man, and Cybernetics 23(2), 400–413 (1993)
20. Nalwa, V.: Automatic On-Line Signature Verification. Proceedings of the IEEE 85(2), 215–239 (1997)
21. Clarke, R.: Human identification in information systems: Management challenges and public policy issues. Information Technology & People 7(4), 6–37 (1994)
22. Moggridge, B.: Designing Interactions, vol. 766. The MIT Press, Cambridge (2007)
23. Coventry, L.A., De Angeli, A., Johnson, G.: Biometric Verification at a Self Service Interface. In: McCabe, P. (ed.) Contemporary Ergonomics 2003, pp. 247–252. Taylor and Francis, London (2003)
24. Coventry, L.A., De Angeli, A., Johnson, G.: Usability and biometric verification at the ATM interface. In: Conference on Human Factors in Computing Systems 2003, Ft. Lauderdale, ACM Press, Florida (2003)
25. Maple, C., Norrington, P.: The Usability and Practicality of Biometric Authentication in the Workplace. In: First International Conference on Availability, Reliability and Security (ARES'06), IEEE, Vienna, Austria (2006)
26. Theofanos, M., et al.: Does habituation affect fingerprint quality? In: Conference on Human Factors in Computing Systems 2006, Montréal, Québec, ACM Press, Canada (2006)
27. Deane, F., et al.: Perceived acceptability of biometric security systems. Computers & Security 14, 225–231 (1995)
28. Deane, F., et al.: Theoretical Examination of the Effects of Anxiety and Electronic Performance Monitoring on Behavioural Biometric Security Systems. Interacting with Computers 7(4), 395–411 (1995)
29. Henderson, R., et al.: Electronic Monitoring Systems: An Examination of Physiological Activity and Task Performance within a Simulated Keystroke Security and Electronic Performance Monitoring System. International Journal of Human-Computer Studies 48(2), 143–157 (1998)
30. Kukula, E., et al.: Biometrics and Manufacturing: A Recommendation of Working Height to Optimize Performance of a Hand Geometry Machine. Purdue University Biometrics Standards, Performance, & Assurance Laboratory. p. 18 (2006)
31. Grandjean, E.: Fitting the Task to the Man: A Textbook of Occupational Ergonomics, 4th edn. Taylor & Francis, London (1988)
32. Rubin, R.: Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. John Wiley and Sons, Inc, New York (1994)
33. Polson, P., Lewis, C.: Theory-Based Design for Easily Learned Interfaces. Human-Computer Interaction 5(2&3), 191–220 (1990)
34. Bailey, R.: Human Performance Engineering: A Guide for System Designers. Prentice Hall, Englewood Cliffs (1982)
35. Booth, P.: An Introduction to Human-Computer Interaction. Lawrence Erlbaum Associates, London (1989)
36. Chignell, M., Hancock, P.: Design orientation and ergonomics. In: Chignell, M. (ed.) IEEE International Conference on Systems, Man, and Cybernetics, 1992, IEEE, Chicago, IL (1992)
37. Gould, J., Lewis, C.: Designing for Usability: Key Principles and What Designers Think. Communications of the ACM 2(3), 300–311 (1985)

38. Intern. Organization for Standardization, ISO 9241: Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability. p. 28 (1998)
39. Pheasant, S.: Bodyspace: Anthropometry, Ergonomics, and the Design of Work. In: Haslegrave, C. (ed.): 3rd edn., vol. 332. Taylor and Francis, Boca Raton (2006)
40. Tayyari, F., Smith, J.: Occupational Ergonomics: Principles and Applications. In: Parsaei, H. (ed.) Manufacturing Systems Engineering Series, vol. 452, Kluwer Academic Publishers, Norwell (2003)
41. Smith, S.: Humans in the Loop: Human-Computer Interaction and Security. IEEE Security and Privacy 1(3), 75–79 (2003)
42. Adams, A., Sasse, M.: Users are Not the Enemy: Why Users Compromise Security Mechanisms and How to Take Remedial Measures. Communications of the ACM 42(12), 41–46 (1999)

Design and Realization of Synthesis Assessment System for Cockpit Ergonomics

Yinxia Li[1], Kaiying La[1], and Xiugan Yuan[2]

[1] School of Mechanical Engineering, Zhengzhou University, Science Road No.100, Zhengzhou Henan 450001, China
liyinxia@buaa.edu.cn
[2] Institute of Human-machine & Environment Engineering, Beihang University, XueYuan Road No.37, HaiDian District, Beijing 100083, China

Abstract. This paper emphasized on how to construct one kind of assessment system based on the standards and specifications for cockpit ergonomic synthesis assessment, and the knowledge of the aviation experts and the experienced pilots. It analyzed the necessity of developing the assessment system, and then presented the basic principal and the function demand of the assessment system. In addition, the key technologies for developing this assessment system were discussed in detail utilizing Visual FoxPro 6.0. This assessment system transforms the assessment model and process into the computer program by the mathematical model, enables the assessment process truly to be operational, and also can reduce the assessment cycle.

Keywords: Cockpit, Ergonomics, Human factors, Human engineering, Fuzzy.

1 Introduction

Along with the unceasing enhancement of airplane performance, the synthesis assessment of cockpit ergonomics more and more becomes an important problem related to enhancing the efficiency and reducing the working load of pilot. Because there is many aspects related to ergonomics in the cockpit, each aspect involves many factors, these factors have both the quantitative factor and the qualitative factor, moreover the evaluation criteria of each factor has the indefinite and fuzzy characteristic. Therefore, it is not an easy process to find out one kind of effective synthesis assessment method of cockpit ergonomics and to develop a practical synthesis assessment system for cockpit ergonomics. A cockpit ergonomic synthesis assessment method blending the modified Delphi method, G_1 method, and fuzzy synthesis evaluation method was proposed in document [1]. Because there are many indexes need to be assessed during synthesis assessment of cockpit ergonomics, when carries on this method, not only needs to inquire massive standard data, but also have the massive indexes characteristic value quantification and the fuzzy relational operation, the so complex assessment process is not obviously to the manual operation.

At present, the soft wares that can be used in cockpit ergonomic assessment mainly are all kinds of computer-aided human modeling systems for cockpit ergonomic

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 915–922, 2007.

analysis [2]. The functions of these soft wares are only limited to analyze the compatibility between cockpit geometry design and pilot anthropometry not to assess cockpit ergonomics completely, thus cannot determine the quality of different cockpit ergonomic design proposals. Along with the technical development of cockpit CAD, digitized prototype, virtual reality and so on, it set the urgent need to research and develop the computer-aided synthesis assessment system for cockpit ergonomics.

The objective of this paper is to research and develop one kind of assessment system with computer technology based on the standards and specifications for cockpit ergonomic synthesis assessment, and the knowledge of the aviation experts and the experienced pilots, that is SASCE (the Synthesis Assessment System for Cockpit Ergonomics). The assessment model and process can be transformed into the computer program by SASCE. This can enable the assessment process truly to be operational, and also can reduce the cycle and expense of assessment.

2 Basic Principle of the Synthesis Assessment Method

The basic principle of synthesis assessment method of cockpit ergonomics presented in document [1] is to use the Delphi method to build the assessment index system, then use the G_1 method to determine the weight factors of all indexes, the person's subjective judgment can be quantified objectively based some scales and the qualitative problem can be analyzed quantitatively by the G_1 method, and finally the assessment result can be obtained by the fuzzy method. The basic principle of synthesis assessment method of cockpit ergonomics shows in figure 1.

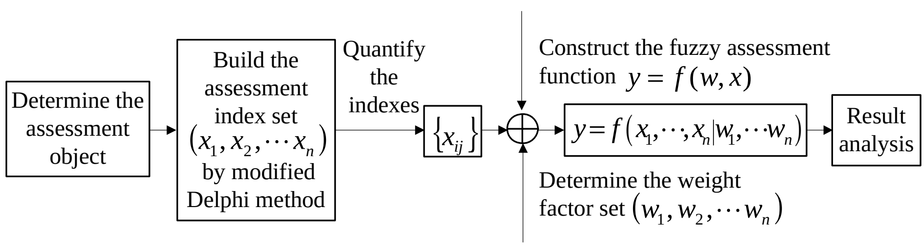


Fig. 1. Basic principal of synthesis assessment of cockpit ergonomic

3 Basic Structure and Functions of SASCE

According to the basic principle of the synthesis assessment method of cockpit ergonomics and the demands of the cooperation department, SASCE is divided into three layers. They are the user layer, the system interior processing layer and the database layer. This layer structure may conveniently separate the data treating processes and the database, and also enable the system to have clear and integrated structure. Figure 2 is the schematic drawing of the assessment system structure. The definitions and functions of three layers are explained in brief as follows.

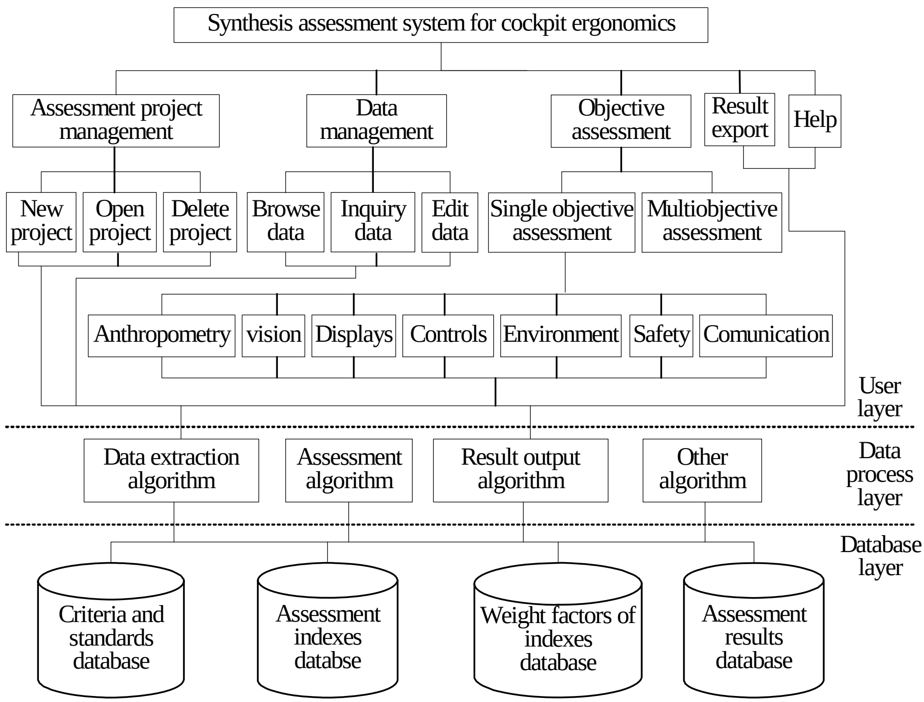


Fig. 2. Schematic drawing of SASCE structure

3.1 User Layer

This layer is the most outer layer and the user interface of SASCE. It contains all the function modules as follows.

Assessment Project Management Module. User can new, open or delete an assessment project of cockpit ergonomics by this module.

Data Management Module. User can browse inquiry or edit the data stored in all kinds of databases by this module, such as assessment standards and specifications database, assessment indexes database, weight factors of indexes database and so on.

Objective Assessment Module. This module can realize the quantification of all indexes characteristic value and make user realize single-objective or multi-objective assessment by input some required data.

Result Export Module. User can export or print the needed information by this module, such as the inquiry results of some assessment standards and specifications, assessment results of indexes and so on.

Help Module. This module provides some helpful information, such as the introduction and application of SASCE.

3.2 Database Layer

This layer is the most inner layer of SASCE. The standards and specifications for cockpit ergonomic synthesis assessment, the knowledge of the aviation experts and the experienced pilots are organized and stored in this database layer. The database structure is flexible. The data stored in database is the foundation stone of the assessment system and can be edited conveniently.

3.3 Data Processing Layer

This layer is the middle layer of SASCE and the bridge connected the user layer and the database layer. The functions of this layer are transferring the data in database and realizing all kinds of algorithms and so on. This layer is the key of realizing the real-time feedback and the intellectualization of the assessment system.

4 Operating Principle of SASCE

Figure 3 is the schematic drawing of operating principle of SASCE. The principle of each constituent is as follows.

4.1 Interface Routine (Internal Interface of SASCE)

Through the image processing technology, the data affected pilot ergonomics can be derived from the digitized cockpit/prototype developed by three-dimensional modeling

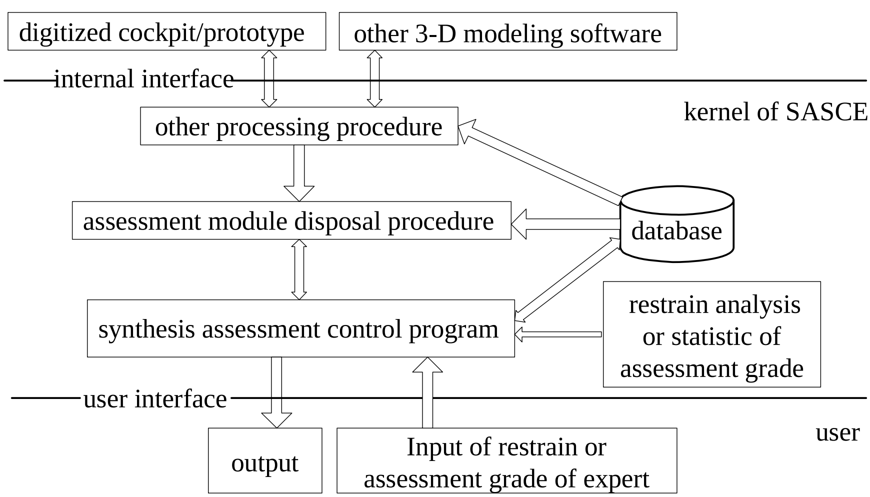


Fig. 3. Schematic drawing of operating principle of SASCE

software. These data may use to assess the cockpit ergonomics after some pretreatments. Moreover, some ergonomic data needed to cockpit design in three-dimensional modeling software can be also obtained from SASCE. Such as some cockpit geometry sizes needed to build the digitized cockpit model, some anthropometrics data used to create the three-dimensional manikin and so on. SASCE and other software, such as the computer-aided analysis system for cockpit layout ergonomics developed by Institute of Man-machine and Environment Engineering of Beihang University, can be integrated to form a computer-aided design and assessment system for cockpit ergonomics with more powerful functions by the internal interface of SASCE. Thus it can realize the real-time assessment of cockpit design, speeds up the cockpit design speed.

4.2 Human-Computer Interactive Interface (User Interface of SASCE)

The function of user interface of SASCE is to realize the dialogue interaction between the user and SASCE. Considered the user's initiative, the user interface of SASCE provides the quite big choice except the convenience for user, such as SASCE operated by menu and mouse, the information needed can be exported as *.txt, *.xls files or printed.

4.3 Main Procedure

Main procedure includes synthesis assessment control program, assessment module disposal procedure, and other processing procedure. Its functions are transforming data and realizing all kinds of algorithms.

4.4 Database

The data stored in database mainly includes the standard and specifications of cockpit ergonomic assessment, the experience knowledge obtained from the experts in aviation department and the experienced pilots, assessment indexes, weight factors of indexes, and some assessment results of single-index or multi-index. These data can be managed by database.

5 Realization Technology and Methods

SASCE was developed using the object-oriented programming method and the powerful function of database development of Visual FoxPro 6.0(Abbr.VFP6) [3-5]. The essential technology and method of each main function realization of SASCE is as follows.

5.1 Creation of Database

Creation of Assessment Index Database. All the indexes needed to assess cockpit ergonomics are stored in the assessment index database named “indexcock”. Table 1 shows the structure of “indexcock”.

Table 1. Structure of "indexcock"

Field name	Field type	Field width	Notes
Index_key	Character	10	Keyword of index
Parent key	Character	10	Keyword of upper-index including the current index
Index name	Character	40	Index name

Creation of Weight Factors Database. The weight factors of all the assessment indexes are stored in the weight factor database named "indexweight". Table 2 shows the structure of "indexweight".

Table 2. Structure of "indexweight"

Field name	Field type	Field width	Notes
Index_ code	Character	4	Sequence number of index
Parent_ name	Character	40	Index name
Index_ weight	Number	(6, 4)	Weight factor of index

5.2 Realization Technology of Display and Edit Function of Assessment Index

The display function of ergonomic assessment indexes of cockpit(that is the index information of"indexcock") was realized using the TreeView Control in VFP6. Thus the index can be displayed as the style of hierarchy structure. Figure 4 is the display

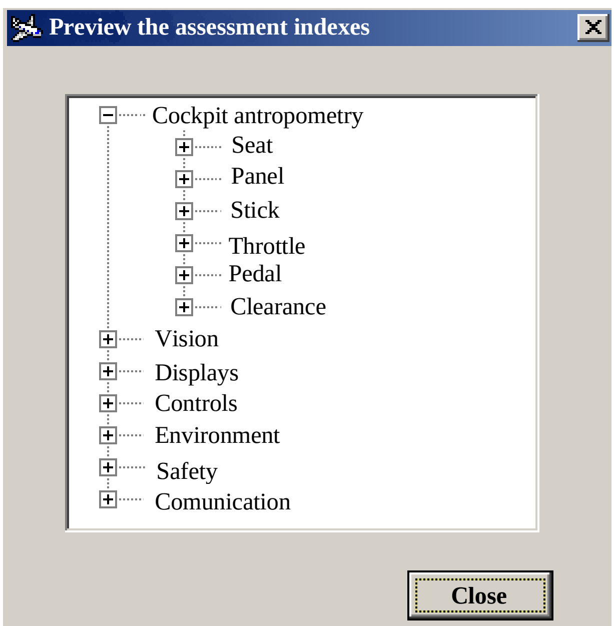


Fig. 4. Display interface of assessment indexes

interface of assessment indexes as the style of hierarchy structure realized by TreeView Control.

In order to enable the assessment index system to have openness, SASCE provides the edition function of index. The edit function interface of index is similar to figure 4, only was increased the revision, the increase and the deletion button to realize edition function on the interface.

5.3 Realization Technology of Assessment Function of Index

The synthesis assessment method of cockpit ergonomics is a multilevel synthesis assessment that starts from the lower level to the upper level (for details see the document [1]). The mathematic model of synthesis assessment method is:

$$y=\sum_{i=1}^{n} w_i \circ \mu_i \ , \ i=1,2,\cdots,n \qquad (1)$$

where w_i is the weight factor of index u_i, μ_i is the membership grade of index u_i, "$\circ$"is the fuzzy operator "$M\,(\bullet,\oplus)$"[1]。 Take the synthesis assessment of cockpit anthropometry as the example, the realization interface of its assessment function like figure 5 shows. The realization of mathematic model of synthesis assessment is by the "Click" event of the "Ok" button.

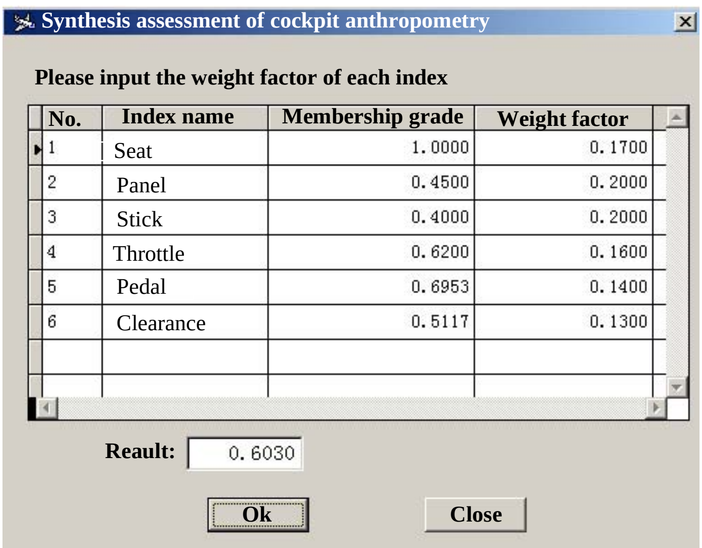


Fig. 5. Interface of synthesis assessment of cockpit anthropometry

6 Conclusions

The Synthesis Assessment System for Cockpit Ergonomics (SASCE) was designed and developed. It integrated the standards and specifications for cockpit ergonomic synthesis assessment, the knowledge of the aviation experts, and the experienced pilots. It transforms the assessment model and process into the computer program, enables the assessment process truly to be operational. SASCE has already used to assess the ergonomics of a new digitization cockpit design proposal in the aeronautical engineering department. The results got approve from the experts and design engineers in aviation department.

References

1. Li, Y., Yang, F., Wang, L., et al.: Study on comprehensive evaluation of cockpit ergonomics and its application [J]. Journal of Beijing University of Aeronautics and Astronautics 35(6), 652–656 (2005)
2. Li, Y., Yuan, X.: Computer-aided human modeling systems for cockpit design [J]. Application Research of Computers, 331–333 (2003)
3. Weng, Z.: Visual FoxPro 6 Database Development tutorial [M]. TsingHua University Press, Beijing (2000)
4. Li, C.: Professional Visual FoxPro 6.0(ActiveX Controls, Multimedia and Internet Applications) [M]. Science Press, Beijing (2001)
5. Jiang, L., Liu, H., Ni, F.: Visual FoxPro 5.0 Commands and Functions [M]. Posts & Telecom Press, Beijing (1998)

Development of a Test-Bed for Synthetical Ergonomics Evaluation of Pilot

Wei Liu[1,2], Xiu-gan Yuan[2], Zhong-qi Liu[2], Rui Ma[2], and Wei-yong Kang[2]

[1] School of automation, Beijing university of Post Technology, Beijing, 100876, China
[2] School of aviation science and engineering, Beijing University of Aeronautics and Astronautics, Beijing, 100083, China
twhlw@163.com

Abstract. The study of pilot perception has evolved from examining simple tasks executed in reduced laboratory conditions to the examination of complex, real-world behaviors. A test-bed for evaluating models of pilot behavior in which a set of infrared video-based eyetrackers to monitor subjects' eye movements while they perform a range of complex tasks such as driving, and manual tasks requiring careful eye-hand coordination was developed. The real cockpit platform is provided to subjects as they pilot in a virtual scene; a dual-haptic interface consisting of two touch-screen display devices and a shift allows free control within the cockpit.

Keywords: test-bed, eye tracking, haptic interface, ergonomics.

1 Introduction

The study of human behavior is challenging for many reasons. Historically, experimenters have attempted to make such studies tractable by reducing the complexity of both the environment and the task under study. As a result, much of what is known about human behavior is restricted to over-simplified tasks, such as single eye movements made in response to the onset of a signal light, or a reach to an isolated target in a dark field[1][2]. We have been striving to expand the range of behaviors to include the kind of complex behavior that makes up flight tasks. In that effort, we are developing a test-bed in which the behavior of pilots can be studied as they perform arbitrarily complex tasks.

A concern about such experiments is the ability to make meaningful conclusions about pilot performance because of the very complexity we are striving to understand. The use of virtual scenes/panels allows us to maintain enough control over the visual and haptic environment to permit meaningful analysis. In addition to traditional metrics of performance (such as reaction time, success rates, etc.), we monitor subjects' eye movements to provide an externally observable marker of attention. Electrocardiograph recordings can also be used to provide multiple indicators of physiological state. In order to accomplish the goal, many component subsystems must be successfully integrated. It is necessary to generate complex, stereo virtual scene/panels. The generated images must be displayed with sufficient fidelity to

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 923–928, 2007.

provide a strong sense of immersion so that natural behaviors can be observed. The visual and haptic simulations comprise the perceptual input to the subject; a critical aspect of the test-bed goal is the ability to monitor subjects' cognitive state as they perform complex tasks in the virtual flight. Each of these elements is discussed in the following sections. See figure 1.

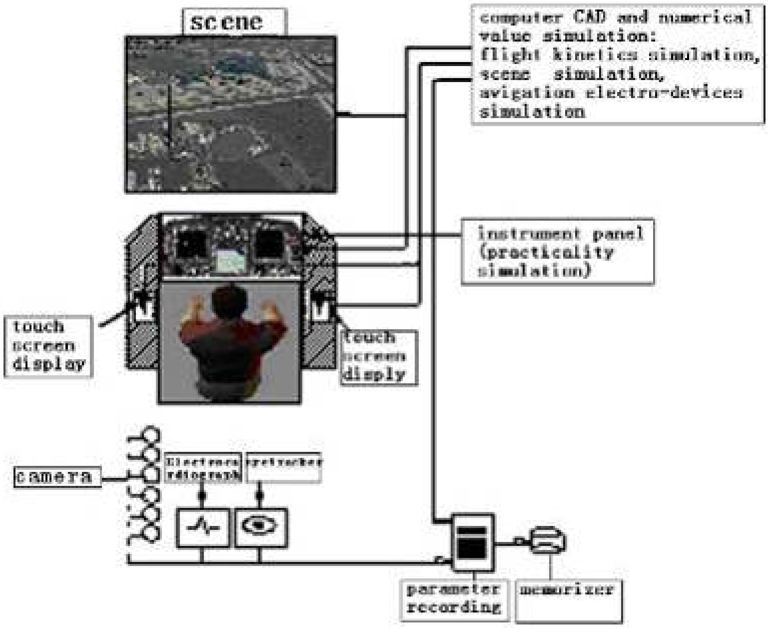


Fig. 1. A test-bed system principle figure of synthetical ergonomics evaluation of pilot

2 Graphics Generation and Display

The heart of the graphics generation system is a double-processor PC equipped with two Pentium 500 processors and two special software Mutigen-Vega. The system generates stereo image pairs at up to 60 Hz, depending on the complexity of the scene. Figure 2 shows a stereo scene.

In addition to generating the graphics, the PC is responsible for interfacing with other computers and equipment in the test-bed. The PC is equipped with a high-speed serial interface board capable of simultaneous communication with haptic interface, eyetrackers, and Electrocardiograph systems. A special-purpose 12-bit analog-to-digital board reads steering, accelerator, and brake signals from the flying simulator.

Central to pilot-based virtual scene systems is the ability to keep the position and orientation of the pilot, in order for the graphics engine to render the appropriate scene given the current viewpoint.

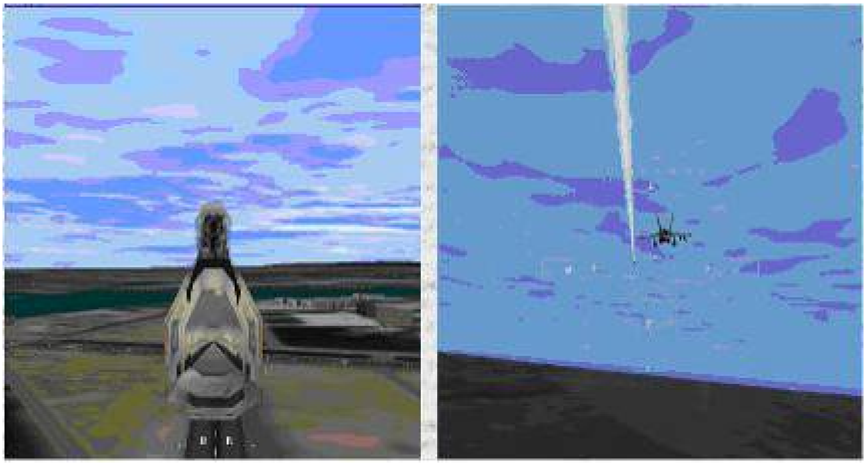


Fig. 2. Flying environment stereo scene

3 Driving Simulator and Motion Platform

3.1 Driving Simulator

One of the components of the virtual scene we use to study complex behavior is a flying simulator. The simulator is based on a frame in which the steering column has been instrumented with a low-noise rotary potentiometer and the accelerator and brake pedals are instrumented with linear potentiometers. A constant voltage supply across each potentiometer produces a variable voltage from each control that is read by a multi-channel sample-and-hold 12-bit A/D board on the bus.

Subject flight in sky, an extensible environment designed by Vega, Figure 2, shows a sample from the environment. We have added fighter planes and scenes to the environment that move along pre-defined or interactively controlled paths. These objects, as well as other planes can be placed in the environment through a configuration file that makes it easy to customize the environment. The visual environment is also easily modified to simulate lighting at any time of day or night, and the addition of fog.

3.2 Motion Platform

To aid our goal of making experiences in the virtual scene as close as possible to those in the real world, the flying simulator is mounted on a real plane cockpit platform. See figure 3. The platform, along with the scene, generates for the subjects the feeling of linear accelerations for heave, surge and sway, and angular accelerations for pitch, roll, and yaw. The platform virtual controller and shift are controlled by a host PC that is connected to the scene via serial interface. The host PC accepts acceleration values and calculates optimum linear and angular values for the visual scene subjects will view.

Fig. 3. Flying on a real plane cockpit platform

4 Experiment System of Instrument/Controller Layout (ESICL)

Due to different layout designs inducing different results of manipulation, layout and/or color of instrument, controller on the panel should be changed by the effect of ergonomics at any moment at the stage of designing. Experiment System of Instrument/controller Layout (ESICL) is a set of simulation systems that adopt computer technology. ESICL can make "Electronical instrument /controller devices system"(or soft instrument/controller panel), and thus the designer can control directly various outlines of panel, shapes of instrument, and layouts of controller on the screen by interactive technology. At the same time, the results of simulation experiment will be achieved by ESICL after designing.

5 Indicators of Cognitive State: Eyatracking

Central to the lab goal of studying complex behavior is the ability to monitor subjects' use of attention and other aspects of their cognitive state. Monitoring eye movements is a powerful tool in understanding behaviors[3]. In previous work with real tasks, we have shown that eye movements are a reliable metric of subjects use of attention and as a gauge of cognitive state[4]. Recorded eye movements provide an externally observable marker of attentional state and of strategies employed in solving complex, multi-step problems. Particularly important is the fact that eye movements can sometimes reveal low-level strategies that are not available to the subject's conscious perception. Complex tasks are often serialized into simpler sub-tasks that are then executed serially[5]. In some cases rapid task-swapping is observed between multiple subtasks[6]. While this is evident in the eye movement record, subjects' self-reports in such tasks reveal that they are unaware of the strategy[7].

Tracking subjects' eye movements in virtual scene/panels presents a special challenge. EMMS system employs infrared video-based eyetrackers that determine the point-of-gaze by extracting the center of the subject's pupil and the first-surface reflections from video fields. Tracking both pupil and first-surface reflections allows the image-processing algorithms to distinguish between eye-in-head movements and motion of the eyetracker with respect to the head. An infrared-emitting diode (IRED) is used to illuminate the eye.

Fig. 4. Subject flying with EMMS

The EMMS system calculates an eye position signal on each video field (i.e. 50 Hz). Single measurements can be read out at that rate or averaged over multiple fields to reduce signal noise. The system also digitizes the rendered image from the PC and overlays a cursor indicating the subject point-of-gaze on that field. In addition to the video record showing gaze on the virtual scene, the eye position signal from the EMMS system is sent via serial port to the PC. Accuracy of the eye position signal is approximately 1 degree in the central 20 degrees of vision and 2 degrees in the periphery.

The EMMS system is capable of tracking eye movements at rates greater than 50 Hz with special eye cameras whose vertical scanning can be controlled externally. By limiting the vertical scan to half the normal image height, two scans can be made in the time usually used for a single scan. The controllers analyze both images in each field, providing a 100Hz sample rate. The increased temporal frequency comes at the cost of vertical field. The higher temporal sampling rate is important for contingent display, while the eyetracker signal is used to initiate changes in the virtual scene/panels based on eye position or eye movements. See figure 4.

The DLL is called from within the Recognition and Biofeedback program. The Recognition and Biofeedback program chooses which data (in a continuous recording) need to be sent for further processing via the Matlab program. This program may send return information to the PC through a serial port interface; saves recognition data; calculates whether correct recognition has actually occurred (using available trigger codes); and may read previously processed data from a Matlab file for a demonstration of recognition speed. We use the Matlab program because it enables the rapid prototyping of algorithms for use in the system and because it allows the easy swapping of algorithms for recognition/feedback/analysis. All routines are Matlab M-files to enable easy use of different recognition routines.

6 Summary

The test-bed provides a valuable tool in understanding complex pilot behaviors because it offers a unique combination of flexibility and control. Virtual scene/panels generation and displays coupled with visual and haptic interfaces create a rich family of virtual environments for psychophysical experiments. The use of integrated eyetrackers and concurrent Electrocardiograph recording gives us a window into subjects' use of attention and their cognitive state as they complete complex, multi-step tasks.

Acknowledgments

The authors want to thank the financial support from China Postdoctor Science Research Fundation (grant Number 2004035033) and National Nature Science Fundation (grant Number 30400137) and National Nature Science Fundation (grant Number 60672182).

References

1. Pelz, J.B.: Visual Representations in a Natural Visuo-motor, Task Doctoral Dissertation, Brain and Cognitive Science, University of Rochester (1995)
2. Geisler, W.S., Perry, J.S.: A real-time foveated multi-resolution system for low-bandwidth video communication. In: Human Vision and Electronic Imaging. SPIE Proceedings 3299 (1998), pp. 294–305 (1998)
3. Ballard, D.H., Hayhoe, M.M., Pelz, J.B.: memory Representations in Natural Tasks. Journal of Cognitive Neuroscience 1, 68–82 (1995)
4. Duchowski, A.T.: Incorporating the Viewer's Point-Of-Regard (POR) in Gaze-Contingent Virtual Environments. In: The Engineering Reality of Virtual Reality, SPIE Proceedings 3295, pp. 332–343 (1998)
5. Land, M.F.: predictable eye-head coordination during driving. Nature 359, 318–320 (1992)
6. Bayliss, J.D., Ballard, D.H.: The Effects of Eyetracking in a VR Helmet on EEG Recording TR 685. University of Rochester National Resource Laboratory for the Study of Brain and Behavior (1998a)
7. Bayliss, J.D., Ballard, D.H.: Single Trial P300 Recognition in a Virtual Environment NRL T.R. 98.1. University of Rochester National Resource Laboratory for the Study of Brain and Behavior (1998b)
8. wei, L.: Development of a Test-Bed for Synthetical Ergonomics Evaluation of Pilot and Study on Situation Cognitive of Pilot[D]. Beijing University of Aeronautics, Astronautics (2003)

Experimental Research of Evaluation of Temperature Ergonomics of EVA Spacesuit Glove

Han Longzhu and Xiu-gan Yuan

Institute of Human-Machine & Environment,
Beihang University, 100083 Beijing, P.R. China
lzhan1978@163.com

Abstract. The EVA spacesuit glove is one of the key components which can assure astronauts to achieve EVA tasks in security. Astronauts are subjected to the influence of low temperature when they are performing tasks outside the spaceship. To establish the standard of the ergonomics of the temperature of the glove work is very important. This paper studies the influence of low temperature on the ergonomics of the glove work when the skin surface temperature of the middle fingertip is at normal temperature, 15.6℃ and 10℃ respectively, in which the strength, fatigue, sensibility and flexibility are made as evaluation targets. The experiment indicates: the influence on the ergonomics of glove is not obvious when the temperature is controlled at 15.6℃. Therefore the temperature of the middle fingertip skin should be controlled above 15.6℃ in glove design; the influence to the ergonomics of glove is relatively obvious when the temperature is about 10℃. The influence on sensibility is the most obvious, and then the influence on fatigue, strength and flexibility is in descending order. The temperature standards of EVA spacesuit glove have practical meaning to design EVA spacesuit.

Keywords: hand, EVA spacesuit glove, ergonomics, low temperature.

1 Introduction

The EVA (Extravehicular activities) space suit glove is the key component to accomplish astronaut's EVA work. The EVA space suit glove usually includes three layers: bladder、restrain layer and TMG（Thermal Micrometeoroid Garment).The glove must have the functions of thermal protection、resist micrometeoroid attack、gas seal and so on, which lead to negative impact on EVA work. So the EVA spacesuit glove ergonomics must be analyzed based on glove's ergonomics characteristics. And the contents of astronauts EVA are design which can make astronauts accomplish EVA work.

Astronauts' outside work can be affected because of the glove's insufficient thermal protection when they are in shadow areas, and when their hands are against sunlight or when they touch cold objects [1] [2]. Russia and America have the same problems, and their astronauts have had to return back to spaceship because of the glove's insufficient thermal protection. Therefore, it's very important to establish a temperature standard which could not affect ergonomics.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 929–938, 2007.

The study of EVA space suit glove's ergonomics was started very early. However, it was not systematically analyzed until 1988 when O'Hara' comprehensive summary appeared. He classified the evaluated standards of glove's ergonomics into six groups: strength、 fatigue、sensing、comfort、movement range and flexibility [3] .Because of the particular request to experiment environment of EVA space suit glove, thus, how to evaluate the ergonomics standards in a simple and effective way is becoming one of the study emphasis at present. After O'Hara, Ram R. Bishu evaluated the flexibility by knotting [4]; Mark S. G. evaluated the strength when different parts of the hand pressed objects [5]. At present, the acknowledged conclusions of the influence of low temperature to glove' ergonomics are: when the temperature of the middle fingertip skin surface is higher than or equal to 15.6℃, the movement of the hand isn't influenced; the temperature of 12.8-15℃ is thought as critical temperature ranges. When the temperature falls into lower than 8℃ or so, the sensibility of hand will be reduced badly and the probability of accidents will be increased greatly. When the temperature falls into 4.4℃, the accurate movement ability of hand will be nearly lost [6, 7]. The conclusions mentioned above are all based on the study of the ergonomics of the naked hand. When hands are put on an EVA space suit glove, the temperature field of the hands differs from the naked hands [8], especially when some parts of fingers are touching cold obejects9]. However, the research conclusions mentioned above were obtained in a simple and low temperature environment on the ground. And the actual space situations are not taken into consideration. Therefore the previous research conclusions can not be applied simply in designing EVA space suit glove.

Based on the research of the hand movement work ergonomics and hand-glove temperature field [8], integrated with the actual space situations, this paper experiments and researches the hand-glove system systemically by using stre-ngth、fatigue、sensing、comfort、movement range and flexibility as ergonomics standards when the temperatures of the middle fingertip skin surface are at normal temperature, 15.6 ℃ and 10℃ respectively and puts forward some new concepts. The systematic experiments aimed at EVA space suit glove technique requests have not reported yet.

2 Experiment

2.1 Experiment Design

Under low temperature, the ability of the different tissues of hands will reduce because of the coldness. As discussed above: 15.6℃ of the middle fingertip skin surface is the critical temperature of the hand movement work. When the temperature of the middle fingertip skin surface is 10℃, it is near to the critical temperature of the receptor. If the temperature continues to decrease, man will not tolerate any more work [4, 8, 10] due to the complexity of individual differences and physiology actions, it is difficult to control the temperature of the skin to a stable degree. Therefore this experiment chooses normal temperature (20-35℃), 15.6℃ and 10℃

for the temperature of the middle fingertip skin surface as experimental points. As these three points have little influence to the movement range and comfort of astronauts, but the experiment doesn't take these two factors into account, and merely focuses on the evaluations of strength, fatigue, sensing and flexibility. The glove used in the experiment is a self-made EVA space suit glove.

2.2 Experiment Equipments

This experiment includes the experiment of four standards of strength, fatigue, sensing and flexibility respectively. The main experiment equipments include:

(1) Experiment tank under low temperature

The main function of the experiment tank under low temperature is to supply the required low temperature environment of this experiment. See drawing 1. The compressed air is expanded and cooled down by passing through the turbine, and then the cool air enters into tank 1 via air inlet. By radiation and thermal conduction, the temperature of tank 2 can be cooled to the needed low temperature. There is a 20cm thermal protective coating on tank 1 to maintain its temperature. Tank 2 is settled into a dodge gate and can be moved by moving dodge gate.

(2) Video camera

See the fig. 1.a. A video camera is settled into the experiment tank under low temperature by a holder and linked to a computer external to the experiment tank by using wire via the wiring gate of the experiment tank. This monitors the situation in the experiment tank.

(3) Hand-dynamometer

The hand-dynamometer includes two parts: a dynamometric device and a reading device. The dynamometric device is in the experiment tank and the reading device is outside. They are linked by a wire, so the experiment data can be read by the reading device external to the experiment tank.

(4) Fatigue measurement device

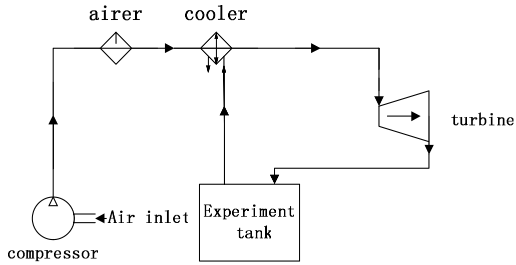
The fatigue measurement device is very much like a spring. There is a hand grip on every side of the spring. The level of hand fatigue can be evaluated by holding the hand grip and pressing the spring. There is a sheet of metal on the end of every side of the hand grip of the fatigue measurement device in the experiment tank. Two wires are used to link the sheet of metal to an indicating lamp and electronic source serially which are outside of the experiment tank. When the hand grips are held tightly, the two sheets of metal will connect, then the circuit will be completed and the lamp will light. This is a successful holding.

(5) Bolts and nuts of different diameters

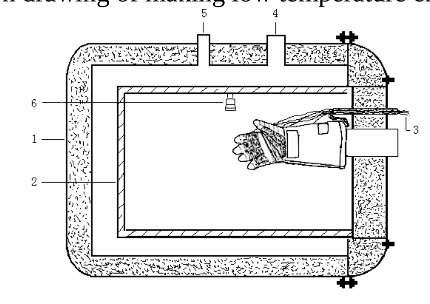
The bolts are fixed in the base plate and the nuts are screwed down the bolts. The bolts' diameters include: 16cm、12cm and 8cm. This device is used to measure the flexibility of the hand.

(6) strength sensing device

The strength sensing device is made up of a reading device and a sensing device. The sensing device is set in the experiment tan and is linked to the reading device outside of the experiment tank by a wire. There is a metal column on the sensing device, when the hand touches the upper side of the column, the strength pressed on



a. the skeleton drawing of making low temperature environment



1-tank1, 2-tank2, 3-wiring gate, 4-air outlet, 5-air inlet, 6-video camera
b. the skeleton drawing of experiment tank

Fig. 1. The simulated experiment tank

the upper side of the column can be transformed to electronic signals by strain foil.Then the data of contact force can be displayed by the reading device. The measurement range of the strength sensing device is from 1 to 1000 gram and the measurement accuracy is 1 gram.

(7) Temperature monitor system

In order to monitor the temperature of the middle fingertip skin surface and assure it in the needed value real-time, brand Tianchen temperature measuring set is used as temperature monitor system.

(8) EVA space suit glove for experiment

The EVA space suit glove for experiment is made by simulating the actual EVA space suit glove, which is made up of gas seal, restrain layer and thermal micrometeoroid garment (TMG). There is a screw style bayonet at the end of glove, by which the glove is linked to the experiment tank.

2.3 Experimented Persons

Six healthy volunteers from Beihang University asisted this glove ergonomics experiment under low temperature. See the table 1 of their detailed parameters.

Table 1. The detailed parameters of experimented persons

experimented persons	A	B	C	D	E	F
sex	male	male	male	male	male	male
age	32	25	28	22	28	25
Height/cm	175	181	170	165	175	172
Weight/kg	65	67	85	55	62	65
Health	excellent	excellent	excellent	excellent	excellent	excellent

2.4 Experiment Method

The persons in the experiment sit on chair quietly, put their hands into the experimental tank and put on the glove which is fixed in the experimental tank. The temperature of the middle fingertip skin surface of them should be control at 15.6℃ or 10℃ by adjusting the airflow from turbine air inlet (it takes about 2-2.5 hours). Ten minutes after the temperature become stable; the experimental subjects stand up and begin to do the ergonomics experiments. Before the next experiment, one or two days' rest for experimental subjects is necessary. The ergonomics experiments for the naked hand are done in the same way outside of the experiment tank under normal temperature

2.4.1 Strength Experiment

Before every experiment, the hand-dynamometer should be zero clearing. We ask the experimental subjects to hold the hand grip of the dynamometric device in the experiment tank with maximum strength, and then record the data in the reading device outside of the experiment tank. After recording the data, zero clearing should be done. In every experiment, the experimental subjects should hold the hand grip three times, and the interval is one minute.

2.4.2 Fatigue Experiment

When experimental subjects hold the hand grip of the fatigue measurement device in the experiment tank tightly, the two sheets of metal on the end of every side of the hand grip will connect. Then the circuit will be made current. After the lamp lights up, the experimental subjects loosens the hand grip. The experimental subjects repeat action above continuously until they can not make the lamp light in two seconds. Record the times of lamp lights.

2.4.3 Sensing Experiment

The experimental subjects should touch the upper side of the column in the sensing device in the experiment tank with their middle finger lightly, and then move their fingers away until they confirm that they have touched it. We can record the strength data in the reading device out of the experiment tank and make it zero clear

2.4.4 Flexibility Experiment

When given starting signal, the experimental subjects set out to remove the nuts which are screwed down the bolts from the bolts. After removing, the nuts turn back quickly and then we can record the times of the entire operation. The previous process should be repeated twice.

3 Experiment Analyses

Evaluating the influence of different temperature to ergonomics by individual comparison is vague (See fig. 2). The average data can avoid the interfering of individual differences and random errors (See fig. 3). Hereinafter the comparisons use the average of the data.

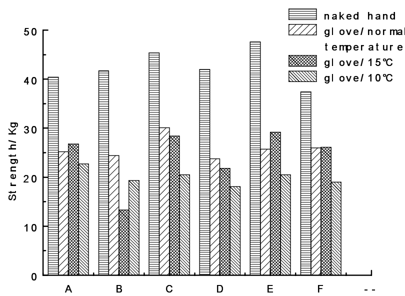


Fig. 2. The maximum grip strength comparison of different experimented persons

3.1 Strength

According to the table 2, the grip strength in the glove falls down 41% compared with that of a naked hand. This result is nearly in accordant with the 45% of reference [3]. When the temperature of the middle fingertip skin surface is 15.6℃, the influence to strength is not obvious. Comparing with normal temperature, the strength only reduced 7%.

Table2. The ergonomics data of EVA space suit glove

	grip strength/kg	fatigue/time	screw the nut/s	sensing/g
naked hand	44.13	159	5.53	3.2
in glove/normal temperature	25.88	73	22.7	15.8
in glove/15.6℃	24.1	52	30.4	16.5
in glove/10℃	20.0	42	27.9	28.8

Note: all the data is average.

When the temperature of the middle fingertip skin surface is 10℃, the influence is relatively obvious; the strength reduces 22.7% comparing with normal temperature (See fig. 3). Although at this time the temperature of muscle tendon cannot be

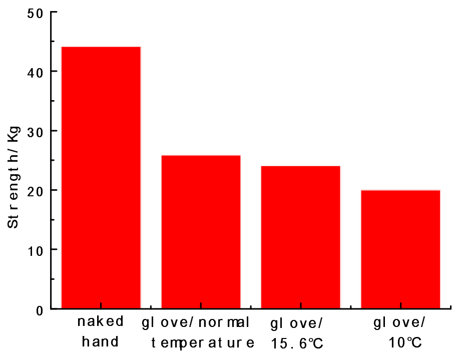


Fig. 3. The comparison of maximum grip strength

experimented, according to reference [8]: under the 10℃ of the middle fingertip skin surface, the temperature of muscle tendon will be below 28℃ and the influence to strength is obvious.

3.2 Fatigue

Compared with naked hand, the number of successful holding of the fatigue measurement device in glove reduces 54% (see fig. 4 and table 2), lowers 10% than the reference [3]. The reason is that the bladder of the experimented glove is relative thick, which influences the repeated operation. When the temperature of the middle fingertip skin surface is about 15.6℃, the ergonomics reduces 29% compared with normal temperature condition. When the temperature of the middle fingertip skin surface is about 10℃, the ergonomics lowers 43% than normal temperature condition. Therefore, lower temperature of the fingertip cast greater influence on the ergonomics. The reason is because of the influence of low temperature to hand strength, and then the fatigue of hand gets greater accordingly.

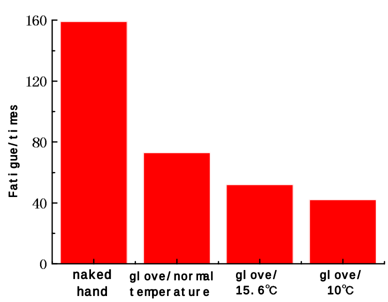


Fig. 4. The comparison of fatigue

3.3 Sensing

Comparing with the naked hand condition, the hand sensing of strength is higher by 394% in glove (see fig. 5 and table2). When the temperature of the middle fingertip skin surface is about 15.6℃, the influence on hand strength sensing is not obvious, which is higher by 4% than normal temperature condition in glove. When the temperature of the middle fingertip skin surface is about 10℃, it's near to the critical temperature of receptor [10]. In practical experiment, the influence is obvious, which is higher by 82% than normal temperature condition in glove. This is aroused by CIVD (cold induced vasodilatation) because of the coldness before controlling the temperature of the finger at 10℃. CIVD makes the temperature of the finger fluctuate and then influences the sensing of finger. In addition, when the temperature of the finger is under 10℃, the influence to the receptor is greater.

3.4 Flexibility

Compared with the naked hand, the time of screwing nuts in the glove increased by 310% (sees fig. 6 and table 2). According to reference [1], the time increased by 163%. There are two main reasons to explain why the experiment results are so different compared with the reference conclusions:

(1) There are differences in materials and constructions between the experimented EVA space suit glove and the glove of reference [3] .This lead to the flexibility of the hand of this experiment become worse than that of reference [3];

(2)There is no non-slip mat in experimented glove, which will influence the work.

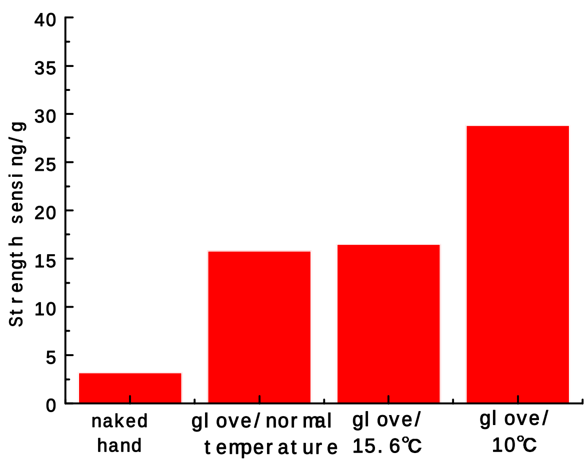


Fig. 5. The comparison of strength sensing

When the temperature of the middle fingertip skin surface is about 15.6℃, the ergonomics of glove reduces by 33.9%. Under 10℃ of the temperature of the middle fingertip skin surface, the ergonomics of the glove lowers by 22.9% compared with

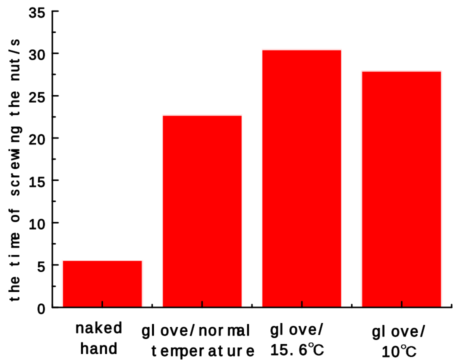


Fig. 6. The comparison of flexibility

normal temperature condition. By the analysis of experiment results, it can be concluded that the influence of the temperature on the hand flexibility is not obvious because of the great influence of glove to hand flexibility. When the temperature of the finger is about 10℃, the flexibility is better on the contrary.

4 Conclusion

According to the analysis above, some conclusions can be obtained as the following:

(1) To control the temperature of the middle fingertip skin surface at 15.6℃ can be a standard for designing EVA space suit glove's lowest temperature. And it inquires that the liquid cooling and ventilation system of EVA spacesuit glove should insures the temperature of the middle fingertip skin surface of glove's internal environment is about 15.6℃

(2) When the temperature of the middle fingertip skin surface is about 10℃, hands can endure the coldness, but this is not permitted to the entire EVA spacesuit internal environment. Because the reduction of ergonomics of the hands is relatively obvious and the veil of the helmet of EVA spacesuit has been fogged.

References

1. Wilson, K.T.: Records of EVA, SPACEFLIGHT[J] 40 (March 1998)
2. Liwei, L.: The record: Space shuttle Discovery was connected with International space station Peace [N], Reference Information, 1995 -2-13 (7)
3. O'Hara, J.M. Brigann, M., Cleland, J. and Winfield, K.: Extravehicular activities limitations study volume II establishment of physiological and performance criteria for EVA gloves-final report[R]. AS-EVALS-FR-8701. NASA: NAS-9-17702.
4. Cheng, P., Jingshan, C.: Medicine and engineering problem of EVA spacesuit temperature[M], In-house data (1999)
5. Ram, R.B., Byungjoon, K., Glen, K.: Force-endurance relationship:does it matter if gloves are donned[J]. Applied Ergonomics 26(3) (1995)

6. Grahne, M.S., Graziosi, D.A., Pauly, R.L.: Benefits of an EMU glove active heating system[R]. SAE 951549 (1995)
7. Fox, W.F.: Human performance in the cold[J]. Human factors 9(3), 203–220 (1967)
8. Li, D.: The study of thermal protection of EVA spacesuit glove[D],An dissertation for Ph.D. degree of Beihang University (2000)
9. Qiuqing, G., Ingvar, H., et al.: Change in contact temperature of finger touching on cold surfaces[J]. International Journal of Industrial Ergonomics 27, 387–391 (2001)
10. Heus, R., Daanen, H.A.M, Havenith, G.: Physiological criteria for functioning of hands in the cold[J]. Appl. Ergon. 26, 5–13 (1995)

Digital Humans for Virtual Assembly Evaluation

Dimitris Mavrikios, Menelaos Pappas, Marios Kotsonis, Vassiliki Karabatsou, and George Chryssolouris*

Laboratory for Manufacturing Systems and Automation, Department of Mechanical Engineering and Aeronautics, University of Patras, Patras 26500, Greece
gchrys@hol.gr, xrisol@lms.mech.upatras.gr

Abstract. This paper studies the use of Virtual Reality and Human Simulation for the ergonomic evaluation of manual assembly processes. A virtual environment has been developed to represent the actual workspace where the assembly task took place. Into the virtual workspace, a digital human /mannequin was imported and programmed to simulate the task, in the same manner as it would be done by the actual worker. Based on the posture-based ergonomics analysis, each posture of the digital human has been "translated" into comfort scores, resulting in conclusions, related to the ergonomic efficiency of the process and in the design of the workstation. The conclusions that have been reached identify the critical points, during the assembly task, and lead to the necessary re-design actions in order for the worker's fatigue as well as the task's execution time to be reduced. A real-life assembly task of a commercial refrigerator has been implemented in order for the capabilities of the proposed process design evaluation method to be demonstrated.

Keywords: Virtual Reality, Digital Humans, Virtual Assembly, Process Design Evaluation, Ergonomics Analysis.

1 Introduction

The minimization of the product cycle time and the fatigue of workers are factors of significant importance [1]. Assembly processing is a part of the manufacturing line that needs to be taken under special consideration, due to the large involvement of human workers. It is therefore essential that a complete ergonomic analysis of the assembly processes be incorporated and its results be used in order for the previously mentioned factors of importance to be improved [2], [3].

The traditional method of ergonomic analysis was the implementation of complete physical mockups of the product, the assembly workspace and the real workers used, in different combinations, in order for the ergonomically optimum set-up to be produced. This was both time consuming, and costly; characteristics that do not conform to the goals of modern industry.

A fast developing method that meets these demands is the combined use of Virtual Reality (VR) and Human Modeling software tools for the ergonomic analysis and optimization of assembly processes [4].

* Corresponding author.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 939–948, 2007.

These technologies enable the integration of the human being's presence and actions into the virtual environment in the form of digital humans (digital mannequins). The ergonomic analyst is provided with a powerful tool, due to the flexibility that such a system can exhibit. The entire analysis is based on a 3D interactive environment supplemented by virtual humans that can be either digital mannequins or immersed humans. Numerous set-ups of the workspace can be tested with only a fraction of the time and cost required by such tasks if done in the traditional way. Furthermore, the quantification and measurement of the ergonomics factors are now more feasible, exact and objective.

Due to the advantages, provided by this type of tools, and considering that the VR technology continues moving towards maturity, an entire new market has emerged offering solutions, in the form of software and hardware. There are also complete analysis platforms with embedded databases that can be directly used for the ergonomic analysis of any given process. Based on these tools, the scientific community has already a number of studies [5-12].

In this paper, a virtual environment has been created and used for the ergonomic analysis of an assembly process. The entire workplace has turned into a virtual environment and a digital mannequin was also integrated so as for the worker, to be simulated. The comfort scores have been used in order to identify any uncomfortable postures and proper modifications are then made for their elimination.

2 Virtual Environment

2.1 Application Environment

The application environment has been created with the use of the Division MockUp 2000i2. This is a Virtual Reality interactive environment developing platform. The Division SafeWork tool has been used for the incorporation of digital humans. The SafeWork provides a wide variety of options regarding the anthropometric data, used by each mannequin. There are six default body types consisting of the 5^{th}, 50^{th} and 95^{th} percentile of the global population, male and female. For more accurate representation of the desired body type, an integrated tool, called Human Builder, can adjust 104 anthropometric variables to the virtual mannequin. The platform runs on an SGI Onyx2 InfiniteReality2 RK Workstation. All geometries employed into the virtual workspace have been created with the use of the PTC's ProEngineer Wildfire2.

2.2 Application Scenario

The simulation has been conducted for an assembly task in the production line of a commercial manufacturer. The primary reasons for choosing the specific workstation were those of fatigue and large time consumption reported by the worker. The investigation was primarily focused on the worker's postures and their impact on the ergonomically oriented factors, throughout the duration of the task. Therefore, the main objective of this work is that the ergonomically critical factors, involved in the assembly task, be identified and the necessary measures be taken in order for them to be reduced.

a b

Fig. 1. Real and simulated assembly task

Generally, this study aims at transferring this task to a controlled environment, and experiment on the current set-up in order for its ergonomically related factors to be quantified. This should minimize the pre-set "point factors", in our case, being the task completion time and the worker's fatigue.

The task to be analyzed consists of the positioning of three shelves supporting the beams on the refrigerator's inner rear wall. Fig. 1a shows a snapshot taken from the video of a real assembly process, while Fig. 1b shows the respective representation of the task in a virtual environment.

Based on the video data of the actual process, the assembly task was divided into five "key frames" chosen to represent the most characteristic and stressful postures, taken by the worker. These five postures were independently analyzed and a posture score was derived for each one. The five postures were as follows:

- *Posture 1*: This is the position the worker takes in order to pick the beams from the lot. The worker is bending forward in order to acquire the proper position. Both hands are bended. His legs are also slightly bended (Fig. 2a).
- *Posture 2*: This is the position the worker takes for placing the two side beams. The worker's right arm is fairly stretched in order to reach the far-off side of the refrigerator, while his left arm is bended. His spine is slightly bended and his legs are straight (Fig. 2b).
- *Posture 3*: This is the worker's position while driving the screw in the right end of the side beam. The worker is leaning over the edge of the refrigerator whereas his right hand is holding the electric screwdriver and his left hand is supporting the screw. His spine is bended not only to the front but also sideways (Fig. 2c).
- *Posture 4*: This is the posture the worker takes while positioning one of the middle screws of the side beams. The worker's back is quite bended, his legs are stretched and his hands have acquired the appropriate position. This gives him leverage in order to use his own weight to speed up the screwing (Fig. 2d).
- *Posture 5*: This is the posture the worker takes for positioning one of the screws in the middle beam. The worker is bending forward while his right hand is raised in order to hold the electric screwdriver more effectively. He uses his left arm and hand as a support against the inner back wall of the refrigerator. His legs are straight (Fig. 2e).

Fig. 2. Key postures of the real assembly task execution

The key frames were transferred into the virtual workspace and analyzed with the appropriate ergonomics functions in the way they will be described hereunder.

3 Ergonomic Analysis

3.1 Digital Human

The Division SafeWork has been used for the simulation of the digital human. This tool offers a wide range of anthropometric data in order to accurately represent any given body type. In this case, the databases were selected to represent the actual worker, performing the assembly task, as accurately as possible. The general body type was selected as the representative of the 50^{th} percentile of the general population. The Division MockUp 2000i2, in conjunction with the SafeWork, uses a type of Inverse Kinematics technology in order to provide the digital humans with a more realistic behavior, in terms of movements and postures. This was used in order for the digital human to be accurately placed in the same 5 key postures the worker took during the execution of the assembly task (Fig. 3).

3.2 Posture Analysis

The Division MockUp 2000i2 incorporates a number of analysis functions for several types of tasks. These include posture analysis, carry analysis, push/pull analysis and

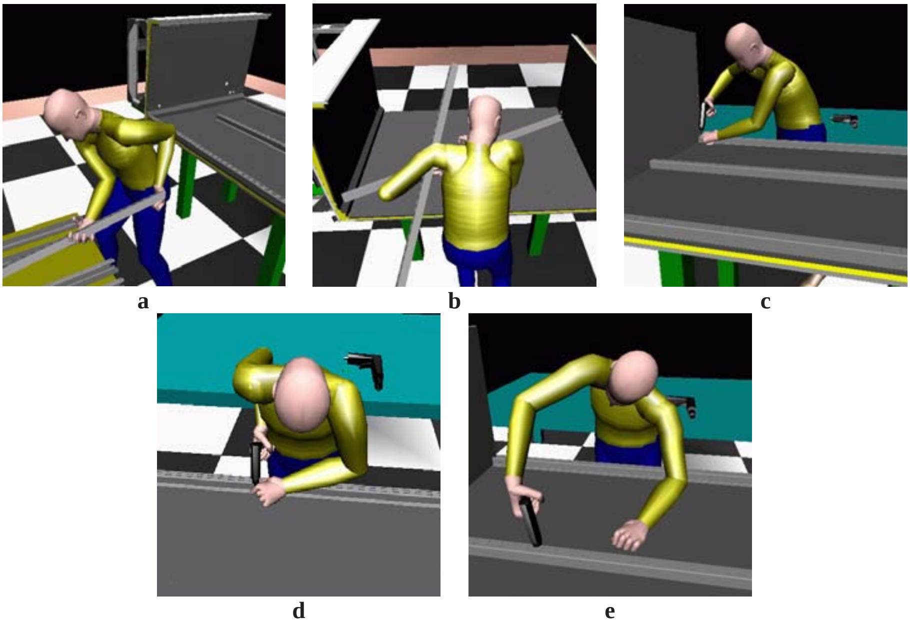


Fig. 3. Key postures of the simulated assembly task execution

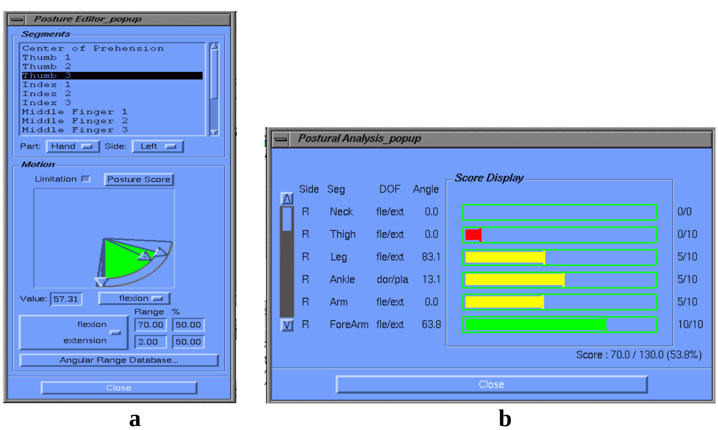


Fig. 4. Posture analysis input and output dialog boxes

lift/lower analysis. As this study has been mainly oriented on the posture induced fatigue, the posture analysis has been selected for the ergonomic evaluation of the task. This was performed for the five key postures previously selected.

The posture analysis function communicates with several databases to acquire the angle limitations for the various limbs of the digital human. In the current work, the GENICOM database, provided via the SafeWork, was used. This function reads the angles of the limbs of the digital human and compares them with the optimum values, found in the database (Fig. 4a).

It produces the results as output, in a graphical way, by assigning a comfort score and a characteristically colored bar to each limb and degree of freedom (DoF) (Fig. 4b). Furthermore, the function sums up the comfort scores of the limbs thus, produces an overall comfort score for the current digital human posture.

The SafeWork provides options for selecting a great number of body parts (limbs) for simulation. The selection covers parts from individual back spine bones to fingers and toes. For the needs of the current study the selected limbs and DoF were (Fig. 5):

- neck in flexion and extension
- thigh (left and right) in abduction and adduction
- thigh (left and right) in flexion and extension
- leg (left and right) in flexion and extension
- arm (left and right) in flexion and extension
- arm (left and right) in abduction and adduction
- forearm (left and right) in flexion and extension
- centre of prehension [wrist] (left and right) in flexion and extension

Based on the GENICOM database and the angles of each limb, the posture analysis function gives a comfort score for each individual part and its respective DoF. This has been done for each one of the 5 key postures. An example of the collected comfort scores is given in Table 1.

As previously mentioned, the posture analysis function can sum up the comfort scores of each limb in order for an overall comfort score of the specific posture to be produced. This was done for each one of the five key postures for constructing the comfort distribution, throughout the length of the entire assembly task (Fig. 6).

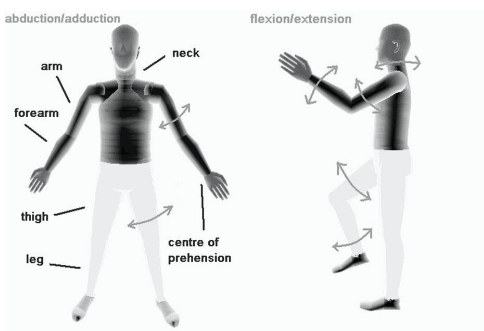


Fig. 5. The digital human and its DoF

3.3 Results

Fig. 6 provides an overall view of the task's comfort score distribution. Careful observation of the distribution, can lead to significant conclusions towards the

Table 1. Comfort Scores for key posture 1

Part	Side	DoF	Score	Part	Side	DoF	Score
Neck	-	fle/ext	10/10	-	-	-	-
Thigh	right	fle/ext	0/10	Thigh	left	fle/ext	0/10
Thigh	right	abd/add	10/10	Thigh	left	abd/add	7/10
Leg	right	fle/ext	2/10	Leg	left	fle/ext	2/10
Arm	right	fle/ext	5/10	Arm	left	fle/ext	0/10
Arm	right	abd/add	0/10	Arm	left	abd/add	0/10
Forearm	right	fle/ext	7/10	Forearm	left	fle/ext	5/10
Wrist	right	uln/rad	5/10	Wrist	left	uln/rad	0/10

ergonomic optimization of the task. It is obvious that the lowest comfort score appears in the key posture 4. This is indicative that the efforts for ergonomical improvement of the task must be focused on the specific posture. By identifying the factors that cause such low comfort scores, on posture 4, they can be eliminated thus the task can be improved.

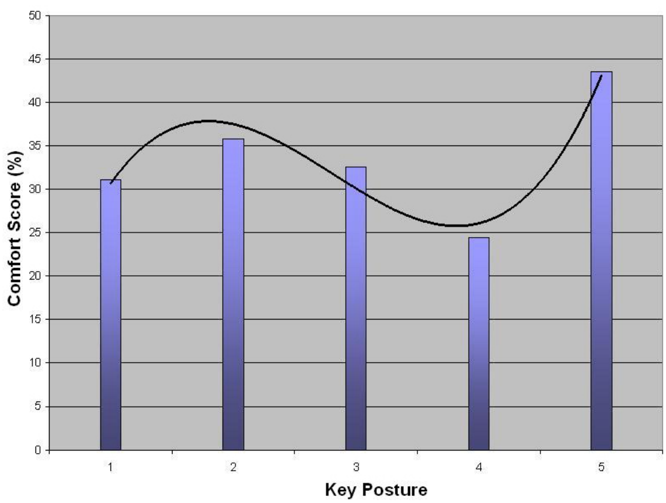


Fig. 6. Comfort score distribution for the several key posture of the studied assembly task

3.4 Optimization

It is obvious, in Fig. 6, that posture 4 is the most stressful for the worker. This could be logically explained by observing the “cramped” position the worker has taken in his effort to “push” the electric screw driver hard enough so as to speed up the process. It is also logical to assume that if the worker stood higher than he does in his current position, he could possibly attain a more comfortable posture.

Based on the previous rough conclusion, an investigation for a layout variant of the workspace has taken place. This includes a standing palette that elevates the worker’s position by about 13 centimeters. After the proposed workplace design modification, a detailed ergonomic analysis was performed, similar to the one presented in the previous sections. The overall comparison between the results of the two studies (analyses) is presented in Fig.7. Different views of the modified workplace (including the standing palette) are depicted in Fig. 8, in which it is obvious that the body posture of the worker has been quantitatively improved (from the ergonomics point of view) in comparison with the initial situation.

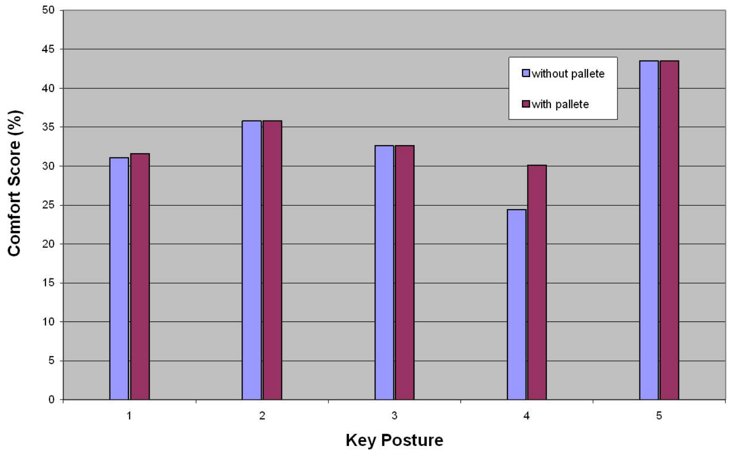


Fig. 7. Comparison of comfort scores before and after the workplace design modification

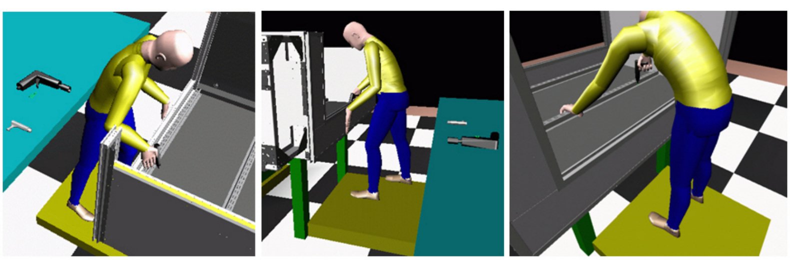


Fig. 8. Different views of the assembly workplace after the proposed design modification

4 Discussion

The traditional workplace design and redesign, based on the results of ergonomics evaluations, are time-consuming and often unfeasible in industry. The computer-aided workplace design simplifies the analysis of postures and allows for a far greater level of detail to be incorporated into the analysis. Several such techniques have been developed including a computer-aided video analysis and a postural analysis software tool for use with a man-model [13, 14]. Moreover, the use of commercially available ergonomics tools, integrated with a VE, provides significantly greater capabilities and they should be used where detailed and complex ergonomics evaluations are required [15].

This research work has examined the use of Virtual Reality and Human Simulation for the ergonomic evaluation of manual assembly tasks. The method proposed, has been validated through a real-life scenario, based on an industry's assembly process of commercial refrigerators. It is fairly generalized thus, making any other ergonomic analysis application feasible. For this analysis, a virtual environment, based on data (e.g. drawings, videos) from the real workspace, needs to be created. By manually adjusting the body posture of a digital human (mannequin) to numerous postures, respecting those that the actual worker took, during the execution of the task, the representation of the entire assembly process in the virtual environment is almost "real". With the use of posture analysis, the body postures are "translated" into comfort scores, which are indicative of the problematic areas of the task. In this specific case, a considerably low comfort posture (key posture 4) has led to the identification of the factor that causes the problem and furthermore, to the necessary redesign actions, required for eliminating it. This has optimized the entire task from the ergonomic point of view.

The time required for one to perform ergonomics analyses of the workplaces has been shortened by using the technique that has been presented. It is also possible for more workplace configurations and design changes to be considered, while the workplace is still in the virtual design stage. The risk of the ergonomics problems that will be occurring later on is therefore reduced since it forms a basic part of an integrated process, required for a concurrent engineering approach of the system design in order for it to work.

References

1. Chryssolouris, G.: Manufacturing Systems: Theory and Practice, 2nd edn. Springer, New York (2006)
2. Chryssolouris, G., Karabatsou, V., Kapetanaki, G.: Virtual Reality and Human Simulation for Manufacturing. CIRP Journal of Manufacturing Systems 32(6) (2003)
3. Chryssolouris, G., Mavrikios, D., Fragos, D., Karabatsou, V.: A virtual reality-based experimentation environment for the verification of human-related factors in assembly processes. Robotics and Computer Integrated Manufacturing 16(4), 267–276 (2000)
4. Chryssolouris, G., Mavrikios, D., Fragos, D., Karabatsou, V., Alexopoulos, K.: A hybrid approach to the verification and analysis of assembly and maintenance processes using Virtual Reality and Digital Mannequin technologies. In: Nee, A.Y.C., Ong, S.K. (eds.) Virtual Reality and Augmented Reality Applications in Manufacturing, Springer, London (2004)
5. Feyen, R., Liu, Y., Chaffin, D., Jimmerson, G., Joseph, B.: Computer-aided ergonomics: a case study of incorporating ergonomics analyses into workplace design. Applied Ergonomics Magazine 31(3), 291–300 (2000)
6. Gomes deSá, A., Zachmann, G.: Virtual reality as a tool for verification of assembly and maintenance processes. Computers and Graphics 23(3), 389–403 (1999)
7. Lu, Y., Shpitalni, M., Gadh, R.: Virtual and Augmented Reality Technologies for Product Realization. Keynote Paper. Annals of the CIRP 48(2), 471–494 (1999)
8. Ottosson, S.: Virtual reality in the product development process. Journal of Engineering Design 13(2), 159–172 (2002)

9. Rajan, V., Sivasubramanian, K., Fernandez, J.E.: Accessibility and ergonomic analysis of assembly product and jig designs. International Journal of Industrial Ergonomics 23(5-6), 473–487 (1999)
10. Sundin, A., Medbo, L.: Computer Visualization and Participatory Ergonomics as Methods in Workplace Design. Human Factors and Ergonomics in Manufacturing 13(1), 1–17 (2003)
11. Allen, C., Karam, K.Z., Le Cam, P., Hill, M., Tindle, T.: Application of virtual reality devices to the quantitative assessment of manual assembly forces in a factory environment. Industrial Electronics. In: Proceedings of the 1995 IEEE IECON 21st International Conference on Control and Instrumentation, vol. 2, pp. 1048–1053 (1995)
12. Gill, S.A., Ruddle, R.A.: Using virtual humans to solve real ergonomic design problems. In: Proceedings of the International Conference on Simulation 1998, pp. 223–229 (1998)
13. Carey, E.J., Gallwey, T.J.: Evaluation of human postures with computer aids and virtual workplace designs. International Journal of Production Research 40(4), 825–843 (2002)
14. De Looze, M.P., Van Rhijn, J.W., Van Deursen, J., Tuinzaad, G.H., Reijneveld, C.N.: A participatory and integrative approach to improve productivity and ergonomics in assembly. Production Planning & Control 14(2), 174–181 (2003)
15. Jayaram, U., Jayaram, S., Shaikh, I., Kim, Y.J., Palmer, C.: Introducing quantitative analysis methods into virtual environments for real-time and continuous ergonomic evaluations. Computers in Industry 57(3), 283–296 (2006)

Foot Digitalization for Last Design and Individual Awareness of Personal Foot Characteristics

Paolo Olivato[1], Manuela Morricone[1], Enrica Fubini[2], and Alessandra Re[1]

[1] Department of Psychology, University of Turin, Via Verdi 10, 10124 Turin, Italy
[2] Department of Animal and Human Biology, University of Turin, Via Accademia Albertina 13, 10123 Turin, Italy
olivato@psych.unito.it

Abstract. The aim of the present work was to gather data concerning the measurements and the shape of individuals' feet of Italians, to design more ergonomic shoe lasts. The unloaded feet of a sample of 316 participants of Italian ethnicity of 20 different towns located in northern, central, and southern Italy and the Italian islands were digitalized. Results showed significant differences by gender, age, and geographical area. The sample's degree of awareness concerning some of their own foot characteristics and interaction with footwear were also investigated. These findings will be used to design footwear that can provide a good fit for the feet of Italians and, consequently, a greater degree of comfort.

1 Introduction

The issue of good shoe fit was posed as early as 1500 B.C. in "Ebers Papyrus", which described a wide range of illnesses resulting from poor foot-shoe interaction. Since then, the question has always gone begging: "Does the shoe fit or give us fit?" [1]. In fact, in many studies, the relation between foot shape and shoe shape is considered a cause of discomfort, foot problems, or even injury [2][3][4][5]. In the 1980's, it was noted that "Despite the emphasis in society and the industry on fashion, we found that problems of size and fit, which are closely interrelated, have far more importance to man and women together than fashion or price and quality"[6]. As a result, there has been a growing necessity for sound, 3D knowledge of feet as a starting point for the design of ergonomic footwear. In fact, various nations have launched foot measurement projects to this specific purpose [7][8][9], and in Italy, two anthropometric surveys with ergonomic aims have been conducted to date — i.e., "*Ente Moda"*[10] and *"L'Italia si misura"[*11]). Both projects did not specifically concern feet, and only foot length and breadth measurements were taken. The present work, conversely, was aimed at acquiring three-dimensional morphometric information on the feet of a sample of individuals of Italian origin, to improve the design of shoe lasts. The study was aimed at defining a set of 3D foot models correlated with different consumer categories subdivided into gender, age, and geographic area of origin. A structured interview was also administered to the same sample to gather data on participants' degree of awareness about their own feet, perceptions, shoes preferences, and buying and use behaviours.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 949–958, 2007.

2 Methods

2.1 The Sample

The study was conducted in 19 mid- to mid-high market footwear shops, on a sample of 316 participants of Italian ethnic origin and nationality, geographically distributed as follows: 106 in Northern Italy, 64 in Central Italy, 71 in Southern Italy, and 75 on the main Italian islands of Sardinia and Sicily. Two shops were selected for each regional survey area: one in either a regional or provincial capital and one in either a small town or village. The sample was selected such that no participant presented pathological foot deformation(s).

2.2 Foot Digitalization

A contact digitalization instrument (3Space Fastrack, Polhemus) was used to measure the participants' feet. By moving a stylus in a transmitter-generated electromagnetic field, researcher sampled 3D foot coordinates with a 120 Hz sample frequency and transferred them in real time to a computer. Measured instrument error was below 0.7 mm for each measurement condition.

Measurement was conducted with the participant seated and wearing approximately 0.6 mm-thick cotton socks. A specific foot board had been set up to receive the participant's foot, in such a way that it could naturally lay at an approximately 100° angle (Fig. 1), and the leg was supported at a fixed height of approximately 50 cm, in extension with the aligned foot. Each foot digitalization lasted approximately 30 seconds.

Fig. 1. The leg support was built of wood with no metal parts (e.g., nails or screws), so as to not interfere with the transmitter's electromagnetic field

The researcher first moved the stylus along the foot board to fix the reference plane and then along the foot lines defined by the morphological model represented in Fig. 2. The procedure was conducted by two different researchers on the entire sample.

The ICC, Intraclass Correlation Coefficient (two way, mixed), for all reported measurements was > 0.8.

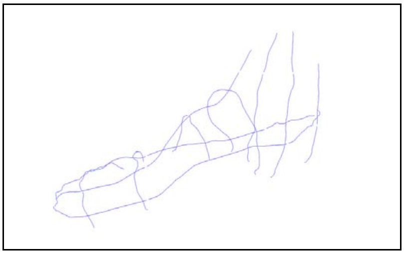


Fig. 2. The model's digitalization lines were made up of 10 points/mm. To trace lines, the stylus was held as perpendicular as possible to the tangential plane to the foot contact point.

The literature reports various types of models to quantify foot shape and size [8][12]. The present work used a model defined by nine digitalization lines: perimetrical, dorsal, posterior, hind-medial, hind-lateral, pivot, navicular, metatarsal and toes line (Fig. 2).

In agreement with Kouchi's [13] model, feet were aligned in the CAD work space along the longitudinal axis passing through the Pternion (the hindmost foot point, set on the posterior line) and the Akropodion (the fore extremity of the II toe, placed on the perimetrical line). A software script was then implemented for the automatic recognition of 11 reference points, as defined in Table 1 here below.

Table 1. Description of automatic recognition criteria of repere points set along reference digitalization lines

n	Point	Reference line	Recognition criterion
1	Pternion	Posterior	The hindmost point
2	Sphyrion tibiale	Hind-medial	The point whose projection on the horizontal plane is the most medially distant from the longitudinal foot axis.
3	Sphyrion fibulare	Hind-lateral	The point whose projection on the horizontal plane is the most laterally distant from the longitudinal foot axis.
4	Medial calcaneous	Perimetrical	The vertical projection of the medial Sphyrion
5	Lateral calcaneous	Perimetrical	The vertical projection of the lateral Sphyrion
6	Pivot	Dorsal	The point at the same height as the Sphyrion tibiale
7	Dorsal	Dorsal	The point of highest convexity (smallest curvature ray) in the metatarsal part of the dorsal line.
8	I Metatarsal	Perimetrical	The point whose projection on the horizontal plane is the most medially distant from the longitudinal foot axis.
9	V Metatarsal	Perimetrical	The point whose projection on the horizontal plane is the most laterally distant from the longitudinal foot axis.
10	Toes	Toes	The highest point.
11	Akropodion	Perimetrical	The most anterior point.

Table 2. All measures were calculated on points projections on the specific reference plane. The points were projected on the horizontal plane for length and breadth measurements, and on the sagittal plane for breadth measurements.

n	Size name	Points, line, or plane involved	Size definition
1	Foot length	Pternion - Akropodion	Distance between the two points on the longitudinal axis
2	Foot breadth	I Metatarsal – V Metatarsal	Distance between the two points on the transversal axis
3	Heel breadth	Medial calcaneous - Lateral calcaneous	Distance between the two points on the transversal axis
4	Malleolus breadth	Sphyrion tibiale - Sphyrion fibulare	Distance between the two points on the transversal axis
5	Medial foot breadth	I Metatarsal – Longitudinal axis	Distance between the point and the transversal axis
6	Lateral foot breadth	V Metatarsal – Longitudinal Axis	Distance between the point and the transversal axis
7	Metatarsal height	Metatarsal line	Heighest point of the metatarsalline
8	Forefoot height	Dorsal line	Mean height of the dorsal line part between Akropodion and Pivot
9	Toe height	Toe line	Heighest point of the toe line

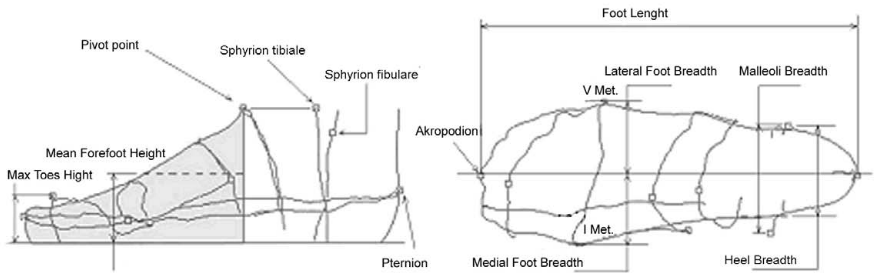


Fig. 3. The mean value of forefoot height was calculated taking into account the gray area shown in the figure. Foot breadth was calculated as the sum of medial foot and lateral foot breadths.

The x,y,z coordinates of the 11 points were first verified by the researcher, and then the data were processed to calculate foot size, as described in Table 2 and represented in Figure 3.

2.3 The Structured Interview

The interview was based on 15 items grouped into four sections: personal data (including region of origin and nationality), self-rating foot evaluation, and problematic aspects concerning foot-shoe interaction (type of preferred shoes, reasons for any footwear-induced pain, frequency with which problematic purchase scenarios occur, etc.) The interview was administered before foot digitalization.

3 Results

3.1 Digitalization Results

Anthropometric data on foot size were obtained through the above-described procedures and statistically processed by ANOVA to identify any differences due to gender, age, and region of origin. Processing was conducted both on the simple measurements sampled and on specific ratios between several of these simple measurements, and specifically: foot length-breadth ratio (L_B); foot length-heel breadth ratio (L_Hb); foot length-mean forefoot height ratio (L_Mforeh); foot breadth-mean forefoot height ratio (B_Mforeh); and lateral-medial foot breadth ratio (Lat_Med).

Gender Differences. As expected, we found that all measures examined in the study differed significantly for men (N = 113) and women (N = 203). This result adds nothing to the understanding of the differences in the foot shapes of men and women, given that is exclusively linked to a question of size: men’s feet (are), on average, are simply higher, longer, and broader than women’s feet are. To further investigate any differences in morphological terms, we examined male-female differences by comparing the means of ratios between measurements. The foot length-mean forefoot height ratio (L_Mforeh) differed significantly for men (mean = 2.24) and women (mean = 2.30). Consequently, given equal foot breadth, women had lower feet. Even the lateral foot breadth-medial foot breadth ratio (Lat_med) significantly differed by gender, with a mean for women of 1.02 vs. 0.96 for men. This finding means that medial foot breadth was on average lower in men. Given that the lateral-medial foot breadth ratio can be correlated with the foot’s abduction and adduction movements, this finding could be interpreted in terms of a tendency for women to have a more abducted foot than men do.

Age Differences. To evaluate the possibility of influences of the age variable on foot morphology, the sample was subdivided into 5 age groups: 15-20 years (N = 17); 21-35 years (N = 117); 36-50 years (N = 103); 51-70 (N = 71); and 71-80 years (N = 5). Data for the first and last groups were not considered due to small group size.

Significant differences by age group were found in the following simple measurements: heel breadth, malleoli breadth, lateral foot breadth, mean forefoot height, and maximum toe height. A comparison of group means showed that main foot breadth and height measurements significantly increased with age (no significant changes were observed for foot length). Even the ratios between measurements confirmed this tendency. In fact, the foot length-heel breadth ratio (L_Hb), the foot length-mean forefoot height ratio (L_Mforeh), and the lateral-medial foot breadth ratio (Lat_Med) significantly differed by age group. Results from Tukey’s test revealed that the 51- to 70-year-old participants (foot length remaining equal) had a significantly broader foot at the heel (mean L_Hb = 3.6105) than both the younger participants (aged 21-35) (mean L_Hb = 3.8410) and the 36- to 50-year-old adults (mean L_Hb = 3.7676). Moreover, given equal foot length, these older adults had a greater mean forefoot height (mean L_Mforeh = 5.8758) than the 21- to 35-year-old did (mean L_Mforeh = 6.1394), but not more than the 36- to 50-year-old age group did (mean L_Mforeh = 6.0030). Lastly, given equal lateral foot breadth, these older adults had a broader medial foot breadth (mean Lat/Med = 1.0600) than both the 21- to 35-year-old (mean Lat/Med = .9740) and the 36- to 50-year-old age groups did (mean Lat/Med = .9898). This latter finding provides support for the hypothesis that feet tend towards adduction with increasing age.

Regional Differences. To investigate the role of geographic area of origin in the foot morphology differences of our participants, we created the following subgroups divided by geographic area in terms of the Italian regions of: Campania and Puglia (South; N = 71); Latium and Tuscany (Centro; N = 64); Sardinia and Sicily (Islands; N = 75); Piedmont, Lombardy, Friuli, and Veneto (North; N = 106).

All simple measures significantly differed in function of participants' geographic region of origin. North-South differences were a constant, with the exception of maximum toe height, which means were not significantly different. We generally found that Northern Italians had longer feet (mean = 25.37) than Italian islanders (mean = 24.51) and Southern Italians (did) (mean = 24.15). The same was true for foot breadth: individuals from Northern Italy had broader feet (mean = 9.74) than both Italian islanders (mean = 9.18) and Southern Italians (mean = 9.06).

Results from the Tukey's test showed that, for the most of the variables examined, the main differences emerged between Northern and Central Italy on one hand and Southern Italy and the islands on the other. Specifically, fewer differences between individuals from Northern and Central Italy emerged than for people of other regions, and this finding was similarly observed for individuals from Southern Italy and the Italian islands. In terms of geographic area, measurements also differed significantly for foot length-breadth (L/B) and foot length-heel breadth (L_Hb), and once again, the Tukey's test pointed to this geographic "divide", in the sense, given equal length, Northern Italians had a broader foot (mean L/B = 2.6075) than Southern (mean L/B = 2,6669) and islander (mean L/B = 2.6746) and a similar result was also found for Central Italians (mean L/B = 2.6421), who did not significantly differ in this regard from Northerners. Moreover, given equal length, individuals surveyed in Northern and Central Italian regions had feet that were also broader at the heel, with mean length-heel breadth (L_Hb) values of 3.6630 and 3.6629, respectively. The other two groups' mean values for this relation were 3.8491 for islanders and 3.9122 for Southern Italians.

3.2 Structured Interview Results

Analysis of participant responses to the 15 interview questions allowed us to reach several conclusions concerning degree of the participants' awareness of their own foot characteristics and certain problematic aspects linked to shoe use.

Awareness of Own Foot Characteristics. As to awareness of their own foot characteristics, the participants were asked to estimate the width of their feet and to try and remember the relative length of their second toe. These estimates were then compared with the participants' actual foot measurements, based on the following procedure. For length, actual measurements were distributed into as many percentile intervals as options available on the interview response scale ("very wide", "wide", "normal", "narrow", "very narrow").

The zero to 20th percentile interval was chosen to represent a very broad foot sole; the 21st to 40th percentile referred to a broad sole; the 41st to 60th was considered normal; the 61st to 80th, narrow; and lastly, the 81st to 100th percentile, was determined to be very narrow.

Table 3 shows the comparison of subjective estimates and objective measurements. Only 25% of the sample responded correctly, in the sense that these individuals' estimates actually corresponded to their foot breadth, whereas 29% of the sample's participants estimated their feet as being broader than they actually were, and 46% thought their feet were narrower than they actually were. We also observed a phenomenon in which each participant tended to "normalize" their own foot toward a standard. In fact, by fixing a mean sole breadth value considered as a "normal" objective value in the sample itself, we observed a strong "migration" in perception of foot breadth towards the more central value of the distribution. Indeed, more than half of the sample (n=147) reported having a "normal" foot breadth—i.e., nearly double the number of individuals with a foot breadth that actually fell into the category labeled "normal" (n=69).

Table 3. Comparison between mean values of actual foot width (calculated as length/breadth ratio) and subjective estimates. Every class represents an interval of 20 percentiles referred to the examined sample.

Perceptions \ Measures	Very breadth (n=55)	Breadth (n=63)	Normal (n=69)	Narrow (n=67)	Very narrow (n=62)
Very wide (n=3)	2%	1%		1%	
Wide (n=121)	69%	32%	39%	27%	10%
Normal (n=147)	29%	27%	51%	57%	51%
Narrow (n=41)		3%	9%	15%	37%
Very narrow (n=2)			1%		2%

As to second toe length in relation to big toe length, 11.1% of the participants responded that they did not remember; 44.9% reported that their second toe was shorter than their big toe; 28.2%, that it was longer than their big toe; and 15.8% that the two toes were the same length. Only 57% of the participants who had a second toe shorter than their big toe (n=222) were actually aware of it. Approximately 79% of the participants with a longer second toe than their big toe (n= 42) responded correctly, but [surprisingly] only 29% of participants having the two toes of the same length (n=50) answered correctly. It was interesting to note that a rather large percentage of participants in this latter category (43%) reported having a second toe longer than their big toe. Altogether, only 55% of the overall sample was able to answer the question concerning the relative length of their second toe correctly.

Foot Sensitivity. The interviewer also asked participants to demonstrate areas of particular sensitivity (if any) for their own feet on chalk models of a right and left foot. Each model was subdivided into 24 areas covering the entire foot. Participants were told that their indications were to refer to "natural" sensitivity foot zones and not to painful areas created by the use of footwear. Approximately 50% of the participants reported having one or more sensitive areas and, approximately 75% of these participants indicated symmetrical areas (i.e., on both feet). The areas that were most frequently reported as being sensitive corresponded to the metatarsal heads and the V metatarsal area, in addition to the lateral/superior part of the III, IV, and V toes (approximately 15% of the sample, with n=188 participants reporting foot sensitivity). Other areas, such as the medial part of the I metatarsal, plantar arch, heel inferior and hind part, showed a lower frequency of reported sensitivity. The remaining foot areas, i.e., the medial and lateral sub-malleolari areas, the superior foot areas, and the areas corresponding to toe pads almost were not reported as sensitive.

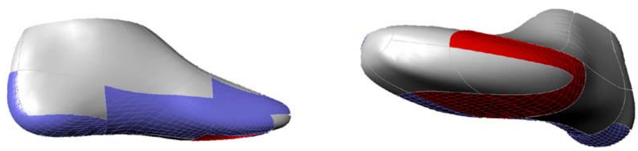


Fig. 4. Areas of reported sensitivity. Areas shown on last on right were reported as being particularly sensitive by approximately 15% of the relative sample, and areas shown on last on left, by approximately 8% of the same sample.

Foot-Shoe Interaction. Participants were asked to indicate the main reasons why the shoes they wear can cause pain and with what frequency it occurs. Participants that reported perceiving their footwear as being too tight at least "sometimes" had feet that were broader on average than the sample mean ($p<0,05$). Nevertheless, in relation with data referred at point 3.2.1., these participants did not seem to be able to link the source of their pain to their own morphological characteristics.

4 Conclusions

The analyses presented herein allowed us to reach some preliminary conclusions about the role of the variables of gender, age, and geographic origin in the morphological and morphometric feet differences of Italian consumers and to more generally speculate about the modifications feet tend to undergo over time.

Our results supported the hypothesis of differences between female and male feet, not only in terms of size, but also in shape, given that men's feet were shown to be higher and slightly more abducted on average than women's feet were. Moreover, age showed an influence, mostly through the observed association of greater breadth (e.g., in the heel and malleolus) with increasing age. This modification most probably can be attributed to an increase in structural load, in terms of the weight gain that frequently occurs as people grow older.

Moreover, the tendency to adduction that we observed in the older participants could be linked to a more general account associated with postural modifications. Similarly, the finding of an increase in foot height (mean forefoot height, but mostly, maximum toe height), suggests that with increasing age, an effect of skeletal, muscular, connective-tissue might develop as an outcome of the constant dynamic load to which the forefoot, particularly, the metatarsal-phalangeal and inter-phalangeal structures, are subjected over the life span.

Lastly, we found that regional differences should be kept in due consideration when accounting for morphometric foot differences.

In the interview responses, participants showed little awareness of the morphometric characteristics of their own feet, in terms both of foot breadth and relative second toe length. Conversely, the high percentage of correct responses by participants with a longer second toe than big toe could point out a relation between second toe length and shoe interaction. Specifically, the fact that participants who had a second toe of the same length as their big toe tended to perceive their second toe as being longer, could be explained by the typical footwear shape, which tends to slant rather drastically downwards from the shoe toe, which usually corresponds to the big toe.

This shape results in contact and/or pressure, considered by the wearer to be derived from his or her own foot shape, rather than from the shoe itself. This phenomenon might have also occurred in participants' self-reports of particularly sensitive foot areas. The metatarsal and antero-lateral zone was indicated by 15% of the sample as a combination of sensitive areas and might in fact be correlated with footwear constriction in the same areas. Moreover, concerning foot-shoe interaction, it was seen that participants reporting the perception of wearing shoes that were too tight had on average feet broader than the sample mean. This finding suggests that participants in the present study were unable to correctly self-represent their own feet morphometric characteristics, unless regarding specific problems determined by foot-shoe interaction.

Our analyses yielded several interesting suggestions requiring further investigation, both through further statistical analysis and through new surveys that will allow to expand the currently available data base. The aim is to obtain useful information to introduce innovations in shoe last design, (by pin-pointing and) defining ergonomic relations among foot measurements, morphometry of shoe lasts and consumers' psychometric variables. Results from the present and the further studies (proposed) will make it possible to enhance ergonomics of footwear both for work and daily use.

References

1. Rossi, W.A.: The Futile Search for the Perfect Shoe Fit. J. of Testing and Evaluation 16(4), 393–403 (1988)
2. Hawes, M.R., Sovak, D., Miyashita, M., Kang, S.J., Yoshihuku, Y., Tanaka, S.: Ethnic differences in forefoot shape and the determination of shoe comfort. Ergonomics 37(1), 187–196 (1994)
3. Mundermann, A., Stefanyshyn, D.J., Nigg, B.M.: Relationship between footwear comfort of shoe inserts and anthropometric and sensory factors. Med. & Science in Sports& Exercise 33(11), 1939–1945 (2001)
4. Cavanagh, P.R. (ed.): The running shoe book. Anderson World, Mountain View, CA (1980)
5. Rossi, W.A.: Footwear: The primary cause of foot disorders. Part 2. Podiatry Management, 129–138 (2001)
6. Collazzo, C.: A 1986-1987 study of consumer problems shopping for footwear, with emphasis on size and fit. J. of Testing and Evaluation 16(4), 421–424 (1988)

7. Rossi, W.A.: The High Incidence of Mismated Feet in the Population. Foot. & Ankle 4(2), 105–112 (1983)
8. Hawes, M.R., Sovak, D.: Quantitative morphology of the human foot in a North American population. Ergonomics 37(7), 1213–1226 (1994)
9. Cheng, F.-T., Perng, D.-B.: A systematic approach for developing a foot size information system for shoe last design. Int. J. of Industrial Ergonomics 25, 171–185 (1999)
10. Ente Italiano della Moda: Le misure antropometriche della popolazione italiana. Franco Angeli Editore, Milano (1979)
11. Fubini, E., Masali, M., Masiero, C., Montanaro, M., Nicolini, G., Riccio, G., Coniglio, I., Pierlorenzi, G.: Le misure antropometriche della popolazione italiana nella progettazione ergonomica europea. Atti del Congresso della Società Italiana di Ergonomia, Palermo, ottobre (1993)
12. Miller, J.E., Nigg, B.M., Liu, W., Stefanyshyn, D.J., Nurse, M.A.: Influence of foot, leg and shoe characteristics on subjective comfort. Foot & Ankle 9, 759–767 (2000)
13. Kouchi, M.: An analysis of foot shape variation based on the medial axis of foot outline. Ergonomics 38, 1911–1920 (1995)

Can We Use Technology to Train Inspectors to Be More Systematic?

Sajay Sadasivan and Anand K. Gramopadhye

Department of Industrial Engineering, Clemson University
sajay@clemson.edu

Abstract. Inspection quality is dependent on the ability of inspectors to weed out defective items. When inspection is visual in nature, humans play a critical role in ensuring inspection quality with training identified as the primary intervention strategy for improving inspection performance. However, for this strategy to be successful, inspectors must be provided with the needed tools to enhance their inspection skills. In this article we outline efforts pursued at Clemson University, focusing on the development of computer-based training systems for inspection training and discuss the results of some of the research briefly.

Keywords: Computer Based Training, Visual Inspection, Search Strategy.

1 Introduction

Visual inspection by humans is a widely used method for the detection and classification of nonconformities in industry. Inspection by humans is not perfectly reliable and it is critical to improve the inspector's performance. Human inspectors have the flexibility to adapt to various tasks and scenarios and improving their inspection process could increase their effectiveness. Visual search and decision making are the two components of a visual inspection process. Visual Search is an important component of the Visual Inspection Process. It is the first step in an inspection task and it involves searching for nonconformities in an item. Past research has shown improvement in visual search behavior with training. The most important aspect influencing visual search is information. Information can be of various types, including defect and location related, performance related, and process. Visual search strategy is a form of process information that can help in improve the accuracy and efficiency of a visual inspection task. Visual search strategy has been categorized as random, which is a memory less process where fixations can occur anywhere in the search field, and systematic, where perfect memory is assumed and no two fixations will occur at the same location. Generally, real world search processes fall in between these two extremes. Inspection performance has been shown to increases when the search strategy tends to be more towards systematic, as the inspection coverage is then exhaustive and no overlap exists between successive fixations. Literature on inspection has shown that moving inspectors from adopting systematic search over random search has significant impact in reducing inspection times and improving

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 959–968, 2007.

defect detection. This paper outlines the use of both low and high fidelity inspection tools in promoting systematic search process leading to superior inspection performance. Specifically the paper discusses a Contact Lens Inspection Training Simulator, VisIns a PC based simulator; Printed Circuit Board simulator; ASSIST - Automated System of Self Instruction for Specialized Training; and finally a VR based inspection tool in achieving the aforementioned objectives.

2 Background

Visual inspection by humans is a widely used method for the detection and classification of nonconformities in industry. Human inspectors have the flexibility to adapt to various tasks and scenarios and improving their inspection process could increase their effectiveness.

Since human inspection remains a central part of quality control, it is necessary to focus on the two functions of human inspection. The visual inspection process are search and decision making[1, 2]. The inspection process in general consists of the visual search process that aims at carefully searching for flaws -- “items that appear differently than the inspector would expect” [3] -- and the decision making process that consists of producing a judgment about the rejectability of the selected item -- “whether it constitutes a fault, an item which meets the criteria for rejection” [3]. Speed and accuracy are the two key variables used in measuring inspection performance. Speed is the measure of the time taken to search for and spot defects or to proceed if none are found while accuracy is the ability of an inspector to correctly identify defects and categorize them.

Inspection processes have been shown to be less than reliable, especially with human inspectors [4]. In an effort to eliminate human error from the inspection system, automated microprocessor-based automated inspection systems were explored [5]. While automation is viewed as the solution to eliminating problems associated with human error, it is not applicable in many industries and processes and it still cannot surpass the superior decision making ability of a human inspector. Human inspectors are more suited for the examination type of inspection tasks wherein the inspector has to search and later decide whether the item is to be accepted or rejected [6]. Training has been shown to be a potentially powerful technique for improving human inspection performance [7, 8].

Feedback training and feedforward training are the predominant training methods in use. Embrey [9] has shown that both knowledge of results (feedback information) and cueing (feedforward information) were effective in improving inspection skills, while Annett [10] found that cueing was equivalent to, or better than, knowledge of results. Megaw [11] mentioned feedback and feedforward information as factors affecting inspection performance while Drury and Prabhu, [12] present evidence supporting the use of feedforward and feedback training in aircraft inspection. Feedback training provides inspectors with information on their past performance or information on the process they adopted in different situations. In a simulated two dimensional airframe component inspection task, process feedback has been shown to improve search strategy [8]. Feedforward training provides prior information, such as information on the defects present, specific locations of defects, or special strategies.

Feedforward training has been found to be effective in past studies providing prior information on faults in industrial inspection.

Visual search is the first step in a visual inspection task and it involves searching for nonconformities in an item. Czaja and Drury [6] have shown improvement in visual search behavior with training. The most important aspect influencing visual search is information. Information can be of various types, including defect and location related, performance related, and process. Visual search strategy is a form of process information that may help to improve the accuracy and efficiency of a visual inspection task. Visual search strategy has been categorized by Megaw and Richardson [13] as random, which is a memory less process where fixations can occur anywhere in the search field, and systematic, where perfect memory is assumed and no two fixations will occur at the same location. Generally, real world search processes fall in between these two extremes. Megaw and Richardson [13] state that inspection performance increases when the search strategy tends more towards systematic, as the inspection coverage is then exhaustive and no overlap exists between successive fixations. They also note that the underlying eye movements of experienced inspectors are far from random. Schoonard et al. [14] found that in chip inspection, trained inspectors adopted a better inspection strategy than novice inspectors. Kundel and LaFollette [15] used fixation analysis to determine that experienced radiologists used a more systematic strategy while inspecting chest radiographs than untrained subjects. It has been shown in the field of radiology that providing an expert's search strategy to a novice can improve the novice's strategy [15]. Wang et al. [16], showed that search strategy can be taught. Graphical cognitive feedback of search strategy has also been shown to enhance visual inspection performance [17].

3 Computer Technology for Training

With computer technology becoming cheaper, the future will bring an increased application of advanced technology in training. In the past, instructional technologists have offered numerous technology-based training devices with the promise of improved efficiency and effectiveness. Many of these training delivery systems, such as computer-aided instruction, computer-based multimedia training, and intelligent tutoring systems, are already being used today, ushering in a training revolution.

In the domain of visual inspection, the earliest efforts to use computers for off-line inspection training were reported by Czaja and Drury [6], who used keyboard characters to develop a computer simulation of a visual inspection task. Similar simulations have also been used by other researchers to study inspection performance in a laboratory setting. Since these early efforts, Latorella et al. [18] and Gramopadhye et al. [17] have used low fidelity inspection simulators with computer-generated images to develop off-line inspection training programs for inspection tasks. Similarly, Drury and Chi [19] studied human performance using a high fidelity computer simulation of a printed circuit board inspection while another domain which has seen the application of advanced technology is the inspection of X-rays for medical practice.

The following section describes some computer based training systems and briefly describes some of the studies conducted using these systems.

4 Technology Based Training Systems – An Overview

Different computer based training systems are discussed below along with research conducted which will illustrate the use of technology to improve human inspection performance.

4.1 Contact Lens Inspection Training Program

This is a low fidelity training program that was used to train inspectors in the contact lens manufacturing industry. The training program was developed using well defined training methodology and images captured from actual contact lenses. When implemented, it resulted in improved inspection performance [20]. Figures 1 and 2 show screenshots of the training system and the details of the program can be found in Gramopadhye et al. [21].

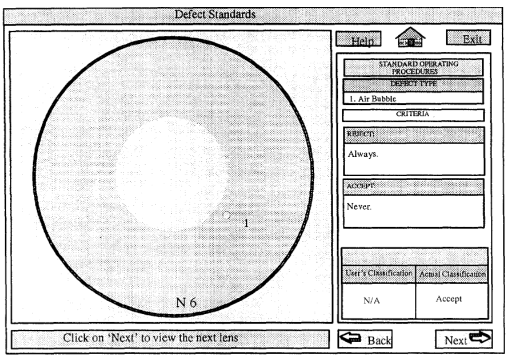


Fig. 1. Inspection screen

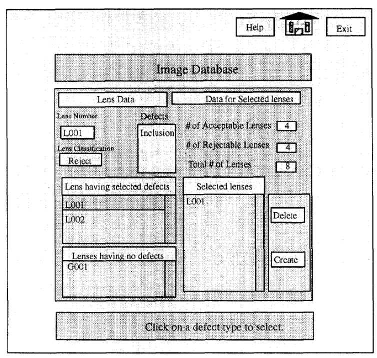


Fig. 2. Results screen

4.2 Visual Inspection Software (VisIns)

The VisInS system is a generic PC based inspection simulator that can be used to simulate the visual search and decision making components of the inspection task. The details of the simulator can be found in Koeniget al. [22]. Screenshots of the simulator are shown below in Figure 3.

Using this simulator various studies were conducted [23, 24, 25, 26] which illustrate the effectiveness of the system in providing inspection training. These studies evaluated different strategies to train inspectors to adopt a systematic inspection strategy. The images below illustrate screenshots of the static (Figure 4) and dynamic (Figure 5) systematic search strategy training.

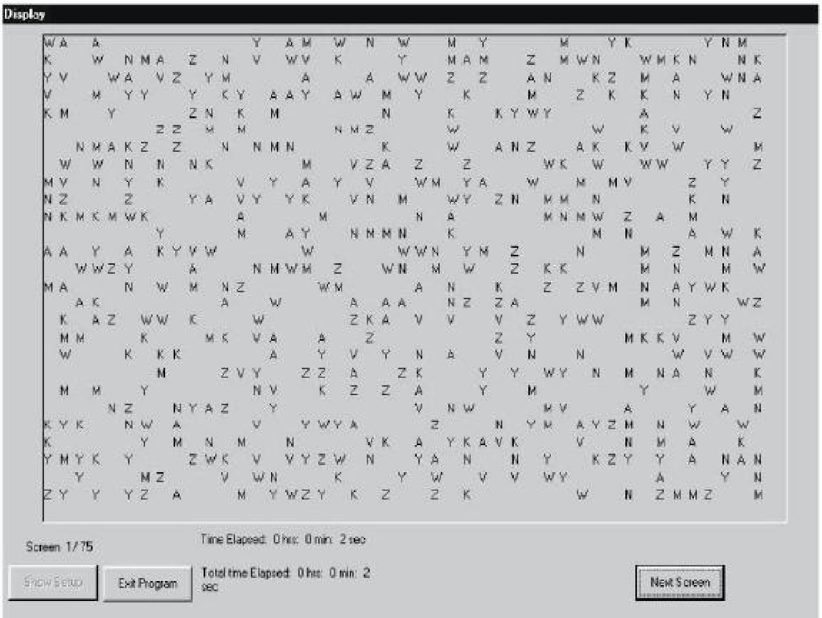


Fig. 3. The VisInS system

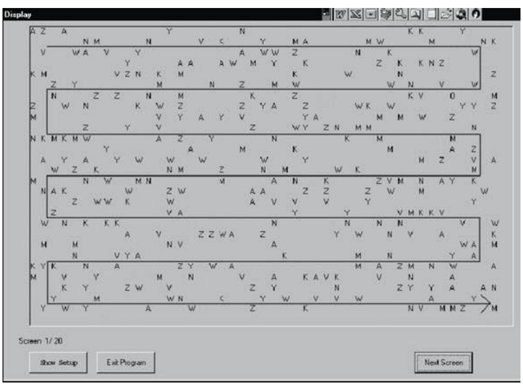


Fig. 4. Static training scenario

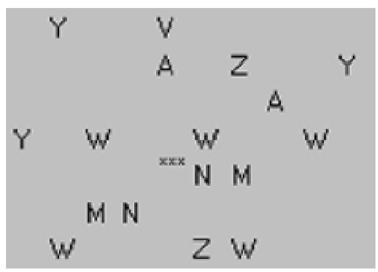


Fig. 5. Dynamic training scenario

4.3 PCB Inspection Simulator

This inspection simulator was developed in order to study function allocation issues in context of a visual inspection Printed Circuit Board (PCB) task. The system can support controlled studies on a printed circuit board inspection task in either human, automated, or hybrid mode. Details of the simulator can be found in Jiang et al. 2002 [27]. The simulator can operate in three separate modes:

1. Human inspection mode: In the human inspection mode, the computer presents the human with a series of PCB images. As each image is displayed, subjects visually search for faults. Once searching is completed, the subjects classify the image based on the severity of the fault(s). Once the image is classified, the inspector can view the next image. The system is designed to operate in two separate modes: with immediate feedback, as in training, and without feedback, as in practice.
2. Computer inspection mode: In the purely automated inspection mode, the operation of the computer parallels that of the human system with the exception that there is no feedback provided. The role of the human is supervisory in nature.
3. Hybrid inspection mode: In the hybrid inspection mode (Figure 6), both the search and decision-making functions are designed so that the functions can be performed cooperatively. The alternatives with parallel human and machine activities enable dynamic allocation of search and decision-making functions.

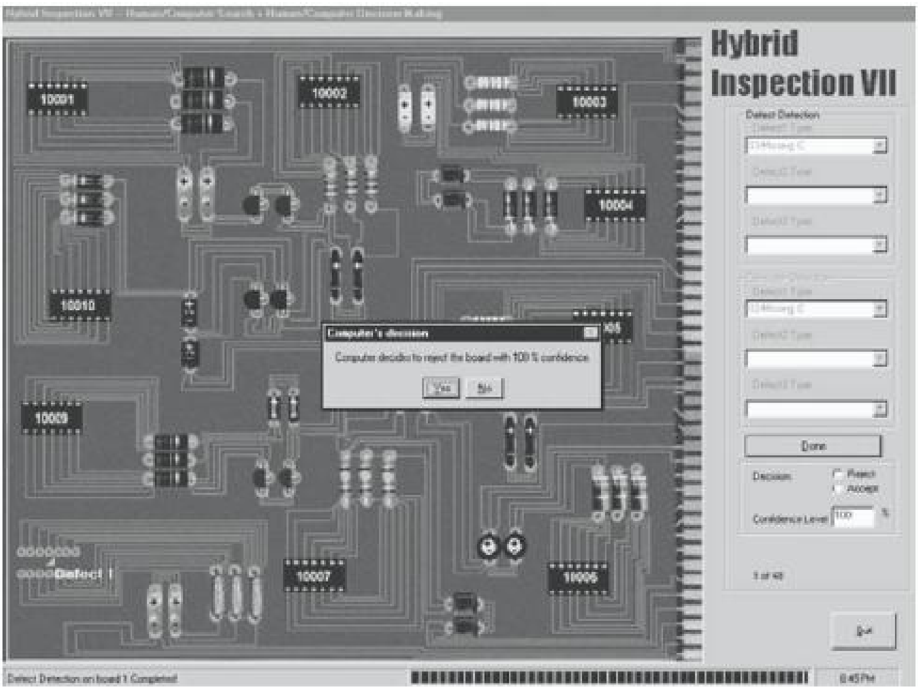


Fig. 6. PCB simulator in the hybrid inspection mode

Studies conducted using the simulator have illustrated its potential in providing training and also have looked at the interaction between human inspectors and computers [28, 29] specifically, the effect of system response bias on inspection quality performance. Search strategy training has also been tested using the PCB simulator using eye tracking information of an expert inspector [30]. The search strategy information was provided to the trainees in three modes – static display

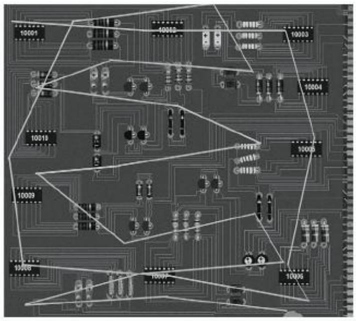


Fig. 7. Static training display

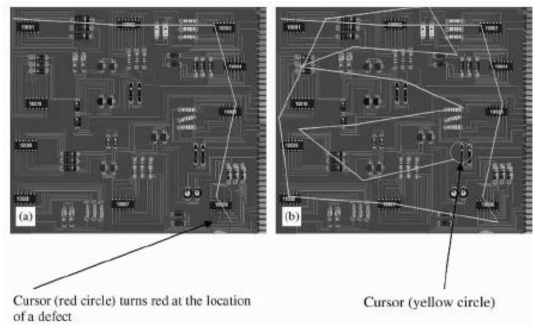


Fig. 8. Hybrid training display

(Figure 7), dynamic display and a hybrid display (Figure 8) of the eye movement information. All three conditions were found to be beneficial in training the novice inspectors to adopt a more systematic search strategy.

4.4 Automated System of Self Instruction for Specialized Training (ASSIST)

ASSIST is a PC based computer simulator for aircraft inspection training. The software consists of three parts: a general module (Figure 9) to provide potential inspectors with general information such as the role of an inspector, factors affecting inspection, safety and inspection procedures; a task simulation module (Figure 11) that allows the trainee to conduct a simulated inspection of an aft cargo compartment; and an instructor's utilities module (Figure 10) that allows an instructor to review a trainee's performance and also configure tasks. The simulated inspection even includes simulated ambient noise such as rivet guns, hangar door horns and other sounds found in a hangar where inspector may be working. During the simulation, the trainee visually searches for and classifies any defects and writes up a computer generated non-routine card. He or she is provided with immediate feedback on his or her performance. Studies [31] using this simulator have demonstrated its effectiveness.

4.5 Virtual Reality Aircraft Inspection Simulator

Despite their advantages, existing multimedia-based technology solutions, including low fidelity simulators like ASSIST, still lack realism as most of these tools use only two-dimensional sectional images of airframe structures and, therefore, do not provide a holistic view of the airframe structure and the complex maintenance/inspection environment. More importantly, the technicians are not immersed in the environment, and, hence, they do not get the same look and feel of inspecting/maintaining a wide-bodied aircraft. To address these limitations, virtual reality (VR) technology has been proposed as a solution. The development of the VR simulator [32] was based on a detailed task analytic methodology [33]. Various scenarios (Figure 12), representative of the aft cargo bay, wing and fuselage of a wide-bodied aircraft, were developed. One of the studies [34] conducted using the VR simulator was to train novice inspectors to adopt an expert inspector's search strategy

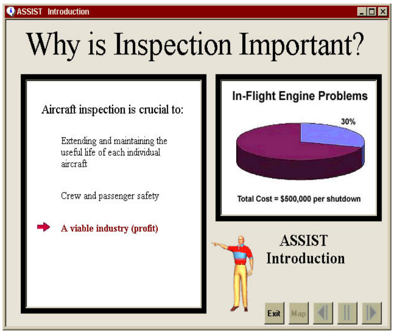


Fig. 9. General inspection module

Fig. 10. Instructor's utility module

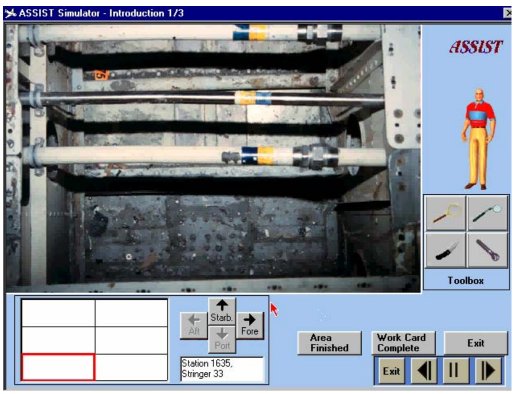


Fig. 11. Simulation training module

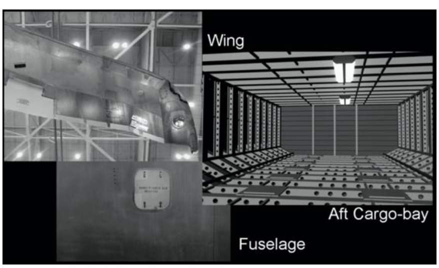


Fig. 12. VR training scenarios

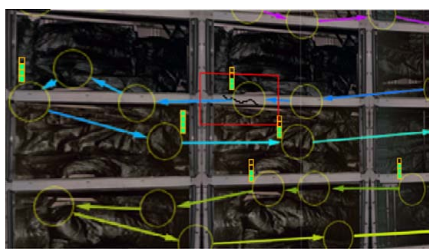


Fig. 13. Search strategy training display

obtained from eye tracking information (Figure 13). The results showed that the training was effective in improving performance and the adoption of the systematic search strategy.

References

1. Drury, C.G.: Integrating human factors in statistical quality control. Human Factors 20(5), 561–570 (1978)
2. Drury, C.G.: Inspection performance. In: Salvendy, G. (ed.) Handbook of Industrial Engineering, vol. 2, Wiley and Sons, New York (1992)
3. Wenner, C., Wenner, F., Drury, C.G., Spencer, F.: Beyond Hits and Misses, Evaluating performance on typical inspection tasks of regional airline inspectors. In: Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting, 1997, vol. 1, pp. 579–583 (1997)
4. Drury, C.G.: Improving inspection performance. In: Salvendy, G. (ed.) Handbook of Industrial Engineering, 4th edn., pp. 8.4.1-14. Wiley, New York (1982)
5. Hou, T., Lin, L., Drury, C.G.: An empirical study of hybrid inspection systems and allocation of inspection functions. International Journal of Human Factors in Manufacturing 3, 351–367 (1993)
6. Czaja, S.J., Drury, C.G.: Training programs for inspection. Human Factors 23(4), 473–484 (1981)
7. Wiener, E.L.: Individual and group differences in inspection. In: Drury, C.G., Fox, J.G. (eds.) Human Reliability in Quality Control, Taylor and Francis, London (1975)
8. Gramopadhye, A.K., Drury, C.G., Prabhu, P.V.: Training strategies for visual inspection. Human Factors and Ergonomics in Manufacturing 7(3), 171–196 (1997)
9. Embrey, D.E.: Approaches to training for industrial inspection. Applied Ergonomics 10(3), 139–144 (1979)
10. Annett, J.: Training for perceptual skills. Ergonomics 9(6), 459–468 (1966)
11. Megaw, E.D.: Factors affecting visual inspection accuracy. Applied Ergonomics 10, 27–32 (1979)
12. Drury, C.G., Prabhu, P.V.: Feedforward and feedback in aircraft inspection (Technical report submitted to FAA), Federal Aviation Administration, Office of Aviation Medicine, Washington (1991)
13. Megaw, E.D., Richardson, J.: Eye movements and industrial inspection. Applied Ergonomics 10(3), 145–154 (1979)
14. Schoonard, J.W., Gould, J.D., Miller, L.A.: Studies of visual inspection. Ergonomics 16(4), 365–379 (1973)
15. Kundel, H.L., LaFollette, P.S.: Visual search patterns and experience with radiological images. Radiology 103, 523–528 (1972)
16. Wang, M.J., Lin, S.-C., Drury, C.G.: Training for strategy in visual search. International Journal of Industrial Ergonomics 20, 101–108 (1997)
17. Gramopadhye, A.K., Drury, C.G., Sharit, J.: Training strategies for visual inspection. International Journal of Industrial Ergonomics 19, 333–344 (1997)
18. Latorella, K.A., Gramopadhye, A.K., Prabhu, P.V., Drury, C.G., Smith, M.A., Shanahan, D.E.: Computer-simulated aircraft inspection tasks for off-line experimentation. In: proceedings of the Human Factors Society 36 th Annual Meeting, 1992, Santa Monica, CA (1992)
19. Drury, C.G., Chi, C.-F.: A test of economic models of stopping policy in visual search. IIE Transactions, 382–93 (1995)
20. Kaufman, J., Gramopadhye, A.K., Kimbler, D.: Using Training to Improve Inspection Quality. Quality Engineering 12(4), 503–518 (2000)
21. Gramopadhye, A., Bhagwat, S., Kimbler, D., Greenstein, J.: The use of advanced technology for visual inspection training. Appl. Ergonomics 29(5), 361–375 (1998)

22. Koenig, S., Nickles, G.M., Kimbler, D., Melloy, B.J., Gramopadhye, A.K.: Visual Inspection Simulator (VisInS): A computer-based inspection simulation tool for off-line experimentation. In: Proceedings of the Industrial Engineering Research Conference, 7th Annual Meeting, 1998 (1998)
23. Nickles, G.M., Sacrez, V., Gramopadhye, A.K.: Can we train humans to be systematic inspectors. In: Proceedings of the Human Factors and Ergonomics Society, 42nd Annual Meeting, 1998, vol. 2, pp. 1165–1169 (1998)
24. Koenig, S., Liebhold, G.M., Gramopadhye, A.K.: Training for systematic search using a job aid. In: Proceedings of the Human Factors and Ergonomics Society, 42nd Annual Meeting, 1998, vol. 2, pp. 1457–1461 (1998)
25. Garrett, S.K., Melloy, B.J., Gramopadhye, A.K.: The effects of per-lot and per-item pacing on inspection performance. International Journal of Industrial Ergonomics 27(5), 291–302 (2001)
26. Koenig, S., Gramopadhye, A.K., Melloy, B.: Use of Job Aid to Promote Systematic Search Under Different Levels of Task Complexity, Human Factors and Ergonomics in Manufacturing (2001)
27. Jiang, X., Bingham, J., Master, R., Gramopadhye, A.K, Melloy, B.: A Visual inspection simulator for hybrid environments. International Journal of Industrial Engineering: Theory, Applications and Practice 9(2), 123–132 (2002)
28. Jiang, X., Khasawneh, M.T., Master, R., Bowling, S.R., Gramopadhye, A.K, Melloy, B., Grimes, L.: Measurement of human trust in a hybrid inspection system based on signal detection theory measures. International Journal of Industrial Ergonomics 34(5), 407–419 (2004)
29. Jiang, X., Gramopadhye, A.K., Grimes, L.: Evaluation of Best System Performance: Human, Automated and Hybrid Inspection Systems. Human Factors and Ergonomics in Manufacturing 13(2), 137–151 (2003)
30. Nalanagula, D., Greenstein, J.S., Gramopadhye, A.K.: Evaluation of the effect of feedforward training displays of search strategy on visual search performance. International Journal of Industrial Ergonomics 36(4), 289–300 (2006)
31. Nickles, G., Marshall, J., Gramopadhye, A.K., Melloy, B.: ASSIST: Training Program for Inspectors in the Aircraft Maintenance Industry. In: International Encyclopedia for Ergonomics and Human Factors, vol. 2, pp. 1178–1180. Taylor and Francis, UK (2001)
32. Vora, J., Gramopadhye, A.K., Duchowski, A., Melloy, B., Kanki, B.: Using Virtual Reality for Aircraft Inspection Training: Presence and Comparison Studies. Applied Ergonomics 39(6), 559–570 (2002)
33. FAA 1991, Human Factors in Aviation Maintenance Phase One Progress Report, DOT/FAA/AM- 91/16, Office of Aviation Medicine, Washington DC (1991)
34. Sadasivan, S., Greenstein, J.S., Gramopadhye, A.K., Duchowski, A.T.: Use of Eye Movements as Feedforward Training for a Synthetic Aircraft Inspection Task. In: proceedings of CHI ’05, Portland, OR, ACM (2005)

Enhancing Sense of Reality by Efficient and Precise Collision Detection in Virtual Environments

Chuan-Jun Su

Department of Industrial Engineering & Management, Yuan Ze University
135, Far-East Rd., Chung-Li, Taiwan, ROC
iecjsu@saturn.yzu.edu.tw

Abstract. As the foundation of user-system interaction in virtual environments, collision detection is a time-consuming process and few real-time interactive algorithms for general objects developed. Most of the existing methods aim for reducing the computation time for some special cases. Collision detection algorithms developed are either not fast enough for practical applications or restricted to a class of specific model. In fact, a general analysis of the performance of collision detection algorithms is extremely difficult because performance is situation specific. The motivation of this work is to satisfy the real-time interaction and high precision requirements of a Virtual Environment (VE) for applications such as virtual design, virtual assembly, virtual training for manufacturing operations and maintenance.

Keywords: Collision Detection, Virtual Environment, virtual reality.

1 Introduction

Collision detection has been extensively used in the fields of robotics operation and process planning for manufacturing. Recently, the research in collision detection has also played an important role in the field of VR. Collision is required specifically to simulate the real-life situations of touching, grasping, moving, and striking. The objective of collision detection in VR is to report accurately all geometric contacts between objects and to respond in real-time. In a VE the user and objects may move with abrupt changes in direction and speed. To provide a sense of reality, a VR system has to efficiently and precisely perform collision checking between virtual objects in a dynamic environment. Due to its widespread importance, there has been an abundance of work on the problem of collision detection. Since Boundary Representation (Brep) and CSG are the most commonly used object representation scheme, the existing collision detection approaches for these two object models are reviewed in this section.

1.1 Collision Detection Approach for Brep-Represented Objects

Spatial decomposition techniques - Octree [1], k-d trees [2], BSP-trees, brep-indices [3], tetrahedral meshes, and regular grids [2] are all examples of spatial decomposition techniques. By dividing the space occupied by the objects, one only

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 969–978, 2007.

needs to check for contact between those pairs of objects (or parts of objects) that are in the same or nearby cells of the decomposition. Using such decomposition in a hierarchical manner can further speed up the collision detection process.

Hierarchical BV techniques - One of the fastest general-purpose polygonal model collision detection system is the "RAPID" system, which is based on OBBtrees, implemented by Gottschalk. Lin, and Manocha [4]. The efficiency of this method is due in part to an algorithm for determining whether two oriented bounding boxes overlap. Another of the fastest approaches publicly available for performing collision detection among arbitrary polygonal models is the QuickCD, which is based on BV-trees, developed by [2].

Distance-based techniques - There has been a collection of innovative work which utilizes Voronoi diagrams [5, 6] to keep track of the closest features between pairs of objects and calculate the distance between them. If the distance between a pair of objects goes below the predefined threshold, a collision is then declared [6]. One popular system, I-COLLIDE [5], uses spatial and temporal coherence in addition to a "sweep-and-prune" technique to reduce the pairs of objects that need to be considered for collision. This software works well for many simultaneously moving convex objects.

1.2 Collision Detection for CSG-Represented Objects

There have been very few approaches dealing with collision detection for CSG-represented objects. Zeiller [7] proposed a three-stage method for detecting collisions among CSG objects using S-bounds [8] and space subdivision. The shortcomings of this method are that at each time step in dynamic simulation, the system needs to recompute new S-bounds for the transformed objects and to perform the spatial subdivision for CSG objects to create an Octree-like data structure that is very time-consuming.

In summary, most of the proposed methods aim for reducing the computation time for some special cases. Most proposed collision detection algorithms are either not fast enough for practical applications or restricted to a class of specific model. In fact, a general analysis of the performance of collision detection algorithms is extremely difficult because performance is situation specific. The motivation of this work is to satisfy the real-time interaction and high precision requirements of a VE for applications such as virtual design, virtual assembly, virtual training for manufacturing operations and maintenance.

In this paper, an efficient and precise collision detection algorithm for rigid objects is proposed. The main idea of our approach is that firstly all Brep objects are converted into CSG representation to take full advantages of CSG's tree structure by using BCSG-1.0 [9]. A localization procedure to detect the potential regions of collision is then performed by testing the interference of several 3-D bounding objects with some non-leaf nodes of two objects' CSG trees. Subsequently, a finer search for true collision is performed within that region using simple and convex bounding representation's face information. CSG's "divide-and-conquer" paradigm, decision rules for sub-tree freezing and result evaluating, distance-aided collision detection for BVs using I-COLLIDE package [5] and adaptive BV selection strategy are used in the localization procedure.

2 Pre-processing

2.1 Brep to CSG and CSG to Brep Conversion

To take full advantages of CSG structure in generating efficient algorithm for the collision detection of virtual objects, all Brep objects are converted into CSG representation in the pre-processing stage using BCSG-1.0 [9]. The procedure induces a set of primitives from a given boundary representation that are used to partition the space into Boolean atoms (cells) that classify either in or out with respect to the solid. The union of those atoms that classify in is optimized using geometric and Boolean algebra techniques to obtain efficient (often minimal) CSG representation of the solid. An example of Brep to CSG conversion performed by BCSG-1.0 is shown in Figure 1.

2.2 Hybrid CSG/Brep Object Model Generation

To fully utilize the strength of both CSG and Brep models for the task of collision detection, a hybrid CSG/Brep object model is constructed in the pre-processing stage. In the construction of a hybrid object model, the input CSG represented object is firstly converted into a Brep model and the links between each object face and its corresponding primitive in the CSG-tree are also established during the conversion. Figure 2 illustrates the procedure of the hybrid CSG/Brep object model generation and Figure 3 depicts the structure of the Hybrid CSG/Brep object model.

2.3 BV Candidates Generation

3-D bounding boxes or spheres are used unambiguously for extremely fast but rough detection of object collision because of their simplicity in intersection computation. To achieve not only fast but also accurate localization for the potential collision region, the BV adopted should be as small as possible and as simple as possible. It is desirable to reduce the unnecessary interference checking among CSG tree nodes by using bounding objects.

One effective way is to minimize the parts of overlapping between the BVs of the target solids. A smaller bounding object implies less interference checking while simpler BV simplifies the required computations such as the minimum distance between objects. Therefore, a list of BVs is generated for each branch node in the CSG tree in the preprocessing stage rather than one fixed type of BV. In the detection stage, the smallest BV for a particular detection case is adaptively selected from the list. In the pre-processing stage of the proposed method, the BVs for each non-primitive CSG node object are selected from the following candidates: sphere, cube, wedge, cylinder, cone, and convex hull. These candidates can be generated by using the covariant analysis based principal axis computation [10] and Qhull package [11]. Figure 4(a), (b), (c), and (d) are the examples of the object-aligned bounding box, bounding sphere, bounding wedge and bounding cylinder of the object respectively. Figure 4(e) illustrates the bounding box of one of the intermediate nodes in the CSG tree of the object.

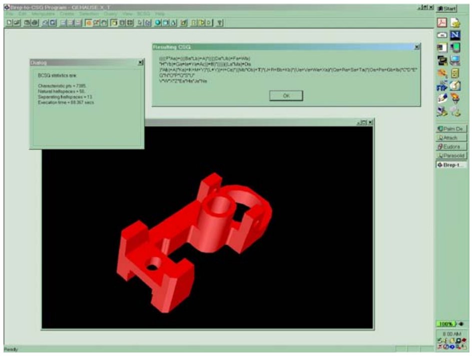


Fig. 1. An example of Brep to CSG conversion by BCSG-1.0

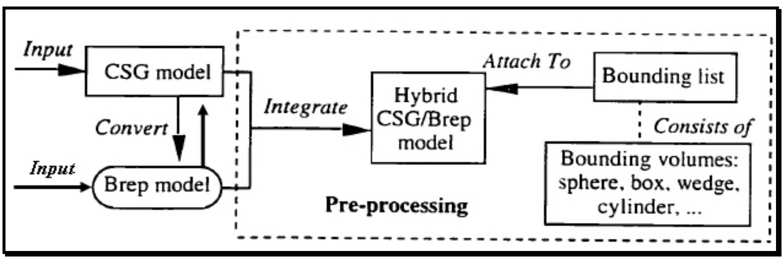


Fig. 2. The procedure of generating the hybrid CSG/Brep object model

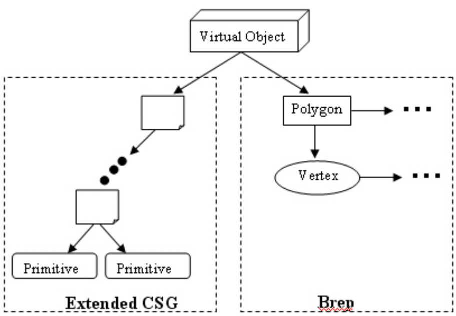


Fig. 3. The Hybrid CSG/Brep object model structure

2.4 BV List Attached to a CSG Tree

Once BV candidates are generated, they are attached to the non-primitive objects in a CSG tree in a list structure called "BV list". The nodes of the BV list are linked in ascending order according to the degree of fitness. Each node in a BV list has four fields:

1. BV type field: ob — object-aligned bounding box: co — bounding cone; cy — bounding cylinder; we — bounding wedge; sp — bounding sphere; ch — convex hull.

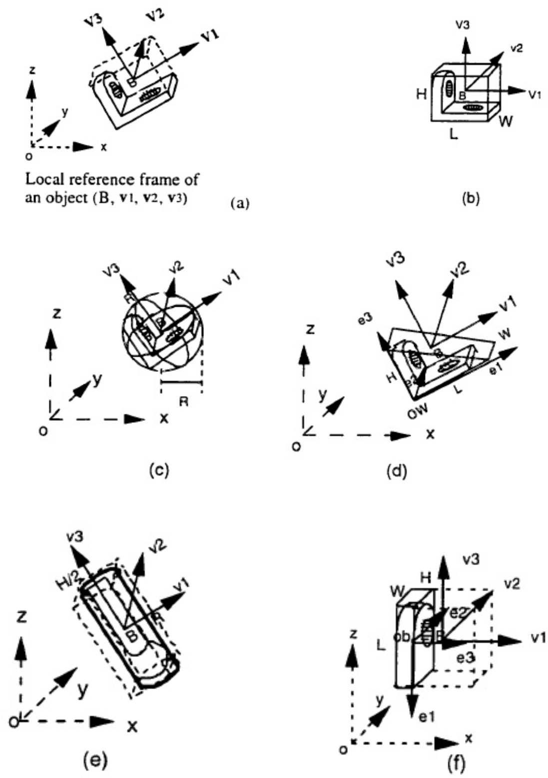


Fig. 4. (a) an example of the basic axis of the object: (B, v_1, v_2, v_3); (b) An example of the object-aligned bounding box of the object A; (c) the bounding sphere of the object A: (d) the bounding wedge of the object A; (e) the bounding cylinder of the object B; (f) the bounding box of one of the intermediate nodes in the CSG tree of the object A.

2. Parameters field: For each type of BV, the corresponding parameters are listed:

(ob, L, W, H; x, y, z; e1, e2, e3): where, L, W, H are the length, width, and height of the bounding box respectively, the point OB(x, y, z) is the central point of the object-aligned bounding box ob, and vectors e1, e2, e3 are the direction vectors of the BV in the local reference system (B, v1, v2, v3);

(co, R, H; x, y, z; e1, e2, e3): where, R, H are the radius, height of the bounding cone: the point OC(x, y, z) is the central point of base circle of the bounding cone co, and vectors e1, e2, e3 are the direction vectors of the BV in (B, v1, v2, v3);

(cy, R, H; x, y, z; e1, e2, e3): where, R, H are the radius, height of the bounding cylinder; the point OY(x, y, z) is the central point of base circle of the bounding cylinder cy, and vectors e1, e2, e3 are the direction vectors of the BV in (B, v1, v2, v3);

(we, L, W, H; x, y, z; e1, e2, e3): where, L, W, H are the length, width, and height of the bounding wedge respectively, the point OW(x, y, z) is the central point of the object-aligned bounding wedge we, and vectors e1, e2, e3, are the direction vectors of the BV in (B, v1, v2, v3);

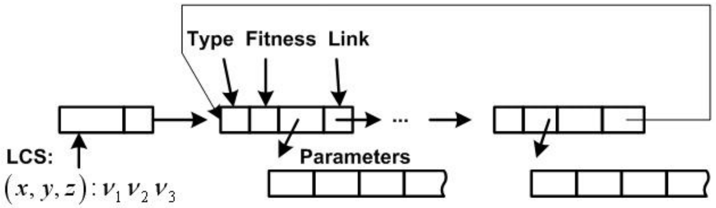


Fig. 5. The data structure of a BV list

(sp, R, x, y, z): where, R is the radius of the bounding sphere sp; the point OS(x, y, z) is the central point of the bounding sphere in (B, v1, v2, v3);

For the convex hull of an object, the parameters are a set of vertices, edges, and facets for the convex hull ch of the object in (B, v1, v2, v3).

3. Fitness degree field: (for simplicity, the volume of BV is stored), and
4. Link field: the pointer linked to next BV node. Figure 5 depicts the data structure of a BV list.

3 Collision Detection Algorithm

After the completion of the pre-processing stage, with the pre-constructed hybrid CSG/Brep object model, a "divide-and-conquer" method can be used to localize the region of interest hierarchically. At each branch node-checking step, it is determined whether there is an intersection between each pair of objects roughly by selecting a most suitable BV from each BV list attached. If no intersection occurs, stop. Otherwise, the CSG operators determine whether further collision checking is needed between children of the current checked node against the current checked node of another object. We devise a set of decision rules for the determination. In this section, adaptive BV selection strategy and decision rules to use the sub-tree freezing approach are presented.

Distance-Based Interference Test for Bounding Objects
Intuitively, if the minimum distance between two solid objects is greater than zero, there is no collision between these two objects. Otherwise, they are touching each other. In this section, we use the distance-based method to detect if two BVs are interfering with each other or not. The suitable BVs of each object are selected from BV lists of each testing object in a CSG tree. In our approach, we have selected sphere, box, cylinder, cone, wedge, and convex hull as BV candidates.

Cohen et al. implemented a package I-COLLIDE for Finding the closest points between two convex polyhedra [5]. It runs almost always in constant time if the previous closest features have been provided and are (on average) linear in the total number of vertices if no special initialization is done. In the membership test process, the running time is proportional to the number of feature pairs traveled. This method was adopted to get the approximate distance between two primitives using the polyhedral representation of each primitive.

If a bounding sphere is selected as a BV from the BV list, a simple and fast method for calculating exactly the distance of a pair of primitives has been derived. A set of closed form equations for precise distance computation by using their relative positions and geometric features were derived.

Adaptive BV Selection

For all possible pairs of BVs, a term c ($0 < c \leq 1$) is defined to express the relative degree of complexity in computing the distance between two objects in a pair. For example, sphere-sphere pair has the lowest c value because checking the distance between two centers and the sum of two radii can easily test the pair's intersection. The cone-cone intersection test may be a relatively complex one to perform. This pair has the higher c value. The BV selection is based on the criterion of minimizing S which is defined as the product of the degree of complexity c and the degree of fitness of two BVs, f1 and f2, of each non-leaf in each CSG tree. That is,

Min. $S = c \times f_1 \times f_2$

Subject to $$fi = 1 - \frac{v_o^{\,i}}{v_b^{\,i}}, i = 1, 2.$$

Where $v_o^{\,i}$ and $v_b^{\,i}$ (i = l, 2) represent the volumes of non-leaf object i and the attached BV i respectively. Before the collision test for two target objects is performed, a bounding volume is adaptively selected for each target object based on this selection criterion. The intersection test between BVs of two branch nodes is then checked.

Decision Rules for Sub-tree Freezing and Detection Result Evaluating

To improve the collision checking for BVs in each CSG tree, sub-tree freezing technique is used to suspend detection of sub-trees (branches). Freezing means that no further actions should be carried out with the "frozen" sub-tree for the object currently being detected at this time step. One branch may be "frozen" or "not frozen" according to the set operation in the node and the detection result of BVs. Detection result of the upper level of node checked can be evaluated. Therefore, decision rules for sub-tree freezing and detection result evaluating are constructed.

Given two objects (or CSG branch nodes) a and b, at the beginning of the procedure of finding collision detection, it can be determined whether there is an intersection between each pair of objects roughly by selecting each BV from each BV list attached to the root of the objects. If no intersection occurs, stop. Otherwise, according to the operator in node a, check first the current level node on the second CSG tree b against the next lower level node(s) on the first CSG tree a. The left child and the right child of node a are denoted as L.a and R.a.

If the operator is UNION, collision between b and L.a is tested. If no collision occurs, checks further the collision between b and R.a. If no collision occurs, stop. If the operator is INTERSECTION, collision between b and L.a is detected first. If no collision occurs, stop; otherwise, check the collision between b and R.a. Table 1 shows the decision rules for union and intersection operation. If the operator is DIFFERENCE, a rough detection between b and R.a is done with BVs. If no collision occurs, check the collision between b and L.a. Otherwise, it is very possible that b is colliding with the complement region of node a. In this time, a polygon intersection test should be executed between the local complement region, b and R.a. If no collision occurs, the result depends on whether there is any collision between b and L.a. Table 2 shows the decision rules for difference operation.

Table 1. Decision rules for union and intersection operators

$UNION(L.a, R.a)$			$INTERSECTION(L.a, R.a)$		
$(L.a, b)$	$(R.a, b)$	(a, b)	$(L.a, b)$	$(R.a, b)$	(a, b)
Yes	Yes	Yes	Yes	Yes	Yes
Yes	No	Yes	Yes	No	No
No	Yes	Yes	No	Yes	No
No	No	No	No	No	No

Table 2. Decision rules for difference operator

$DIFFERENCE(L.a, R.a)$		
$(L.a, b)$	$(R.a, b)$	(a, b)
Yes	Yes	Yes
No	Yes	Yes
Yes	No	Yes
No	No	No

Table 3. Computation time required for each processing cycle of collision detection between two objects with different complexity

Object one (O1)	**Hand**	**Hand**	**Hand**	**Panel**	**Workpiece**	**Probe**	**Probe**
O1 triangles	86	86	86	524	164	1268	3268
Object two (O2)	**Fixture**	**Probe**	**Tool**	**Hand**	**Probe**	**Hand**	**Tool**
O2 triangles	1120	3268	11294 i	86	1268	186	11294
# of Steps	1000	1000	1000	1000	1000	1000	1000
# of contacts	323	360	320	186	162	356	279
Average time (in msec.)	0.05	0.35	0.21	0.025	0.22	0.25	0.31

Recursive Algorithm for Collision Detection

With the pre-constructed hybrid CSG/Brep object model, a "divide-and-conquer" method, adaptive BV selection strategy and decision rules are used to enable fast and accurate localization of the potential collision regions of each object. We can see that the above procedure is recursive. After several branch node-checking steps, the leaf nodes are checked. For two positive leaf nodes generating union or intersection operations, the distance-aided collision detection approach is used to perform the intersection test. For the complemented part(s) generated by subtraction, all of corresponding faces in Brep model are found first to quickly rule out unnecessary polygon-based collision checking, and then polygon-based intersection test is used for the remaining faces.

Implementation

The algorithm proposed in this paper was implemented in C on a 200 MHz R4400 CPU SGI Onyx Reality Engine. The Qhull package [11] for computing convex hulls was used to construct convex hulls of the branch nodes and the I-COLLIDE package was incorporated in the scheme. As a general assessment for the algorithm, the average processing time required for detecting collision between two objects in a typical Virtual CNC Training setting is depicted in table 3.

4 Conclusion

We presented an efficient and precise collision detection algorithm for CSG-represented objects in a VE. The objects are assumed to be rigid solids and would not be deformed during and after collisions. In summary, this method has following advantages:

1. Using natural hierarchical CSG tree structure and adaptive selection strategy for bounding volumes, it realizes an effective and efficient intersection or contact region localization. An intersection or contact can be localized into two leaf nodes of each object after limited steps of simple bounding volume intersection test.
2. The hybrid CSG/Brep representation facilitates the combination of the advantages of the faster, less accurate BVs with accurate distance-assisted collision detection and precise polygon-based intersection test.
3. Our methods have been implemented. Effectiveness and correctness have been proven by experiments with various object pairs with different geometrical complexity. For complex object pairs, our method is significantly more effective.

References

1. Moore, M., Wilhelms, J.: Collision detection and response for computer animation. Comput. Graphics. In: SIGGRAPH ’88 Proc vol. 22, pp. 289–298 (1998)
2. Held, M., Klosowski, J., Mitchell, J.:Real-time collision detection for motion simulation within complex environments. In: SIGCRAPH’96 Visual Proc. (1996), vol. 151 (1996)
3. Vanecek, G.: A Data Structure for Analyzing Collisions of Moving Objects. In: Proc. of the 24th Annual Hawaii Int. Conf. on Sys. Sci. (1991), pp. 671–680 (1991)
4. Gottschalk, S., Lin, C., Manocha, D.: OBBTree: A hierarchical structure for rapid interference detection. Comput. Graphics 30, 171–180 (1996)
5. Cohen, J., Lin M., Manocha D., Ponamgi M.: I-COLLIDE: An Interactive and Exact Collision Detection System for Large-scale Environments. In: Proc. ACM Interactive 3D Graphics Conf. (1995), pp. 189–196 (1995)

6. Lin, M.C., Manocha, D.: Fast interference detection between geometric models. The Visual Computer 11, 542–561 (1995)
7. Zeiller, M.: Collision detection for objects modeled by CSG. Visualization and Intell. Design in Eng. and Architecture, 165–180 (1993)
8. Cameron, S.: Efficient Bounds in Constructive Solid Geometry, IEEE Computer Graphics & Applications, 68–74 (1991)
9. Hartquist, E.: BCSG-1.0: A Practical Implementation of Boundary to CSG Conversion,Tech. Rep. CPA93-3, Sibley School of Mech. & Aerospace Eng., Cornell Univ. (1994)
10. Wu, X.L.: A linear-time simple bounding volume algorithm. In: Graphics Gems HI, pp. 301–306. Academic Press, Inc, San Diego (1992)
11. Barber, B., Dobkin, D., Huhdanpaa, H.: The Quickhull Algorithm for Convex Hull. ACM Trans, on mathematical Software 22(4), 469–483 (1996)

Dynamic Generation of Human-Populated VR Models for Workspace Ergonomic Evaluation

Tien-Lung Sun[1], Wen-Yang Feng[2], and Chin-Jung Chao[3]

[1] Dept. of Industrial Engineering and Management, Yuan-Ze University, 135 Yuan-Tung Rd., Tao Yuan 320, Taiwan, R.O.C.
tsun@saturn.yzu.edu.tw
[2] Dept. of Industrial Engineering, Chung Yuan Christian University, 200 Chung Pei Rd., Chung Li 320, Taiwan, R.O.C.
wyf9556@yahoo.com.tw
[3] Dept. of Industrial Engineering, National Chung-Tung University, 1001 Ta Hsueh Road, Hsinchu, Taiwan 300, R.O.C.
c356028@yahoo.com.tw

Abstract. Human-populated VR models have been employed by many researchers to evaluate workspace occupational health and safety (OHS) problems. Previous researches in VR-based ergonomic evaluation have taken a model-based, quantitative approach where human posture data are extracted from virtual space and sent to ergonomic models to calculate numeric results. Such model-based approach requires the VR model to be integrated with complex human models and ergonomic evaluation models. Construction of the virtual workspace thus becomes complicated and difficult, which often stops industrial adoption of the VR technology in ergonomic evaluation. To address this problem, this paper presents an easy-to-construct human-populated VR model to support visualization-based, qualitative workspace OHS evaluation. Details of the workspace layout design and the operation procedure design, and the dynamic VRML modeling to support the workspace design are discussed. Based on the proposed method, a human-populated virtual workspace is implemented to evaluate OHS problem associated with the design of a ship operation room.

Keywords: Virtual reality, VRML, virtual prototype, digital human, occupational health and safety.

1 Introduction

Human-populated VR models have been employed by many researchers to evaluate occupational health and safety (OHS) problems in airplane maintenance, car operation, factory cell operations, and so on. The virtual workspace allows domain experts to evaluate layout designs and operation or maintenance procedures. VR-based OHS evaluation is particular useful at the early design stage where it is infeasible or impossible to build a physical workspace to evaluate OHS problems associated with different design ideas.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 979–987, 2007.

Previous researches in VR-based ergonomic evaluation have taken a model-based, quantitative approach which requires the VR model to be integrated with a complex human model and an ergonomic evaluation model. The construction of the virtual workspace thus becomes complicated and difficult, which often stops the industrial adoption of the VR technology in ergonomic evaluation. Even the ergonomic experts or product designer know that they could benefit from VR in the long run, they still hesitate to adopt VR because of the difficulty and complexity to build a VR mode. To address this problem, this paper presents an easy-to-construct human-populated VR model to support visualization-based, qualitative workspace OHS evaluation.

2 Literature Review

2.1 Human-Populated Virtual Environment for Ergonomic Evaluation

Most of previous researches focus on demonstrating the feasibilities and advantages of conducting ergonomic evaluations within a virtual space. The process to construct the underlying VR model that supports the ergonomic evaluation is seldom addressed. Chedmail *et al*. [2] integrate a manikin or a robot in a virtual environment to access the suitability of a product, its shape, and functions. De Antonio *et al*. [4] develop a human-populated virtual environment for inspection, maintenance, operation and repair of nuclear power plants. Kuo and Chu [10] develop a web-based, human-populated virtual environment for online ergonomic evaluation of car interior design. Zülch and Grieger [20] describe a human-populated virtual environment for macro-ergonomic occupational health and safety (OHS) evaluation. Jayaram *et al*. [7] propose two approaches to link human-populated virtual environments (VE) with quantitative ergonomic analysis tools for real time occupational ergonomic studies.

2.2 Dynamic Scene Graph Construction

Due to the large number of possible configurations a custom product may have, it is infeasible to pre-define the VR models for custom products. Recently several researches start to investigate the so-called dynamic-content VR model for visualizing custom products. These systems employ a template-based approach in which a VRML template is used as the generic VR model for a custom product and specific VR models are derived from the template for different customization configurations. The VRML models derived, however, are behavior-exclusive or have very limited behaviors. Yen and Ng [18] develop a dynamic catalog and evaluate its advantages over the static catalog. The VR models of the objects that will be used to assemble a product are stored as VRML files. Jezernik and Hren [8] present a virtual prototyping system that allows design engineers to examine different configurations of a design prototype. Varlamis *et al*. [15] develop a dynamic-content VR system for interactive design and visualization of a room. Wojciechowski *et al*. [16] discuss dynamic VR model generation in the domain of museum exhibition.

2.3 Human Behavior Programming

Human behavior programming is an important technical challenge for creating human-populated virtual environments. Monzani *et al*. [13] propose a method to create virtual humans that could take decisions by themselves. They separate the low-level action control from the high-level logical management of behaviors. Chittaro and Serra [3] employ probabilistic automata to develop a goal-oriented approach to program virtual human behaviors. Lee and Williams [11] expanded the parameterized action representation (PAR) architecture to create a feedback mechanism so that a digital human can produce messages about commands back to the user. Dinerstein and Egbert [5] describe a technique that lets virtual characters to quickly adapt online according to their interaction with a human user. Yang *et al*. [17] describe a three-layer hierarchical system for human behavior control. Zhuang *et al*. [19] describe a framework to program human movements and generate navigation animations in virtual environment.

3 Visualization-Based Workspace OHS Evaluation

The workspace OHS evaluation considered in this paper is a visualization-based, qualitative approach. Most of previous VR-based ergonomic evaluation researches have taken the model-based, quantitative approach where human posture data are extracted from the VR model and sent to ergonomic models to calculate numeric results. The qualitative evaluation approach considered in this work does not rely on ergonomic models to perform evaluation. Instead, domain experts examine the operation or maintenance procedures performed in the virtual workspace to visually identify possible OHS problems. This approach takes advantage of human's natural strength in rapid and flexible visual pattern recognition ability. Visual perception is well known for its creativity and flexibility. As such, human experts could recognize OHS problems that are hard or even impossible to be precisely defined by ergonomic models.

The visualization-based qualitative evaluation could server as a good supplementary to the model-based quantitative evaluation methods. In certain situations, it may be even favored over the quantitative methods. For example, at the early design stage, the designer often does not want to spend too much time and effort to conduct a precise, complete OHS analysis. In this case, the visualization-based qualitative approach is sufficient and is favored over the quantitative methods.

4 Structured VR Model Construction Based on VRML Modeling Scheme

The VR model used in this work will be constructed using the Virtual Reality Modelling Language (VRML), an international standard for defining Internet-based virtual environments [14]. The structured data elements defined in VRML makes it more suitable than the lower level graphic language such as DirectX [12] for structured construction and analysis of VR model.

The building blocks of a VRML model are nodes [14], which are simply objects that contain data and methods. VRML has classified the composing nodes of a virtual environment into eight categories, i.e., geometry, appearance, grouping, environment, viewing, animation, interaction, and miscellaneous.

Nodes in a VRML model are organized in a tree structure called scene graph. In addition to specifying the aggregation/decomposition relationships of the objects in the scene, the scene graph also defines the parent-child relationships of the objects during their movement. The transformation matrices applied to a parent node in the scene graph will also be applied to its children, as such, the children objects will move together with their parent.

The tree structure of the scene graph is unable to specify the time-dependent relationships between nodes. To amend this problem, a process flow diagram called animation circuit is used to describe the execution sequence and logics of the behavior nodes in the scene graph.

5 Constructing a Human-Populated Virtual Workspace for Visualization-Based OHS Evaluation

The human-populated virtual workspace for visualization-based OHS evaluation is dynamically constructed as shown in Fig. 1. The major component is a working VRML model that represents the workspace. At start, this VRML model contains only the geometry and behavior data for the default objects in the workspace, i.e., objects that will remain unchanged during workspace layout configuration. As the user configures the workspace, the content of the VRML model will be updated.

The optional object that could be inserted to the workspace at run time has its VRML model stored as a separate file. When an optional object is selected into the workspace, its VRML model is integrated with the VRML model of the workspace. The scene graph of the object is inserted to scene graph of the workspace to become its sub-tree. To insert the scene graph of the object to the correct place, the workspace scene graph uses a dummy node to mark the place where the optional object should be inserted to. The next step is to integrate the animation circuits. Assuming that the behavior nodes in the animation circuits do not interact with each other, then the animation circuits of the optional object could be appended to the animation circuit set of the workspace. This assumption, of course, may not be true when object behaviors become complicated so that they interact with each other. Integrating animation circuits whose composing nodes interact with each other is a complex task and will be left as future research of this work.

In addition to the optional objects, another optional element that could be inserted to the workspace at run time is the virtual human of different sizes, e.g., 50%ile, 95%ile, etc.. Each different size of human model has its VRML model stored as a separate file. There are many different ways to model the virtual human and to control it movement [9]. In this work, the human model is animated by pre-recorded motion data (discussed later). The scene graph therefore is sufficient to represent the human model. The scene graph of the human model is a set of hierarchically organized limb meshes whose movements are controlled by forward kinematics. Integration of the human model's scene graph with the workspace scene graph is the same as described earlier. The human model's VRML model only contains a scene graph but no animation circuit. The behavior nodes and animation circuits that define the behavior of a human model will be dynamically generated as discussed in next section.

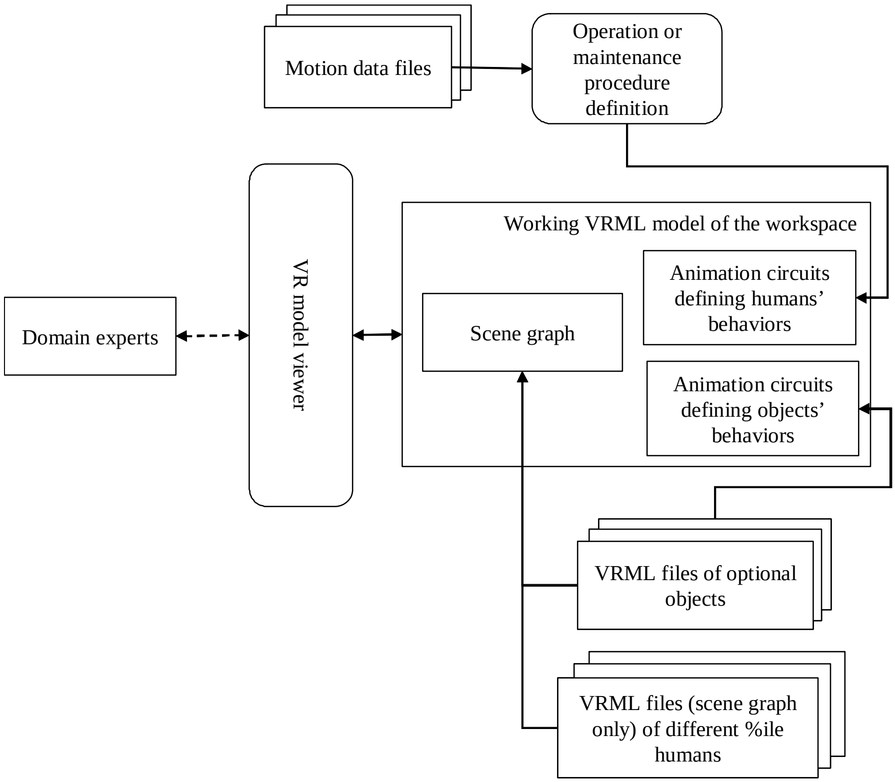


Fig. 1. Construction of a human populated virtual workspace

6 Specifying Human Model Behaviors in the Workspace

After loading the optional objects and human models into the workspace, the next step in constructing a virtual workspace for OHS evaluation is to specify the behaviors of the human models to let them perform operation or maintenance procedures in the workspace. In general, programming the human model is a complex and tedious task. Considering the fact that the domain experts who use the virtual workspace are usually not familiar with human model programming, in this work the human model programming is restricted to loading motion data files to perform pre-defined actions. The users move the human model to a place in the workspace and then load the motion data files to perform the pre-defined actions. Although such approach does not allow general and flexible behavior programming, it is a simple and suitable for this work. The animation data of the pre-defined actions are generated using a motion capturing system and stored in a set of motion data files (see Fig. 1). Each motion data file represents a different action.

The operation procedure is defined as a sequence of actions. These actions either move the virtual human along a 2D path in the workspace or perform a pre-defined action at a place. To move the human model, the user defines a 2D path in the workspace, the movement type (e.g., walk or run), and the distance between two steps. The animation data is then automatically generated. As the walk or run action has repeated human posture pattern, automatic generation of such action is rather easy. Actually, commercial software such as 3ds Max [1] already has built-in support for this function. To perform a pre-defined action at a particular place, the user simply loads the corresponding motion data file to the virtual human model.

Fig. 2 shows an example of defining an operation procedure in a workspace. This operation procedure consists of seven actions. The first action performs a pre-defined action named ‘step over the door’. The second action is a ‘walk’ action that moves the human model along a 2D path defined by the user. After walking to the desired place, the virtual human performs three pre-defined actions, i.e., ‘sit down’, ‘stand up’, and ‘stretch’. Then the virtual human walk to another point to perform a pre-defined action, i.e. ‘bend’.

After defining an operation or maintenance procedure in the workspace, the VRML model of the workspace will be updated to add the human model’s behaviors to the VRML model. Each action specified in the operation procedure is converted to a behavior node that is added to the scene graph under the human model. The sequential relationship between the actions is specified by an animation circuit. The animation circuit is then appended to the workspace VRML model.

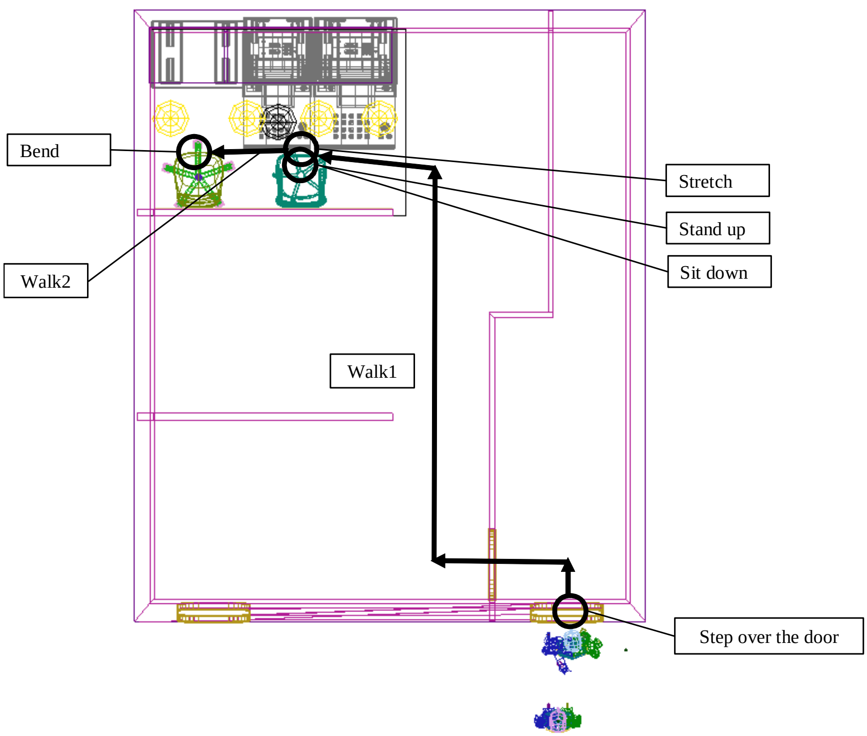


Fig. 2. An operation procedure example defined for a virtual human

When adding the human model's animation circuit to the workspace, it is assumed that the behavior nodes in the animation circuits do not interact with each other. That is, the human models in the workspace do not interact with each other, and they do not interact with the objects in the workspace either. This assumption, of course, makes the human behavior less realistic. But it is sufficient for initial OHS evaluation purpose. Animation circuit integration considering the human-human and human-object interaction is a complex task and will be left as future research.

7 Implementation and Case Study

Based on the proposed method, a virtual workspace has been implemented and used in a project we did for a sail boat manufacturing company in Taiwan to evaluate OHS problems associated with the operation room design. The 3D geometric model of the workspace is created in a CAD system. The geometric model is converted to a polygonal model and loaded to 3ds Max. In 3ds Max, materials and lights are added to improve the appearance of the workspace. The biped skeleton, which is a 3ds Max built-in human model, is employed to create virtual humans in the workspace. The workspace constructed in 3ds Max is then exported to a commercial VR authoring tool called EON Studio [6]. The graphic user interface in Eon facilitates the construction and management of the scene graph and animation circuits in a VRML model. The working VRML model of the workspace and the VRML files for the optional objects and human models are prepared in Eon.

The motion data to animate the pre-defined actions in the workspace are prepared using VICON motion capture system (www.vicon.com). These data are stored in a set of motion data files where each motion data file stores the animation data for a pre-defined action. The motion data for movement actions such as walk or run are generated using 3ds Max. The user specifies the 2D path for the movement and then call the built-in functions in 3ds Max to generate the animation data.

The virtual workspace is visualized using the freely distributed Eon viewer. The virtual operation room allows the sailing boat designer, together with the ergonomics experts, to dynamically configure the operation room such as changing the layout, assigning different numbers of workers, changing the behaviors of the workers, and so on. Observing multiple humans working together in the control room, domain experts could discover possible OHS-related problems associated with early design concepts in the cyberspace, before the expensive physical prototype is created.

Fig. 3 shows several screen dumps captured from the visualization-based OHS evaluation process when the virtual human performs the operation procedure defined in Fig. 2. For business secrete concern, the objects in the workspace have been modified so that they do not show the actual sizes and shapes. The visualization allows designers to evaluate whether the current design of the operation room has available space to conduct an operation. Also visualization from the virtual human's point of view allows designers to evaluate whether the virtual human has difficulty to see the top of the console while doing the operation.

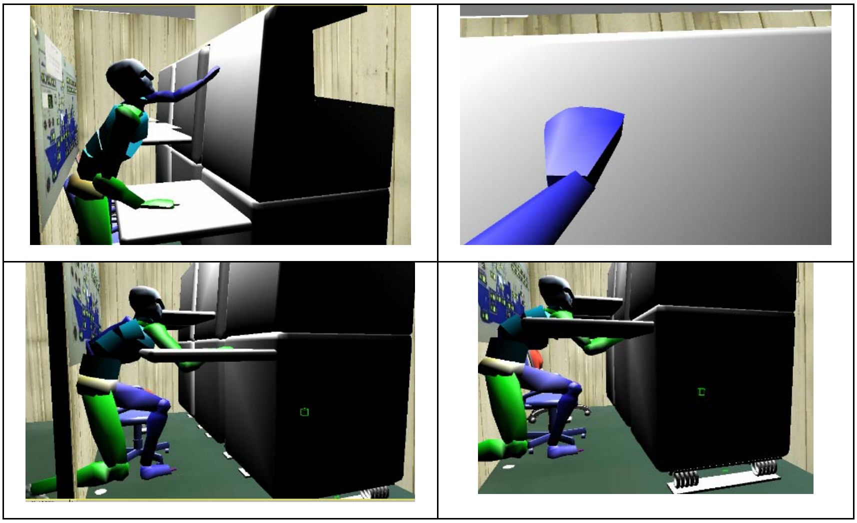


Fig. 3. Examples of visualization-based OHS evaluation process

8 Conclusion and Future Research

This paper discusses the process to configure a virtual workspace for visualization-based, qualitative OHS evaluation and the dynamic VR modeling techniques to support this process. Based on the methods described in this paper, a virtual workspace has been implemented and used for visualization-based OHS evaluation of a ship operation room in a project we did for a sail boat manufacturing company in Taiwan. The prototype demonstrates the usage of VR-based dynamic visualization in helping ship designer evaluate OHS problems at the early design stage.

It is assumed in this work that the humans and objects in the workspace behave independently. As such, integration of the animation circuits during VRML model updating process does not consider the interaction between the behavior nodes. This assumption makes the behaviors simple and unrealistic. Although the simple behavior modeled in this work is already sufficient for early OHS evaluation purpose, more complicated behaviors where the human will interact with objects in the workspace should be considered in the future research.

Acknowledgements. This research is partially supported by National Science Council, Taiwan, R.O.C. under contract No. 95-2221-E-155-003-MY3. The author appreciates the financial support of NSC, R.O.C.

References

1. Autodesk: 3DS Max 8.0 User's Reference, Autodesk, Inc. (2005)
2. Chedmail, P., Chablat, D., Le Roy, C.: A Distributed Approach for Access and Visibility Task with a Manikin and a Robot in a Virtual Reality Environment. IEEE Transactions on Industrial Electronics 50(4), 692–698 (2003)

3. Chittaro, L., Serra, M.: Behavioral Programming of Autonomous Characters based on Probabilistic Automata and Personality. Computer Animation and Virtual Worlds 15, 319–326 (2004)
4. de Antonio, A., Imbert, R., Ramírez, J., Ferré, X.: Usability Issues in the Design of an Intuitive Interface for Planning and Simulating Maintenance Interventions Using a Virtual Environment. Virtual Reality 7, 212–221 (2004)
5. Dinerstein, J., Egbert, P.K.: Fast Multi-Level Adaptation for Interactive Autonomous Characters. ACM Transactions on Graphics 24(2), 262–288 (2005)
6. Eon Reality: Eon 5.5 User guide, Eon Reality, Inc. (2006)
7. Jayaram, U., Jayaram, S., Shaikh, I., Kim, Y., Palmer, C.: Introducing Quantitative Analysis Methods into Virtual Environments for Real-Time and Continuous Ergonomic Evaluations. Computers in Industry 57, 283–296 (2006)
8. Jezernik, A., Hren, G.: A Solution to Integrate Computer-Aided Design (CAD) and Virtual Reality (VR) Databases in Design and Manufacturing Processes. International Journal of Advanced Manufacturing Technology 22, 768–774 (2003)
9. Kim, G.J.: Designing Virtual Reality Systems – The Structured Approach. Springer, London (2005)
10. Kuo, C.F., Chu, C.H.: An Online Ergonomic Evaluator for 3D Product Design. Computers in Industry 56, 479–492 (2005)
11. Lee, J.R., Williams, A.B.: Behavior Development Through Task Oriented Discourse. Computer Animation and Virtual Worlds 15, 327–337 (2004)
12. Luna, F.D.: Introduction to 3D Game Programming with DirectX 9.0, 1st edn. Wordware Publishing Inc. (2003)
13. Monzani, J.-S., Caicedo, A., Thalmann, D.: Integrating Behavioural Animation Techniques, Eurographics, 20(3) (2001)
14. Nadeau, D.R.: Building Virtual Worlds with VRML. IEEE CG&A 19(2), 18–29 (1999)
15. Varlamis, I., Lazaridis, I., Papageorgiou, M., Panayiotopoulos, T.: Distributed Virtual Reality Authoring Interfaces for the WWW: The VR-Shop Case. Multimedia Tools and Applications 22(5-30), 5–30 (2004)
16. Wojciechowski, R., Walczak, K., White, M., Cellary, W.: Building Virtual and Augmented Reality Museum Exhibitions. In: Proc. of 9th International Conference on 3D Web Technology 2004, Monterey, California, pp. 135–144 (2004)
17. Yang, X., Petriu, D.C., Whalen, T.E., Petriu, E.M.: Hierarchical Animation Control of Avatars in 3-D Virtual Environments. IEEE Transactions on Instrumentation and Measurement, 54(3) (2005)
18. Yen, B.P.C., Ng, Y.M.: Development and Evaluation of Dynamic Virtual Object Catalogs. Information & Management 40(4), 337–349 (2003)
19. Zhuang, Y., Xiao, J., Wu, Y., Yang, T., Wu, F.: Automatic Generation of Human Animation Based on Motion Programming. Computer animation and virtual worlds 16, 305–318 (2005)
20. Zülch, G., Grieger, T.: Modelling of Occupational Health and Safety Aspects in the Digital Factory. Computers in Industry 56, 384–392 (2005)

Effecting Validity of Ergonomics Analysis During Virtual Interactive Design

Renran Tian[1,3], Vincent G. Duffy[2], and John McGinley[1]

[1] Center of Advanced Vehicular Systems
[2] Schools of Industrial Engineering, Purdue University
[3] Department of Industrial Engineering, Mississippi State University

Abstract. This paper focuses on validating the performance of virtual interactive design (VID) environment with dynamic ergonomics analysis. Previous studies have only validated VID for posture-based static ergonomics analysis, and applied the methodology on several studies. While since dynamic information is important for ergonomics analysis, this study will investigate the performance of VID environment for ergonomics analysis considering dynamic information such as velocity, which uses motion instead of posture as analysis target.

Keywords: virtual interactive design, ergonomics analysis.

1 Background

A main reason for considering ergonomic issues in workplace design or redesign is to reduce the occurrence of Work-Related Musculoskeletal Disorders (WMSDs). Chaffin has identified potential benefits of using DHM early in the design process, such as reducing the cost and time of the entire design process [3]. Digital human models (DHM) can help predict improved safety and performance for the ergonomic issues related to product design. They are comprised of both the visualization and the underlying fundamentals of math and science. Among the research applications incorporating DHM in design, virtual interactive design (VID) focuses on integrating motion capture systems (MOCAP) with the visualization part of DHM as well as the underlying ergonomic analysis tools that are currently included in commercial software. This enables real-time ergonomic analysis to justify modifications to human-machine interface designs. Li et al. initially introduced VID and examined seated reaching behavior for accessing an ATM by applying this methodology [5]. Validity and reliability of the integrated systems have also been initially verified [7]. However, two problems need to be further considered: (1) current research and application of VID is based on the assumption of static postures; and (2) dynamic aspects of tasks that may affect the validity and reliability of VID have not been reported in the literature. This study will seek a more detailed investigation of the comparisons of real-world analyses and virtual analyses of designs as well as the use of dynamic aspects of tasks information in VID methodology using integrated dynamic ergonomics analysis tools.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 988–997, 2007.

2 Research Scope

In an effort to avoid current deficiencies in simulating realistic human motion and complex interactive behavior between specific workers and devices, motion capture systems have been used to drive digital manikins for the evaluation of workplaces. This was called the Virtual Build methodology [1]. Based on that, Li et al. constructed a real-time virtual interactive design assessment methodology to improve the capability for justifying human-machine interface design modifications [5]. A successful case of using VID for ergonomic evaluations in industry was reported by Du et al [4].

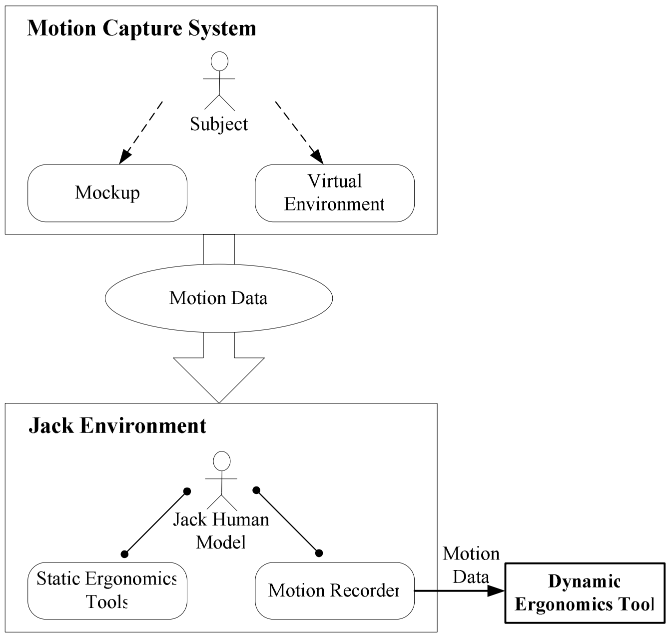


Fig. 1. Scheme of virtual interactive design shows the structure composed of motion capture platform, digital human model and virtual environment, and static and dynamic ergonomics analysis tools

The current scheme for VID environments is shown as figure 1. A dynamic VID environment includes motion data generation, DHM, and related ergonomics analysis tools (Jack embedded static tools and external dynamic tool). The motion data generation part is a MOCAP platform, in which real subjects may act in both physical mockups and immersive virtual environments. Their actions are captured and exported by the MOCAP system. The DHM currently used in our VID system is UGS-Jack^{TM} system, composed of a human manikin driven by MOCAP output, virtual environment, and some tools like imbedded ergonomics tools and motion recorder. It can offer animation of human motion, immersive virtual environment for subjects, and manikin-related static ergonomics analysis. Besides static ergonomics analysis tools embedded in Jack environment as SSP, NIOSH and RULA, an external dynamic tool was used in this study.

An experiment was conducted at the Center for Advanced Vehicular Systems (CAVS) at Mississippi State University to study the performance of VID methodology. An immersive virtual environment (VE) and an identical physical mockup were constructed. In each environment, 36 subjects performed six load transfer tasks with varying external load and rotation of lift. Each task was performed twice by each subject in each environment (VE and mockup). All the movements were recorded by a Motion AnalysisTM MOCAP system operating at 60Hz. Motion data were synchronized, and input into UGS-JackTM to drive human manikin in real-time.

Wu et al. initially studied the validity and reliability of the VID methodology assuming static posture for the analysis [7]. He used the ergonomics analysis tools embedded in UGS-JackTM to analyze the risk of each task based on the postures of the Jack human manikin, and compared results of the NIOSH lifting equation and SSP. Postures taken from interaction with the VE and the mockup were compared to manually-controlled postures specified in the software.

Dynamic analyses were not involved in Wu's validation and the influence of different tasks on the validity assessment also was not previously reported in the literature. The current research will apply and validate the use of dynamic analysis in VID methodology. The Job Risk Classification Model (JRCM) [6] has been shown to work well with MOCAP output [2]. This model uses various trunk motion and workplace factors to predict probability of high-risk job classification, and it will be integrated into the VID environment. Motion data from the above experiment will be used to drive the overall assessment, and the new dynamic VID will be validated. Also, with the help of dynamic VID, different tasks and subtasks can be categorized into low and high risk while incorporating the velocity information not currently included in traditional design assessments. Their effect on validity of VID will be studied.

3 Method

3.1 Dynamic VID Working Process

Process to evaluate task risk through dynamic VID environment includes four steps, and motion data is transferred and processed through all parts to get information, shown as Figure 2; these information are input into JRCM, and job risk value is calculated using equation in that model.

1. Load motion data of special task into MOCAP system. In MOCAP system, task movement of one subject is recorded in a file as a series of postures, where each posture is described as positions of all the markers put on that subject; these files of all subjects can be loaded and drive rigid human model to act.
2. Connect Jack manikin to MOCAP system, and let Jack human model act as MOCAP rigid model acts. An interface between UGS-Jack™ system and Motion Analysis™ system is used here to keep Jack manikin moving synchronically with the rigid human model in MOCAP system by calculating translations and rotations of the more than 100 joints although Jack DHM based on the position of those pre-defined markers.

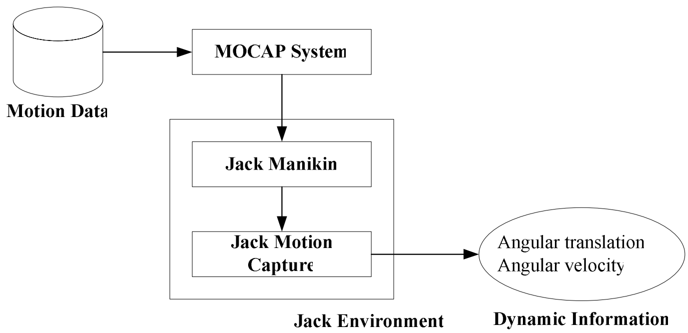


Fig. 2. Data flow is started from saved motion data output from MOCAP system. By replaying those motions in MOCAP system, Jack manikin can be driven synchronously. Embedded motion recorder in Jack environment will record the motion again and output motion data files. By analyzing these motion files, information including angular translation and angular velocity in different anatomical planes can be calculated.

3. Capture the movement of manikin using motion recorder embedded in Jack system. Jack environment offers a motion recorder tool to capture movement of Jack manikin in real time, and real and virtual markers can be added on that manikin to record their position information during capture process.
4. Calculate necessary information for JRCM by analyzing Jack output motion file. The Jack output motion file consists of two sections: the first section includes all the rotation information about body joints; the second section includes all the position information of markers attached on the body. Positions of necessary markers related to trunk motion are investigated in this study to calculate JRCM-needed values, while data-smooth filers are applied during the calculation.

3.2 Description of JRCM

JRCM is a multiple regression model which can discriminate high and low risk lift task by calculating a risk number for each task. This model mainly consider trunk motion and some other related information during lift actions, parameters and corresponding coefficients used are shown in following table 1 [6].

So for each task, a risk value will be calculated and normalized using the equation 1 shown below:

$$R = -3.8 + 0.0014 * LR + 0.024 * MM + 0.02 * MSF + 0.061 * ATV + 0.036 * MLV \quad (1)$$

In order to normalize the risk value, estimated logistic probability can be calculated using the equation 2:

$$\hat{R} = \frac{e^{-R}}{1 + e^{-R}} \quad (2)$$

Greater number represents more danger for tasks.

Table 1. Essencial part of JRCM is a muitiple regression equation consisting of five parameters, corresponding coefficients and the constant is shown here

Parameter	**Coefficient**
Constant	-3.80
Lift Rate (LR)	0.0014
Maximum Moment (MM)	0.024
Maximum Sagittal Flexion (MSF)	0.020
Average Twisting Velocity (ATV)	0.061
Maximum Lateral Velocity (MLV)	0.036

3.3 Calculation of Trunk Angles

Trunk angles are the basis to calculate trunk angular translation and velocity in three planes, and in this study, trunk angles in three planes are calculated from markers. Since Jack marker set is applied, following markers are used. PSIS_L and PSIS_R can position the bottom of trunk, and REAR_NECK_BASE is used to position top of trunk. Trunk rotations in sagittal and lateral planes are calculated by these two points. While trunk rotation in transverse plane is calculated by trunk end point (average of PSIS_L and PSIS_R) and load position (average of PALM_CENTER_L and PALM_CENTER_R).

The angular translation is calculated by simple geometric calculation. Using projection of markers in planes to calculate plane angle is not an exactly accurate method since angle in space is different from angle in plane. But in frontal lift task, most rotation happens in sagittal plane; and in side lift task, different segments of the whole moving cycle have their own main rotation planes. Therefore, this approximation of simply calculating angles using projections in planes is reasonable.

After the angles are calculated, they should be smoothed firstly before the calculation of velocity. First of all, relative angle towards the initial posture is used to reduce errors, and then, a seven-point smoothing routing using normal distribution weighting is used to account for noise in the translation data.

At last, required average and maximum values are calculated.

3.4 Hypothesis to Be Tested for Validation

In VID environment, subjects perform their actions in both real mockup and virtual environment. Mockup-actions can represent actual cases more realistically and accurately. In order to validate the dynamic VID, risk number calculated from special tasks performed in MOCKUP will be compared with those from static analysis in previous studies.

Hypothesis 1: $\text{Risk}_{\text{Dynamic-Mockup-Task}} = \text{Risk}_{\text{Static-Task}}$

Although virtual environment can greatly help to reduce time and cost for design and redesign, subjects may change their actions from those they perform in actual mockup. Wu (2006) compared analysis results based on postures performed in mockup and virtual environment and concluded that using virtual environment works well in some sense. In this study, motions for different tasks from mockup and virtual environment were input into dynamic VID and risk of tasks were evaluated; by comparing results between mockup-based and virtual environment-based, we can validate the performance of VID with virtual environment.

Hypothesis 2: Risk $_{\text{Mockup-Task}}$ = Risk $_{\text{Virtual-Task}}$

4 Result and Conclusion

4.1 Comparison of Dynamic Analysis Results and Static Analysis Results

The purpose of using VID is to evaluate risk of different tasks, which is best represented by mean risk value of one task calculated across all subjects. Table 2 shows the mean risk values based on different ergonomics tools.

Table 2. Mean risk values for lift tasks calculated from all ergonomics analysis tools

	JRCM	SSP	NIOSH
Front Lift (1 lb)	0.11446	203.89	0.0408
Front Lift (20 lb)	0.49110	212.69	0.8440
Side Lift (1 lb)	0.63751	196.19	0.0412

In the table, JRCM is calculated using the above formula. SSP is FLEX/EXT Moment for final Posture and NIOSH is calculated using the following equation 3:

$$\text{LI (Lifting Index)} = \text{Load Weight} / \text{RWL} \quad (3)$$

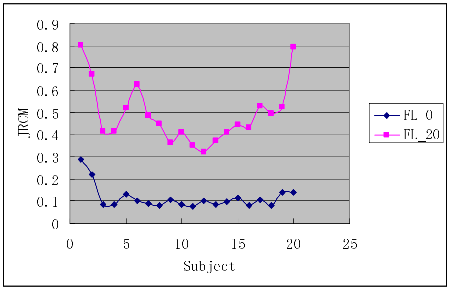
From the table, we can figure out that for the three different lift tasks, JRCM gives a clear discrimination of the risk of them: Side Lift > Front Lift with 20lb > Front Lift with 1lb. From SSP and NIOSH analysis, we can also draw the conclusion that Front Lift with 20lb has higher risk rather than Front Lift with 1lb clearly. But for Side Lift, SSP and NIOSH give a similar value as 1lb front lift, which shows that NIOSH is lack of enough consideration of rotations in transverse plane, and more SSP attributes are needed in order to analyze actions with much rotation. On the other hand, JRCM has quite round consideration of movement in all planes.

For different tasks, following figures show the trends of risk across all subjects evaluated by different tools.

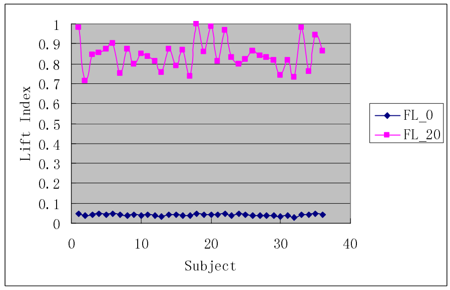
From the below figures about risk value of each subject for the two frontal lift tasks with different load weight, we can figure out that nearly all the subjects have a higher risk value for 20lbs lift whatever tools are applied, and JRCM can accurately figure out this situation.

RULA vs. JRCM. RULA and JRCM were used to evaluate reach actions for all subjects. RULA can output Grand Score from 1 to 7 representing the risk value. In order to check compare RULA and JRCM output, JRCM output was firstly indexed, and then JRCM output was categorized based on corresponding RULA Grand Score. For example, JRCM output for reach action of subject 3 is 0.28975, which is the 10th

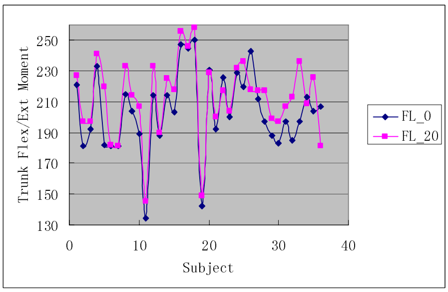
JRCM:



NIOSH:



SSP:



if all the JRCM output is sorted from smallest to the highest, so its JRCM index is 10; higher index means higher risk. Since the RULA output for subject 3 is 7, so for the category whose RULA Grand Score equal to 7, there is a JRCM index which is 10. Table 3 shows all the categories.

Table 3. JRCM index categorized by RULA Grand Score. For each existant RULA Grand Score of a task trial, all the corresponding JRCM Index value are shown in table.

RULA	JRCM Index	Average
3	2, 4, 5, 6, 7, 8, 9, 13, 14, 17	8.5
4	3, 19	11
5	11	11
6	1, 12, 15	9.33(13)
7	10, 16, 18 20	16

Firstly, we can figure out that with the increase of RULA Grand Score, generally the JRCM index is increased. While since the JRCM analysis is influenced by maximum velocity in some planes, so it may have some abnormal points, like subject 18 whose RULA output is 6 while JRCM index is 1. If we remove this abnormal point, the average JRCM index is increased with the increase of RULA for all the subjects.

To sum up, we can say that JRCM based on motions captured based on mockup can discriminate risk level of tasks clearly; and based on the comparison with RULA, its analysis result is valid.

4.2 Validation of VE Based Dynamic VID

Analysis of Mean Value. When subjects performed their actions in immersive virtual environment, they may change their regular actions due to some reasons like limited view field, missing of force back, etc. Following table 4 shows all the mean value of JRCM output.

Table 4. Mean Value of JRCM output for each task and overall case are calculated based on both mocakup-based and VE-based movements. Difference between them are also calculated.

	Mockup	VE	Difference
FL_1	0.115	0.155	-0.04
FL_20	0.491	0.396	0.095
SL	0.638	0.423	0.215
Reach	0.323	0.178	0.145
Overall	0.391	0.288	0.103

Table 5. Results of two-tail pair t-test between mockup-based and VE-based experiment results for each task

FL_1	t Stat	-8.4E-06		
	P(T<=t) two-tail	0.999993	t Critical two-tail	2.093024
FL_20	t Stat	-8.4E-06		
	P(T<=t) two-tail	0.999935	t Critical two-tail	2.093024
SL	t Stat	-7.5E-05		
	P(T<=t) two-tail	0.999941	t Critical two-tail	2.093024
Reach	t Stat	-0.00016		
	P(T<=t) two-tail	0.999873	t Critical two-tail	2.093024
Overall	t Stat	-4.3E-06		
	P(T<=t) two-tail	0.999997	t Critical two-tail	1.99045

By examining all the mean value of JRCM output, we can figure out that mockup-based JRCM output is not equal to VE-based one. For overall comparison, mockup-based output is greater than VE-based output; this is because when subject was performing their actions in virtual environment, generally they will reduce the moving velocity greatly due to unfamiliar with feeling and scene seen in head-mounted displayer, reduced view field, and less feed back. Since JRCM is sensitive to velocity very much, the lower velocity will decrease the JRCM output value. Same situation happens for most tasks except Frontal Lift with 1 lb box (FL_1). For FL_1 task, the reason to get a bigger VE-based JRCM output is that the mockup-based JRCM output is relatively too small due to quite low load weight and less rotation, side bend, and flexion; since original value of those parameters are low, the effect of VE to reduce them are limited.

Also, by investigating the mean difference of lift tasks, we can find that with the increase of risk value, the difference increases. This trend mostly represents the influence from limitation of virtual environment. When we are processing the data analysis, nearly all the subjects used a quite lower velocity to perform more complex tasks. The reason to make reach task out of the trend is that reach is different type of movement, and load weight is not used for this motion.

Validation of VE-based Analysis. The mean value of VE-based analysis is different from mean value of mockup-based analysis, but the trend is still need to be studied. Paired t-test is applied to test the following hypothesis:

$$H0: \mu_1 - \mu_2 = \Delta_0 \quad H1: \mu_1 - \mu_2 \neq \Delta_0$$

Results are summarized in table 5.

Based on the results, we can figure out that H0 can not be rejected for all tasks and overall situation. This means for all tasks and situation, VE-based VID output and mockup-based VID output are correlated, and beyond the internal error caused by limitation of VE, two methods can offer similar analysis results. Figure 3 shows all the JRCM results across trials, and we can figure out that the trends of two different methods are similar.

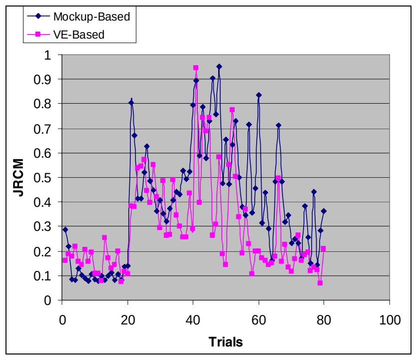


Fig. 3. JRCM output data

5 Discussion

From the investigation on validating mockup-based and VE-based dynamic VID system using JRCM, we can conclude that mockup-based dynamic VID system can discriminate the risk value of special task clearly, and it has better performance to analyze risk when treating with complex movement (side lift) due to round consideration of motions in all planes. Another advantage of dynamic VID is that time is an important issue now to analyze risk, and obviously this will make the analysis more realistic.

For VE-based dynamic VID environment, there are some bias exist due to the limitation of VE itself, but the method still can bring a similar risk trend across tasks comparing to mockup-based analysis.

There are some work needs to be considered and implemented in the future. Firstly, VE itself need to be improved, including extension of view field, more accurate and sensitive animation, and better feed-back. Also, in order to furthermore understand the performance of dynamic VID in ergonomics analysis, more tasks are needed to be investigated.

References

1. Brazier, J.B., Tomko, B., Brow, T.: The car that Jill built. The 6th Annual Applied Ergonomics Conference, March 11-15, 2003, Dallas, Texas, USA (2003)
2. Cappelli, T. and Duffy, V.G.: Motion capture for job risk classifications incorporating dynamic aspects of work. In: Proceedings of the Society of Automotive Engineers, Conference on Digital Human Modeling for Design and Engineering, SAE-DHM-2006-01-2172, July 4-6, 2006, Lyon, France (in press)
3. Chaffin, D.B., (ed.): Digital Human Modeling for Vehicle and Workplace Design. Society of Automotive Engineers, Warrendale, PA (2001)
4. Du, C.J., Williams, S.N., Duffy, V.G., Yu, Q., McGinley, J., Carruth, D.: Using computerized ergonomics analysis tools for assembly task assessment. In: HAAMAHA proceedings, San Diego, CA (2005)
5. Li, K., Duffy, V.G., Zheng, L.: Universal Accessibility assessments through Virtual Interactive Design. Human factors modeling and simulation (in press, 2006)
6. Marras, W.S, Lavender, S.A., Leurgans, S.E., Rajulu, S.L., Allread, W.G., Fathallah, F.A., Ferguson, S.A.: The role of dynamic three-dimensional trunk motion in occupationally-related low back disorders. Spine 18(5), 617–628 (1993)
7. Wu, L.T., Duffy, V.G., White, K.M., McGinley, J.A.: Are DHM and MOCAP reliable and valid for ergonomics. In: The 9th Annual Applied Ergonomics Conference, March, 2006, Orlando, Florida, USA (2006)

Advanced Technology Training for Operating the Microlithography Panel Printer

Bharat Upadrasta, Mohammad T. Khasawneh, and Sarah S. Lam

Department of Systems Science and Industrial Engineering
State University of New York at Binghamton
Binghamton, New York 13902, USA
mkhasawn@binghamton.edu

Abstract. Electronics packaging plays a crucial role in the manufacturing of advanced microelectronics, with an increased interest in technology that produces greater power functionality in smaller spaces. The recent advances in this field have led to advanced manufacturing techniques, with the insertion of components in the micro- and nano-scales. In particular, the microlithography panel printer uses photolithography to create fine image patterns through reticles on substrate boards. The need for advanced features, updated software skills, and the high cost of human error places an increasing requirement for more efficient operator training at a reduced financial investment. Therefore, this research demonstrates the usefulness of virtual reality in modeling the microlithography machine using high fidelity simulations for training purposes. The development process and the structure of the training simulator is presented, with an outline of its effectiveness in supporting training needs, and a description of its reliability assessment through validation and testing.

Keywords: Virtual reality, electronics packaging, training, modeling, simulation.

1 Introduction

The recent advances in electronics packaging have led to advanced manufacturing techniques of printed circuit boards, with the insertion of components in the micro- and nano-scales. Photolithography is one such technique in the electronic packaging process that uses light rays to create, emboss and integrate image patterns on circuits. In particular, the microlithography panel printer uses photolithography to create fine image patterns through reticles on substrate boards. The need for advanced features, updated software skills, sophisticated technology, and the high cost of human error places an ever increasing requirement for more efficient operator training, to ensure successful operation of the system at a minimum financial investment. Configuring and setting up machines in any environment, including electronics manufacturing, is generally affected by both process and human factors issues [2,3]. Process issues include configuring and setting up the machine for a particular process, while human factors issues include decision making and adaptation behavior.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 998–1007, 2007.

In a roll-to-roll manufacturing environment, operators of the microlithography panel printer have to deal with adaptation to an advanced lithography machine and requirements of sufficient knowledge to operate it using a computerized system, which motivated this research. In fact, knowledge of the SUN Operating System (SUN OS or MAC) is a requirement for printer operation. In addition to dealing with a sophisticated and expensive system with utmost care, there is demand for production and throughput to be high during the processing of substrates for optimized and effective performance.

Training has always been the foremost method to improve operator performance in a variety of settings as it provides operators with the necessary skill and knowledge to perform the job goals, thereby improving efficiency and minimizing potential errors [3,8,9]. An effective training program can, not only improve performance of experienced operators for a variety of task conditions, but also bring novices to performance levels beyond that of those experienced in a short period of time. The two types of typical training methods are computer-based training (CBT) and on-the-job training (OJT). OJT generally suffers from inherent drawbacks, such as large investments, set up times, and operational costs. In addition, distance learning and standardization of training makes CBT a favorable choice over OJT. With regard to these factors, a cost effective and complete solution has to ensure consistency as compared to the existing training setup.

CBT has many favorable aspects that explain its prevalence in many fields, such as marine safety training [7], CNC milling training and education [6], radio frequency identification training [11], logistics training [12], emergency handling in the medical industry [15], hurricane evaluation [16], nuclear reactors training [18], astronaut training [17], and complex machine maintenance [19]. Over the years, CBT has been proven to be effective, with significant improvements in the operational understanding of many systems through realistic training environments. In fact, advances in CBT, particularly through the use of virtual reality (VR) technology, is seen as one of the promising solutions that will provide a cost-effective training solution in the electronics industry [2,3].

Though VR-based systems suffered initial setbacks, increasing advancement in technology and globalization has sparked their resurgence. VR-based simulators differ from traditional simulators in the manner they command environments, with applications that employ immersive headsets and head mounted displays (HMDs), desktop workstations, or special purpose projection systems [20]. Over the years, simulation-based training has been found to be effective, improves operator understanding and provides a realistic training environment that can aid operator's adaptation to unfamiliar situations. VR technology enhances such training by integrating it with high end 3-D graphics and real-time interactions to provide a user-friendly, realistic, informative and advanced training solution.

The current training program for the microlithography panel printer is normally conducted on-site. However, a sizeable investment in terms of cost, effort and time is needed to provide effective training for operators who have no experience in training or handling such sophisticated equipment and up-to-date programming software. Therefore, this research is focused on the development of a VR-based training simulator for the microlithography panel printer. This simulator will provide the operator with an understanding of the printer, its operation and everything that he/she needs to know to enhance knowledge and translation of performance to the actual operational setup. VR-based simulation training for operating the panel printer is expected to score over OJT and traditional CBT, with an increase in overall safety through an effective simulation of the real environment. Even though virtual machine models of electronics assembly exist [e.g., 5], to the best of the authors' knowledge, there is no existence of a VR-based training simulator for the microlithography panel printer, which motivated this research.

2 Methodology

VR is a 3-dimensional computer simulation of a real world environment, which interacts with the user and transfers knowledge through real-time experience. A VR system generally comprises both software (i.e., modeling and interface development software) and hardware (i.e., trackers, mouse, and/or gloves) components. Software includes modeling software and interface/application development software [4]. Hardware devices consist of tools for user interaction, input, and tracking, which are used to locate and pinpoint the position of the user in the virtual environment (VE) in relation to the actual real world operational environment. There are various types of trackers, such as acoustic, electromagnetic, mechanical, optical, sensor, and wave trackers, which can be integrated in the system depending on the environment simulated and user requirements.

A primary aspect of any VR training system is the kind of immersion that it provides its users with [10]. Based on the degrees of immersion that a user can experience, Kalawsky [13] and Stuart [22] classified systems as non-immersive, partially immersive, and fully immersive systems. Non-immersive systems offer simulation at a non-immersive level of experience for the user. Desktop computation applications and computing devices, such as the mouse and joystick are examples of non-immersive systems. Partially immersive systems comprise larger projections on a bigger scale involving tracking devices, such as gloves, depending on the type of environment used. Fully immersive systems, on the other hand, provide the user with a 360-degree rotational, translational, and spatial experience, making it the closest as one can get to reality in a VE.

The development of the VR-based training simulator for the microlithography panel printer is comprised of a series of steps. The first step towards the development of a simulator is to understand and analyze the printer and its operational procedure with respect to the existing training program through task analysis. Task analysis aids in understanding the series of tasks involved in each goal of a system procedure and can be used to evaluate operator performance and provide suggestions for optimal performance. Task analysis for the microlithography panel printer has been formulated on the lines of Kirwan and Ainsworth [14], a framework later adopted by Gramopadhye et al. [8,9].

To understand the panel printer and its operation, hierarchical task analysis (HTA) and procedural task analysis (PTA) has been already conducted [24]. HTA represents tasks at different levels of detail highlighting indispensable characteristics in the process, and has been performed for operations such as printer startup and operation process which are critical for operating the printer. PTA is a sequence of tasks that are performed in a specific order to complete a system goal, and has been performed for operations such as maintenance and troubleshooting of the panel printer, which is important for its functioning. This was accompanied by a human error taxonomy that gives a list of all possible errors for tasks listed in task analysis and aids in understanding occurrence of errors and ways to minimize them during operation. The methodology adopted has been discussed in detail by Upadrasta et al. [24], forming the basis for the development of the simulator.

The structure of the training simulator is on the lines of the framework adopted by Su et al. [23] and later by Bhuvanesh et al. [2,3], which essentially focuses on modeling the machine and its features, animating the features and integrating hardware devices to create a virtual environment for operator training. The choice of the final software is normally determined based on the client's needs. Inclusion of a number of interactive screens will help a novice user get acquainted with the machine, which, along with operational procedures, will be adopted in the development of simulator. The simulator will allow operators to understand all aspects of the printer and undergo tests in the form of pre-training and post-training tests.

3 Training Program Prototype

The training program prototype is based on the outcomes of task analysis and can be designed to meet custom requirements of operators during training. The training program prototype can be used as a measure of the effectiveness of the training procedure and thus used to test the validity of the simulator in achieving the objective of providing enhanced training procedure as compared to traditional training. The simulator development process consists of the following phases:

3.1 Virtual Machine Modeling

This phase is used to develop 3D models of the microlithography panel printer. The output at the end of this phase is primarily a high end, accurate and precise representation of the machine, called a virtual machine model (VMM). The model represents the exact real world machine to the remotest features. This model can be used for investigation and analysis of the machine for intricate details. Maya by Alias Software, 3D Studio Max by Autodesk, LightWave 3D by NewTek are leaders in modeling/animation software. In this phase, 3D models of the entire printer and its layout, and its components are developed. First, a basic CAD structure of the printer that paves way for the development of the 3D model, and aids in understanding the cross section view of the printer is developed, after which a 3D model is created. 3D modeling allows for the creation of materials and textures to enhance the model effects for different cross-sectional views of all printer components. Materials and textures give a realistic and complete look to the model created. All models created are very fine in technical detail and the material and texture only add to the realistic aspect of the model. Rich textures provide insight into finer aspects of modeling and provide a better understanding of the model layout. The materials, texture and design capability provided by the modeling language adds to the entire layout of the model and can be used for background and canvas creation for the model.

3.2 Virtual Prototype

In this phase, the VMMs are animated and illustrated. Animation helps in understanding the smaller aspects of the model under observation or study to recognize the behavioral aspects of the model, even at the component level. This allows the user to understand the operation of the panel printer and its components for all specific functions. Animations give the user a sense of better understanding with respect to the model. Since animations are nothing but illustrations from a user's point of view, they are used to describe a model in complete detail. Animations pave the way for the development of video and graphic illustrations with respect to the functioning of the machine. These 3D models are incorporated in web-pages to create animated films/videos using 3D plug-ins, such as Shockwave 3D plug-ins and support capabilities provided either by VRML or Java script. They can be loaded onto the webpage and thereby viewed through Windows AVI or Apple's QuickTime player. The VP development phase also allows for using other features, such as rendering, lighting, shadowing, and immersion capability provided by the modeling language (i.e., MAYA). Rendering allows for an aesthetic sense and fine tuning to a model and gives maximum effect to the animations that are developed.

Lighting and camera capability help in rendering the animation and, along with shadowing, adds to the canvas of the animation and provide an effective background and texture. It increases the clarity and the transparency of the model and the frame for the viewer. It also increases lucidity of the view and provides a better sense of immersion. Cameras, when placed at the optimum human level, provide a realistic perspective and their movement is defined in the line path of the 3D space. The whole content of the training program prototype described in sections above is then incorporated into the development of the virtual prototype (VP) to create the heart of the training simulator.

3.3 Virtual Environment

Once the VP has been developed it needs to be integrated with hardware to facilitate user interaction and a sense of reality for the user. The user interacts with the simulator through a keypad or joystick. The keypad usually has specific functions for each key which the user is told once logged into the simulator. A clear illustration of the keys to be used to operate the simulator guides the user to proceed interacting in the VE. Joystick support in addition to the keypad creates a much more comfortable and interactive kind of environment for the user with specific functions to the joystick buttons [21]. Both the keypad and joystick button functionalities are integrated with the prototype developed to create the complete VE for user interaction. Functions of the keypad and the joystick are programmed to integrate the hardware to create the VE through interface development software. Input devices (i.e., 3D mouse) can be used to provide the tracking effect in the VE, where the location of the user with respect to the environment can be judged. If the user actually loses track in the VE, the tracking device can be used to locate his/her position. In addition, tracking devices are also used to track user performance as well. HMDs and 3D goggle support can also be programmed in the development of the VE to create a sense of immersion during interaction.

3.4 Virtual Machine Tour

VR tours are very popular in conveying the basic concept behind a training application. VR tours usually provide a complete description of the application for which the simulator is developed. These tours are generally viewed using QuickTime plug-ins and are applied at the very outset of a training program to provide an idea about the layout of the object or system. Tours are shown to operators so they can develop a basic understanding of the operation before training takes place. A virtual machine tour of the microlithography panel printer is then developed after the VE is created. The virtual tour can be developed by integrating the 3D models and animations onto a web-page using Shockwave 3D plug-ins. This is similar to the creation of video/animated movies that guides the user through the printer. The user can use QuickTime player by Apple or the Microsoft AVI to view the virtual tour. This is an extremely important feature since the user can click and view the entire layout and features of the printer by a mere click of a mouse. The virtual tour can be developed using VRML or 3D Java script that provides runtime plug-in support.

4 Simulator Development and Validation and Testing

The simulator structure (as shown in Figure 1) incorporates the training program content in addition to the overall design and machine modeling. Instruction-based training is a very important part in the development of the simulator. Information content is available to the user through an interactive and user friendly environment in the form of screens and web pages for user navigation. The user is greeted by a “LOGIN” screen that requires users’ identity for authentication and retrieval of personal information. Furthermore, menus allow for the user to choose from a range of options that include various features and services that the user can avail. Ranging from only an explanation of the layout of the printer to understanding even its minor components, menus serve the main purpose of user interaction through a number of available options. Users can choose from different cross-sectional views, layout of the printer, each individual component and its function, operation of the printer based on task analysis, and handling of the touch panel. This form of training is accomplished through text display on the page, as well as audio and video. Each component of the printer is labeled and illustrated in a lucid fashion to aid the understanding of the user, making the process of learning easy and interactive.

The operator training program in the simulator includes tests that users can take at varying difficulty levels, depending on the level of training that the user needs. Availability of menus that ask users the level of training undergone are followed by the level of difficulty of the test that users wish to take. Each question is rated and performance feedback is given to the user at the end of the test. Users are also instructed whether to proceed or not to the next section, depending on performance. If performance drops below average, the user would be required to take further tutorial sessions well before the test can be taken again, with more customized feedback and continuous evaluation at various stages. This design is aided by a database, as shown in Figure 2, which allows each user who logs in and takes the test/quiz to save records for future reference, or is automatically saved so that users can retrieve incomplete

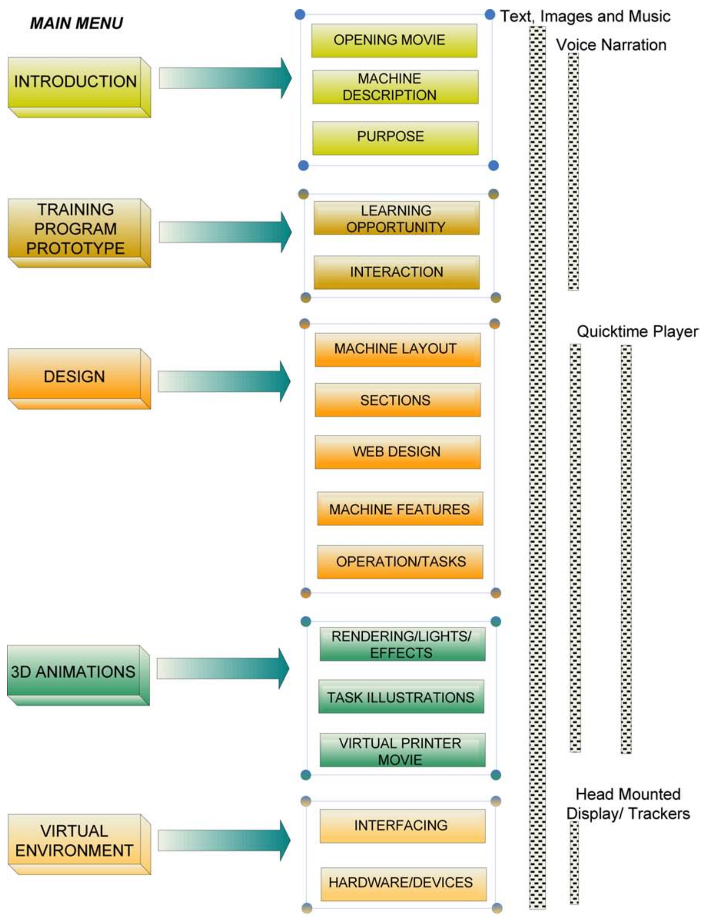


Fig. 1. Structure and development process of the training simulator

and saved records the next time they access the simulator. At the end of all tests undertaken, scores of all the tests are provided with a feedback suggesting suitable improvement actions. Database support provides effective retrieval and query processing capability in the simulator, and can be used for the creation and maintenance of user records.

Once the simulator has been developed, testing and validation is by far the most important phase that a modeler encounters. Verification is the process of ensuring that the model operates as intended and validation ensures that the machine represents reality as best as it can. Subjective face validation is conducted by domain experts based on tests, such as the Turing test [1], subjecting the outputs to make a distinction between a real and a false output to assess model performance. Objective quantitative validation is accomplished by statistical analysis of the pre-test and post-test scores of an operator who undergoes simulated training. Various forms of hypothesis testing is used to analyze the performance metrics before and after undergoing the training to ensure that there is some degree of disparity in the understanding. This proves whether or not the simulator has achieved its objective of providing effective training to the operators subject to the training.

5 Discussions and Future Work

This paper describes the development of a VR-based training simulator for the microlithography panel printer. Task analysis and error taxonomy provide an understanding of the printer and the tasks that an operator would perform from start-up to operation of the printer. The advent of VR simulators has given the cutting edge in CBT with highly realistic training environments, real-time interactive experience and complete knowledge acquisition to effectively translate into the operational environment. As described already, the foray of VR to develop training simulators is a motivating factor for this research, particularly since the existence of training simulator for the microlithography panel printer (electronics packaging in general) has not yet been explored. VR-based simulations provide flexibility, improved performance and brings with it efficiency and effectiveness. Operators can gain a complete understanding of the machine and its layout and apply the knowledge gained during training. This research also provides a cost cutting and effective management solution to conducting training programs. Work is currently being performed to effectively design and develop the simulator for operator training. The potential of the simulator for training more advanced versions of the panel printer is immense, and is implemented using a completely flexible design for customization for future versions of the panel printer.

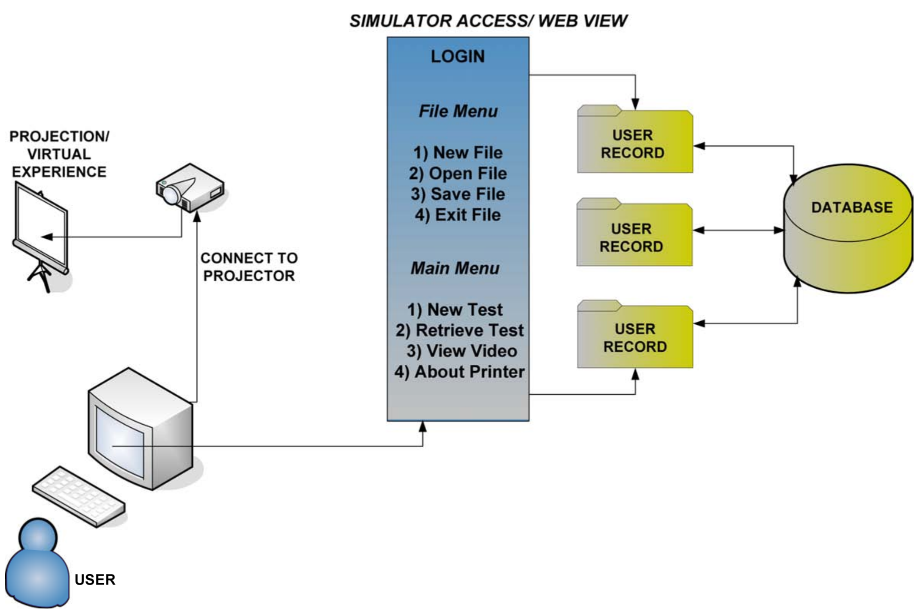


Fig. 2. Interaction and operation of the training simulator

Acknowledgements. The authors would like to acknowledge the Center for Advanced Microelectronics Manufacturing (CAMM) and the Small Scale Systems Integration and Packaging (S3IP) Center of the State University of New York at Binghamton for providing funding to support this research through a grant from NASA under award number NNG06GA06G. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Aeronautics and Space Administration.

References

1. Ahn, V.L., Blum, M., Langford, J.: Telling Humans And Computers Apart Automatically: How Lazy Cryptographers Do AI. Communications of The ACM 47(2), 57–60 (2004)
2. Bhuvanesh, A., Khasawneh, M.T., Lam, S., Srihari, K.: Can Virtual Reality be used to Measure and Train Operators in Electronics Manufacturing? In: Proceedings of the 10th Annual International Conference on Engineering Theory, Applications and Practice, Clearwater, FL, December 4-7, 2005, pp. 94–101 (2005)
3. Bhuvanesh, A., Khasawneh, M.T., Lam, S., Srihari, K.: Task Analysis for SMT Placement Machine Setup for Virtual Reality Based Training: Methodology and Findings. In: Proceedings of the 15th Annual Industrial Engineering Research Conference, May 16-20, Orlando, FL (2006)
4. Bierbaum, A., Just, C.: Software Tools for Virtual Reality Application Development. In: SIGGRAPH 98 Course 14, Applied Virtual Reality (1998)
5. Bodner, D.A., Damrau, M., Griffin, P.M., McGinnis, L.F., McLaughlin, A.: Virtual Machine Models of Electronics Assembly. In: Proceedings of the Deneb International Simulation Conference and Technology Showcase, September 29-October 3, Troy, MI (1997)
6. Chung, C.A., Donaghey, C.E.: Use of Preoperation CNC Mill Training Simulators for Engineering Education. Journal of Engineering Systems Simulators 1(4), 2–10 (2004)
7. Chung, C.A., Martin, V.: Marine Safety Equipment Training Simulator. Journal of Engineering Systems Simulators 2(3/4), 2–7 (2004)
8. Gramopadhye, A.K., Melloy, B.J., Chen, S., Jiang, X., Khasawneh, M.T., Kaewkuekool, S., Bowling, S.R.: Using High Fidelity Simulator for Aircraft Inspection Training: Development and Evaluation. Human Factors and Aerospace Safety 3(2), 121–146 (2003)
9. Gramopadhye, A.K., Melloy, B.J., Khasawneh, M., Kaewkuekool, S., Duchowski, A., Melin, E., Cournia, N., Bowling, S.R.: Using Training to Improve Performance of Inspectors on the Hangar Floor. In: Proceedings of the 16th Human Factors in Aviation Maintenance Symposium, April 2–4, San Francisco, CA (2002)
10. Jerome, C.J., Witner, B.G.: Human Performance in Virtual Environments: Effects of Presence, Immersive Tendency, and Simulator Sickness. In: Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, September 20- 24, New Orleans, LA (2004)
11. Jones, E.C., Anantakrishnan, G.: A Training Simulator for Radio Frequency Identification. Journal of Engineering Systems Simulators 1(4), 44–51 (2004)
12. Jones, E.C.: Design and Development of a Logistics Training Simulator. Journal of Engineering Systems Simulators 1(4), 38–43 (2004)
13. Kalawsky, R.S.: The Science of Virtual Reality and Virtual Environments. Addison-Wesley Ltd, Workingham, UK (1993)

14. Kirwan, B., Ainsworth, L.K.: A Guide to Task Analysis. Taylor and Francis Ltd, London, UK (1993)
15. Lai, F., Entin, E., Dierks, M., Raemer, D., Simon, R.: Designing Simulation-Based Training Scenarios for Emergency Medical First Responders. In: Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, September 20- 24, New Orleans, LA (2004)
16. Okoye-Eze, O.: Hurricane Evaluation Management Simulator. Time Warner Cable (Advanced Technology Group) (2003)
17. Ronkko, J., Markkanen, J., Launonen, R., Ferrino, M., Gaia, E., Basso, V., Patel, H., D'Cruz, M., Laukkanen, S.: Multimodal Astronaut Virtual Training Prototype. International Journal of Human-Computer Studies, 182–191 (2006)
18. Sebok, A., Nystad, E., Droivoldsmo, A.: Improving Safety and Human Performance in Maintenance and Outage Planning through Virtual Reality-based Training Systems. In: Proceedings of Institute of Electrical and Electronics Engineers, Inc., 7th Human Factors Meeting, 2002, Scottsdale, AZ (2002)
19. Sharma, R.: Design and Development of A Complex Office Machine Maintenance Training Simulator. Journal of Engineering Systems Simulators 2(3/4), 8–15 (2005)
20. Stedmont, A.W., Stone, R.J.: Re-viewing reality: human factors of synthetic training environments. International Journal of Human Computer Studies 55, 675–678 (2001)
21. Stone, R.J.: Virtual Reality for Interactive Training: An Industrial Practitioner's Viewpoint. International Journal of Human-Computer Studies 55, 699–711 (2001)
22. Stuart, R.: The Design of Virtual Environments. McGraw Hill, Fairfield, PA (1996)
23. Su, C.J., Duffy, V.G., Hang, Y.M.: Toward the Virtual Factory: A Design Planning Tool for the Future. In: Proceedings of the 22nd International Conference on Computers and Industrial Engineering, Egypt, Dectember 20-22, 1997, pp. 36–39 (1997)
24. Upadrasta, B., Khasawneh, M.T., Lam, S.S.: Virtual Reality Based Training for the Microlithography Panel Printer: Task Analysis and Methodology. In: Proceedings of the 2007 Industrial Engineering Research Conference, May 19-23, Nashville, TN (2007)

Ramsis – The Leading Cad Tool for Ergonomic Analysis of Vehicles

Peter van der Meulen and Andreas Seidl

Human Solutions GmbH
Europaallee 10
67657 Kaiserslautern
Germany

Abstract. Early integration of ergonomics in the vehicle design process requires appropriate CAD tools. 15 years ago, the German car industry developed a new, three-dimensional tool for computer-aided ergonomics and occupant packaging, called RAMSIS. Its goal was to overcome the limitations of two-dimensional human templates, as well as to provide methods for predicting driver postures and comfort. The core of RAMSIS is a highly accurate three-dimensional human model that can be made to simulate occupants with a large variety of body dimensions and based on anthropometry databases from around the world. Extensive research was conducted on driver postures and comfort, which resulted in a probability-based posture prediction model. No subjective manikin manipulation by the user is necessary, so that fast, realistic and consistent analysis results are ensured at all times. An assessment of comfort allows designers to optimize packages with respect to driver comfort early in the design process. RAMSIS offers a number of other analysis tools, for example for vision, reach, force and seat belt studies. Over the years, new research projects have resulted in more sophisticated RAMSIS functions, such as a force-based posture and comfort prediction model, seat belt certification, compatibility with full body laser scanners, simulation of the interaction between seat and occupant and simulation of ingress and egress.

1 Introduction

Manufacturers around the world recognize that occupant comfort and safety contribute significantly to a car's economic success. They are major areas in which manufacturers can distinguish them selves from the competition. This is demonstrated by the emphasis on ergonomics, comfort and safety in today's car brochures, as well as by the increasing attention car magazines devote to ergonomics and safety in test reports.

To ensure adequate comfort and safety levels, it is essential that manufacturers systematically address ergonomics and safety throughout the car design process. Today, a major part of that design process is digital. To study occupant comfort and safety in 3D CAD, an accurate representation of the occupant in the digital world is necessary. However, human beings come in many shapes and sizes, have many different

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 1008–1017, 2007.

preferences and can adopt many different postures. Standardized representations of the human body for automotive design, such as the widely used SAE J826 drafting template, do not describe the human body in its full complexity.

In the 1980s, the German car industry recognized that the ergonomic tools available then were insufficient, as they were two-dimensional, not based on up-to-date research and not integrated in 3D CAD. The German *Forschungsverein für Automobiltechnik* (FAT) initiated development of a new tool that would provide German car engineers with an accurate, three-dimensional representation of the human body in their CAD environment [Seidl]. An Industry Advisory Panel (IAP) was created, consisting of the manufacturers Audi, BMW, DaimlerChrysler, Ford, Opel, Porsche and Volkswagen, as well the seat suppliers Johnson Controls and Keiper, that oversaw development of the human model. The Technical University of Munich conducted the necessary research and Human Solutions GmbH of Kaiserslautern (then called Tecmath) was responsible for software implementation.

The tool was named RAMSIS, an acronym for *Rechnergestütztes Anthropmetrisches Mathematisches System zur Insassensimulation*, which translates as computer-aided anthropometrical mathematical system for occupant simulation. After its introduction to the German automotive industry in the early 1990s and subsequent adoption by the American and Asian automotive industry, RAMSIS has become the world's most sophisticated and comprehensive system, if not the de-facto standard, for ergonomic vehicle layout and occupant packaging. Over 70% of the world's major vehicle manufacturers use RAMSIS.

This paper highlights some of the research that has been and is conducted as part of the continuous effort by the IAP, Human Solutions, the Technical University of Munich and others to provide vehicle designers around the world with a state-of-the-art software tool that allows them to address ergonomics and ensure proper occupant accommodation from earliest stages of the design process.

Fig. 1. The digital human model RAMSIS

2 Benefits of Digital Human Models

The use of digital human models like RAMSIS offers many advantages over traditional ergonomics methods, such as guidelines, tables, two-dimensional templates or user clinics. Porter et al. distinguish time-related, cost-related and accuracy-related advantages.

Time

Detailed evaluation of designs with user questionnaires, clinics or mock-ups can take weeks or months. Digital human models allow designers to simulate user and task with CAD data only. Without a digital human model, designers must often wait for a mock-up to conduct ergonomic studies, which causes delays to the design process, or, more likely, the design process continues without the benefit of timely ergonomics input.

Cost

In addition to being time-consuming, the production of mock-ups is an expensive process. The cost of a digital human model can be less than the costs of making one full size mock-up. Most important, digital human models enable ergonomics input to be provided much earlier in the design process, which reduces the likelihood of expensive or unfeasible modifications being necessary later on.

Accuracy

3D human models offer far more accuracy than guidelines, two-dimensional templates or numerical tables. The human body is highly complex and a large variety of combinations of and correlations between body dimensions exists. Three-dimensional human models are able to reflect this complexity.

3 Initial Approach for Posture Prediction and Discomfort Assessment

The IAP decided that for effective use, it was essential that every user in every design centre produce realistic, reproducible and consistent analysis results with RAMSIS. Thus, the creation of manikins, the process of positioning a manikin in a CAD model and the analyses had to be non-ambiguous and require as little subjective judgment by the user as possible.

For the manikin positioning process, an automatic posture prediction model was developed. RAMSIS can predict, with minimal user interference, how a person with given anthropometric features will most likely perform typical driver tasks in the car that is being designed. Based on such predicted postures, for example, possible collision with the interior can be detected, vision studies can be conducted and feasible reach zones can be generated.

For the development of the posture prediction model, postures of a sample of test subjects, representative of the German driver population, were recorded in three simulated package configurations, one representing a sports car, one a sedan and one a minivan. The subjects performed typical driving tasks, including normal driving, using pedals, shifting gear, looking in different directions and reaching for a number of target points. Using a photographic analysis system that allowed projecting a RAMSIS manikin with the same size as the subject on the recorded images, the joint angles of the subjects in their respective postures were extracted.

From the extracted joint angles, multi-dimensional probability distribution were derived, i.e. for each joint it was determined which angle range is most likely to be used for driving tasks.

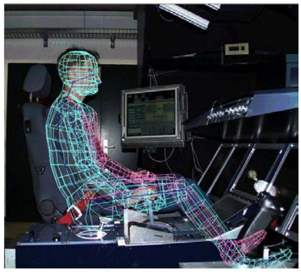


Fig. 2. Photographic analysis of driver posture

To position a RAMSIS manikin in a CAD model, the user must give the manikin a "task" in the form of a number of simple geometric constraints. For example, the user can define connections between body parts and geometry parts (put hands on steering wheel), force the manikin to stay on a particular side of a geometry part (keep head below roof trim) or assign directions to body parts (look 5 degrees downwards). The automatic posture prediction then finds the posture in which all constraints are fulfilled, while at the same time, each joint angle is within a range of high probability. Thus, at all times, the user can be sure that the occupant postures that RAMSIS predicts for a given task are as probable or realistic as possible.

In addition, the German manufacturers wanted to be able to evaluate and compare package designs with respect to driver comfort early in the design process, a possibility not available until then.

After the subjects had performed a typical driving task for 10-15 minutes and their postures were recorded, they were requested to rate self-perceived discomfort levels for a number of body parts on a scale, developed especially for this purpose [Krist].

The (subjective) discomfort levels on each subject's questionnaire were statistically correlated to their (objective) posture data using a multi-linear regression analysis. These correlations are used to predict postural discomfort levels for occupants in new car designs. Thus, designers can study the effect of design changes on driver posture and comfort and optimize packages with respect to driver comfort early in the design process.

4 Force-Controlled Posture and Comfort Prediction

In the late 1990s, the IAP decided to enhance RAMSIS' posture and comfort prediction capabilities with a force-based model. This would allow prediction of postures and comfort levels in terms of joint load for situations in which drivers exert or are exposed to external forces, such as pushing a pedal or shifting gear. Additionally, such a model would allow prediction of postures and comfort levels for non-driving and even non-vehicle-related tasks.

Research resulted in the Force Controlled Comfort and Posture Prediction (FOCOPP) [Seitz]. The FOCOPP assumes that humans try to minimize the relative load in each of their joints when performing a task. Thus, for a given task and combination of external forces, the FOCOPP calculates a manikin posture in which the relative load in each joint is as small as possible.

Relative joint load is defined as the ratio between effective torque in a joint and the maximum torque that that joint can generate/absorb. Maximum torque for all joints of the human body was determined through experiments at the Technical University of Munich, in which subjects' joints were isolated and maximum torque was measured for a range of excursions around the joints' axes of rotation. It was found that maximum torque changes with joint angles and thus with posture.

There is a linear correlation between relative joint load and joint discomfort. Thus, from the relative joint loads, it is possible to associate a comfort rating with a posture.

Currently, the FOCOPP is integrated in RAMSIS in the form of a prototype. Initial validation of the FOCOPP has shown plausible results [Seitz]. A full implementation of the FOCOPP in RAMSIS is expected to be available in 2008.

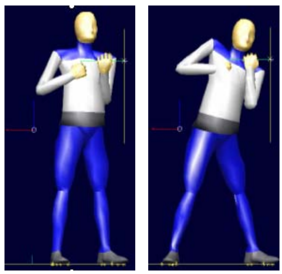


Fig. 3. postures for drill forces of 10N and 100N

5 Enhanced Anthropometry

RAMSIS manikins can be made to represent occupants with a large variety of body dimensions and be based on a number of anthropometry databases from around the world. Currently, databases for Germany, France, USA, South America, Mexico, Japan, Korea and China are available.

In the past, anthropometry data were collected by hand, measuring subjects in standing and seated postures with special measurement tools. Obviously, manual

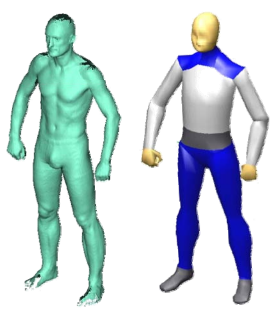


Fig. 4. A RAMSIS manikin based on a full body laser scan

measurement is prone to inaccuracies. Therefore, digital measurements methods, such as full body laser scanners, are increasingly being used for anthropometry surveys. The full body scans that they produce are point clouds that can not be used and animated in CAD. RAMSIS offers the possibility to automatically transfer body dimensions of a scanned subject to the RAMSIS manikin, that *can* be used and animated in CAD. This enables direct application of new anthropometry data to product design, as well as the creation of "individual" RAMSIS manikins, for design of race cars or customized vehicles.

6 Interaction Between Occupant and Seat

Similar to SAE packaging regulations, in RAMSIS, the connection between manikin and CAD model is the H-point. However, when RAMSIS was developed, it was found that real occupants' hip centers are not located exactly on the H-point, but at an offset that correlates to an occupant's height, weight and gender. For a number of seats, H-point offset was measured and statistical correlations with occupant height, weight and gender were determined. A RAMSIS manikin' H-point is not located in its hip centre, but at an offset from the hip centre, that correlates with the manikin‘s anthropometry. This H-point offset approach requires designers to use statistical correlations found on existing seats to predict occupant postures in other or new seats. The properties of the new seat are not considered in the manikin positioning process.

In many cases, it is useful to study the effect of seat geometry, stiffness or adjustability on occupant postures (and thus visibility, reach, etc.) in an early design phase as well. In the late 1990s, the project RAMSIS Sitzt (RAMSIS Sits) was initiated by the IAP, with the goal to develop a method for predicting an occupant's exact position in a seat, based on the seat's geometry, stiffness and kinematics. The project consisted of three phases:

1. Simulation of torso posture and load distribution in seats
2. Modelling of a detailed contact surface between (loaded) seat and occupant
3. Posture simulation in CAD seats

For simulation of torso postures, the kinematic coupling between pelvis and thorax found in the ASPECT program is used [Hubbard et al.]. The ASPECT model assumes that pelvis and thorax rotate evenly, but in opposite direction, when moving from a kyphotic to a lordotic posture.

Research for phase 2 was conducted by Michigan State University in the USA and resulted in a method for predicting load distribution in a seat [Radcliffe], as well as an accurate surface representation of the human back and buttocks when seated [Zhenyu]. Both the load distribution and surface are posture-dependent.

The position and posture of a manikin in a seat is based on a force-equilibrium. The back-and-buttocks-surface is pressed into the seat, with local forces according to the load distribution of phase 2. The seat pushes back with a force that is related to its stiffness or force-deflection behaviour.

Rather than using FEA or similar methods to model seat stiffness, it was decided to measure seat stiffness locally in a number of specified points with a force-deflection gauge and transfer the measured stiffness to the CAD surfaces that represent the seat's cushion and back.

Fig. 5. Accurate surface representation of human back and buttocks

The model was validated in a posture experiment [Wirsching et al.]. Real postures of 20 people in 2 seats were compared to RAMSIS posture simulations. The validation showed that the model consistently simulates the observed driving postures. Significant systematic deviations are caused by the fact that the model cannot properly simulate a phenomenon from reality: people do not move their pelvis to the backrest as closely as possible. Hence the observed horizontal offset between pelvis and backrest must be analyzed and integrated into an enhanced simulation. The new approach was also compared with the H-Point vector method. The models show equal accuracy and predictive power. Nevertheless, the simulation approach opens the door to further research on the effect of seat parameters on driver postures and, possibly, seat comfort.

7 Seat Belt Simulation

Throughout the 1970s, 1980s and 1990s, in attempt to reduce the risk of accident injuries caused by inappropriate belt fitting, Transport Canada developed the Belt-fit Test Device (BTD) [Balzulat et al.]. The BTD is a hardware measuring device that tests whether the lap and torso belt are appropriately positioned with respect to the bony structures of the pelvis and rib cage of a restrained occupant. The premise is that belts that meet the fit criteria established by Transport Canada adequately restrain the occupant in a crash, without causing serious injuries to soft tissue and organs from belt forces.

In 1995, amendments were proposed to two Canadian Motor Vehicle Safety Regulations relating to seat belts: CMVSR 208 (Seat Belt Installations) and CMVSR 210 (Seat Belt Anchorages). The proposed amendment to CMVSR 210 included provisions for abdominal protection using BTD test procedures.

To overcome the deviations of physical hardware tests and to enable review of belt design in early design phases, Transport Canada, the Canadian Vehicle Manufacturers' Association (CVMA) and the Association of International Automobile Manufacturers of Canada (AIAMC) agreed to participate in a joint working group to develop a computer simulation of the BTD test procedures. The development of this "electronic" Belt fit Test Device (eBTD) was funded by the Alliance of Automobile Manufacturers, a leading US advocacy group for the automobile industry on a range of public policy issues. Because of the widespread use of its products by the global automotive industry, Human Solutions was selected as the partner for developing the eBTD software.

In 2006, Transport Canada and vehicle manufacturers that sell cars in Canada signed a Memorandum of Understanding (MOU), called “Seat Belt Fit Evaluation for the 50th Percentile Male ATD in Front Outboard Seating Positions”. By doing so, the manufacturers made a statement of commitment to identify to Transport Canada each new vehicle model equipped with front outboard seat belt systems that have been evaluated in accordance with the fit criteria established by Transport Canada.

The eBTD is one of the world’s first digital certification tools.

The seat belt routing and anchorage simulation developed for the eBTD project can also be applied to RAMSIS manikins, so that besides checking compliance with Transport Canada’s belt fit criteria, seat belt systems can be evaluated from an ergonomic point of view.

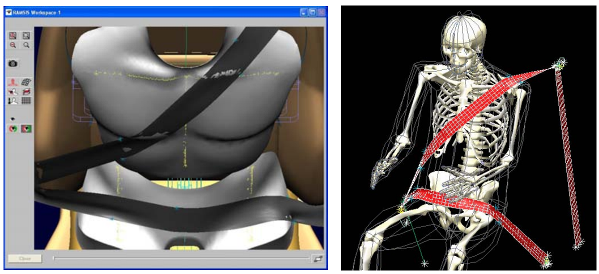


Fig. 6. eBTD and belt simulation on a manikin

8 Ingress/Egress

A major area of attention of human factors and vehicle engineers is the effort required by a person to get into and out of a car. Styling and aerodynamics have pushed the A-pillar and roofline down, while new vehicle concepts, such as minivans, SUVs and cross-overs, have unusual entry opening configurations. At the same time, the customer population is ageing and therefore less flexible or willing to sacrifice comfort for styling.

People use a variety of ingress strategies, such as head-first, foot-first or buttocks-first. A person’s ingress strategy correlates to the geometry of the door opening, his or her body dimensions, as well as to seat and steering wheel position, age, and habit.

A number of RAMSIS customers is conducting research on ingress and egress. Their goal is to be able to predict how people get in and out of vehicle based on the vehicle’s geometry on the one hand and occupant parameters, such as body dimensions, age and habits, on the other. Accurate prediction of ingress/egress motion in RAMSIS allows for early optimization of vehicle geometry.

For these studies, the parameters that affect ingress/egress motion are varied systematically while subjects repeatedly get in and out of a vehicle. Their motion is recorded and analyzed using motion capture systems, such that statistical relations between ingress/egress strategy and the relevant parameters can be established.

An example of an ingress/egress study for which RAMSIS was used in conjunction with motion tracking is the REAL MAN project [Lestrelin et al.].

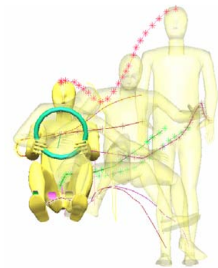


Fig. 7. Ingress postures measured with a motion tracking system

9 Conclusion

Since the early 1990s, the digital human model RAMSIS has become the preferred CAD tool for ergonomic vehicle layout and occupant packaging. Developed by a consortium of German companies, it is now used by the majority of car manufacturers worldwide.

The primary function of RAMSIS is to provide designers with an accurate representation of occupants in their CAD model, both in terms of anthropometry and posture, so that they can ensure proper accommodation of these occupants right from the start of the design process.

Extensive research on anthropometry and driver postures was conducted to build the core of RAMSIS. However, the system has been and is continuously enhanced, as science and experience provide new information, data and insights all the time.

A focus for further development of RAMSIS will be on cognitive ergonomics. In today's vehicles, a large amount of information is presented to the driver. Board computers, navigation systems, car phones, and an increasing number of operating elements all compete for a driver's attention. At the same time, the influx of information from outside the vehicle increases, too, due to intensified traffic, more complex road situations and an abundance of traffic signals. For driver comfort and safety, the way in which all this information is presented is of great importance. Proper ergonomic design of displays, HMIs and operating elements is one of the major future challenges for human factors and vehicle engineers.

Obviously, for car companies as much as for anyone, only a system that keeps pace with science and technology is a justified investment.

References

1. Seidl, A.: Das Menschmodell RAMSIS – Analyse, Synthese und Simulation dreidimensionaler Körperhaltungen des Menschen, Dissertation, Technical University Munich, Munich (1994)
2. Porter, J.M., Case, K., Freer, M.T., Bonney, M.C.: Computer-aided ergonomics design of automobiles. In: Peacock, B., Karwowski, W. (eds.) Automotive Ergonomics, Taylor & Francis, London (1993)

3. Krist., R.: Modellierung des Sitzkomforts - eine experimentelle Studie, Lehrstuhl für Ergonomie, Technical University Munich, Munich (1993)
4. Seitz, T., Recluta, D., Zimmermann, D., Wirsching, H.J.: FOCOPP – An Approach for a Human Posture Prediction Model Using Internal/External Forces and Discomfort, 2005, SAE Paper Number 2005-01-2694 (2005)
5. Hubbard, R., Gedraitis, C.: Initial Measurements and Interpretations of Seat Factors for the ASPECT Program, 1999, SAE Paper Number 1999-01-0958 (1999)
6. Radcliffe, N.: Kinetic Modeling and Seat Factors relating to Manikin-Seat Interaction in Automobiles. Thesis for the Master of Science Degree, Department of Materials Science and Mechanics, Michigan State University (2001)
7. Zhenyu, L.: Measurement and Analysis of Human Posterior Back and Thigh Contours for Automotive Seat Design. Dissertation, Department of Mechanical Engineering, Michigan State University (2005)
8. Wirsching, H.J., Junker, J., Fritzsche, F.: An Integrated and Evaluated Simulation of Human- Seat Interaction 2006, SAE Paper Number 2006-01-2361(2006)
9. Balzulat, J., Wirsching, H.J., Hassan, J.E., Noy, I., Garder, W., Shewchenko, N.: The electronic Belt Fit Test Device – Methodology, Results, and Prospects 2004, SAE Paper Number 04DHM-80 (2004)
10. Lestrelin, D., Trasbot, J.: The REAL MAN Project: Objectives, Results and Possible Follow-up 2005, SAE Paper Number 2005-01-2682 (2005)

Using Multimodal Technologies to Enhance Aviation Maintenance Inspection Training

Carl Washburn[1], Paris Stringfellow[2], and Anand Gramopadhye[2]

[1] Greenville Technical College, Department of Aviation Maintenance Technology, Greenville, South Carolina, USA
[2] Clemson University, Department of Industrial Engineering, Clemson, South Carolina, USA
Carl.Washburn@gvltec.edu, {fparis,agramop}@clemson.edu

Abstract. Recent collaborative efforts between Greenville Technical College's Aviation Maintenance Technology (Greenville, South Carolina, USA) training facility and Clemson University (Clemson, South Carolina, USA) have lead to significant improvements in Aviation Maintenance Technician training through the use of advanced computer technology. Such applications included: 2.5D and 3D virtual environments of a large-bodied aircraft cargobay with interaction modalities ranging from fully immersive (using a head-mounted display and 6 degrees-of-freedom mouse) to semi-immersive (using a spatially-tracked suspended, touch-sensitive window display) to non-immersive (using a basic desktop computer and mouse); and 3D virtual environments of turbine engine blades where nondestructive inspection methods (e.g. borescoping) could be practiced. This paper discusses the integration of these technologies into an existing educational curriculum and provides insight as to how such programs might be implemented and evaluated.

Keywords: Education, Virtual Reality, Computer-Based-Training, Aviation Maintenance, Multimodal Instruction.

1 Introduction

Sound aircraft inspection and maintenance are an essential part of safe and reliable air transportation. Aircraft inspection is a complex system interrelating the human and machine. A number of previous aircraft crashes have been attributed to faulty maintenance procedures. Visual inspection by a trained human inspector forms an important part of the maintenance procedure, contributing to almost 90% of the visual inspection of an aircraft. Training has been identified as the primary intervention strategy to improve the quality and reliability of aircraft inspection. Traditionally, the aircraft inspector obtained on-the-job training (OJT), which helped bridge the gap from the classroom teaching to practical workplace environment. This, however, may not always be the best method of instruction [1]. Some of the limitations inherent to OJT include the lack of feedback, the high cost of aircraft exposure, and the limited defect exposure.

Older, more experienced maintenance technicians are retiring from commercial aviation maintenance and are being replaced by a much younger workforce coming

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 1018–1026, 2007.

directly from schools. Often, these new graduates have not been exposed to the complex wide-bodied aircraft maintenance environment and are often not fully prepared to make a smooth transition to the workplace. Also, these students may never receive hands-on inspecting experience and, as a result, are not adequately prepared for the transition to the workplace.

Traditionally, aircraft maintenance schools have concentrated on teaching students the theoretical aspects of the inspection procedure. One such program is the Aircraft Maintenance School at Greenville Technical College in South Carolina, USA. The program provides two year aircraft maintenance course aimed at training aircraft maintenance technicians for the workforce. A major limitation of the program has been the inability to provide actual hands-on training and the practical experience needed to work in a complex aircraft maintenance environment, especially in wide-bodied aircraft. Usually, neither can the aircraft maintenance schools afford the prohibitive costs of acquiring wide-bodied aircraft, nor do they have the hangar facilities to store such aircraft. This leads to training students on smaller aircrafts. However, such training on smaller aircraft may not transfer well to the wide-bodied aircraft. Thus, students trained via traditional methodologies are confronted with on-the-job situations that require them to provide quick and correct responses to stimuli in environments where they have no previous experience and in a situation where inspection and maintenance errors can be costly and at times catastrophic. To alleviate some of the shortcomings of these traditional methods, computer technology has been proposed to improve the curriculum and provide practical experience to the students.

Computer-based technologies have been developed that have promised improved efficiency and effectiveness of the inspection procedure. Such training devices are being applied to a variety of technical training applications, including computer-based simulation, interactive videodiscs, and other derivatives of computer-based applications. Computer-aided instruction, computer-based multi-media training and intelligent tutoring systems are already being used today in classrooms to promote active learning. Computer-based simulators have been previously used to provide much needed practical experience to novice students, ranging from surgery training to flight simulators.

In visual inspection training, the earliest effort using off-line inspection training was by Czaja and Drury [2], who used keyboard characters to develop a computer simulation of a visual inspection task. Low fidelity inspection simulators with computer-generated images to develop off-line inspection training programs have been used by Latorella et al. [3] and Gramopadhye et al.[4] for inspection tasks. Drury and Chi-Fen [5] studied human performance using a high fidelity computer simulation of a PCB inspection task. Kundel et al. [6] have applied advanced technology to the inspection of X-rays for medical practice. However, such tools are often not widely adopted due to the low-fidelity and limited interaction in such simulators. Virtual reality (VR) simulators have been proposed to overcome these shortcomings.

Virtual reality has been defined as a computer simulated 3D environment in which the objects can be manipulated by the user through a standard input device such as a keyboard or mouse. The display technology can range from the computer monitor to specialized head-mounted displays (HMDs). A model of the object is created using a 3D modeling software and rendered on the screen. Due to advances in the computer graphics industry, it is now possible to render and interact with high polygon count objects in real-time (>30fps). Using a VR simulator, we can more accurately represent complex aircraft inspection and maintenance situations, enabling students to experience the real hangar-floor environment. The instructor can create various inspection and maintenance scenarios by manipulating various parameters – for example, defect types, defect mix, defect severity, defect location, defect cues -- reflective of those experienced by a mechanic in the aircraft maintenance hangar environment. As a result, students can inspect airframe structure as they would in the real world and initiate appropriate maintenance action based on their knowledge of airframe structures and information resources such as on-line manuals, airworthiness directives, etc. Their performance in tackling these scenarios can be tracked in real-time with the potential for immediate feedback. Students will be able to grasp the links between various visual cues presented, the need for specific inspection items and potential maintenance solutions. Repeated exposure to various scenarios along with classroom teaching will help them link theoretical scientific knowledge, for example, physical and chemical characteristics of structures, to various engineering solutions. The result is an innovative curriculum application, one in which the student has the added advantage of simulator experience in addition to the theoretical knowledge.

This paper outlines recent efforts at Greenville Technical College and Clemson University on the use of advanced virtual reality technology to upgrade Aircraft Maintenance Technician (AMT) inspection skills, reduce human error during such inspections and ultimately improve aviation safety. This paper is organized as follows: Section 2 details the aircraft inspection training simulators developed for use in this project. Section 3 discusses the process of integrating these technologies into the current educational curriculum, while Section 4 provides further discussion on this process followed by concluding statements.

2 Instructional Tool Development

The integration of technology and theoretical aspects of learning to aid training has been used for decades. The earliest simulators built were for military and research purposes. The computers used to run these simulations were expensive and hence these simulators were limited to research facilities. With the improvements in the commodity graphics market, we can now render more complex environments within these simulators for a fraction of the price. The visual realism and real-time interaction make it feasible to develop a desktop simulator for the aircraft inspection

Fig. 1. Hardware used to develop and test the simulators including the HMD and 6DoF mouse (left), the HMD eyetracking system (center) and the touch screen WindowVR unit (right)

tasks. The current hardware used to develop the simulator (Figure 1) consists of a Dual Xeon processor machine with GeForce 6800 video card and 1 GB of RAM, all for a total cost of less than $3000.

The first aircraft inspection simulator developed at Clemson University was called Automated System of Self Instruction for Specialized Training (ASSIST). The simulator consisted of 2D interfaces and a collection of photographs that were presented with instructions to a novice trainee inspector. The results of the follow-up study [7] conducted to evaluate the usefulness and transfer effects of ASSIST were encouraging with respect to the effectiveness of computer-based inspection training, specifically in improving inspection performance.

To add more realism and interactivity to the inspection simulations, an immersive simulator called INSPECTOR was developed. INSPECTOR uses photographs obtained from an actual aircraft as textures in the cargo bay inspection simulation. The main environment consists of a simple wire-frame cube, texture-mapped to resemble the aft cargo-bay of the aircraft. The simulator along with the real environment it is representing is shown in Figure 2. Note that the simulator does not have any depth cues or shadows due to the use of textures in the image. The software is capable of interacting with a variety of input and output devices from fully immersive head mounted displays, to the keyboard and mouse. The instructor can customize the training scenarios using simple image manipulation software such as Adobe Photoshop. The instructor can modify the textures used in the scenarios to include defects such as cracks, corrosion broken conduits in the simulator. The performance of the subject in the simulator can be recorded and stored for later offline analysis. Process and performance measures such as accuracy, defects detected, defects missed and time to completion can be recorded and the subject can obtain feedback on their performance in the simulator [8].

Fig. 2. Actual cargobay(left) with the texture mapped cargobay(right)

In addition to using texture mapped environments in the simulator, it is also capable of rendering 3D models of the environment. The 3D model of the environment is built using Alias Wavefront Technologies [9] software called Maya and rendered on the desktop using OpenGL [10] and C++. The graphical user interface (GUI) presented to the user was developed using SDL and uses simple scripts to control the behavior of the simulator. Figure 3 shows the rendering of the 3D cargobay. Notice that the use of 3D models leads to a more realistic environment with shadow-casting lights and depth to the environment.

Another important simulator developed during this period is the borescope simulator. The borescope is a tool used to inspect the internal parts of the engine for defects such as cracks, stress fractures and corrosion. The borescope consists of a handheld unit and a long, flexible, fiber-optic cable as shown in Figure 3 (center). The handheld unit consists of a full color LCD screen and a mini-joystick. The fiber-optic cable is connected to the tip of the handheld unit and has a camera and light source at the other end. The joystick controls the articulation of the fiber-optic tip attached to the unit. Preliminary user testing using the standard presence questionnaire [11] showed that the simulator, Figure 3 (right), is visually similar to the actual device [12]. Figure 4 shows the model of the engine blades developed for use in the borescope simulator.

Figure 5 shows the hardware setup of the lab at Greenville Tech with a participant using the inspection simulator. The hardware components include a Virtual Research

Fig. 3. Enhanced VR cargobay model (*left*), actual video borescope (*center*) and virtual borescope simulator (*right*)

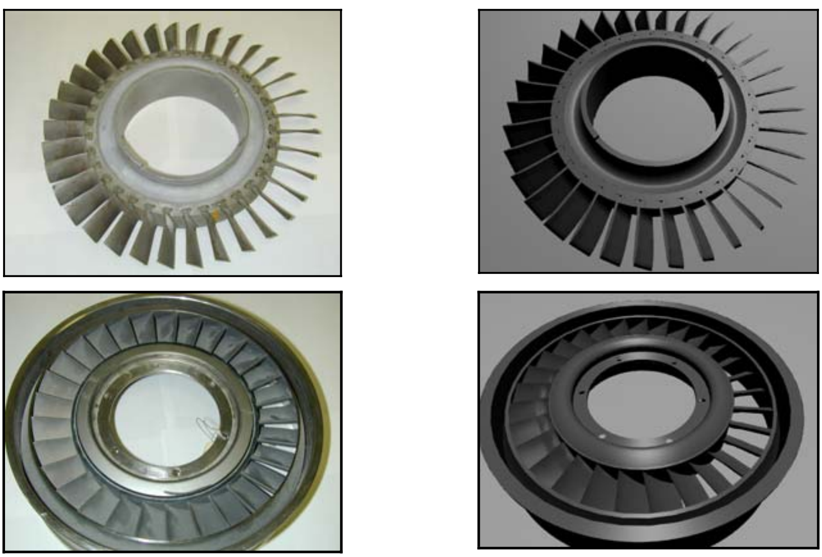


Fig. 4. Engine blades modeled (right) from the actual engine (left)

HMD and 6 degree-of-freedom Flock of Birds tracker integrated with the helmet. User motion in the immersive environment is through button presses in the hand-held 6-DOF mouse. The mouse is also used to select and highlight defects within the environment, which are stored for later offline analysis of the subjectÊs performance in the simulator. The WindowVR consists of a 17ÊÊflat-panel, touch-sensitive display suspended from an immobile stand. The WindowVR consists of a 2-way directional joystick to control the motion of the subject in the VR environment and a 6DOF tracker embedded within the display to obtain the orientation. Notice that the simulator can be run on either the HMD or the WindowVR depending on the level of immersion required and the experience level of the student trainee.

Fig. 5. User testing of the training scenarios; **a:** Window VR, **b:** HMD

3 Curriculum Evaluation

The regulation of Aviation Maintenance Technology schools in the U.S. is overseen by the Federal Aviation Administration (FAA). Current FAA regulations mandate the objectives and minimum requirements of all AMT program curriculums. Subsequently, the design and effectiveness of these curriculums is validated by conformance to published regulations and by annual site visits from FAA personnel. Any change in curriculum, including the addition of new pedagogical tools, must therefore 1) adhere to FAA regulations and 2) be approved by FAA personnel.

To integrate the different training simulators with the Aircraft Maintenance Technology curriculum, it was necessary to devise a curriculum development and assessment plan. A presently accepted form of pedagogical evaluation is Bloom's Taxonomy [14], [15]. This evaluation tool considers the level of learning at which students are required to participate through a given educational material. It argues that there are 3 types of learning: cognitive (knowledge), affective (attitude) and psychomotor (skills). In the case of AMT training, the cognitive and the psychomotor domains will be of primary focus. The taxonomy identifies six learning levels within the cognitive domain, ranging from basic recognition of facts and information to more complex integration of information through judgment and assessment. Similarly, the psychomotor domain includes levels of learning from basic imitation of skills to the ability to adapt movements and skills to new and changing environments.

The methodology used to begin to scientifically and systematically incorporate the VR technologies into the current curriculum is based upon the findings and results of the Bloom's evaluation. Blooms will be used to first identify significant gaps in the original curriculum, and then evaluate the quantitative impacts of incorporating VR technologies into this curriculum. A graphical flow of this methodology is presented in Figure 6. It is hypothesized that the incorporation of the VR training technologies into the original curriculum will help improve the curriculum's capacity to meet the specified learning objectives.

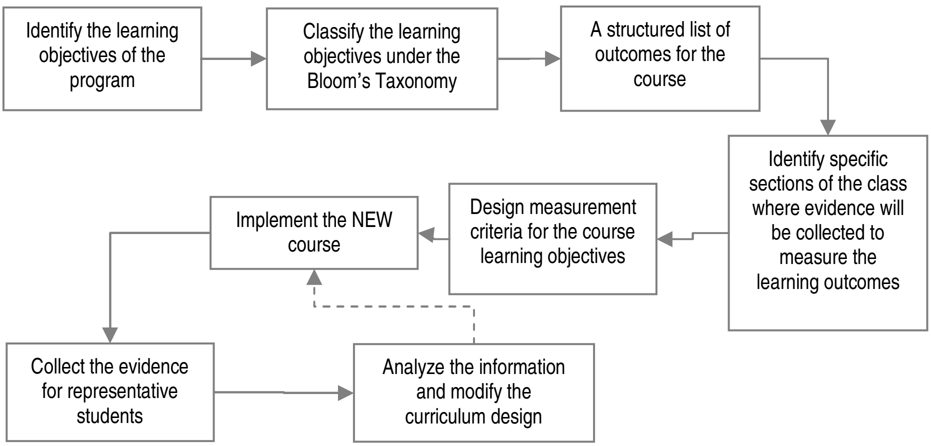


Fig. 6. Methodology for VR – Curriculum integration using Bloom's Taxonomy

Using course objectives, Blooms taxonomy descriptors were developed for a prototypical course, *Airframe Inspection*. These descriptors identified the degree to which each of the course objectives were met using the available curriculum. For example, the original curriculum (without the integration of the computer-based training technologies) met the course learning objective "Exhibits knowledge of

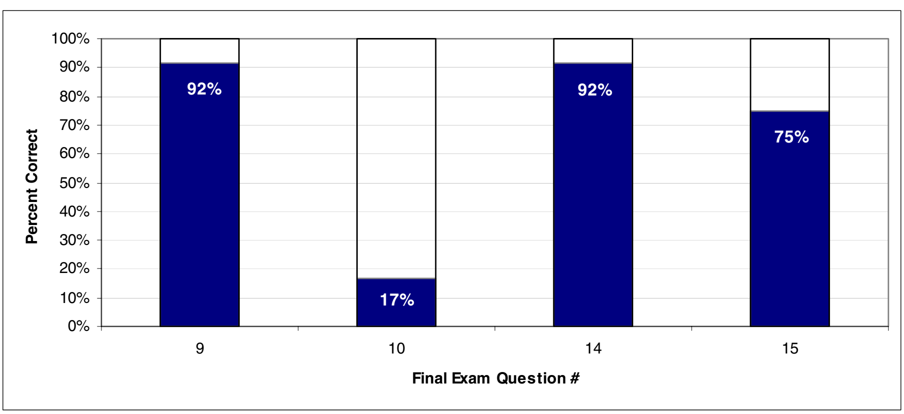


Fig. 7. Aggregate student performance on *Airframe Inspection* Final Exam questions 9, 10, 14 and 15

requirements for complying with airworthiness directives” in the *knowledge* and *comprehension* components of the cognitive domain, but not in the psychomotor domain. The curriculum’s capacity to meet this objective was evidenced by certain questions administered on the course’s final exam. Further, quantifiable data on how well the course met this objective was gathered through student performance on the selected evidence (test questions). The figure below (Figure 7) illustrates students’ performance on the test questions which support learning objective 1 for the course *Airframe Inspection*. Subsequently, it is evident that student performance was lacking in the successful completion of question 10.

4 Summary and Conclusions

Existing multimedia approaches, while valuable, have not been able to mimic accurately the complexity of the aircraft maintenance process, reporting limited transfer capabilities and student preparedness for the workplace. As a result, students trained with this methodology face a steep learning curve in on-the-job situations where mistakes can have potentially catastrophic consequences. Providing students with the total immersion afforded by virtual reality will give them a more realistic, hence accurate, view of the airframe structure and complex inspection/maintenance environment, one that is more effectively internalized and transferred to the workplace. The integration of this technology in a systematic and scientific manner will allow us to study the impact of the different enhancements to the curriculum. This innovative research represents a much-needed first effort to extend tested VR technology to the aircraft maintenance technology curriculum in a two-year college allowing for a better understanding of the use of VR as a pedagogical tool. The successful completion of this research shows promise in fulfilling a state and national need for well-prepared technicians entering the aircraft maintenance industry. Most importantly results of this research can be extended to a broader range of science, mathematics, engineering, and technology areas thereby making significant impact on student learning.

Acknowledgement and Disclaimer. This material is based upon work supported by the National Science Foundation under Grant Number 0302780. ("Integrating Asynchronous Technology and Virtual Reality to Support Education in Aircraft Maintenance Technology"). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. Federal Aviation Administration: Written Comments; Public Responses to the July 6, 1998, Revision of Certification Requirements: Mechanics and Repairmen, Notice of Proposed Rule Making (Notice No. 98-5) FAR PART 66 NPRM (1999)
2. Czaja, S.J., Drury, C.G.: Training Programs forIinspection. Human Factors 23(4), 473–484 (1981)

3. Latorella, K.A., Gramopadhye, A.K., Prabhu, P.V., Drury, C.G., Smith, M.A., Shanahan, D.E.: Computer-simulated Aircraft Inspection Tasks for Off-line Experimentation. In: Proceedings of the Human Factors Society 36th Annual Meeting 1992, Santa Monica, CA, pp. 92–96 (1992)
4. Gramopadhye, A.K., Drury, C.G., Sharit, J.: Feedback Strategies for Visual Search in Airframes Structural Inspection. International Journal of Industrial Ergonomics, 333–344 (1997)
5. Drury, C.G., Chi, C.F.: A Test of Economic Models of Stopping Policy in Visual Search. IIE Transactions 27, 382–393 (1995)
6. Kundel, H., Nodine, C., Krupinski, E.: Computer displayed Eye Pposition as a Visual Aid to Pulmonary Nodule Interpretation. Investigative Radiology 25, 890–896 (1990)
7. Nickles, G., Marshall, J., Gramopadhye, A.K., Melloy, B.: ASSIST: Training program for inspectors in the aircraft maintenance industry. In: International Encyclopedia for Ergonomics and Human Factors 2001. Taylor and Francis, UK (2001)
8. Vora, J., Nair, S., Medlin, E., Gramopadhye, A.K., Duchowski, A.T., Melloy, B.J.: Using Virtual Reality to Improve Aircraft Inspection Performance: Presence and Performance Measurement Studies. In: The Human Factors and Ergonomics Society Annual Meeting Minneapolis, MN (2001)
9. Alias Wavefront, Maya 3D modeling tool (Last accessed 18 January, 2006) www.alias.com
10. OpenGL Graphics API (Last accessed 18 January, 2006) www.opengl.org
11. Witmer, B.G., Singer, M.J.: Measuring Presence in Virtual Environments: A Presence Questionnaire. Presence 7(3), 225–240 (1998)
12. Vembar, D., Sadasivan, S., Duchowski, A., Stringfellow, P., Gramopadhye, A.: Design of a Virtual Borescope: A Presence Study. In: Proceedings of 11th International Conference on Human-Computer Interaction 2004, Las Vegas, NV (2005)
13. Jacob, R.J., Raina, S., Regunath, S., Subramanian, R.C., Gramopadhye, A.K.: Improving Inspector's performance and reducing errors - The General Aviation Inspection Training Systems (GAITS). In: proceedings of the Human Factors and Ergonomics Society Meeting, New Orleans, LA (2004)
14. Bloom, B.S.: Taxonomy of educational objectives: The classification of educational goals. Handbook I, Cognitive Domain. Longmans, Green, New York, Toronto (1956)
15. Anderson, L.W., Krathwohl, D.R.: A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Longman, New York (2001)

Colored Petri Net Based Formal Airport Control Model for Simulation and Analysis of Airport Control Processes

Bernd Werther, Christoph Moehlenbrink, and Michael Rudolph

German Aerospace Center, Institute of Flight Guidance, Lilienthalplatz 7,
38108 Braunschweig, Germany
{Bernd.Werther,Christoph.Moehlenbrink,Michael.Rudolph}@dlr.de

Abstract. The development of the experimental Remote Tower Operation Human Machine Interface and the new Remote-Controller work position is supported by a cognitive work and task analysis (CWA) of the presently existing work environment and decision processes at airport Leipzig. This paper presents a formal approach for the description of the whole Human Machine System. It is shown how the results of a cognitive work analysis on a medium size airport are transferred into a formal executable human machine model for simulating the controllers work processes in relation to the airport processes. The model is implemented with Colored Petri Nets. The mathematical basis of Petri Nets allows a formal analysis of whole systems. Critical system states and inconsistencies in the human machine system are identified through comparison of knowledge states of the controllers with process states of the airport system by using State Space analysis. The represented formal work process model provides a valuable support for the communication between domain experts and system developers.

Keywords: Airport control model, Human Machine System, Colored Petri Net, State Space, Cognitive work analysis.

1 Introduction

Remote Tower Operation (RTO) describes the goal of remote control of small airports and of movement areas of large airports which are not directly visible from the tower. This is achieved by means of an augmented vision video panorama. In 2005 the DLR project RapTOr [1] was initiated in order to realize an RTO experimental system as extension of the Advanced Surface Movement Guidance and Control System (A-SMGCS) at the Braunschweig research airport.

Analysis and simulation of the tower work procedures as introduced in this paper support the design and development of the future tower work positions. The design process for a RTO work environment relies on interviews from domain experts (controllers) of the German Air Traffic Control Organization (DFS), in particular with respect to the work analysis and the validation of the work process model [2]. Initial design goal of RapTOr (Remote airport Tower Operation research) is the integration of the RTO work position into an existing tower work environment of a medium size airport in order to simultaneously control one or more neighboring small airports. The

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 1027–1036, 2007.

project encompasses as a major research and development goal, the simulation of operator decision making within the Tower work positions. The starting point of the model based design process is a formal description of current tower and ground control processes in a generic Petri net model. With the homogeneous description of the human machine system based on Colored Petri Nets it is possible to investigate the consistency of the human and the process model based on formal analysis as suggested in [3].

In section 2 the methodical approach and initial results of a tower work and task analysis is described, following a systematic procedure developed by Vicente [4]. The development of a formal Colored Petri Net model for description of controllers work processes is introduced in section 3, realized with CPN-Tools [5, 6]. The model serves for simulating the operator's decision making processes with the work analysis data as input. Section 4 presents first steps in validation of the work process model. Section 5 provides a conclusion and outlook with regard to the design process of future RTO work position.

2 Methodical Approach

The design and development of the experimental RTO HMI is supported by a cognitive work and task analysis (CWA) of the presently existing work environment and decision processes on airport Leipzig. The formalized results serve as input data for a human machine model for the simulation of the controller decision making processes at the tower work positions. This section introduces the work analysis method and the work process model approach.

The CWA is based on a formal procedure suggested by Vicente [4], separating the analysis into five phases: (1) work domain analysis, (2) control task analysis, (3) strategy analysis, (4) analysis of social organization and cooperation, (5) operator competency analysis (the latter however not being considered in this phase of the RapTOr project). Details of the CWA as well as the cognitive modeling and simulation using Colored Petri Nets are described in [4, 7 and 8].

The work domain analysis aims at analyzing the aircraft movements. For this purpose the air-to-air process which describes the complete movements from arrival to departure is treated separately for the different control areas (e.g. approach, runway, taxi and apron). Acquired information and possible actions are attributed to corresponding control areas. Accessible information from the different sources (e.g. visual view from the tower windows, approach radar) and possible actions via the corresponding interaction devices (e.g. radio, telephone) is acquired without considering the controllers tasks in this first phase. In the control task analysis phase the tasks are identified which have to be completed. Here decision and support processes are treated separately. The task description follows a well defined structure which covers the triggering event, the preconditions, the task containing coordination, and the post-condition. The strategy analysis is the most laborious phase. This is because controllers to a large extent use implicit knowledge which is hard to extract. In an empirical study [9] Sperandio e.g. detected strategy differences, dependent on workload. The development of strategies depends extensively on the handling of goals under restricted cognitive resources [10]. This is one important motivation for the resource based Petri net modeling technique. An important aspect of multiple task situations as typical for controllers is the relative weighting of different simultaneous goals with respect to each other. Action strategies evolve due to limited human processing capacity [8]. Phase 5, the analysis of cooperation and social organization yields a clear tasks and functions allocation for the two controllers (ground controller (PG), tower controller (PL)) within the current air traffic control procedures. The future RTO workplace, however, represents a significant change of this situation. On the one hand the augmented vision video panorama offers revolutionary new possibilities for the support of air traffic controllers. On the other hand the integration of remotely located control areas within the present day tower environment represents a completely new work condition.

3 Formalization of Cognitive Work Analysis Results with Colored Petri Nets

In this section the different levels of the Petri net architecture resulting from the work analysis process will be introduced in detail. To graphically indicate the hierarchical structure and different levels of the model, transition framed sub networks are replaced by transitions with light grey boxes (see Fig.1) on next higher hierarchy level. Places which are connected with a higher hierarchical level are marked with grey boxes on the lower network level (Fig.2).

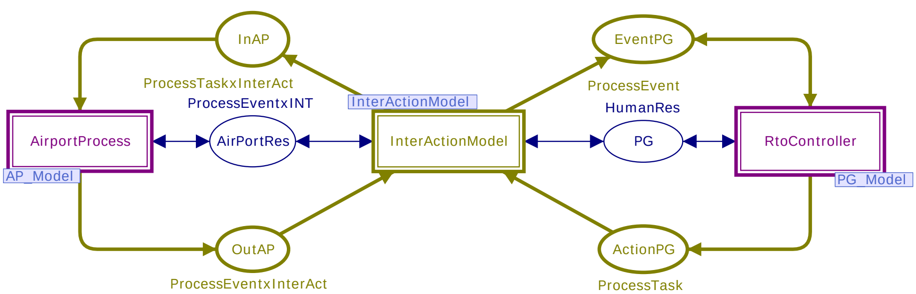


Fig. 1. Controlled airport system as Human Machine System is mapped on the highest hierarchy level of a Colored Petri Net. Transitions (rectangles) with grey boxes include subnets.

3.1 Cooperation Between Process Model and Controller Model

The results of the fourth CWA phase (cooperation) are fed into the highest hierarchical level of the CPN structure. Here the distribution of roles and functions among the different human operators and their technical support systems is defined. On this level the work process is described in a holistic manner whereas on the lower levels focus is put on the single work positions. Following an approach of Cacciabue

[11] a human machine model can be separated into three different Model types: Human Model, Interaction Model and Machine Model. An Interaction Model manages the connection between Human Model(s) and Machine Model(s). So Human Model(s) and Machine Model(s) can work independently from each other for some time periods. Fig.1 shows the realization of the airport control process with a focus on RTO controller work positions. The transition framed subnet *RtoController* describes the cognitive behavior of the RTO controller. The Interaction Model (*InterActionModel* as replacement transition, Fig.1) defines the controller interactions and includes networks for description of information resources, such as radio and far view.

The airport process model (*AirportProcess*, in Fig.1) describes the movement process of aircrafts. The state of the process model determines the type and content of information which can be acquired by the controller. This can be perceptible information about the status of airport resource usage (token on place *AirPortRes*) such as usage of taxiways or the occurrence of a specific event like aircraft is started up. Other possible information would be a request of a pilot via radio like *request start-up clearance*.

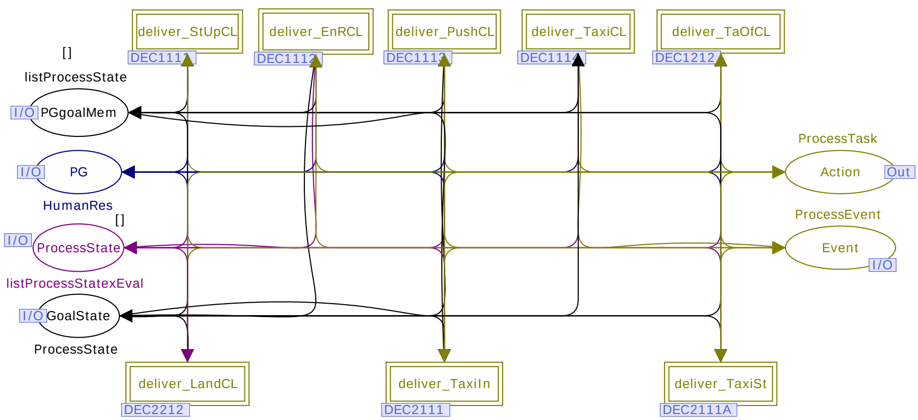


Fig. 2. In the mental model implemented tasks

3.2 Formal Description of Tasks

On the next lower level of the controller model (*RtoController* in Fig.1) within the Petri net architecture knowledge representation about the airport process (mental model) and the goal driven actions fulfilled by controllers are implemented. Control tasks are described explicitly and in great detail in operation instruction manuals of Air Traffic Control (ATC) organizations. There are also general rules for handling of more than one task or task object, such as the first come first serve rule. Depending on the situation the task coordination could be very different.

Fig.2 shows the tasks which are implemented in the model. In the current version of the FAirControl model the following tasks are implemented: Deliver (Start-up, Push back, En-route, Taxi (In), Taxi (Stand), Taxi (Out), Landing and Take Off) Clearance. This is the lowest set of tasks for description of the reduced inbound-outbound process on an airport. These transitions in Fig.2 labeled with light grey boxes represent those tasks and are sub nets on the next lower hierarchy level. The execution and processing of tasks are controlled by the active goals as tokens on the black place *GoalState* and knowledge about the current airport process state *ProcessState* on the purple place and the detected process events (place *Event* on the left side of Fig. 2).

The knowledge and information about the current work process (mental states) are represented by tokens. The results of the task processing would be situated on the place *Action* on the right side of Fig.2. The place *PG* considers the used human resource RTO controller which is responsible for these tasks. Different tasks in Fig.2 will be described on the next lower level by subnets. As an example the sub net within replacement transition *DEC1111* is shown in Fig.3.

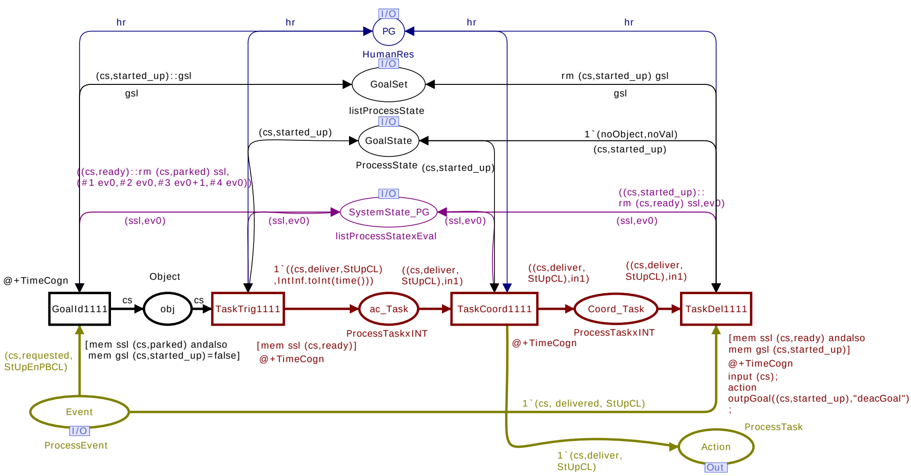


Fig. 3. Description of decision and support tasks, e.g. DEC1111 in Fig.2: task *deliver StartUp clearance*

The control tasks identified in CWA phase two are modeled with regard to the actions to be performed and the required and created information. As mentioned above control tasks are depicted in operation instruction manuals of Air Traffic Control (ATC) organizations. Generally decision and support tasks can be distinguished. Tasks are separated into those representing preconditions, others which coordinate the task (decision process, support process) and those actions which complete the task and lead to post conditions [12].

Fig.3 shows the net which is contained in transition *DEC1111* (Fig.2) for description of task *deliver start-up clearance*. By switching the transition *GoalId1111* the goal to deliver the start-up clearance will be added to the goal set. The task can be triggered if the transition *TaskTrig1111* fired by internal intended goal states, identified system states or when external events are detected in the airport process (cp. Fig.3). The Task from type *1`(LH120,deliver,StUpCL)* will be activated. This means that the Start-Up Clearance should be delivered to the pilot of an aircraft. If the transition *TaskCoord1111* fires all actions which are necessary for fulfilling the task are prepared. Through processing of these actions the modeled controller collects information about the work process und tries to find out if the conditions for executing the decision tasks are given. These actions are processed on next deeper level (action level) in the hierarchy. The lower hierarchy levels are not significant for this paper. More information is found in [12].

4 Identification of Controllers' Strategies Supported by Simulations

The work process structure introduced in this paper was developed together with domain experts in the CWA. During first validation steps based on defined scenarios the correctness of the logical process flow and ability of the controller models to separate the aircrafts was successfully tested in the model. In the following section the logic flow between work process model of the controllers and simulated airport process model are described. This is followed by the representation of simulations using more realistic traffic flows which should be controlled and guided. These generated scenarios are the basis for a model review and model refinement with domain experts. For a better communication with the domain experts the graphical visualization of the airport processes for a generic airport was realized, as shown in Fig. 4.

The next section starts with a short introduction of the FAirControl visualization, followed by the first results of an evolution of specific traffic scenarios.

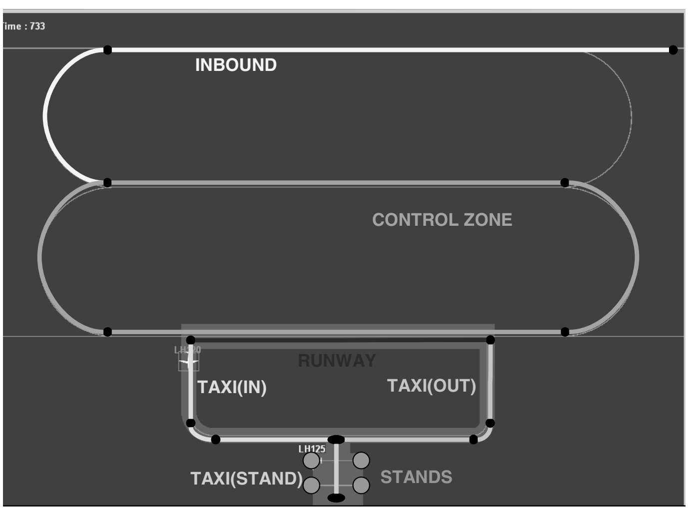


Fig. 4. Graphical visualization of the simulated controlled inbound/outbound sequence of a generic small airport with two controlled aircrafts

4.1 Visualization of Controlled Airport Processes

The executable model introduced in this paper should support the identification of controllers' strategies in organization of task and pursuance of goals. The traffic flow as represented in the airport process model has to move through different control areas which are inbound area, control zone, runway area, taxi fields and stand, each of those having a certain resource limitation. Fig.4 shows the control areas for a generic small Airport. An aircraft movement within the airport area is described by the following sequence: approach, landing, taxiing, going around, push-back, taxiing, take off and departure. The aircrafts (pilots) described in the process model need clearances from the controller to pass through the different control areas. So the controller model has to collect information about the airport resource usage via observation of the movement areas and also by messages and requests via radio (Fig. 5). These two channels are realized in the interaction model and enable the communication between airport process model and controller model.

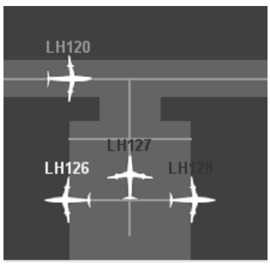


Fig. 5. Represents the graphical visualization of the pilot communication. The representation of label *LH127* (dark grey) shows, that the pilots send a request to the controller. The label *LH120* (grey) visualizes a message and *LH126* (light grey) represents no communication activity.

4.2 Evaluation of Specific Traffic Scenarios

After implementing the work analysis results into a Petri net the FAirControl visualization was used to discuss the correctness of the implementation with two domain experts. 14 static scenarios of the FAirControl airport, representing different possible global states of this microworld were presented to the controllers. Scenarios were discussed in respect to closeness to reality and task sequences. In order to find out about task sequences the experts anticipated the next four clearances that would be handled, assuming normal operation within each scenario. To give an example scenario 11 is shown in figure 6. First airplane LH125 will receive a take-off clearance, before LH132 will receive a taxi-in clearance. Then LH122 or LH124 will be cleared to be pushed before LH123 can land.

An insufficient understanding of information processing, memory, behavioral strategies and decision making for tower controllers is claimed in the literature [13]. Interviewing the tower controllers about action possibilities within each scenario is one first step to get an idea, how they interact with their environment. Going back to

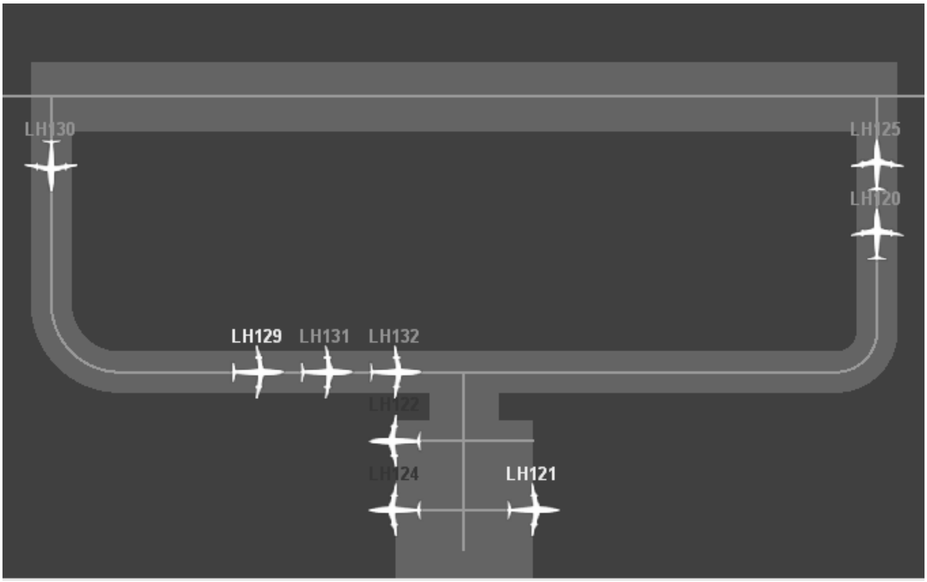


Fig. 6. A Cut-out of scenario 11: The airplane will be handled in the following order: 1) LH125 2) LH132 3) LH122/LH124 4) LH123 (not present in this cut-out)

the interviewer and discussing ones' understanding of the work analysis enables an iterative process in model development. Therefore we can conclude that a subject executing the FAirControl Model has several degrees of freedom to interact with the microworld comparable to a simplified tower control position.

The advantage of a Petri Net based mircoworld lies in the comparison of State Space results for each scenario with the subjects' behavior within the mircoworld. On the one hand the circumstances under which a controller gave a clearance can be reported. In addition, all action alternatives within the mircoworld can be depicted by the State Space. Figure seven represents part of the State Space for scenario 11, which was discussed with the controllers. A resource conflict is apparent on the apron area. LH132 asks for this resource to taxi-in, while LH122 and LH124 intend to use the same resource for push-back and taxi-out. Figure 7 enables an evaluation which clearance has to be given first, to get an optimal throughput in respect to time. To clear LH132 (taxi-in apron) first will result in a delay of 35 TS (timesteps) for LH124, but LH132 will reach the state "taxied-out" without delay (first case: t_{LH124}=110 TS, t_{LH132}=380 TS). Clearing LH124 (push-back, taxi-out) first is optimal for this airplane, but LH132 then will be delayed by 70 TS (second case: t_{LH124}=75 TS, t_{LH132}=310 TS).

Time critical decisions in respect to the most efficient throughput always occure, when the tower controller has to deal with resource conflicts. The formal State Space analysis offers an objective perspective to find out about optimal task sequences to handle the airport processes. For our example scenario 11 we can conclude, that the interviewed controllers suggested the best solution.

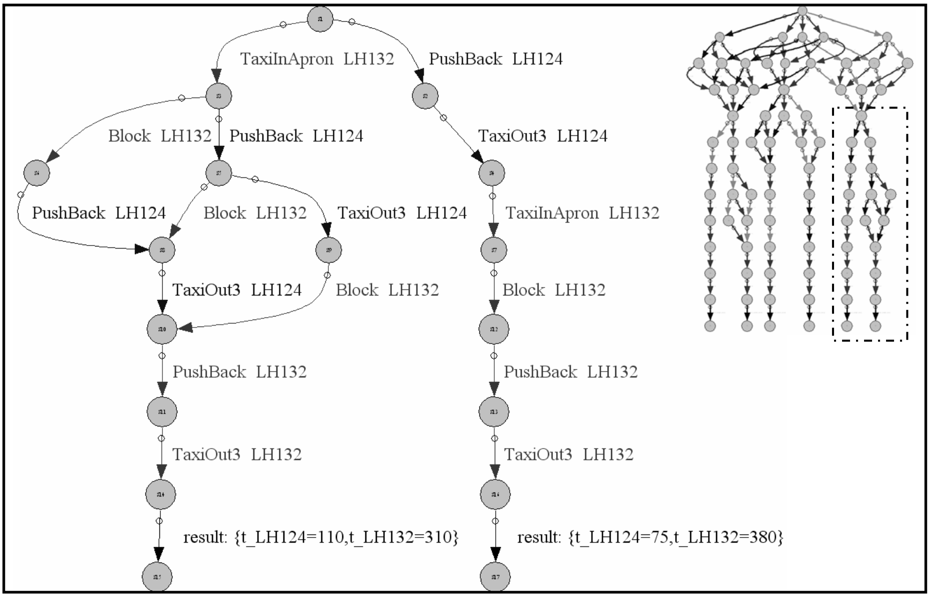


Fig. 7. The small State Space graph on the right side presents the possible traffic flow alternatives in scenario 11 (Fig. 6). The reduced State Space on the left side shows the conflict of two aircraft LH124, LH132 by usage of the same resource *Apron* as presented in scenario 11 more in the detail.

5 Conclusion and Outlook

The most important information source of each tower controller is currently the visual view out of the tower window. Under remote tower operations this critical information source will likely be replaced by an augmented video panorama. However, especially augmented tower vision (ATV) concepts, i.e. the superposition of additional information on the far view like weather data and aircraft labels with object tracking also raise questions regarding controllers attention and perception processes.

With the homogeneous description of the human machine system based on Colored Petri Nets it is possible to investigate the consistency of the controllers' strategies and actions and of the process model of airport movements. This is achieved by a formal analysis and comparison of the State Spaces of the internal mental representation of airport process in the mental model of the controllers on the one hand and simulated airport process in the identically named model on the other hand. Critical work situations can be detected and analyzed in an early phase of the system design and alternative solutions can be investigated by means of model based simulations.

The transfer of the CWA results into the formal controller models was introduced in this paper. The resulting model represents a basis for monitoring psychological parameters of the operators (e.g. work load), for deriving the operator requirements and for uncovering missing situational awareness by means of simulations and State

Space analysis in the future by using the method presented in [8]. Furthermore the graphically represented formal work process model provides a valuable support for the communication between domain experts and system developers.

Acknowledgments. We are indebted to Detlef Schulz-Rückert, Holger Uhlmann and Dieter Bensch of German Air Traffic Control (DFS) for continuous support of the Virtual Tower and Remote Tower Operation projects during the recent years, in particular for providing detailed domain expert knowledge for the tower work analysis.

References

1. Schmidt, M., Rudolph, M., Werther, B., Fürstenau, N.: Remote Airport Tower Operation with Augmented Vision Video Panorama HMI. In: Proc. 2nd Int. Conf. Research in Air Transportation ICRAT 2006, Belgrade, Serbia, June 24.28, 2006, pp. 221–230 (2006)
2. Werther, B., Uhlmann, H.: Ansatz zur modellbasierten Entwicklung eines Lotsenarbeitsplatzes. In: Zustandserkennung und Systemgestaltung, Fortschritt Berichte VDI, vol. 22, pp. 291–294 (2005)
3. Degani, A., Heymann, M.: Formal verification of human-automation interaction. Human Factors 44, 28–43 (2002)
4. Vicente, K.J.: Cognitive Work Analysis. Lawrence Erlbaum Associates, Mahwah/NJ (1999)
5. Jensen, K.: Coloured Petri Nets. Springer, Berlin (1997)
6. CPN Tools: http://wiki.daimi.au.dk/cpntools/cpntools.wiki
7. Werther, B., Schnieder, E.: Formal Cognitive Resource Model: Modeling of human behavior in complex work environments. In: Proc. Int. Conf. Computational Intelligence for Modelling, Control & Automation, CIMCA 2005, Wien, pp. 606–611 (2005)
8. Werther, B.: Kognitive Modellierung mit farbigen Petrinetzen zur Analyse menschlichen Verhaltens, PhD Dissertation, DLR-Inst. of Flight Guidance and Technische Universität Braunschweig (2006)
9. Sperandio, J.C.: The regulation of working methods as a function of work-load among air traffic controllers. Ergonomics 21(3), 195–202 (1978)
10. Dörner, D.: The Logic of Failure: Recognizing and avoiding error in complex situations. Metropolitan Books, New York (1996)
11. Cacciabue, P.C.: Modelling and simulation of human behaviour in System control. Springer, London (1998)
12. Werther, B.: Colored Petri net based modeling of airport control processes. In: Proc. Int. Conf. Computational Intelligence for Modelling, Control & Automation, CIMCA, Sydney (2006)
13. Hagemann, T.: Belastung, Beanspruchung und Vigilanz in der Flugsicherung unter besonderer Berücksichtigung der Towerlotsentätigkeit. PhD Dissertation, Europäische Hochschulschriften, 2000, Reihe 6, Psychologie; Bd. 664 (2000)

Research on Modeling of Complicate Traffic Simulation System

Jiankun Wu[1], Linpeng Huang[1], Jian Cao[1], Minglu Li[1], Dejun Wang[1], and Mingwen Wang[2]

[1] Department of Computer Science, Shanghai Jiaotong University, Shanghai, 200240, P.R. China
[2]School of Computer Information Engineering, Jiangxi Normal University, Nanchang 330027, P.R. China
jkwu@sjtu.edu.cn, {huang-lp,cao-jian,li-ml}@cs.sjtu.edu.cn, wangdejun@sjtu.edu.cn, mwwang@jxnu.edu.cn

Abstract. With the increasing of traffic complexity, traffic simulation becomes an important approach to deal with the complicated traffic problems; meanwhile, system modeling plays a more and more important role in the simulation systems. Cellular automata provide a simple discrete deterministic mathematical model for physical, biological, and computational systems and are shown to be capable of complicated behavior and generate complex and random patterns, which are very suitable for the description of complicate traffic environment [7]. A simulation model based on agent technology, HLA/RTI technology and expanded cellular automaton is presented in this paper. The simulation model makes the platform expandable and flexible, at the same time, it can provide high-capable computing resources to solve the complex traffic issues. In the traffic entity model aspects, the expanded cellular automata and agent technology were adopted to model the behaviors of passengers, vehicles, traffic signal lights and so on. The optimal scheme for evacuation of traffic disaster, superintendence of large scale activities and design of traffic environment will be obtained through the simulation model.

Keywords: Traffic simulation, Agent, HLA/RTI, Complexity system, Cellular automata.

1 Introduction

With the traffic complexity rapidly increasing, it becomes more and more difficult to obtain the solution of traffic problems in reality, so traffic simulation becomes an effective tool for analyzing traffic states and problems. Based on simulation results, good decisions are made in traffic project and effective traffic guidance are designed.

Currently, traffic simulation research has been carried out in some countries, such as American, England, Germany, Japan and so on [6]. Although with respective advantages on special areas, it is difficult for users to customize the software per their

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 1037–1046, 2007.

own needs. Especially, it is not enough in the research on simulation of evacuation of traffic disaster and superintendence of large-scale activities aspects. From the point of view of resource sharing and inter-operation, the description of traffic entity behavior and inter-operation between the sub-systems are inadequate. To solve the above problems, it is significant to propose a new simulation model to make simulation systems flexible and reused.

The accurate description of traffic entities behaviors and system architecture design play an important role in the traffic simulation. In complicated traffic environment, traffic systems become nonlinear due to the existence of random factors like persons. Currently, cellular automata are used as an effective method to solve self-organization and non-balance problems. Any system with many identical discrete elements undergoing deterministic local interactions may be modeled as a cellular automaton. Many complex systems including traffic system can be broken down into identical components, each obeying simple laws. These components that make up the whole system act together to yield very complex behavior [7]. So it is significant to adopt cellular automata with some additional conditions to build abstract traffic simulation model.

In order to make the simulation application and RTI (run-time infrastructure) independent and developed separately, the HLA (High Level Architecture) [1] is selected as the simulation architecture. However, it is somewhat difficult to simulate complicated objects like persons just using cellular automata. Agent is characteristic with autonomy, adaptability, and intelligence, so it is an effective tool that simulates the living objects and other objects that conforms some changing laws [2][15]. That agent behavior based on expanded cellular automata model is selected in this paper is due to two factors: (1) Cellular automata is a simple model that can generate pseudorandom patterns and can be efficiently implemented by hardware. The research on traffic flow and passenger flow based on cellular automata has obtained some useful theoretical models and related analysis; (2) Agent is more flexible and convenient than cellular automata in art intelligence. The model in this paper uses cellular automata to describe the rules of traffic environment and adopts the agent to implement the traffic entities; as a result, the factuality had been improved greatly; (3) Since the behavior description of traditional cellular automata is too simple to describe non-linear behaviors like person's activities, friction and repulsion are introduced in cellular automata to simulate traffic entities more accurately. The simulation technology combined with cellular automata, agent and HLA is effective to simulate the complicate micro traffic.

This research is based on Shanghai Traffic Information Service Grid and the simulation platform and the model built in the project can efficiently solve the problems confronted in traditional simulation systems. The optimal scheme for evacuation of traffic disaster, superintendence of large-scale activities and design of traffic environment will be obtained through the simulation model. The remainder of this paper is organized as follows: In section 2 the system architecture model is described; Section 3 discussed the traffic entity model in detail; Section 4 given the verification of the model; Finally, the summary and future work are given.

2 Simulation System Model

The simulation model presented in this paper is focus on the evacuation of traffic disaster and superintendence of large-scale activities. In the micro traffic simulation, it must pay attention to not only the whole traffic flow state but also the model of traffic entity behavior such as vehicle, passenger, signal light, control center et al. In the following section, all subjects will be discussed.

2.1 System Architecture

Due to the huge number of roads, vehicles and passengers, distributed simulation architecture based on HLA is adopted to improve the expandability and flexibility. In complicated traffic environment, traffic entities have characteristics of autonomy, adaptability, intelligence, so agent is adopted to implement these entities. With HLA and agent focus on different aspects, the simulation application sub systems act as plug-and-play components and the RTI acts as the soft bus.

To describe the model in more details, the simulation architecture and layered sub simulation system will be described respectively. The general architecture of model is showed in Fig. 1.

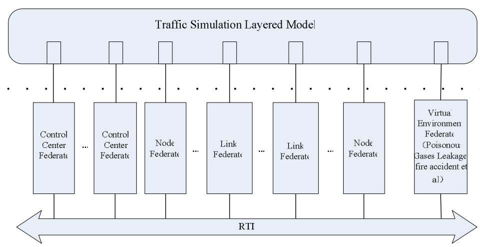


Fig. 1. System simulation architecture

As shown in figure 1, RTI is the run-time infrastructure of HLA. The federation that includes many federates is above the RTI. The federate presented in this paper mainly include control center, road (node federate, link federate) and virtual environment which simulate the poisonous gases leakage and fire accident et al. Publish/ subscribe mechanism is adopted as communications between different federates. The top is the concrete simulation application system.

2.2 Layered Model

The layered micro traffic simulation model is the implementation model of the simulation application system. The federation in the general architecture is the high level abstract of it. The layered model mainly include AE (Agent Entity) layer, road layer and control layer and is shown as Fig. 2.

As shown in figure 2, AE layer represents the traffic entities such as vehicle, passenger and signal light and so on. Those entities have characteristics of autonomy, adaptability, intelligence, which should be considered in the simulation.

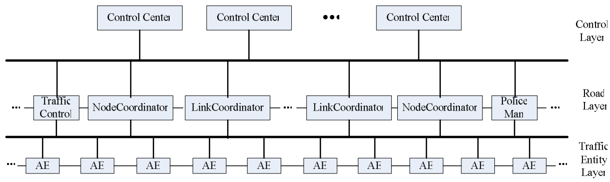


Fig. 2. Layered micro traffic simulation model

The road layer is the most complicated layer, which is responsible for the following: (1) collect the information from the AE layer. Link Coordinator will collect simulation results of link entities such as speed, traffic flow, density, queue length of road cross and so on, Node coordinator will record the crossing time delay. All the information will be transmitted to higher federate through asynchronous communication and exchanged between the federate at the same layer. (2) Another function is responsible for transferring the signal light state information. Signal lights are half- autonomous, so the signal lights are implemented as agent, which can adjust the rule according to traffic states or change the light color per commands from the control center. (3) The third special function is to convey the policeman's command to traffic entities and report traffic environment information to policeman.

The top layer is control layer, which is used to monitor and control the traffic state. It can adjust the signal lights' rule according the traffic flow and inform policeman when traffic accident happens.

2.3 Communication Model and Time Management

There are two communication styles in the model: (1) communications between different layers and different objects at same layer will conform to the agent communication protocol (ACL); (2) Communications between different federates and simulation applications will use the HLA/RTI, which includes synchronization in both time management and simulation result information between federates, the latter is achieved by publish/subscribe style [1].

Time management plays an important role in the simulation, it's important to keep simulation time consistent between components to get accurate simulation results. In our model, HLA implements time management and use Time-Stepped advance style. The simulation procedure will advance to next step when all the simulation computing complete and step length could vary between different steps, it showed as figure 3.

2.4 Abstract Agent Model

An abstract agent model was presented to describe traffic entities and makes the development of agent. The architecture is shown as figure 4.

Abstract model works as following: (1) acquire the environment information through communication; (2) convert the information to the knowledge that could be understood; (3) input the knowledge and self-desire to the BDI reasoning machine

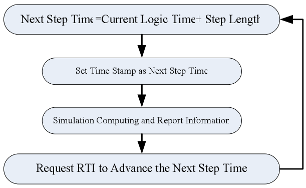


Fig. 3. Advance strategy based on time step

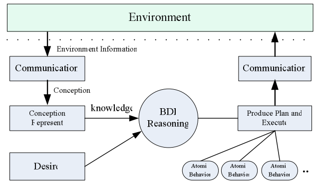


Fig. 4. Abstract model of Agent

[8]; (4) BDI reasoning machine will give a plan which includes a sequence of agent atomic behavior; (5) execute the plan and report the simulation state.

BDI reasoning machine is the core and activity engine of agent. The autonomy, sensation and intelligence characteristics of agent depend on the BDI reasoning machine. The BDI reasoning module makes use of the interior knowledge, self-desire and environment information based on CA to produce a behavior sequence called a plan. CA abstract model will be discussed in section 3.

3 Abstract Theory Model of Traffic Entity

3.1 Introduction of Model Context

The understanding of the traffic entities' behavior holds the balance in building perfect evacuation model and rules. We refer to several related models [9][10][12][13][14] to improve the accuracy of the CA evacuation model. The friction and repulsion are introduced into the CA traffic urgency evacuation model; it is shown as figure 5 and figure 6.

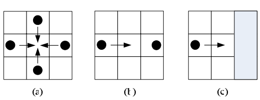


Fig. 5. Three repulsion type

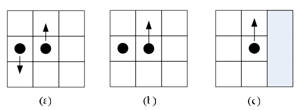


Fig. 6. Three friction types

In the model, complicate traffic environment is divided into uniform grids and each traffic entity's moving obeys simple local rule. The traffic entity is implemented as agent and the behavior of it is adjusted depending on surrounding environment.

3.2 Cellular State Transition Model

Every cellular is occupied by the traffic entity or idle, some traffic entities can occupy more than one cellular. The traffic entity can move to four neighbor cellular with some probability or stay in the original cellular, as shown in figure 7 (downward direction is the progress). When the four direction probabilities are not equal, the phenomenon is called moving with tendency and the entity is called tendency entity.

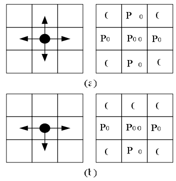


Fig. 7. Mobile styles of traffic entity in cellular automata

In this paper, Von Neumann Neighborhood is used to represent entities' visual field. Each entity tries to reach the destination as quickly as possibly and always avoid the congestion place. We present a hypothesis that traffic entity would abandon competition and select the hypo-optimal place when there are more than two entities to compete to enter same place.

The cellular state aggregate is S and the element is $s_{i,j}, i \in [0, N], j \in [0, M]$, N and M are the grid number of abscissa and ordinate, i, j are the subscript coordinates of abscissa and ordinate. $s_{i,j}$ is an five-unit group:

$$s_{i,j} = (is\_occupied, p_{i,j-1}, p_{i+1,j}, p_{i-1,j}, p_{i,j+1})$$

$is\_occupied$ is a flag indicate if the cellular is occupied. $p_{i,j-1}$ is the probability that traffic entity enter the cellular from the right direction and other three probabilities in the group are similar. The state of each cellular changes as the following format:

$$s_{i,j}^{(t+1)} = f(s_{i-1,j}^{(t)}, s_{i,j-1}^{(t)}, s_{i,j}^{(t)}, s_{i+1,j}^{(t)}, s_{i,j+1}^{(t)})$$

Since the number of traffic entity is very huge and the areas occupied by vehicle and passenger are different, the cellular number occupied by different traffic entities are different. The function is implemented by agent in this research project.

3.3 Mobile Probability of Traffic Entity

As showed in Fig. 6(a), when m(m≥2) traffic entities move to the same object, the repulsion rate which presented by Kirchner et al[11] was considered and the moving probabilities of traffic entity is determined by following formula:

$$p_i = (1 - r_1)/m \qquad (i = 1, 2 \ldots m),$$

The traffic entity stays in the same cellular with probability r_1 and forward to another cellular with probability $1 - r_1$.

For the case of figure 5(b) and figure 5(c), because it is similar to (a) and just the speed of traffic entity and traffic environment obstacle are different, we adopted following formula to determine the moving probability:

$$p_i = 1 - r_2$$

In the formula, the traffic entity stays in the same cellular with probability r_2 and moves with probability $1 - r_2$.

Similarity, we use the following formula to determine the traffic entity moving probability of figure 5(c):

$$p_i = 1 - r_w$$

It states that the traffic entity remains the same point with probability r_w and moves forward to next point with probability $1-r_w$. If the value of r is determined, we could quantitively analyze the repulsion between traffic entity and traffic entity or between traffic entity and traffic environment obstacle.

In order to quantitively analyze the effects of friction, we introduce the friction probability f. In case of figure 6(a), the moving probability is defined as formula :

$$p_i = 1 - f_1$$

It shows that traffic entity will stay in the same point with probability f_1 and move to the next point with probability $1-f_1$.

In the case of figure 6(b), because the traffic entity is affected by the static object's friction, the moving probability is defined as following formula:

$$p_i = 1 - f_2$$

It shows that traffic entity will stay in same point with probability f_2 and move to the next point with probability $1-f_2$.

As considered for figure 6(c), the traffic entity is affected by the static obstacle's friction, the formula for the moving probability will be defined as:

$$p_i = 1 - f_w$$

It means that the traffic entity will stay in the same point with probability f_w and move to the next point with probability $1-f_w$.

Based on the above analysis, we can conclude that if the friction probability f is determined, the affect of friction could be quantitively analyzed.

3.4 Selection of Repulsion Probability and Friction Probability

The existence of repulsion is attributed to the traffic entity's behavior of avoiding hurt. What's the extent depends on the relative speed between traffic entities or between traffic entity and environment obstacle. The traffic entity would be hurt more heavily when relative speed increases and it will tend to avoid with higher probability. So the repulsion probability increases direct proportion with repulsion and will go up when relative speed increases. There is an entity sustainment limitation and once the limitation exceed, the repulsion probability will tend to be 1. As considered that avoiding behavior is reaction of nervous system, the well known Sigmoid function of artificial neural network is adopted to define the repulsion. The repulsion is defined as formula:

$$r = \frac{1 - e^{-\alpha V}}{1 + e^{-\alpha V}}$$

V is relative speed and defined as follows: when there are multi traffic entities moving to each other: V= 2*v; in case of one moving and another static: $V = v$; in case of moving traffic entity and environment obstacle: $V = v$. $a \in [0, \infty]$ is a hardness coefficient and it shows the hurt extent of entity to entity or environment obstacle to entity, such as the wall hardness and roughness.

Friction depends on the contiguity extent, relative speed and friction coefficient between traffic entities or between traffic entity and environment obstacle. In the CA model, because the cellular is uniform, the contiguity content is same when the traffic entities are neighbors. So the friction is defined as formula:

$$f = \theta * V$$

V is relative speed and defined as follows: when there are multi traffic entities move to each other: V= 2*v; in case of one moving and another static: $V = v$; in case of moving traffic entity and environment obstacle: $V = v$. $\theta \in [0, \infty]$ is friction coefficient. It reflects the friction extent between traffic entities or between traffic entity and environment obstacle.

4 Experimental Verification

The simulation platform is based on the grid environment. In our simulation environment, the hard resource includes the Itanium Servers, IBM690 and IBM E1350 for complicate computation. We adopt GLOBUS Toolkit 4.0[5] as the grid middleware to build the runtime environment and implement the grid service conform to WSRF [4]. The entity environment is showed in figure 8. When the environment is built, all the resources include the agent platform register to the WS MDS [5] server.

Jade [3] platform was adopted to implement the traffic entity agent and all the implementations are using java programming language.

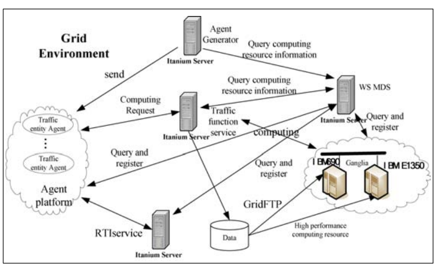


Fig. 8. High performance grid-based simulation environment

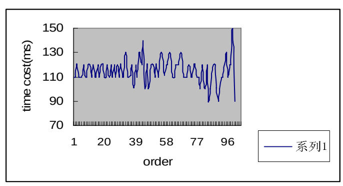


Fig. 9. Synchronization time performance test

In order to test the architecture proposed in this paper, the synchronization time performance test was given and one thousand traffic entities distributing in fifteen sub areas were considered in the test. The computing was carried by ten experimental nodes and the performance curvilinear is showed in figure 10. As showed in figure 9, every synchronization delay time is about 0.1 second. It is satisfied the traffic simulation requirement. From the curvilinear of figure 10, we could conclude that the distributed environment improves the performance greatly.

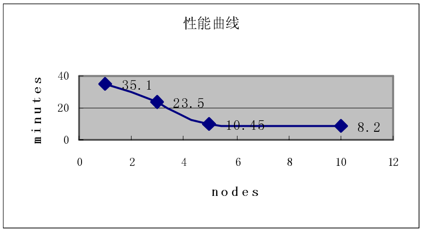


Fig. 10. Performance curvilinear

5 Summary and Future Work

A simulation model of complicate environment is presented in this paper. HLA/RTI and agent technology are the implement carrier of the model and the BDI reasoning machine of agent uses the expanded CA as theory base. Evacuation of traffic disaster, superintendence of large scale activities and design of traffic environment could be solved by the simulation model. Finally, a grid-based simulation environment is given with shanghai south railway station as the simulation background; the model is validated by the simulation platform and obtains well effects.

The future work we plan to do is to enrich the CA-based model; in additional to friction and repulsion, we will consider other factors. We could improve the simulation veracity through the enriched CA-based BDI reasoning machine.

Acknowledgments. This paper is supported by Grand 60673116 of National Natural Science Foundation of China, Grand 2006AA01Z166 of the National High Technology Research and Development Program of China (863), Grand SYSKF0607 of the Foundation of Laboratory of Computer Science, Institute of Software, Chinese Academy of Science.

References

1. IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA)-Frame and Rules [S], IEEE Std 1516-2000, September 2000 (2000)
2. Pan, X., Han, C.S., Law, K.H.: A Multi-agent based Simulation Framework for the Study of Human and Social Behavior in Egress Analysis (2006) http://eil.stanford.edu/ egress/ publications/ASCE_2005_Pan_Han_Law.pdf
3. Jade - Java Agent Development Framework (2006) http://jade.tilab.com/
4. WSRF-The WS-Resource Framework (2006) http://www.globus.org/wsrf/
5. Globus Toolkit 4.0 (2006) http://www.globus.org/
6. Bing, S.H., li, C.Y.: An Overview of Computer Simulation Research on Traffic Systems. Computer Simulation 21(3), 101–104 (2004)
7. Mamei, M., Roli, A., Zambonelli, F.: Emergence and Control of Macro-Spatial Structures in Perturbed Cellular Automata, and Implications for Pervasive Computing Systems. IEEE Transactions on Systems, Man, and Cybernetics 35(3), 337–348 (2005)

8. Bauer, B., Müller, J.P., Odell, J.: Agent UML: A Formalism for Specifying Multi-agent Software Systems. International Journal of Software Engineering and Knowledge Engineering 11(3), 1–24 (2001)
9. Helbing, D., Farkas, I., Vicsek, T.: Simulating dynamical features of escape panic. Nature 407, 487–490 (2000)
10. Kirchner, A., Schadschneider, A.: Simulation of evacuation process using a bionics-inspired cellular automaton model for pedestrian dynamics. Physica A 312, 260–276 (2002)
11. Kirchner, A., Nishinari, K., Schadschneider, A.: Friction effects and clogging in a cellular automaton model for pedestrian dynamics. Physical Review E 67, 56–112 (2003)
12. weiguo, S., yanfei, Y., Tao, C.: Influences of Exit Conditions on Pedestrian Evacuation. Fire Safety Science 12(2), 100–104 (2003)
13. weiguo, S., yanfei, Y., weicheng, F., heping, Z.: An evacuation model based on cellular automaton which pays attention to friction and repulsion. Science in China Ser. E Engineering & Materials Science 35(7), 725–736 (2005)
14. weiguo, S., yanfei, Y., heping, Z.: Evacuation Analysis of a Large Shopping Mall. Engineering Science 7(10), 78–83 (2005)
15. weifeng, F., lizhong, Y., rui, H.: A Multi2agent Model Based on Cellular Automata for Simulating Human Behavior to Assess Building Design Performance. Engineering Science 5(3), 67–71 (2003)

Design and Implementation of Ergonomics Evaluation System of 3D Airplane Cockpit

Libo Zhang [1,2], Xiugan Yuan[1], Lijing Wang[1], and Dayong Dong[1]

[1] Institute of Man-Machine-Environment Engineering, Beihang University, Beijing, China
[2] School of Measurement-control Technology & Communications Engineering, Harbin University of Science & Technology, Harbin, China
{Libo Zhang,Xiugan Yuan,Lijing Wang,Dayong Dong, bjlbzhang}@163.com

Abstract. According to ergonomics factors referred by airplane design department and corresponding standards, structure and function of ergonomics evaluation system of 3D cockpit was designed. Digital human model, based on anthropometry database, comprises 66 segments, 65 joints and 131 degree of freedom. The ergonomics design of man-machine interface can be evaluated in terms of vision, reach and working posture comfort analysis methods and evaluation rules. Interior and exterior visual field of pilot can be achieved with the aid of vision analysis. The comfort of working posture and joints motion can be assessed by reference to joint angles for any selected posture. The details regarding system implementation with technology of OpenGL are discussed at last. The system can be computer-aided tool for airplane designer considering its convenience in using excellent model data interface with other 3D software.

Keywords: 3D airplane cockpit, Digital human model, Ergonomics, Working postural, Comfort.

1 Introduction

Ergonomics design of airplane cockpit is an important part of overall designs. Some ergonomics evaluation methods have been used, such as using two-dimension body template, physics emulator, or appropriative ergonomics evaluation software etc. Evaluation software is adopted extensively among above methods because of the reduction of development costs and easy modification. Since the end of 1960s, American air force started to conduct computer-aided ergonomics evaluation of airplane cockpit. With the progress of computer technology, evaluation scene presented by software has changed from two dimensions to three dimensions. Efficiency and availability of evaluation are strengthened owing to the improvement in visualization level.

Most design departments in China are still using some ergonomics software developed by other countries, such as SAFEWORK, JACK and RAMSIS, because of the comparatively laggard research work. Considering the difficulty in modifying these software' core technologies for their particular application field and staff with

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 1047–1055, 2007.

special anthropometric data, and high cost, it is of great necessity to develop a 3D ergonomics evaluation system of airplane cockpit in china.

2 System Structure and Function

This system was composed of five parts. They are 3D scene building, ergonomics analysis and evaluation, human motion simulation, database support and management of evaluation file. The database based its source on the latest national and military standards and the results of correlative experiments. Human motion simulation is performed through reading the data files exported by motion capture system. It can analyze the interference between human and equipments and carry out ergonomics evaluation at the key frame of motion process. Evaluation data of different phase is saved in files and evaluation scene can be built again through reading these files. By comparing the evaluation results of different phase, the optimization of layout design of airplane cockpit can be achieved.

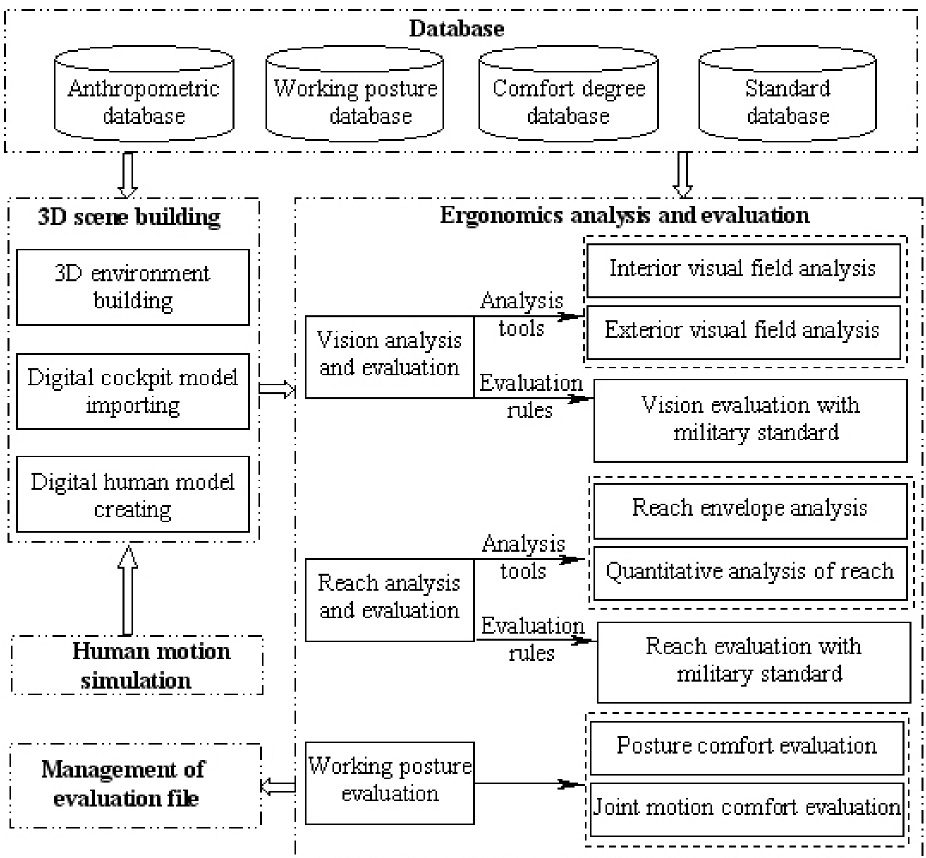


Fig. 1. Structure and function of ergonomics evaluation system of 3D airplane cockpit

3 Digital Human Model

Whether the anthropometric parameters of digital human model are accurate will directly affect the availability of ergonomics evaluation. The common methods of human geometric modeling are stick model, surface model, body model and layer model [1]. Stick model is very simple and only can express length of body. But more anthropometric parameters usually are needed in doing ergonomics evaluation. Layer model is the most complicated modeling method, but it can increase the modeling workload and lower the running speed of software. Owing to this, an entity human model with lifelike surface is created as an improvement over the above-mentioned models. This human model is composed of framework, surface, connection of joint and limits of joint motion. Anthropometric data of model are mainly from the pilot measures of fighter in GJB 4856-2003[2]. The human model is divided into 66 segments, including torso (comprising buttocks and 12 thorax vertebras and 5 lumbar), left and right thorax-shoulder, neck, head, left and right upper arm, left and right forearm, left and right hand (comprising metacarpus and 15 phalanges), left and right thigh, left and right calf, left and right feet (comprising 2 segments). The human model is composed of 65 joints and 131 degree of freedom and can act freely. The biggest movable scope of joints is restricted to ensure that the activity of the human model can't outrun the physiology extreme limit of body. The digital human model can be located into the 3D airplane cockpit according to design eye position and seat reference position.

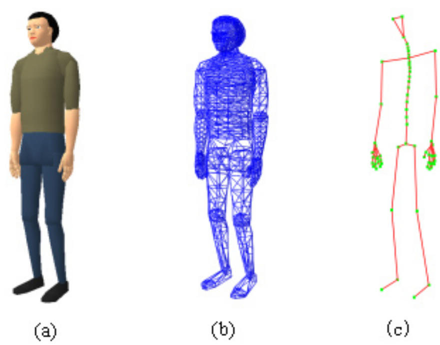


Fig. 2. Digital human model (a) entity model (b) grid structure (c) framework

4 Ergonomics Analysis and Evaluation

When the space layout design of display and control equipments in cockpit is planning, it must ensure that pilot can see and reach them. So, vision and reach analysis and evaluation methods are developed in the software according to national military standard and other relative standards. The comfort level of working posture can also be evaluated according to this software for its part in the efficiency and availability of working.

4.1 Vision Analysis and Evaluation

Eighty percent of all information that human being gets was obtained through eyes. Therefore, there are strict provision about the vision aspect of cockpit design in military standards, including interior visual field and exterior visual field. Interior visual field means interior cockpit scope that pilot can see with natural vision line and the visual datum mark of pilot is design eye position. Exterior visual field means exterior cockpit scope that pilot can see without interference of airplane and cockpit structure from design eye position of pilot [3].

(1) Analysis of interior visual field

Vision cone and vision window can be used to analyze the interior visual field of pilot in this software. Both the methods can be found in other ergonomics software such as SAFEWORK and JACK. Between them, vision window can guarantee a view that pilot can see directly. Vision cone is a cone whose apex is on the design eye position. In other ergonomics software, the hemline of cone is circular, which can't accurately describe the visual field. But the hemline of cone referred in this paper is an enclosed anomalous curve. Angles of the vision cone are defined according to the optimal and maximum visual areas angle under the circumstance that head turns, eyes turns and both head and eyes turns in GJB 2873-97 [4]. The vision cone is drawn with cubic B-spline curve of enclosed period, which can approximatively reappear the visual areas of human eyes.

(2) Analysis of exterior visual field

Although the structure of fighter and aerobus is different, definition mode of the exterior visual field is alike. The exterior visual field can be described by a two dimension Cartesian coordinate system whose x-axis represent the angle that head turns to left or right from the sagittal plane of pilot and whose y-axis represent the minimum visual angle that pilot can watch upwards or downwards from the horizontal plane which cross the design eye position. A general analysis tool is developed in this software. The actual curve of exterior visual field can be easily drawn through imputing some consecutive points in the window frame of cockpit. Exterior visual field can be directly evaluated by comparing the actual curve and standard curve of exterior visual field in the figure.

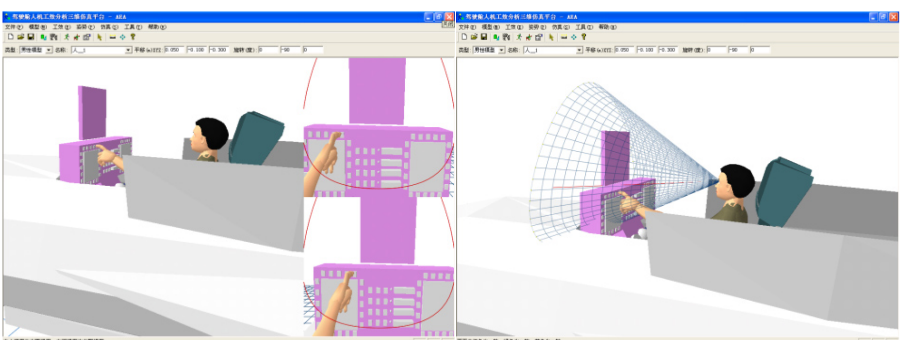


Fig. 3. Analysis methods of vision window and vision cone

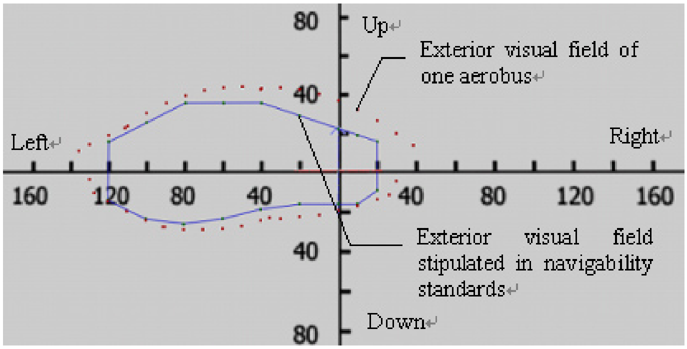


Fig. 4. Evaluation result of exterior visual field in the cockpit of one aerobus

(3) Vision evaluation in military standards

Designers usually are uncertain about how to evaluate the ergonomics design in the process of human machine engineering design of cockpit. So, a rules database of vision evaluation is set up in this software. The rules in the database are in accordance with provisions about the interior and exterior visual field in national military standards. Some important equipments and parts can be evaluated by using above analysis tools according to these rules.

4.2 Reach Analysis and Evaluation

There are respective reach analysis methods for upper limbs and lower limbs. Reach analysis tools for upper limbs are as follows.

(1) Analysis of reach envelope

According to the provision in GJB 35A-93 [5], reach area refers to three types as follows: the first area is the minimum function reach area under the circumstance of restricted shoulder strap being locked; the second area is the maximum function reach area under the circumstance of restricted shoulder strap being locked and the third

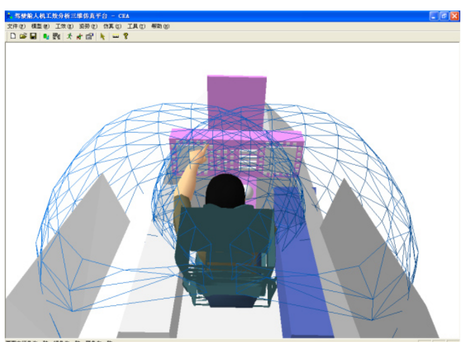


Fig. 5. Analysis of reach envelope for upper limbs

area is the maximum function reach area under the circumstance of restricted shoulder strap being unlocked. Since there is no standard data of reach envelope for upper limbs in china, these data is obtained by converting the data of USA fighter pilot, referring to the reference [6]. The reach envelope is approximately drawn with these data in this software.

(2) Quantitative analysis of reach

Analysis of reach envelope doesn't give the quantitative data of reach because it is a qualitative analysis method. But designers often want to know the actual unreachable value in order to modify the scheme of design. Therefore, Quantitative analysis of reach is developed in this software. As long as designer appoints the position of target, the unreachable value for left and right hand will be automatically calculated.

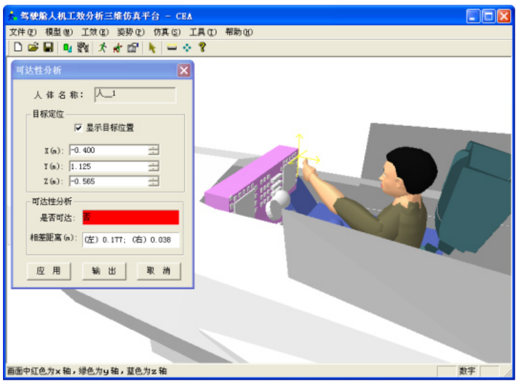


Fig. 6. Quantitative analysis of reach for upper limbs

(3) Reach evaluation with military standards

In the same way, a rules database of reach evaluation is set up in this software. The rules in the database are obtained by referring to the provision about the reach of equipments or parts in national military standards.

Two kinds of method could be used to analysis the reach of lower limbs. One is to simulate the motion of lower limbs through which the reach can be directly judged. The other method is to compare the actual excursion of manipulating pedal and the calculated excursion. The latter is calculated according to the method in reference [6].

4.3 Working Posture Comfort Evaluation

Adopting appropriate and healthy posture is usually regarded as a basic principle of design in various manuals and standards. But as to which postures are good and healthy and which postures are disadvantageous to health, there is no explicit specification for them. So, evaluation methods of working posture comfort are developed in this software. Working postures comfort of pilot can be evaluated and classified in the method in reference [7]. At the same time, the comfort of eight main joints (including wrist, elbow, shoulder, neck, spine, hip, knee and ankle) can also be evaluated with the method in reference [8]. Both the evaluation results will help the designers to make the modification.

5 Implement of 3D Analysis and Evaluation System for Airplane Cockpit

5.1 System Environment Requirement

This system is developed with the aid of Visual C++6.0 and can independently run in the Windows system. Considering that 3D model of airplane cockpit is usually very big and complicated, the memory of computer must be above 256MB. And it is better to use independent display card in order that 3D scene or 3D models can be transformed smoothly.

5.2 Interface of 3D Model File

Software used by the airplane design department usually used includes CATIA, UG and SolidWork etc. Ergonomics evaluation could be carried out effectively only when the cockpit model created in above-mentioned software is imported into this software accurately and integrally. As all the above software can generate model file whose extension of filename is ‘wrl’, an interface program for this type of model files is developed in this software. It can process and manage all model information about geometry, lighting, material and texture etc. Cockpit model can be reconstructed easily by making use of OpenGL technique in this software.

5.3 Key Technique of OpenGL

(1) LOD (Level of Details) technique

Model of airplane cockpit need bigger memory of computer and longer time to draw for the complication of the model’s structure. LOD technique is usually adopted to acquire ideal visual impression and processing speed. The technique can create 3D model with accuracy and can display with different accuracy according to the position of viewpoint.

(2) Collision detection technique

Collision detection is an essential part in 3D simulation system. It can ensure that the simulation system physics are relatively realistic. There will be encountered in the two objects and mutual insert in the absence of collision detection. Rough or precise detection can be adopted according to the demand of simulation. Rough collision detection is a method that can achieve rapid detection of collisions through building bounding boxes of objects and judging whether they are overlapping. The bounding box is the smallest cuboid surrounded object and its edges are parallel with coordinate axis. When accurate collision detection is in need, precise collision detection can be used. In the use of virtual objects that surround boxes borders, it is also used to indicate bounding boxes with polygon parcels. Clearly, the smaller polygon is divided, more precise their borders will be. Thus, the borders of a three-dimensional virtual object can be expressed by a group of polygons. To judge whether the two objects intersect, it first needs to confirm that two bounding boxes intersect before judging whether two polygons assembly intersect.

(3) Display list

The display list of OpengGL isn't a dynamic database buffer but a high-speed buffer. Therefore, it can not only advance the speed of drawing but also optimize the capability in many aspects, such as matrix manipulation, raster bitmap and image, lighting, material character, lighting model, texture mapping and polygon stipple mode.

(4) Pick mechanism

In the 3D evaluation system, objects in the screen usually need to be identified in order to perform some operation, such as translation, rotation, modification or deletion. If not, after doing these operation or transformation, it would usually be difficult to pick the object again in the screen without especial mechanism. Using pick mechanism of OpenGL will save the trouble. It will be clear that which objects are being drawn in the appointed region of screen and which objects are being picked with mouse in the region.

6 Conclusion

In this paper, an ergonomics evaluation system of 3D airplane cockpit is proposed, together with an introduction of its structure, function, and the detail of system implementation. Considering that the data and evaluation rules in the system all base on national military standards and relative research achievements, the evaluation results are credible. In addition, the system has excellent model data interface with other 3D software and its running environment is simple. Therefore, the system can be an effective computer-aided tool of ergonomics analysis and evaluation for airplane designer.

References

1. Wang, Z.Q.: Study on Synthesis of Virtual Human. Journal of the Graduate School of the Chinese Academy of Sciences, Beijing, China 17(2), 89–98 (2000)
2. Chinese PLA General Armament Department: GJB 4856-2003 Human Dimension of Chinese Male Pilot Population. The People's Republic of China Military Standards
3. Commission of Science Technology and Industry for National Defense: GJB 307-87 Visual Field of Fighter Cockpit and Visual Areas of the Main Instrument Panel. The People's Republic of China Military Standards
4. Commission of Science Technology and Industry for National Defense: GJB 2873-97 Human Engineering Design Criteria for Military Equipment and Facilities. The People's Republic of China Military Standards
5. Commission of Science Technology and Industry for National Defense: GJB 35A-93 Aircrew Station Geometry for Fighter. The People's Republic of China Military Standards

6. Wang, L.J.: The Study of Computer-aided Ergonomic Analysis and Assessment System for Cockpit of Military Aircraft. Beihang University, Beijing (2002)
7. Zhang, L.B., Yuan, X.G., Yang, C.X., et al.: Comfort Evaluation of Working Posture of Upper Limb in Sitting Posture. China Safety Science Journal, Beijing, China 16(9), 51–55 (2006)
8. Zhang, L.B., Yuan, X.G., Wang, L.J., et al.: Evaluation of Joint Motion Based on Perceived Discomfort. Space Medicine and Medical Engineering, Beijing, China, vol. 19, 412–416 (2006)

A Novel Method for Cloth-Body Collision Detection

Xiaolong Zhu[1,2], Shihong Xia[1], Yong Yu[1,2], and Tianlu Mao[1,2]

[1] Institute of Computing Technology, Chinese Academy of Sciences,
100080 Beijing, China
[2] Graduate School of the Chinese Academy of Sciences, 100080 Beijing, China
{zhuxiaolong,xsh,yyu,ltm}@ict.ac.cn

Abstract. This paper presents a novel cloth-body collision detection method by using the generalized cross-sectional contour technique, which has two main steps. During preprocessing step, the so-called skin hierarchical structure (Skin-H) of the body is constructed by using the improved generalized cross-sectional contour technique, which doesn't need to be updated in subsequent step. During runtime step, the cloth vertices are projected onto Skin-H structure efficiently, and then the exact collision detection can be done by a ray-triangle test technique at the lowest level of the structure. The simulation result demonstrates that the proposed method has some advantages in algorithm's efficiency, accuracy as well as practicability.

Keywords: Collision Detection, Hierarchical Approaches, Image-Based Methods, Cloth Simulation, Animation.

1 Introduction

Cloth simulation has been an area of active research in the past 15 years due to its increasing importance for commercial applications such as virtual try-on [2,3,9,20]. Unfortunately, cloth-body collision detection is the crucial but expensive technology and proved to be the bottleneck of dynamic cloth simulation. Motivated by this fact, numerous approaches have been proposed to achieve efficiency and accuracy as well as practicability. These approaches can be classified into: (1)*object-space interference test*[2,3,4,11,12,19], and (2)*image-space interference test*[1,5,6,7,8,14,20]. Most of object-space methods use bounding volume hierarchy (BVH) to accelerate interference computations. For example, An OBB-Hierarchy is used in [19] to represent polyhedra with triangulated boundaries. The overlaps between OBBs are determined by performing 15 simple axis projection tests, however, it associates with high cost of building and updating hierarchy. [11] improved the efficiency of BVH by proposing adapted techniques for building and traversing hierarchies, and [4] introduced the Bounded Deformation Tree to perform collision detection for reduced deformable models at similar costs to standard algorithms for rigid objects. In addition, with the recent advent of high performance GPU, the GPU-based methods have opened new avenues for improving the efficiency of collision detection. The methods are based on projecting the object geometry onto the image plane and performing the analysis in a dimensionally reduced space with the depth map

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 1056–1063, 2007.

maintained in an image buffer such as the Z-buffer. The first application of these methods to dynamic cloth simulation has been presented in [20]. In this approach, an avatar is rendered from a front and a back view to generate the depth information for collision query. Recently, more and more hybrid methods which combined object-space methods with GPU-based ones were presented, such as [1, 8, 13]. In [8], it introduced $(\pi,\beta)-\mathrm{surface}$ to dynamically generate a hierarchy of cloth bounding boxes in order to perform object-level culling and GPU-based intersection tests using conventional graphics hardware support. Another kind of GPU-based methods is recoding the geometry information into textures and then computing the collision results in the fragment program [1, 5]. There are many other notable approaches [15, 16] proposed to improve the efficiency and accuracy of collision detections. For example, a chromatic mesh decomposition is used to develop a linear time detection algorithm in [15].

By comparison, GPU-based collision detection methods are easy to implement and most of which involve no preprocessing and therefore can be applied to both rigid and deformable models. Meanwhile, it has a good potential to exploit the hardware-assisted rendering features of the current graphics boards. The main problem of GPU-based methods, however, is that it suffers from limited precision which is due to the limited viewport resolution and GPU float-point errors [6, 14]. Most object-space methods work well for rigid objects; nevertheless, they need frequently update the hierarchy especially for deformable models. In this paper, a novel hierarchy named skin hierarchical structure (Skin-H) is presented, which is different from the conventional BVH . It can avoid time consuming hierarchy updating because the new hierarchical building method considers the special structure of human and cloth.

2 Collision Detection Using Skin-H

In this section, we give an overview of our collision detection algorithm. It includes the building of Skin-H during precomputing and collision detection in runtime. Without loss of generality, we assume that the body model is seamless and can be represented as triangular meshes, which are driven by the corresponding skeleton. We firstly proposed a so-called skin hierarchical structure (Skin-H), according to the driving principle of virtual human with seamless model and the fact, that the displacement of skin is very small related to the innate bone. The Skin-H is invariant while virtual human moving. It’s needless to update the hierarchy frequently. This is because each human skin vertex is tied to a certain joint to generate joint block which can be treat as a rigid object. Although there are skin deformations, they can also be handled well in this paper. Subsequently each cloth vertex is projected onto Skin-H, and ray-triangle collision test can be used during runtime.

2.1 Skin Hierarchical Structure (Skin-H)

The design choice of the type of bounding volume (BV) mainly determines the efficiency of traversing and updating the BVH. People have tried many types of BV such as OBBs and DOPs attempted to quickly update or refit the BVH after a deformation has taken place. However, the hierarchy should be updated after each frame for animation cloth.

To solve this problem, we design our hierarchy based on the principal of skeleton driven deformation and the structure of the special human body shape. The skeleton deformation uses vertex weights which tie the joints and the mesh together. This could help dividing the human skin triangles into several **Blocks** according to different joints influence different skin vertices. For each **Block**, the local coordinate system is built and then each block is partitioned further into **Segments** and **Sections**.

Block. Block is the top of Skin-H. Two methods (Fig.1) for segmenting the human skin into blocks have been tried in our experiment. One is to select several main joints and take them as the boundaries of the blocks which are built according to the coordinate of each skin vertex. Another method for classifying the vertices is based on the seamless human model, which is driven by the articulated skeleton, wherein each skin vertex is affected by the corresponding joints. So we can classify all of the skin vertices by their effective weights.

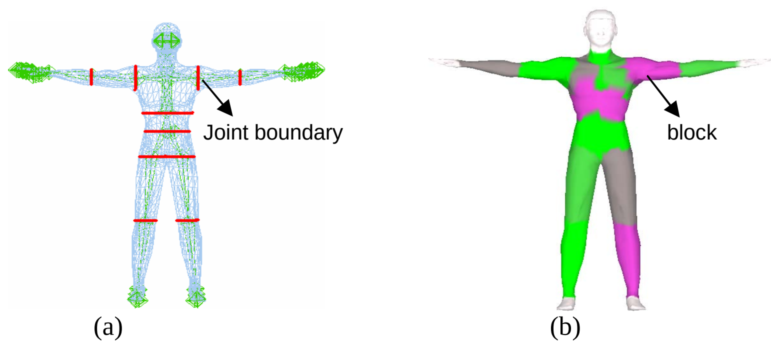


Fig. 1. Two methods for generating Blocks: (a) take the joints as the boundaries of blocks;(b) classify vertices based on effective weights

The former method mainly have two shortcomings: 1) it's difficult to project the cloth vertex into the corresponding block when the human is in motion because there are some boundary planes would overlap each other then; 2)it doesn't consider the skin deformation which could result in missing collisions near the boundary. This paper therefore adopts the latter method which classifies the skin vertex based on the influence weights. Generally, there are tens of joints in a human body skeleton. It is needless to divide the human skin into so many blocks. The 13 main joints are selected for efficiency consideration. For each main joint, because its translation and rotation transformation will affect a number of skin vertices, a weight threshold δ_{weight} is set in order to determine each vertex belonging to which block. As a result, the joint block generation algorithm is described as follows.

Algorithm 1. Generating Block

For each main joint joint_i
  For each skin vertex vertex_j affected by joint_i
    If weight(joint_i, vertex_j) >= δ_{weight}
      AppendToBlock(joint_i, GeTriangle (vertex_j));

In the algorithm 1, GeTriangle (vertex_j) returns the triangle which contains the vertex vertex_j and the threshold δ_{weight} is determined keeping to the following two rules:

1) Eliminate the influenced vertices far away from the joint;
2) Generate redundancy in the adjacent blocks against missing collisions during detection.

Segment. The Skin-H is built by using top-down strategy. The segment is on the second level of the Skin-H. There are usually hundreds of triangles in a joint block. For reducing the times of the ray-triangle collision detection, each joint block is divided further into segments.

The method of generating segments is based on the generalized cross-sectional contour technique [10,18]. For each block, we choose the joint center as the origin and set the direction of y-axis along the corresponding bone to build local coordinate system (Fig. 2). As a result, we can easily get the local coordinate and the polar angle of the center of each triangle which belongs to this joint block. The local y-coordinate helps us divide the block into several segments.

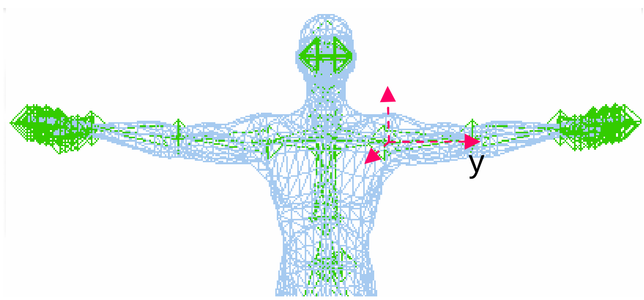


Fig. 2. Building local coodinate system for each segment

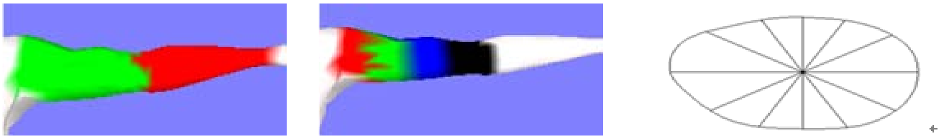
Section. Sections are in the bottom of the Skin-H and built according to polar angle. Every segment could be considered as a column so that the triangles in them can be classified into several sections in term of the polar angles.

In pretreatment, all of the skin triangles are classified into sections. Each level of the Skin-H stores the indices of the corresponding skin triangles. For avoiding missing the collisions result from skin deformation, a certain degree of redundancy is kept within each level. These measures guarantee it's needless to refit or update the Skin-H during runtime.

As shown in Fig.3, the Skin-H has three levels (see figure 1): *Block*, *Segment* and *Section*.

2.2 Project Cloth Vertices on Skin-H

In runtime, our cloth-body collision detection algorithm includes coarsely localizing the joint blocks and the segments and further the sections which potentially collide with cloth vertices and at last performing the ray-triangle intersection test.



(a) Block (b) Segment (c) Section

Fig. 3. The three levels of the Skin-H

There is possibility that each cloth vertex collides or even penetrates the skin triangles. So we traverse each cloth vertex and search skin triangles which potentially collide with the vertex. These triangles must be in a certain section, and located via projecting the vertex to the Skin-H.

The first step of searching the potential colliding skin triangles is that locate the potential joint blocks (PCB). In our experiment, we choose 13 main joints and classify the all skin triangles into 13 joint blocks in pretreatment. Meanwhile, the maximal distance $dis_{\text{bone\_i}}$ from the triangles to the corresponding bone is calculated for each joint block. During collision detection, the distance $dis(\text{vertex\_i}, \text{bone\_i})$ from cloth vertex to each main bone is computed and the joint block is added to the PCB if $dis(\text{vertex\_i}, \text{bone\_i})$ is less than $dis_{\text{bone\_i}}$. Locate the potentially colliding segments and sections in Skin-H just need to transform the global coordinate of the cloth vertex to the local coordinate under the joint block's coordinate system. The local y-coordinate and the local polar angle can help to find the potential Segments and Sections. The whole projection or location method is described as follows.

Algorithm 2. Cloth-vertex Projecting

For each cloth vertex VCloth

- *Computing the distance from cloth vertex to each block*
- *Traverse the blocks to find blocks satisfying* $dis(\text{vertex\_i}, \text{bone\_i}) < dis_{\text{bone\_i}}$ *and add them to PCB*

  For each $block \in PCB$

  - *Vcloth_localCoord = GetLocalCoord(VCloth, Block_i)*
  - *segment = GetSegment(Vcloth_localCoord, Block_i)*
  - *section = GetSection(Vcloth_localCoord,segment)*

    For each triangle Tri_i in section

    - *RayTriangleCollisionTest(VCloth, Tri_i)*

In this algorithm, *GetLocalCoord(VCloth, Block_i)* calculate the local coordinate includes polar angle of the cloth vertex, and then, *GetSegment(Vcloth_localCoord, Block_i)* and *GetSection(Vcloth_localCoord, segment)* locate the segment and the section in Skin-H according to the local y-coordinate and polar angle.

2.3 Ray-Triangle Intersection Detection

For each triangle in the section above, Moller's method [21] is used for Ray-Triangle intersection test and computing the penetration depth. The difficulty is that obtain a

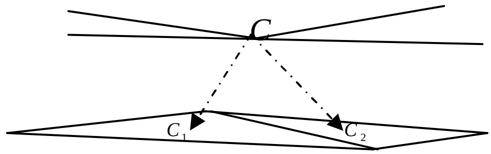


Fig. 4. C_i is defined as the center of the skin triangle, and then the direction from cloth vertex C to C_i is defined as the ray direction for ray-triangle intersection test

proper direction of the rays which has a significant impact on the accuracy of the collision detection method. Generally, there are three different methods to generate ray direction. The first method is set the normal direction of the cloth as the ray direction. It needs recalculating the normal of every cloth vertex for each frame and furthermore it can't deal with the situation that there is a drape in the collision area. The ray will intersect with a skin triangle in another section so the collision would be missed in this case. The second method takes the vertical direction from the cloth vertex to the corresponding bone as the ray direction, which is unsuitable for some areas such as shoulder block. In this paper, the direction (Fig.4) from the cloth vertex to the center of each skin triangle in the section is calculated as the ray direction. This method is easy to implement and has no collisions missed. The ray-triangle intersection test algorithm is discussed in [21].

3 Performances

Our algorithm has been implemented on a PC running Windows XP with a 2.2 GHz Pentium Ⅳ CPU, an NVIDIA GeForce 5200 GPU, and 512M of Main Memory. We

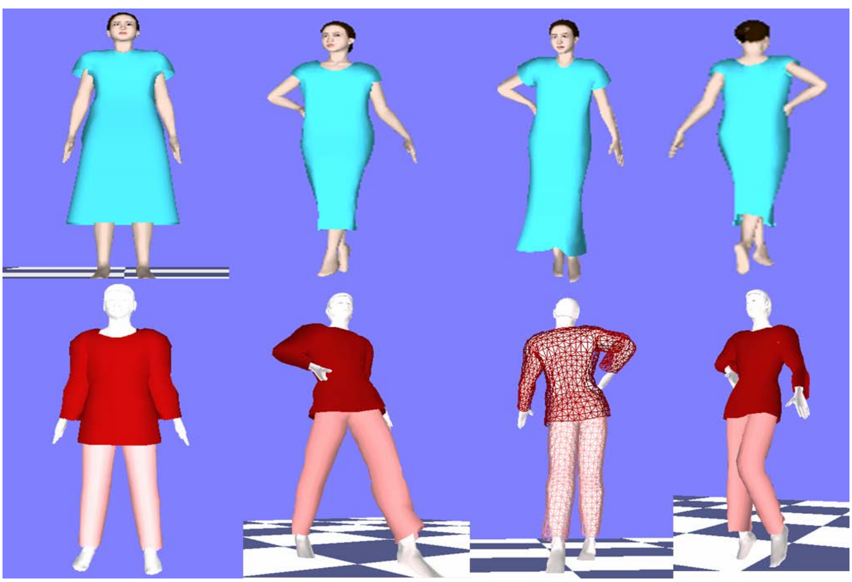


Fig. 5. a sequence of cloth animation (Cloth: 8,000 – 10,000 Tir ; Human: 7,500 - 9,000 tri ;Time : 50-100msec)

use it in cloth simulation and a virtual try-on system which will soon be introduced. We highlight the results on two cloth simulations shown in Fig.5. The number of triangles in the mesh used to model the body and clothes vary from 8k to 10k. The time to check for cloth-collisions is in the range of 50-100msec. The performance depends on the input complexity. During our implementation, the skin triangles are classified into thirteen joint blocks. Each block is divided into four segments and each segment is further divided into eight to ten sections. So there are only twenty skin triangles or so in a section. As a result, our algorithm obtains speedup due to no Skin-H updating operation and less number of elementary tests between the primitives.

Our approach has demonstrated that simulating dressed virtual human with middle-range PC is feasible. It can deal with several kinds of clothes such as skirts and suits. For further accelerate the method, some inactive cloth areas such as the areas near shoulder and waist can be bound to the corresponding joints in order to reduce the number of the collision tests. However, no self-collision is calculated in this paper, therefore, it can't simulate the wrinkles well.

4 Conclusions

To our knowledge, it's the first time that the generalized cross-sectional contour technique[10,18] is used for cloth-body collision detection. The proposed method has the following advantages.

Efficiency. The method can achieve near-real-time because most of the time-consuming operation was done during preprocessing step and the hierarchical update was avoided.

Accuracy. The method belongs to the hierarchical ones. The advantage of accuracy of the Hierarchical methods is inherited in a natural way.

Practicality. The method is very simple. It is easy to implement and can be applied to virtual modeling, virtual human animation and other applications.

Acknowledgements

This work is supported in part by NSFC-60403042, NSFC-60573162, NSFC-60533070, NSFC-60603082, CISTC-2005DFA11060, BNSFC-4051004, BNSFC-4062032, HTRDPC-2006AA01Z336.

References

1. Benes, B., Villanueva, N.G.: GI-COLLIDE: Collision Detection with Geometry Images. In: SCCG'05:Proc. of the Sping Conference on Computer Graphics, pp. 291–312 (2005)
2. Cordier, F., Magnenat-Thalmann, N.: Real-time Animation of Dressed Virtual Humans. Eurographics, Blackwell publishers, 21(3), 327–336 (2002)
3. Cordier, F., Hyewon, S., Magnenat-Thalmann, N.: Made-to-Measure Technologies for an Online Clothing Store. IEEE CG&A 23, 38–46 (2003)

4. James, D.L., Pai, D.K.: BD-Tree: Output-Sensitive Collision Detection for Reduced Deformable Models. ACM Trans. Graphics 23(3) (2004)
5. Fan, Z.W., Wan, H.G., Gao, S.: Streaming Real Time Collision Detection Using Programmable Graphics Hardware. Journal of Software(in Chinese with English abstract). 15(10), 1505–1514 (2004)
6. Baciu, G., Wong, W., Sun, H.: RECODE: An Image-based Collision Detection Algorithm. In: Proceedings of the 6th Pacific Conference on Computer Graphics and Applications, p.125, October 26-29, 1998 (1998)
7. Baciu, G., Wong, W.: Hardware-Assisted Self-Collision for Deformable Surfaces. In: Proc. ACM Symp. Virtual Reality Software and Technology, pp. 129–136, November 2002 (2002)
8. Baciu, G., Wong, W.: Image-Based Collision Detection for Deformable Cloth Models. IEEE Trans.Visualization and Computer Graphics 10(6), 649–663 (2004)
9. Chittaro, L., Corvaglia, D.: 3D Virtual Clothing: from Garment Design to Web3D Visualization and Simulation. In: Proceedings of Web3D 2003: 8th International Conference on 3D Web Technology, pp. 73–84. ACM Press, New York (March 2003)
10. Yan, L., Zhaoqi, W., Tianlu, M.: A Method of Virtual Human Skin Deformation based on Generalized Cross-Sectional Contour. Computer Science(in Chinese with English abstract) 32(1), 190–193 (1) (2005)
11. Mezger, J.S., Kimmerle, O.: Etzmus. Hierarchical Techniques in Collision Detection for Cloth Animation. Journal of WSCG, 11(1) (2003)
12. Lin, M., Gottschalk, S.: Collision Detection between Geometric Models: A Survey. In: Proc. of IMA Conference on Mathematics of Surfaces 1998 (1998)
13. Teschner, M., Kimmerle, S., Zachmann, B., Raghupathi, L., Fuhrmann, A., Cani, M.-P., Faure, F., Magnenat-Thalmann, N., Strasser, W., Volino, P.: Collision Detection for Deformable Objects. Eurographics State-of-the-Art Report, pp. 119–139 (2004)
14. Naga, K., Govindaraju, M.: Lin and Dinesh Manocha. Fast and Reliable Collision Detection Using Graphics Processors. In: Symposium on Computational Geometry, pp. 384–385 (2005)
15. Naga, K., Govindaraju, D., Knott, N., Jain, I., Kabul, R., Tamstorf, R., Gayle, M.C.: Interactive Collision Detection between deformable Models Using Chromatic Decomposition. ACM Trans Graphics 24(3), 991–999 (2005)
16. Volino, P., Magnenat-Thalmann, N.: Resolving Surfaces Collisions through Intersection Contour Minimization. In: ACM Transactions on Graphics (SIGGRAPH 2006 proceedings), vol. 25(3), pp. 1154–1159. ACM Press, New York (2006)
17. Rodriguez, J., Sainz, M., Susin, A.: Fast Body-Cloth Simulation with Moving Humanoids. Short presentations EG'05, pp. 85–88 (2005)
18. Jianhua, S., Magnenat-Thalmann, N., Thalmann, D.: Human Skin Deformation from Cross Sections. In: Proceedings of Computer Graphics International CGI'94 (1994)
19. Gottschalk, S., Lin, M.C., Manocha, D.: OBBTree: A Hierarchical Structure for Rapid Interference Detection. In: Computer Graphics (SIGGRAPH'96), pp. 171–180, ACM (August 1996)
20. Vassilev, T., Spanlang, B., Chrysanthou, Y.: Fast Cloth Animation on Walking Avatars. In: Computer Graphics Forum Proc. Of Eurographics (2001)
21. Moller, T., Trumbore, B.: Fast, Minimum Storage Ray-Triangle Intersection. Journal of Graphics Tools 2(1), 21–28 (1997)

Author Index

Abdel-Malek, Karim 490, 500
Adam, Clayton 655
Ait El Menceur, Mohand Ouidir 23
Anderson, Glen 392
Andreoni, Giuseppe 3
Aoki, Hiroyuki 894
Armitage, Alistair 120
Aromaa, Susanna 847
Arora, Jasbir 490

Beauregard, Russell 325
Beck, Steve 490
Benyon, David 120
Beurier, Georges 13
Blum, Rainer 453
Bramlett, Brian W. 392
Branemark, Rickard 655
Bubb, Heiner 66, 779

Cai, Fuwen 633
Cairns, Nicola 655
Cao, Chunmei 794
Cao, Jian 1037
Carruth, Daniel W. 205, 333
Case, Keith 110
Chang, Ju-Ming 56
Chao, Chin-Jung 979
Chateauroux, Elodie 215
Chen, Hao 130
Chen, JiuYun 434
Chen, Lei 747
Chen, Xiaoping 794
Cheng, Zhiqiang 294
Chevalot, Nicolas 13, 215
Cho, Jungbum 874
Choi, Hyung-Il 646
Choi, Hyung Yun 813
Choi, Young Lim 803
Choi, Kueng Mi 803
Chouvarda, Ioanna 694
Chryssolouris, George 939
Chuang, Ming-Chuen 343
Chung, Min Suk 353
Clapworthy, Gordon J. 549
Colombo, Giorgio 559
Corriveau, Philip 325
Coveney, P. 549
Cowan Novak, Julie 710
Cui, Gao-Yu 530
Cui, Ming Hai 803

Dai, Guozhong 747
Debril, Jean-François 23
DeLaurentis, Daniel 363
Demirel, H. Onan 372, 824
Diao, Wei-He 243
Diao, Weifeng 294
Ding, Li 33, 856
Dionysiou, Dimitra D. 569, 703
Dong, Dayong 42, 1047
Dong, Liju 314
Du, Junmin 48
Du, Yingzi 460
Duffy, Vincent G. 197, 372, 460, 824, 904, 988

Elliott, Stephen J. 904
Endo, Yui 833
Esteves, Marcela 576
Evans, John 655

Falkmer, T. 843
Fan, Shu-Kai S. 56
Fang, Lidan 640
Fang, Xin 253
Feng, Hailan 747
Feng, Wen-Yang 979
Fenner, John 549
Feyen, Robert G. 382
Filippi, Stefano 559
Frey-Law, Laura 490
Fritzsche, Florian 66
Frohriep, Susanne 586
Frossard, Laurent 655
Fubini, Enrica 949
Furuta, Kazuo 511

Gao, Ji 392
Gao, Lei 74

Ge, Bao-zhen 82
George, Jennifer 595
Georgiadi, Eleni 703
Geraci, Filippo 606
Gnanayutham, Paul 595
Godil, Afzal 92
González Ballester, Miguel A. 767
Gorce, Philippe 23
Gramopadhye, Anand K. 959, 1018
Gregerson, Hans 549
Guo, Qiang 101
Guo, Yan-Li 530

Häberling, Sascha 453
Han, Jiwon 813
Hanson, L. 843
Hao, Zhixiu 616
Harrison, Chad 500
Haug, Eberhard 813
He, Huogao 399
Helin, Kaj 847
Heng, Pheng Ann 530, 623
Högberg, Dan 110
Holden, A.V. 659
Hose, D.R. 549
Hu, Huimin 856
Hu, Xiao 483, 633, 727
Hu, Ying 668
Huang, Chaoyu 399
Huang, Linpeng 1037
Hwang, Hojin 884
Hwang, Su-Hwan 813

Ji, Linhong 253, 794
Jia, Xiaohong 633, 640, 727
Jiang, Tianzi 409
Jin, Dewen 253, 259, 268, 483, 616, 633, 640, 727
Jones, Trevor D. 120

Kanai, Satoshi 833
Kanezashi, Ohki 865
Kang, Rui 130
Kang, Wei-yong 923
Karabatsou, Vassiliki 939
Khakzar, Karim 453
Kharitonenko, Igor 314
Khasawneh, Mohammad T. 998
Kim, Bohyun 874
Kim, Gye-Young 646
Kim, Hongtae 884
Kim, Joo H. 490
Kim, Kyung Min 813
Kim, Seok-Hwan 813
Kim, Yoon Hyuk 136
Kim, Youngjun 874
Kishinami, Takeshi 833
Kitauchi, Ryo 415
Knapheide, Claus 536
Kohl, Peter 549
Komischke, Tobias 576
Kotsonis, Marios 939
Kouchi, Makiko 833
Kubo, Mitsunori 894
Kukula, Eric P. 904
Kuroda, Tomohiro 415

La, Kaiying 915
Lam, Sarah S. 998
Lawford, P. 549
Lawson, Shaun W. 120
Lee, Chang-Min 884
Lee, Inhyeok 813
Lee, Kunwoo 874
Lee, Kwang No 813
Lee, Na-Young 646
Lee, Sang-Wook 157
Lee, Winson 655
Leoncini, Mauro 606
Lepoutre, François Xavier 23
Li, Bo 189
Li, Chao 74
Li, Fang 727
Li, Kai 530
Li, Kang 144, 157
Li, Minglu 1037
Li, Qi-Yu 530
Li, Sheng 147
Li, Shuyu 409
Li, Wanqing 314
Li, Xiaobo 409
Li, Yong 33
Li, Yinxia 915
Li, Zheng 243
Li, Zhizhong 161, 171
Liao, Rong-Xia 530
Lin, James J.W. 423, 536
Liu, Dahai 474
Liu, Fei 668
Liu, Hailong 467

Liu, JianGuo 434
Liu, Jihong 659, 668
Liu, Mingyu 467
Liu, Wei 923
Liu, Yuncai 304
Liu, Yuqing 101
Liu, Zheng-Jin 530
Liu, Zhong-qi 923
Lo, Inglen 343
Longzhu, Han 929
Lu, YanSheng 434
Luo, Jia-Wei 678

Ma, Lizhuang 276
Ma, Rui 923
Ma, Weina 668
Ma, Xin 684
Määttä, Timo 847
Maglaveras, Nicos 694
Mao, Tianlu 286, 1056
Mao, Xia 243
Marca, Christian 813
Margaritis, Thanasis 703
Marsh II, Delbert 392
Marias, Kostas 569, 703
Maris, Thomas G. 703
Marler, Tim 500
Marler, Timothy 490
Mathai, Anith 490
Mavrikios, Dimitris 939
McCormack, K. 549
McGinley, John A. 205, 988
Meng, Qinglei 189
Miyata, Natsuki 833, 865
Mochimaru, Masaaki 833
Moehlenbrink, Christoph 1027
Monnier, Gilles 13
Montangero, Manuela 606
Montmayeur, Nicole 813
Montonen, Jari 847
Morais, Alex 205, 333
Morricone, Manuela 949
Murphy, Chris 490

Nam, Yunja 803
Niu, Jianwei 161, 171

Ogunbona, Philip 314
Oh, Insuk 443
Ohboshi, Naoki 415
Okamoto, Kazuya 415
Olivato, Paolo 949
Ota, Jun 865

Palmer, Ryan 392
Pappas, Menelaos 939
Park, Jin H. 884
Park, Jin Seo 353
Parker, Donald E. 423
Pearcy, Mark 655
Pedotti, Antonio 3
Pei, Xiaomei 521
Pellegrini, Marco 606
Petzel, Jan 586
Pinney, D. 549
Pudlo, Philippe 23
Pyun, Jong-Kweon 813

Qiao, Mu 179
Qiu, Ming-Guo 530

Rabuffetti, Marco 3
Rahmatallah, Salam 490, 500
Rapala, Kathryn 710
Re, Alessandra 949
Renda, M. Elena 606
Rissone, Paolo 559
Rizzi, Caterina 559
Robbins, Bryan 205, 333
Rudolph, Michael 1027

Sadasivan, Sajay 959
Sah, Sungjin 813
Salvendy, Gavriel 161, 171
Seidl, Andreas 1008
Shi, Feng 409
Shi, Haiwen 48
Shu, Yufei 511
Soltysinski, Tomasz 717
Song, Jianbin 189
Stamatakos, Georgios S. 569, 703
Stone, Matthew 443
Stringfellow, Paris 1018
Su, Chuan-Jun 969
Sun, Tien-Lung 979
Sun, Yu-chen 82
Sutherland, Jackie 197
Swan, Colby 490

Tai, Changfeng 467
Takemura, Tadamasa 415
Tan, Li-Wen 530
Terauchi, Fumio 894

Thomas, Mark.D. 205, 333
Thomas, S.R. 549
Tian, Qing-guo 82
Tian, Renran 988
Tollis, Ioannis G. 703
Trasbot, Jules 13

Upadrasta, Bharat 998
Uritsky, Vadim 474

van der Meulen, Peter 1008
Van Sint Jan, S. 549
Viceconti, Marco 549
Viitaniemi, Juhani 847
Volz, Andreas 453

Wan, MingXi 147
Wang, Chuncai 794
Wang, Dejun 1037
Wang, Jing 736
Wang, Jue 467, 757
Wang, Lee-ming 460
Wang, Lijing 42, 1047
Wang, Mingwen 1037
Wang, Rencheng 253, 259, 268, 483, 633, 640, 727
Wang, SuPin 147
Wang, Xiang-Bo 757
Wang, Xuguang 13, 215
Wang, Zhaoqi 286
Wang, Zixi 794
Wanliss, James 474
Washburn, Carl 1018
Waters, S. 549
Weiss, Antje 576
Werther, Bernd 1027
Wu, Di 276
Wu, Fangfang 259, 268, 483, 633, 727
Wu, Jiankun 1037
Wuergler, Michael 474

Xia, Shihong 286, 1056
Xiang, Jian 224, 234
Xie, Yong-Ming 530
Xiong, Zhang 74
Xu, ChaoXiong 130
Xu, Guanghua 736
Xu, HongBo 130
Xu, Jian-Xue 757
Xu, Xinhe 668
Xue, Yu-Li 243

Yan, Nan 467
Yang, Chunxin 33, 179, 856
Yang, Feng 33
Yang, Jiankun 253
Yang, Jingzhou 490, 500
Yang, Jinzhu 668
Yang, Li 678
Yang, Yiyong 259, 268, 483, 727
Yao, Hong 757
Yao, Li 276
Yong, L. 843
Yoshihara, Hiroyuki 415
Young, K. David 82
Youshimura, Noboru 483
Yu, Xiaolin 521
Yu, Yong 286, 1056
Yuan, Xiu-gan 33, 42, 48, 179, 856, 915, 923, 929, 1047

Zacharopoulou, Fotini 703
Zapf, Salome 536, 576
Zhang, Chi 747
Zhang, Fengjun 747
Zhang, Chong 521
Zhang, Guang-Jun 757
Zhang, Hao 747
Zhang, Henggui 659
Zhang, Jichuan 253, 616, 633, 640, 727
Zhang, JiChuang 483
Zhang, Libo 1047
Zhang, Ming 259, 268, 640
Zhang, Qing 736
Zhang, Shao-Xiang 530
Zhang, Xi-Zhen 678
Zhang, Xiaodong 294
Zhang, Xudong 144, 157
Zhang, Yu 616
Zhao, Xu 304
Zheng, Chongxun 521
Zheng, Gang 314
Zheng, Guoyan 767
Zheng, Shuangxi 633
Zheng, Xianjun Sam 536
Zhou, Dawei 253
Zhu, Chengjun 74
Zhu, HongLi 224
Zhu, Hongli 234
Zhu, Junming 736
Zhu, Wanlin 409
Zhu, Xiaolong 1056