

James S. Duncan
Guido Gerig (Eds.)

LNCS 3749

Medical Image Computing and Computer-Assisted Intervention – MICCAI 2005

8th International Conference
Palm Springs, CA, USA, October 2005
Proceedings, Part I

1
Part I


MICCAI

 Springer

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

University of Dortmund, Germany

Madhu Sudan

Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos

New York University, NY, USA

Doug Tygar

University of California, Berkeley, CA, USA

Moshe Y. Vardi

Rice University, Houston, TX, USA

Gerhard Weikum

Max-Planck Institute of Computer Science, Saarbruecken, Germany

James S. Duncan Guido Gerig (Eds.)

Medical Image Computing and Computer-Assisted Intervention – MICCAI 2005

8th International Conference
Palm Springs, CA, USA, October 26-29, 2005
Proceedings, Part I

Volume Editors

James S. Duncan

Yale University

Departments of Biomedical Engineering and Diagnostic Radiology

333 Cedar Street, P.O. Box 208042, New Haven, CT 06520-8042, USA

E-mail: james.duncan@yale.edu

Guido Gerig

University of North Carolina

Department of Computer Science

CB# 3175, 219 Sitterson Hall, Chapel Hill, NC 27599-3175, USA

E-mail: gerig@cs.unc.edu

Library of Congress Control Number: 2005933259

CR Subject Classification (1998): I.5, I.4, I.3.5-8, I.2.9-10, J.3, J.6

ISSN 0302-9743

ISBN-10 3-540-29327-2 Springer Berlin Heidelberg New York

ISBN-13 978-3-540-29327-9 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springeronline.com

© Springer-Verlag Berlin Heidelberg 2005

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper SPIN: 11566465 06/3142 5 4 3 2 1 0

Preface

The 8th International Conference on Medical Imaging and Computer Assisted Intervention, MICCAI 2005, was held in Palm Springs, California, USA, at the Riviera Resort, October 26–29, 2005.

MICCAI has become a premier international conference with in-depth papers on the multidisciplinary fields of medical image computing, computer-assisted intervention and medical robotics. The conference brings together clinicians, biological scientists, computer scientists, engineers, physicists and other researchers and offers them a forum to exchange ideas in these exciting and rapidly growing fields.

The impact of MICCAI increases each year and the quality and quantity of submitted papers this year was very impressive. We received a record 632 full submissions (8 pages in length), an increase of 22% from 2004, from 36 different countries and 5 continents (see fig. 2). Based on a decision of the MICCAI board, this year's conference employed a double-blind review procedure on a trial basis. Our Program Committee was made up of 11 area chairs, each of whom supervised the review of almost 60 papers. Four reviews were generated for each paper from 262 reviewers and the area chairs. A final paper selection meeting took place during two days in early June 2005 in Chapel Hill, North Carolina. We are especially grateful to Elizabeth Bullitt, Polina Golland, David Haynor, Rasmus Larsen, Greg Hager and Daniel Rückert, who attended this meeting and helped us make the final selections. Martin Styner provided valuable help with information management and the Web-site, and James Stewart is acknowledged for reliable and timely support of the Web-based reviewing system. We are grateful to everyone who participated in the review process; they donated a large amount of time and effort to make these volumes possible and insure a high level of quality. Because of the overall quality of the submissions and because of the limited number of slots available for presentation, paper selection was especially challenging. The MICCAI 2005 Program Committee finally accepted 236 full papers. The normal mode of presentation at MICCAI 2005 was as a poster; in addition, 46 papers were chosen for oral presentation. All of the full papers accepted are included in these proceedings in 8-page format. We also accepted 34 short communications (2 pages) which were presented as posters but not included in the proceedings.

The first figure below shows the distribution of the 236 full paper contributions by topic; the topics are defined by the primary keyword of the submission. The second figure illustrates the distribution of full paper submissions (a total of 632) by region.

We note that this year's program included some new features, including a session on Cellular and Molecular Imaging and Analysis. We hope that all who

attended the 2005 meeting felt as we do that the program was both strong and diverse, within the range of topics covered by MICCAI.

It was our pleasure to welcome this year's MICCAI 2005 attendees to Palm Springs. Sitting in lush farming land, Palm Springs does not conform to any typical image of the desert, embodying a mix of Spanish Colonial and mid-twentieth century modern styling. Ever since Hollywood stars first came here in the 1930s, laying claim to ranch-style estates, holing up in elite hotels, and enjoying the clean dry air and sunshine, Palm Springs has been a special place to visit. We hope that the attendees, in addition to visiting the conference, took the opportunity to enjoy the hospitality and amenities of the Riviera Resort, and to explore the city, the desert region, and other parts of Southern California. For those unable to attend, we trust that these volumes will provide a valuable record of the state of the art in the MICCAI disciplines.

We also want to thank both our sponsors who are listed below and our two keynote speakers, Profs. Scott Fraser from Caltech and Arthur Toga from UCLA for excellent and stimulating lectures.

Finally, we note that this year a landmark event occurred in the life of MICCAI, namely the formation of the Medical Image Computing and Computer-Assisted Intervention Society (the MICCAI Society) which was officially announced on December 9, 2004. The main focus of the society is our annual international conference series (www.miccai.org) which has become the premier conference in the field of medical image computing and computer-assisted interventions, including biomedical imaging and robotics. The society is governed and administered by the MICCAI Board of Directors. The society will continue to publish the proceedings of the annual MICCAI conference in a prestigious scientific series. Having a paper accepted for publication in this series is highly meritorious and on a par with publication in highly regarded peer-reviewed journals in the field. The society is negotiating with three journals in the field of MICCAI themes, each to become "an affiliated MICCAI journal". These journals will offer significant benefits to members, including sharply discounted rates for paper subscriptions and access to on-line content. The society will continue to develop, enrich, and maintain a dynamic website with exclusive content for members (www.miccai.org).

We look forward to welcoming you to MICCAI 2006, to be held October 2–4, 2006 in Copenhagen, Denmark, and chaired by Mads Nielsen.

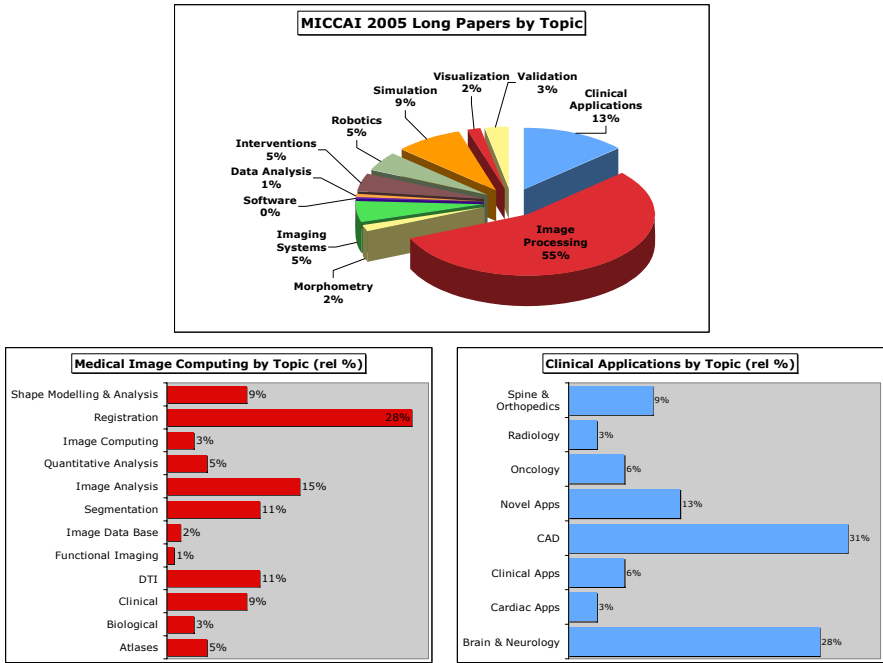


Fig. 1. View at a glance of MICCAI 2005 full paper contributions based on the declared primary keyword. A total of 236 full papers were presented

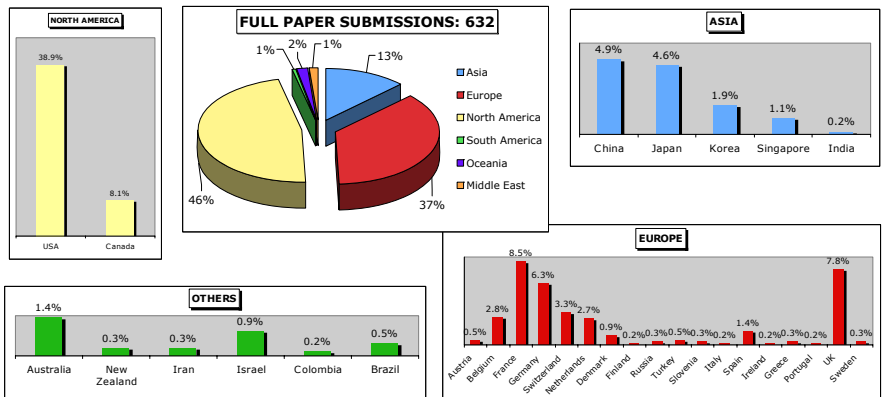


Fig. 2. Distribution of MICCAI 2005 full paper submissions (632 in total) by region

MICCAI Student Awards

Every year MICCAI awards outstanding work written and presented by students. Both oral and poster presentations are eligible for the awards, and the awards are presented to the winners in a public ceremony. Student awards at MICCAI 2003 and 2004 were sponsored by Northern Digital Incorporation (NDI), and NDI will also be the sponsor for the MICCAI 2005 awards.

MICCAI 2003 Student Awards

Robotics: Hashimoto, Ryuji: A Transurethral Prostate Resection Manipulator for Minimal Damage to Mucous Membrane

Segmentation: Pichon, Eric: A Statistically Based Surface Evolution Method for Medical Image Segmentation: Presentation and Validation

Image Guided Therapy Surgery: DiMaio, Simon: Needle Steering and Model-Based Trajectory Planning

Medical Image Analysis: Fillard, Pierre: Quantitative Analysis of White Matter Fiber Properties Along Geodesic Paths

Medical Image Processing and Visualization: Arsigny, Vincent: Polyrigid and Polyaffine Transformations: A New Class of Diffeomorphisms

MICCAI 2004 Student Awards

Image Segmentation and Processing: Dikici, Engin: Quantification of Delayed Enhancement MR Images

Image Registration and Analysis: Perperidis, Dimitrios: Spatio-temporal Free-Form Registration of Cardiac MR Image Sequences

Image Guided Therapy and Robotics: Stoyanov, Danail: Dense 3D Depth Recovery for Soft Tissue Deformation During Robotically Assisted Laparoscopic Surgery

Image Simulation and Display: Valtorta, Davide: Dynamic Measurements of Soft Tissue Viscoelastic Properties with a Torsional Resonator Device

Organization

Executive Committee

James Duncan (General Chair), New Haven, USA
Guido Gerig (Program Chair), Chapel Hill, USA
Christian Barillot, Rennes, France
Mads Nielsen, Copenhagen, Denmark
Terry Peters, London, Canada

Program Committee

Nicholas Ayache, Sophia-Antipolis, France
Elizabeth Bullitt, Chapel Hill, USA
Christos Davatzikos, Philadelphia, USA
Polina Golland, Boston, USA
Gregory D. Hager, Baltimore, USA
David Haynor, Seattle, USA
Leo Joskowicz, Jerusalem, Israel
Rasmus Larsen, Copenhagen, Denmark
Keith Paulsen, Hanover, USA
Daniel Rückert, London, UK
Ross Whitaker, Salt Lake City, USA

MICCAI Board

Alan Colchester (General Chair), Canterbury, UK
Nicholas Ayache, Sophia Antipolis, France
Christian Barillot, Rennes, France
Takeyoshi Dohi, Tokyo, Japan
James Duncan, New Haven, USA
Karl Heinz Höhne, Hamburg, Germany
Anthony Maeder, Brisbane, Queensland, Australia
Mads Nielsen, Copenhagen, Denmark
Terry Peters, London, Ontario, Canada
Richard Robb, Rochester, USA
Max Viergever, Utrecht, The Netherlands

MICCAI Society

Society Officers

President and Chair	Alan Colchester, Canterbury, UK
Executive Director	Richard Robb, Rochester, USA
Vice Chair (2005)	Christian Barillot, Rennes, France
Executive Secretary	Nicholas Ayache, Sophia Antipolis, France
Treasurer	Terry Peters, London, Ontario, Canada

Society Staff

Membership Coordinator	Gabor Székely, Zurich, Switzerland
Publications Coordinator	Nobuhiko Hata, Boston, USA

Tutorial Chair

Larry Staib, New Haven, CT, USA

Information Management and Web Administration

Martin Styner, Chapel Hill, NC, USA

Poster Coordination

Isabelle Corouge, Chapel Hill, WA, USA

Sponsors and Exhibits Coordination

Tina Kapur, Seattle, USA

Student Awards Coordination

Karl Heinz Höhne, Hamburg, Germany

Local Conference Management/Secretariat

Carolyn Meloling, New Haven, CT, USA
Delphine Bull, Chapel Hill, NC, USA

Sponsors

We are grateful to the following organizations for their generous financial support of MICCAI 2005:

- Northern Digital Inc. of Waterloo, Ontario
- Springer Lecture Notes in Computer Science (LNCS)
- GE Healthcare
- Medtronic Navigation
- Siemens Corporate Research
- NIBIB

Furthermore, the following organizations put up financial support: Claron Technology, Ascension Technology Corporation, BrainLAB and Elsevier.

Reviewers

Alberola-López, Carlos	Cheng, Lishui
Archip, Neculai	Christensen, Gary
Asano, Takehide	Chung, Moo
Ashburner, John	Colchester, Alan
Audette, Michel	Collins, D. Louis
Ayache, Nicholas	Colliot, Olivier
Aylward, Stephen	Comaniciu, Dorin
Azar, Fred S.	Corouge, Isabelle
Bardinet, Eric	Coulon, Olivier
Barillot, Christian	Crouch, Jessica
Barratt, Dean	Dam, Erik
Bello, Fernando	Davatzikos, Christos
Berger, Marie-Odile	Davis, Brad
Bhotika, Rahul	Dawant, Benoit
Boctor, Emad	de Bruijne, Marleen
Bouix, Sylvain	Delingette, Hervé
Bucholz, Richard	Descoteaux, Maxime
Bullitt, Elizabeth	Desvignes, Michel
Burghart, Catherina R.	Dillenseger, Jean-Louis
Cahill, Nathan	DiMaio, Simon
Camara-Rey, Oscar	Duchesne, Simon
Castro, Alfonso	Duncan, James S.
Castro-Pareja, Carlos	El-Baz, Ayman
Cen, Feng	Ellis, Randy
Chang, Hsun-Hsien	Falcão, Alexandre
Chang, Sukmoon	Fan, Yong
Chen, Jian	Faugeras, Olivier

Fenster, Aaron
 Fichtinger, Gabor
 Forest, Clément
 Frangi, Alejandro
 Friedman, Allan
 Friman, Ola
 Gan, Rui
 Gee, James
 Gerig, Guido
 Gering, David
 Gerritsen, Frans
 Gibaud, Bernard
 Golland, Polina
 Gonzalez Ballester, Miguel Angel
 Gooding, Mark
 Grau, Vicente
 Grimson, Eric
 Haber, Eldad
 Hager, Gregory D
 Hahn, Horst
 Haidar, Haissam
 Haigron, Pascal
 Hamarneh, Ghassan
 Harders, Matthias
 Hashizume, Makoto
 Hassouna, Mohamed
 Hastenteufel, Mark
 Hastreiter, Peter
 Hata, Nobuhiko
 Hawkes, David
 Haynor, David
 He, Ying
 Hedjidemetriou, Efstathios
 Hellier, Pierre
 Heng, Pheng Ann
 Hibbard, Lyn
 Holmes, David
 Höhne, Karl Heinz
 Hong, Byung-Woo
 Hong, Helen
 Hornegger, Joachim
 Howe, Robert
 Huang, Heng
 Iseki, Hiroshi
 Jannin, Pierre
 Jaramaz, Branislav
 Jomier, Julien
 Joshi, Sarang
 Joskowicz, Leo
 Kang, MyungJin
 Karron, DB
 Kaus, Michael
 Keeve, Erwin
 Kerrien, Erwan
 Khamene, Ali
 Kikinis, Ron
 Kim, Sun I.
 Kissi, Adelaide
 Kronreif, Gernot
 Kruggel, Frithjof
 Lötjönen, Jyrki
 Lange, Thomas
 Lapeer, Rudy
 Larsen, Rasmus
 Leahy, Richard
 Lee, Sang-Chul
 Li, Rongxin
 Li, Shuo
 Liang, Jianming
 Liao, Hongen
 Liao, Rui
 Lienard, Jean
 Linguraru, Marius George
 Liu, Alan
 Liu, Huafeng
 Liu, Jundong
 Liu, Tianming
 Loog, Marco
 Lorensen, William
 Lorenzen, Peter
 Macq, Benoit
 Madabhushi, Anant
 Maddah, Mahnaz
 Maeder, Anthony
 Maes, Frederik
 Makram-Ebeid, Sherif
 Malandain, Gregoire
 Mancas, Matei
 Mangin, Jean-Francois
 Martin-Fernandez, Marcos

Mattes, Julian
Maurer, Calvin
McInerney, Tim
Metaxas, Dimitris
Meyer, Chuck
Miga, Michael I.
Mori, Kensaku
Muragaki, Yoshihiro
Navab, Nassir
Nicolau, Stephane
Niessen, Wiro
Noble, Alison
Nolte, Lutz
O'Donnell, Thomas
Okada, Kazunori
Olabarriaga, Silvia
Olsen, Ole Fogh
Ordas, Sebastian
Ortmaier, Tobias
Ourselin, Sebastien
Palm, Christoph
Papademetris, Xenophon
Paragios, Nikos
Patriciu, Alexandru
Paulsen, Keith
Peitgen, Heinz-Otto
Pelegriani-Issac, Mélanie
Pennec, Xavier
Peters, Terry M.
Pham, Dzung
Pichon, Eric
Pizer, Stephen
Pluim, Josien
Pommert, Andreas
Prager, Richard
Prastawa, Marcel
Prima, Sylvain
Prince, Jerry L.
Pujol, Sonia
Robb, Richard A.
Rohlfing, Torsten
Rohling, Robert
Rohr, Karl
Rosenthal, Michael
Rosiene, Joel
Rueckert, Daniel
Ruiz-Alzola, Juan
Russakoff, Daniel
Sato, Yoshinobu
Sauer, Frank
Schnabel, Julia
Shen, Dinggang
Shi, Pengcheng
Shimizu, Akinobu
Skrinjar, Oskar
Sonka, Milan
Stewart, James
Stokking, Rik
Studholme, Colin
Styner, Martin
Subramanian, Navneeth
Suetens, Paul
Sun, Hai
Syeda-Mahmood, Tanveer
Székely, Gabor
Tannenbaum, Allen
Tao, Xiaodong
Tasdizen, Tolga
Taylor, Chris
Taylor, Russell
Tek, Huseyin
ter Haar Romeny, Bart M.
Terzopoulos, Demetri
Thiran, Jean-Philippe
Thirion, Bertrand
Tohka, Jussi
Tosun, Duygu
Troccaz, Jocelyne
Vaillant, Regis
van Assen, Hans
van Ginneken, Bram
Van Leemput, Koen
Vandermeulen, Dirk
Vermeer, Koen
Vetsa, Yethiraja Sampath K.
Vos, Frans
Vosburgh, Kirby
Vossepoel, Albert
Wai, Lionel C.C.
Wang, Jianzhe

Wang, Yalin
Wang, Yongmei Michelle
Wang, Yuanquan
Warfield, Simon
Westin, Carl-Fredrik
Whitaker, Ross
Wildermuth, Simon
Williams, James
Wong, Wilbur C.K.
Wu, Chia-Hsiang
Wu, John Jue
Wu, Ziji
Wyatt, Chris
Xu, Xiaoyin

Xue, Zhong
Yan, Pingkun
Yang, Guang-Zhong
Yang, Jing
Yaniv, Ziv
Yoo, Terry
Yuan, Xiaohui
Yushkevich, Paul
Zeng, Jianchao
Zhao, Zheen
Zheng, Guoyan
Zhu, Chaozhe
Zou, Kelly
Zrimec, Tatjana

Table of Contents – Part I

Image Analysis and Validation

Classification of Structural Images via High-Dimensional Image Warping, Robust Feature Extraction, and SVM <i>Yong Fan, Dinggang Shen, Christos Davatzikos</i>	1
Bone Enhancement Filtering: Application to Sinus Bone Segmentation and Simulation of Pituitary Surgery <i>Maxime Descoteaux, Michel Audette, Kiyoyuki Chinzei, Kaleem Siddiqi</i>	9
Simultaneous Registration and Segmentation of Anatomical Structures from Brain MRI <i>Fei Wang, Baba C. Vemuri</i>	17
Synthetic Ground Truth for Validation of Brain Tumor MRI Segmentation <i>Marcel Prastawa, Elizabeth Bullitt, Guido Gerig</i>	26

Vascular Image Segmentation

Automatic Cerebrovascular Segmentation by Accurate Probabilistic Modeling of TOF-MRA Images <i>Ayman El-Baz, Aly A. Farag, Georgy Gimel'farb, Stephen G. Hushek</i>	34
A Segmentation and Reconstruction Technique for 3D Vascular Structures <i>Vincent Luboz, Xunlei Wu, Karl Krissian, Carl-Fredrik Westin, Ron Kikinis, Stéphane Cotin, Steve Dawson</i>	43
MRA Image Segmentation with Capillary Active Contour <i>Pingkun Yan, Ashraf A. Kassim</i>	51
Spatial Graphs for Intra-cranial Vascular Network Characterization, Generation, and Discrimination <i>Stephen R. Aylward, Julien Jomier, Christelle Vivert, Vincent LeDigarcher, Elizabeth Bullitt</i>	59

Image Registration I

Surface Alignment of 3D Spherical Harmonic Models: Application to Cardiac MRI Analysis <i>Heng Huang, Li Shen, Rong Zhang, Fillia Makedon, Bruce Hettleman, Justin Pearlman</i>	67
Unified Point Selection and Surface-Based Registration Using a Particle Filter <i>Burton Ma, Randy E. Ellis</i>	75
Elastic Registration of 3D Ultrasound Images <i>Pezhman Foroughi, Purang Abolmaesumi</i>	83
Tracer Kinetic Model-Driven Registration for Dynamic Contrast Enhanced MRI Time Series <i>Giovanni A. Buonaccorsi, Caleb Roberts, Sue Cheung, Yvonne Watson, Karen Davies, Alan Jackson, Gordon C. Jayson, Geoff J.M. Parker</i>	91
Generalised Overlap Measures for Assessment of Pairwise and Groupwise Image Registration and Segmentation <i>William R. Crum, Oscar Camara, Daniel Rueckert, Kanwal K. Bhatia, Mark Jenkinson, Derek L.G. Hill</i>	99
 Diffusion Tensor Image Analysis	
Uncertainty in White Matter Fiber Tractography <i>Ola Friman, Carl-Fredrik Westin</i>	107
Fast and Simple Calculus on Tensors in the Log-Euclidean Framework <i>Vincent Arsigny, Pierre Fillard, Xavier Pennec, Nicholas Ayache</i>	115
3D Curve Inference for Diffusion MRI Regularization <i>Peter Savadjiev, Jennifer S.W. Campbell, G. Bruce Pike, Kaleem Siddiqi</i>	123
Fiber Tract-Oriented Statistics for Quantitative Diffusion Tensor MRI Analysis <i>Isabelle Corouge, P. Thomas Fletcher, Sarang Joshi, John H. Gilmore, Guido Gerig</i>	131
White Matter Tract Clustering and Correspondence in Populations <i>Lauren O'Donnell, Carl-Fredrik Westin</i>	140

76-Space Analysis of Grey Matter Diffusivity: Methods and Applications <i>Tianming Liu, Geoffrey Young, Ling Huang, Nan-Kuei Chen, Stephen TC Wong</i>	148
Fast Orientation Mapping from HARDI <i>Evren Özarslan, Timothy M. Shepherd, Baba C. Vemuri, Stephen J. Blackband, Thomas H. Mareci</i>	156
An Automated Approach to Connectivity-Based Partitioning of Brain Structures <i>P.A. Cook, H. Zhang, B.B. Avants, P. Yushkevich, D.C. Alexander, J.C. Gee, O. Ciccarelli, A.J. Thompson</i>	164
Deformable Registration of Diffusion Tensor MR Images with Explicit Orientation Optimization <i>Hui Zhang, Paul A. Yushkevich, James C. Gee</i>	172
A Hamilton-Jacobi-Bellman Approach to High Angular Resolution Diffusion Tractography <i>Eric Pichon, Carl-Fredrik Westin, Allen R. Tannenbaum</i>	180
Automated Atlas-Based Clustering of White Matter Fiber Tracts from DTMRI <i>Mahnaz Maddah, Andrea U.J. Mewes, Steven Haker, W. Eric L. Grimson, Simon K. Warfield</i>	188
MR Diffusion-Based Inference of a Fiber Bundle Model from a Population of Subjects <i>V. El Kouby, Y. Cointepas, C. Poupon, D. Rivière, N. Golestani, J.-B. Poline, D. Le Bihan, J.-F. Mangin</i>	196
Knowledge-Based Classification of Neuronal Fibers in Entire Brain <i>Yan Xia, And U. Turken, Susan L. Whitfield-Gabrieli, John D. Gabrieli</i>	205
A Physical Model for DT-MRI Based Connectivity Map Computation <i>Erdem Yörük, Burak Acar, Roland Bammer</i>	213
Image Segmentation and Analysis I	
A Novel 3D Partitioned Active Shape Model for Segmentation of Brain MR Images <i>Zheen Zhao, Stephen R. Aylward, Eam Khwang Teoh</i>	221

Cross Entropy: A New Solver for Markov Random Field Modeling and Applications to Medical Image Segmentation <i>Jue Wu, Albert C.S. Chung</i>	229
Semi-automated Basal Ganglia Segmentation Using Large Deformation Diffeomorphic Metric Mapping <i>Ali Khan, Elizabeth Aylward, Patrick Barta, Michael Miller, M. Faisal Beg</i>	238
Particle Filters, a Quasi-Monte Carlo Solution for Segmentation of Coronaries <i>Charles Florin, Nikos Paragios, Jim Williams</i>	246
Hybrid Segmentation Framework for Tissue Images Containing Gene Expression Data <i>Musodiq Bello, Tao Ju, Joe Warren, James Carson, Wah Chiu, Christina Thaller, Gregor Eichele, Ioannis A. Kakadiaris</i>	254
Fully Automatic Kidneys Detection in 2D CT Images: A Statistical Approach <i>Wala Touhami, Djamel Boukerroui, Jean-Pierre Cocquerez</i>	262
Segmentation of Neighboring Organs in Medical Image with Model Competition <i>Pingkun Yan, Weijia Shen, Ashraf A. Kassim, Mubarak Shah</i>	270
Point-Based Geometric Deformable Models for Medical Image Segmentation <i>Hon Pong Ho, Yunmei Chen, Huafeng Liu, Pengcheng Shi</i>	278
A Variational PDE Based Level Set Method for a Simultaneous Segmentation and Non-rigid Registration <i>Jung-ha An, Yunmei Chen, Feng Huang, David Wilson, Edward Geiser</i>	286
A Tracking Approach to Parcellation of the Cerebral Cortex <i>Chris Adamson, Leigh Johnston, Terrie Inder, Sandra Rees, Iven Mareels, Gary Egan</i>	294
Cell Segmentation, Tracking, and Mitosis Detection Using Temporal Context <i>Fuxing Yang, Michael A. Mackey, Fiorenza Ianzini, Greg Gallardo, Milan Sonka</i>	302

A Unifying Approach to Registration, Segmentation, and Intensity Correction <i>Kilian M. Pohl, John Fisher, James J. Levitt, Martha E. Shenton, Ron Kikinis, W. Eric L. Grimson, William M. Wells</i>	310
Automatic 3D Segmentation of Intravascular Ultrasound Images Using Region and Contour Information <i>Marie-Hélène Roy Cardinal, Jean Meunier, Gilles Soulez, Roch L. Maurice, Éric Thérasse, Guy Cloutier</i>	319
Automatic Segmentation of the Articular Cartilage in Knee MRI Using a Hierarchical Multi-class Classification Scheme <i>Jenny Folkesson, Erik Dam, Ole Fogh Olsen, Paola Pettersen, Claus Christiansen</i>	327
Automatic Segmentation of the Left Ventricle in 3D SPECT Data by Registration with a Dynamic Anatomic Model <i>Lars Dornheim, Klaus D. Tönnies, Kat Dixon</i>	335
Intravascular Ultrasound-Based Imaging of Vasa Vasorum for the Detection of Vulnerable Atherosclerotic Plaque <i>Sean M. O'Malley, Manolis Vavuranakis, Morteza Naghavi, Ioannis A. Kakadiaris</i>	343
Parametric Response Surface Models for Analysis of Multi-site fMRI Data <i>Seyoung Kim, Padhraic Smyth, Hal Stern, Jessica Turner</i>	352
Clinical Applications – Validation	
Subject Specific Finite Element Modelling of the Levator Ani <i>Su-Lin Lee, Ara Darzi, Guang-Zhong Yang</i>	360
Robust Visualization of the Dental Occlusion by a Double Scan Procedure <i>Filip Schutyser, Gwen Swennen, Paul Suetens</i>	368
Segmentation of Focal Cortical Dysplasia Lesions Using a Feature-Based Level Set <i>O. Colliot, T. Mansi, N. Bernasconi, V. Naessens, D. Klironomos, A. Bernasconi</i>	375

Effects of Healthy Aging Measured By Intracranial Compartment Volumes Using a Designed MR Brain Database
Bénédicte Mortamet, Donglin Zeng, Guido Gerig, Marcel Prastawa, Elizabeth Bullitt 383

Predicting Clinical Variable from MRI Features: Application to MMSE in MCI
S. Duchesne, A. Caroli, C. Geroldi, G.B. Frisoni, D. Louis Collins 392

Finite Element Modeling of Brain Tumor Mass-Effect from 3D Medical Images
Ashraf Mohamed, Christos Davatzikos 400

STREM: A Robust Multidimensional Parametric Method to Segment MS Lesions in MRI
L.S. Ait-Ali, S. Prima, P. Hellier, B. Carsin, G. Edan, C. Barillot 409

Cross Validation of Experts Versus Registration Methods for Target Localization in Deep Brain Stimulation
F. Javier Sánchez Castro, Claudio Pollo, Reto Meuli, Philippe Maeder, Meritzell Bach Cuadra, Olivier Cuisenaire, Jean-Guy Villemure, Jean-Philippe Thiran 417

Localization of Abnormal Conduction Pathways for Tachyarrhythmia Treatment Using Tagged MRI
G.I. Sanchez-Ortiz, M. Sermesant, K.S. Rhode, R. Chandrashekar, R. Razavi, D.L.G. Hill, D. Rueckert 425

Automatic Mammary Duct Detection in 3D Ultrasound
Mark J. Gooding, Matthew Mellor, Jacqueline A. Shipley, Kathy A. Broadbent, Dorothy A. Goddard 434

Automatic Segmentation of Intra-treatment CT Images for Adaptive Radiation Therapy of the Prostate
B.C. Davis, M. Foskey, J. Rosenman, L. Goyal, S. Chang, S. Joshi 442

Inter-Operator Variability in Perfusion Assessment of Tumors in MRI Using Automated AIF Detection
Edward Ashton, Teresa McShane, Jeffrey Evelhoch 451

Computer-Assisted Deformity Correction Using the Ilizarov Method
A.L. Simpson, B. Ma, D.P. Borschneck, R.E. Ellis 459

Real-Time Interactive Viewing of 4D Kinematic MR Joint Studies <i>Heinrich Schulz, Kirsten Meetz, Clemens Bos, Daniel Bystrov, Thomas Netsch</i>	467
Computer-Assisted Ankle Joint Arthroplasty Using Bio-engineered Autografts <i>R. Sidler, W. Köstler, T. Bardyn, M.A. Styner, N. Südkamp, L. Nolte, M.Á. González Ballester</i>	474
Prospective Head Motion Compensation for MRI by Updating the Gradients and Radio Frequency During Data Acquisition <i>Christian Dold, Maxim Zaitsev, Oliver Speck, Evelyn A. Firle, Jürgen Hennig, Georgios Sakas</i>	482
Harmonic Skeleton Guided Evaluation of Stenoses in Human Coronary Arteries <i>Yan Yang, Lei Zhu, Steven Haker, Allen R. Tannenbaum, Don P. Giddens</i>	490
Acquisition-Related Limitations in MRI Based Morphometry <i>Arne Littmann, Jens Guehring, Christian Buechel, Hans-Siegfried Stiehl</i>	498
Combining Classifiers Using Their Receiver Operating Characteristics and Maximum Likelihood Estimation <i>Steven Haker, William M. Wells III, Simon K. Warfield, Ion-Florin Talos, Jui G. Bhagwat, Daniel Goldberg-Zimring, Asim Mian, Lucila Ohno-Machado, Kelly H. Zou</i>	506
Two Methods for Validating Brain Tissue Classifiers <i>Marcos Martin-Fernandez, Sylvain Bouix, Lida Ungar, Robert W. McCarley, Martha E. Shenton</i>	515
Comparison of Vessel Segmentations Using STAPLE <i>Julien Jomier, Vincent LeDigarcher, Stephen R. Aylward</i>	523
Validation Framework of the Finite Element Modeling of Liver Tissue <i>Hongjian Shi, Rachid Fahmi, Aly A. Farag</i>	531
A Complete Augmented Reality Guidance System for Liver Punctures: First Clinical Evaluation <i>S.A. Nicolau, X. Pennec, L. Soler, N. Ayache</i>	539

Imaging Systems – Visualization

A Novel Approach to High Resolution Fetal Brain MR Imaging <i>F. Rousseau, O. Glenn, B. Iordanova, C. Rodriguez-Carranza, D. Vigneron, J. Barkovich, C. Studholme</i>	548
Respiratory Signal Extraction for 4D CT Imaging of the Thorax from Cone-Beam CT Projections <i>Simon Rit, David Sarrut, Chantal Ginestet</i>	556
Registering Liver Pathological Images with Prior In Vivo CT/MRI Data <i>Huadong Wu, Alyssa M. Krasinskas, Mitchell E. Tublin, Brian E. Chapman</i>	564
Support Vector Clustering for Brain Activation Detection <i>Defeng Wang, Lin Shi, Daniel S. Yeung, Pheng-Ann Heng, Tien-Tsin Wong, Eric C.C. Tsang</i>	572
Inter-frame Motion Correction for MR Thermometry <i>S. Suprijanto, M.W. Vogel, F.M. Vos, H.A. Vrooman, A.M. Vossepoel</i>	580
Adaptive Multiscale Ultrasound Compounding Using Phase Information <i>Vicente Grau, J. Alison Noble</i>	589
3D Freehand Ultrasound Reconstruction Based on Probe Trajectory <i>Pierrick Coupé, Pierre Hellier, Noura Azzabou, Christian Barillot</i>	597
Self-Calibrating Ultrasound-to-CT Bone Registration <i>Dean C. Barratt, Graeme Penney, Carolyn S.K. Chan, Mike Slomczykowski, Timothy J. Carter, Philip J. Edwards, David J. Hawkes</i>	605
A Hand-Held Probe for Vibro-Elastography <i>Hassan Rivaz, Robert Rohling</i>	613
Real-Time Quality Control of Tracked Ultrasound <i>Emad M. Boctor, Iulian Iordachita, Gabor Fichtinger, Gregory D. Hager</i>	621
Fully Truncated Cone-Beam Reconstruction on Pi Lines Using Prior CT <i>Krishnakumar Ramamurthi, Norbert Strobel, Rebecca Fahrig, Jerry L. Prince</i>	631

C-arm Calibration - Is it Really Necessary? <i>Ameet Jain, Ryan Kon, Yu Zhou, Gabor Fichtinger</i>	639
Laser Needle Guide for the Sonic Flashlight <i>David Wang, Bing Wu, George Stetten</i>	647
Differential Fly-Throughs (DFT): A General Framework for Computing Flight Paths <i>M. Sabry Hassouna, Aly A. Farag, Robert Falk</i>	654
Panoramic Views for Virtual Endoscopy <i>Bernhard Geiger, Christophe Chefd'hotel, Sandra Sudarsky</i>	662
Computer Assisted Diagnosis	
Toward Automatic Computer Aided Dental X-ray Analysis Using Level Set Method <i>Shuo Li, Thomas Fevens, Adam Krzyżak, Chao Jin, Song Li</i>	670
Exploiting Temporal Information in Functional Magnetic Resonance Imaging Brain Data <i>Lei Zhang, Dimitris Samaras, Dardo Tomasi, Nelly Alia-Klein, Lisa Cottone, Andreana Leskovjan, Nora Volkow, Rita Goldstein</i>	679
Model-Based Analysis of Local Shape for Lesion Detection in CT Scans <i>Paulo R.S. Mendonça, Rahul Bhotika, Saad A. Sirohey, Wesley D. Turner, James V. Miller, Ricardo S. Avila</i>	688
Development of a Navigation-Based CAD System for Colon <i>Masahiro Oda, Takayuki Kitasaka, Yuichiro Hayashi, Kensaku Mori, Yasuhito Suenaga, Jun-ichiro Toriwaki</i>	696
A Prediction Framework for Cardiac Resynchronization Therapy Via 4D Cardiac Motion Analysis <i>Heng Huang, Li Shen, Rong Zhang, Fillia Makedon, Bruce Hettleman, Justin Pearlman</i>	704
Segmentation and Size Measurement of Polyps in CT Colonography <i>J.J. Dijkers, C. van Wijk, F.M. Vos, J. Florie, Y.C. Nio, H.W. Venema, R. Truyen, L.J. van Vliet</i>	712
Quantitative Nodule Detection in Low Dose Chest CT Scans: New Template Modeling and Evaluation for CAD System Design <i>Aly A. Farag, Ayman El-Baz, Georgy Gimel'farb, Mohamed Abou El-Ghar, Tarek Eldiasty</i>	720

Graph Embedding to Improve Supervised Classification and Novel Class Detection: Application to Prostate Cancer
Anant Madabhushi, Jianbo Shi, Mark Rosen, John E. Tomaszewski, Michael D. Feldman 729

Quantification of Emphysema Severity by Histogram Analysis of CT Scans
Paulo R.S. Mendonça, Dirk R. Padfield, James C. Ross, James V. Miller, Sandeep Dutta, Sardar Mal Gautham 738

Cellular and Molecular Image Analysis

Efficient Learning by Combining Confidence-Rated Classifiers to Incorporate Unlabeled Medical Data
Weijun He, Xiaolei Huang, Dimitris Metaxas, Xiaoyou Ying 745

Mosaicing of Confocal Microscopic *In Vivo* Soft Tissue Video Sequences
Tom Vercauteren, Aymeric Perchant, Xavier Pennec, Nicholas Ayache 753

Segmentation and 3D Reconstruction of Microtubules in Total Internal Reflection Fluorescence Microscopy (TIRFM)
Stathis Hadjidemetriou, Derek Toomre, James S. Duncan 761

Physically-Based Modeling

Ligament Strains Predict Knee Motion After Total Joint Replacement – A Kinematic Analysis of The Sigma Knee
Elvis C.S. Chen, Joel L. Lanovaz, Randy E. Ellis 770

A Boundary Element-Based Approach to Analysis of LV Deformation
Ping Yan, Ning Lin, Albert J. Sinusas, James S. Duncan 778

Reconstruction of Cerebrospinal Fluid Flow in the Third Ventricle Based on MRI Data
Vartan Kurtcuoglu, Michaela Soellinger, Paul Summers, Kevin Boomsma, Dimos Poulidakos, Peter Boesiger, Yiannis Ventikos 786

Schwarz Meets Schwann: Design and Fabrication of Biomorphic Tissue Engineering Scaffolds
Srinivasan Rajagopalan, Richard A. Robb 794

Robotics and Intervention I

Automatic Detection and Segmentation of Robot-Assisted Surgical Motions <i>Henry C. Lin, Izhak Shafran, Todd E. Murphy, Allison M. Okamura, David D. Yuh, Gregory D. Hager</i>	802
DaVinci Canvas: A Telerobotic Surgical System with Integrated, Robot-Assisted, Laparoscopic Ultrasound Capability <i>Joshua Leven, Darius Burschka, Rajesh Kumar, Gary Zhang, Steve Blumenkranz, Xiangtian (Donald) Dai, Mike Awad, Gregory D. Hager, Mike Marohn, Mike Choti, Chris Hasser, Russell H. Taylor</i>	811
Design and Control of In-Vivo Magnetic Microrobots <i>K. Berk Yesin, Philipp Exner, Karl Vollmers, Bradley J. Nelson</i>	819
3D Needle-Tissue Interaction Simulation for Prostate Brachytherapy <i>Orcun Goksel, Septimiu E. Salcudean, Simon P. DiMaio, Robert Rohling, James Morris</i>	827
Development and Application of Functional Databases for Planning Deep-Brain Neurosurgical Procedures <i>Ting Guo, Kirk W. Finnis, Andrew G. Parrent, Terry M. Peters</i>	835
 Medical Image Computing for Clinical Applications	
Gaze-Contingent Soft Tissue Deformation Tracking for Minimally Invasive Robotic Surgery <i>George P. Mylonas, Danail Stoyanov, Fani Deligianni, Ara Darzi, Guang-Zhong Yang</i>	843
Registration and Integration for Fluoroscopy Device Enhancement <i>James C. Ross, David Langan, Ravi Manjeshwar, John Kaufhold, Joseph Manak, David Wilson</i>	851
Computer Aided Detection for Low-Dose CT Colonography <i>Gabriel Kiss, Johan Van Cleynenbreugel, Stylianos Drisis, Didier Bielen, Guy Marchal, Paul Suetens</i>	859
Photo-Realistic Tissue Reflectance Modelling for Minimally Invasive Surgical Simulation <i>Mohamed A. ElHelw, Stella Atkins, Marios Nicolaou, Adrian Chung, Guang-Zhong Yang</i>	868

Biological Imaging - Simulation and Modeling I

Motion Tracking and Intensity Surface Recovery in Microscopic Nuclear Images <i>Lopamudra Mukherjee, Mingen Lin, Jinhui Xu, Ronald Berezney</i>	876
Towards Automated Cellular Image Segmentation for RNAi Genome-Wide Screening <i>Xiaobo Zhou, K.-Y. Liu, P. Bradley, N. Perrimon, Stephen TC Wong</i>	885
Adaptive Spatio-Temporal Restoration for 4D Fluorescence Microscopic Imaging <i>Jérôme Boulanger, Charles Kervrann, Patrick Bouthemy</i>	893
Kinematic Geometry of Osteotomies <i>Erin J. Smith, J. Tim Bryant, Randy E. Ellis</i>	902
Predictive Camera Tracking for Bronchoscope Simulation with CONDensation <i>Fani Deligianni, Adrian Chung, Guang-Zhong Yang</i>	910
Experimental Validation of a 3D Dynamic Finite-Element Model of a Total Knee Replacement <i>Joel L. Lanovaz, Randy E. Ellis</i>	917
An In Vitro Patient-Tailored Model of Human Cerebral Artery for Simulating Endovascular Intervention <i>Seiichi Ikeda, Fumihito Arai, Toshio Fukuda, Makoto Negoro, Keiko Irie, Ikuo Takahashi</i>	925
Author Index	933

Table of Contents – Part II

Robotics, Image-Guided Surgery and Interventions

Sensor Guided Ablation Procedure of Left Atrial Endocardium <i>Hua Zhong, Takeo Kanade, David Schwartzman</i>	1
A Method to Evaluate Human Spatial Coordination Interfaces for Computer-Assisted Surgery <i>M.A. Cardin, J.X. Wang, D.B. Plewes</i>	9
3D TRUS Guided Robot Assisted Prostate Brachytherapy <i>Zhouping Wei, Mingyue Ding, Donal Downey, Aaron Fenster</i>	17
Invisible Shadow for Navigation and Planning in Minimal Invasive Surgery <i>Marios Nicolaou, Adam James, Benny P.L. Lo, Ara Darzi, Guang-Zhong Yang</i>	25
A Navigation System for Minimally Invasive CT-Guided Interventions <i>Markus Nagel, Gerd Schmidt, Ralf Petzold, Willi A. Kalender</i>	33
Passive Markers for Ultrasound Tracking of Surgical Instruments <i>Jeffrey Stoll, Pierre Dupont</i>	41
Optimal Trajectories Computation Within Regions of Interest for Hepatic RFA Planning <i>Caroline Villard, Claire Baegert, Pascal Schreck, Luc Soler, Afshin Gangi</i>	49
Effects of Latency on Telesurgery: An Experimental Study <i>Reiza Rayman, Serguei Primak, Rajni Patel, Merhdad Moallem, Roya Morady, Mahdi Tavakoli, Vanja Subotic, Natalie Galbraith, Aimee van Wynsberghe, Kris Croome</i>	57
A Novel Phantom-Less Spatial and Temporal Ultrasound Calibration Method <i>Ali Khamene, Frank Sauer</i>	65
Electromagnetic Tracker Measurement Error Simulation and Tool Design <i>Gregory S. Fischer, Russell H. Taylor</i>	73

Compact Forceps Manipulator Using Friction Wheel Mechanism and Gimbals Mechanism for Laparoscopic Surgery <i>Takashi Suzuki, Youichi Katayama, Etsuko Kobayashi, Ichiro Sakuma</i>	81
Spatial Motion Constraints for Robot Assisted Suturing Using Virtual Fixtures <i>Ankur Kapoor, Ming Li, Russell H. Taylor</i>	89
Contact Force Measurement of Instruments for Force-Feedback on a Surgical Robot: Acceleration Force Cancellations Based on Acceleration Sensor Readings <i>Shigeyuki Shimachi, Fumie Kameyama, Yoshihide Hakozaki, Yasunori Fujiwara</i>	97
Development of the Needle Insertion Robot for Percutaneous Vertebroplasty <i>S. Onogi, K. Morimoto, I. Sakuma, Y. Nakajima, T. Koyama, N. Sugano, Y. Tamura, S. Yonenobu, Y. Momoi</i>	105
Laparoscope Self-calibration for Robotic Assisted Minimally Invasive Surgery <i>Danaïl Stoyanov, Ara Darzi, Guang Zhong Yang</i>	114
A Hand-Eye Robotic Model for Total Knee Replacement Surgery <i>Fanhuaì Shi, Jing Zhang, Yuncai Liu, Zijian Zhao</i>	122
Robot-Assisted Image-Guided Targeting for Minimally Invasive Neurosurgery: Planning, Registration, and In-vitro Experiment <i>R. Shamir, M. Freiman, L. Joskowicz, M. Shoham, E. Zehavi, Y. Shoshan</i>	131
Soft-Tissue Motion Tracking and Structure Estimation for Robotic Assisted MIS Procedures <i>Danaïl Stoyanov, George P. Mylonas, Fani Deligianni, Ara Darzi, Guang-Zhong Yang</i>	139
Image Registration II	
Mass Preserving Registration for Heart MR Images <i>Lei Zhu, Steven Haker, Allen Tannenbaum</i>	147
Liver Registration for the Follow-Up of Hepatic Tumors <i>Arnaud Charnoz, Vincent Agnus, Grégoire Malandain, Clément Forest, Mohamed Tajine, Luc Soler</i>	155

Maximum a Posteriori Local Histogram Estimation for Image Registration <i>Matthew Toews, D. Louis Collins, Tal Arbel</i>	163
Dynamic 3D Ultrasound and MR Image Registration of the Beating Heart <i>Xishi Huang, Nicholas A. Hill, Jing Ren, Gerard Guiraudon, Derek Boughner, Terry M. Peters</i>	171
Learning Best Features for Deformable Registration of MR Brains <i>Guorong Wu, Feihu Qi, Dinggang Shen</i>	179
Stochastic Inverse Consistency in Medical Image Registration <i>Sai Kit Yeung, Pengcheng Shi</i>	188
A Novel Incremental Technique for Ultrasound to CT Bone Surface Registration Using Unscented Kalman Filtering <i>Mehdi Hedjazi Moghari, Purang Abolmaesumi</i>	197
Automatic 4-D Registration in Dynamic MR Renography Based on Over-Complete Dyadic Wavelet and Fourier Transforms <i>Ting Song, Vivian S. Lee, Henry Rusinek, Manmeen Kaur, Andrew F. Laine</i>	205
Model of a Vascular C-Arm for 3D Augmented Fluoroscopy in Interventional Radiology <i>S. Gorges, E. Kerrien, M-O. Berger, Y. Troussel, J. Pescatore, R. Anxionnat, L. Picard</i>	214
2D/3D Deformable Registration Using a Hybrid Atlas <i>Thomas S.Y. Tang, Randy E. Ellis</i>	223
Reconstruction-Based 3D/2D Image Registration <i>Dejan Tomažević, Boštjan Likar, Franjo Pernuš</i>	231
Comparison of Simultaneous and Sequential Two-View Registration for 3D/2D Registration of Vascular Images <i>Chetna Pathak, Mark Van Horn, Susan Weeks, Elizabeth Bullitt</i>	239
Interpolation Artefacts in Non-rigid Registration <i>P. Aljabar, J.V. Hajnal, R.G. Boyes, D. Rueckert</i>	247
Learning Based Non-rigid Multi-modal Image Registration Using Kullback-Leibler Divergence <i>Christoph Guetter, Chenyang Xu, Frank Sauer, Joachim Hornegger</i>	255

Deformable Registration of Brain Tumor Images Via a Statistical Model of Tumor-Induced Deformation
Ashraf Mohamed, Dinggang Shen, Christos Davatzikos 263

Myocardial Motion Estimation in Tagged MR Sequences by Using α MI-Based Non Rigid Registration
E. Oubel, C. Tobon-Gomez, A.O. Hero, A.F. Frangi 271

Iterative 3D Point-Set Registration Based on Hierarchical Vertex Signature (HVS)
Jun Feng, Horace H.S. Ip 279

Automatic Patient Registration for Port Placement in Minimally Invasive Endoscopic Surgery
Marco Feuerstein, Stephen M. Wildhirt, Robert Bauernschmitt, Nassir Navab 287

Hybrid Formulation of the Model-Based Non-rigid Registration Problem to Improve Accuracy and Robustness
Olivier Clatz, Hervé Delingette, Ion-Florin Talos, Alexandra J. Golby, Ron Kikinis, Ferenc A. Jolesz, Nicholas Ayache, Simon K. Warfield 295

Automatic Registration and Fusion of Ultrasound with CT for Radiotherapy
Wolfgang Wein, Barbara Röper, Nassir Navab 303

Medical Image Computing - Atlases - Shape I

Lung Deformation Estimation and Four-Dimensional CT Lung Reconstruction
Sheng Xu, Russell H. Taylor, Gabor Fichtinger, Kevin Cleary 312

Automatic Parameter Optimization for De-noising MR Data
Joaquín Castellanos, Karl Rohr, Thomas Tolxdorff, Gudrun Wagenknecht 320

Towards a Dynamic Model of Pulmonary Parenchymal Deformation: Evaluation of Methods for Temporal Reparameterization of Lung Data
Tessa A. Sundaram, Brian B. Avants, James C. Gee 328

4D MR Imaging Using Internal Respiratory Gating
M. von Siebenthal, Ph. Cattin, U. Gampfer, A. Lomax, G. Székely 336

Anatomically Constrained Surface Parameterization for Cortical Localization
C. Clouchoux, O. Coulon, D. Rivière, A. Cachia, J.-F. Mangin, J. Régis 344

Multiresolution Parametric Estimation of Transparent Motions and Denoising of Fluoroscopic Images
Vincent Auvray, Jean Liénard, Patrick Bouthemy 352

Plaque and Stent Artifact Reduction in Subtraction CT Angiography Using Nonrigid Registration and a Volume Penalty
Dirk Loeckx, Stylianos Drisis, Frederik Maes, Dirk Vandermeulen, Guy Marchal, Paul Suetens 361

Respiratory Motion Correction in Emission Tomography Image Reconstruction
Mauricio Reyes, Grégoire Malandain, Pierre Malick Koulibaly, Miguel A. González Ballester, Jacques Darcourt 369

Optimal Embedding for Shape Indexing in Medical Image Databases
Xiaoning Qian, Hemant D. Tagare 377

Retrospective Cross-Evaluation of an Histological and Deformable 3D Atlas of the Basal Ganglia on Series of Parkinsonian Patients Treated by Deep Brain Stimulation
Eric Bardinet, Didier Dormont, Grégoire Malandain, Manik Bhattacharjee, Bernard Pidoux, Christian Saleh, Philippe Cornu, Nicholas Ayache, Yves Agid, Jérôme Yelnik 385

Anatomical and Electrophysiological Validation of an Atlas for Neurosurgical Planning
M. Mallar Chakravarty, Abbas F. Sadikot, Jurgen Germann, Gilles Bertrand, D. Louis Collins 394

Construction of a 4D Statistical Atlas of the Cardiac Anatomy and Its Use in Classification
Dimitrios Perperidis, Raad Mohiaddin, Daniel Rueckert 402

Unbiased Atlas Formation Via Large Deformations Metric Mapping
Peter Lorenzen, Brad Davis, Sarang Joshi 411

Least Biased Target Selection in Probabilistic Atlas Construction
Hyunjin Park, Peyton H. Bland, Alfred O. Hero III, Charles R. Meyer 419

Automatic Selection of DBS Target Points Using Multiple Electrophysiological Atlases
Pierre-Francois D’Haese, Srivatsan Pallavaram, Ken Niermann, John Spooner, Chris Kao, Peter E. Konrad, Benoit M. Dawant 427

Nonrigid Shape Correspondence Using Landmark Sliding, Insertion and Deletion
Theodor Richardson, Song Wang 435

Statistical Face Models for the Prediction of Soft-Tissue Deformations After Orthognathic Osteotomies
Sebastian Meller, Emeka Nkenke, Willi A. Kalender 443

Fully Automatic Shape Modelling Using Growing Cell Neural Networks
Luca Ferrarini, Hans Olofsen, Mark A. van Buchem, Johan H.C. Reiber, Faiza Admiraal-Behloul 451

Multiscale 3D Shape Analysis Using Spherical Wavelets
Delphine Nain, Steven Haker, Aaron Bobick, Allen R. Tannenbaum 459

Structural and Functional Brain Analysis

Discriminative Analysis of Brain Function at Resting-State for Attention-Deficit/Hyperactivity Disorder
C.Z. Zhu, Y.F. Zang, M. Liang, L.X. Tian, Y. He, X.B. Li, M.Q. Sui, Y.F. Wang, T.Z. Jiang 468

Finding Landmarks in the Functional Brain: Detection and Use for Group Characterization
Bertrand Thirion, Philippe Pinel, Jean-Baptiste Poline 476

Topology Correction Using Fast Marching Methods and Its Application to Brain Segmentation
Pierre-Louis Bazin, Dzung L. Pham 484

New Ratios for the Detection and Classification of CJD in Multisequence MRI of the Brain
Marius George Linguraru, Nicholas Ayache, Miguel Ángel González Ballester, Eric Bardinet, Damien Galanaud, Stéphane Haïk, Baptiste Faucheur, Patrick Cozzone, Didier Dormont, Jean-Philippe Brandel 492

Model-Based Image Analysis

Statistical Representation and Simulation of High-Dimensional Deformations: Application to Synthesizing Brain Deformations <i>Zhong Xue, Dinggang Shen, Bilge Karacali, Christos Davatzikos</i>	500
Model-Based Parameter Recovery from Uncalibrated Optical Images <i>S.J. Preece, I.B. Styles, S.D. Cotton, E. Claridge, A. Calcagni</i>	509
MRI Tissue Classification with Neighborhood Statistics: A Nonparametric, Entropy-Minimizing Approach <i>Tolga Tasdizen, Suyash P. Awate, Ross T. Whitaker, Norman L. Foster</i>	517
Image-Guided Intervention: Simulation, Modeling and Display	
Robotic Assisted Radio-Frequency Ablation of Liver Tumors – Randomized Patient Study <i>A. Patriciu, M. Awad, S.B. Solomon, M. Choti, D. Mazilu, L. Kavoussi, D. Stoianovici</i>	526
New Approaches to Catheter Navigation for Interventional Radiology Simulation <i>S. Cotin, C. Duriez, J. Lenoir, P. Neumann, S. Dawson</i>	534
Hybrid Bronchoscope Tracking Using a Magnetic Tracking Sensor and Image Registration <i>Kensaku Mori, Daisuke Deguchi, Kenta Akiyama, Takayuki Kitasaka, Calvin R. Maurer Jr., Yasuhito Suenaga, Hirotsugu Takabatake, Masaki Mori, Hiroshi Natori</i>	543
Toward Robotized Beating Heart TECABG: Assessment of the Heart Dynamics Using High-Speed Vision <i>Loïc Cuvillon, Jacques Gangloff, Michel de Mathelin, Antonello Forgione</i>	551
Data-Fusion Display System with Volume Rendering of Intraoperatively Scanned CT Images <i>Mitsuhiro Hayashibe, Naoki Suzuki, Asaki Hattori, Yoshito Otake, Shigeyuki Suzuki, Norio Nakata</i>	559

Simulation and Modeling II

A Hybrid Cutting Approach for Hysteroscopy Simulation <i>M. Harders, D. Steinemann, M. Gross, G. Székely</i>	567
Hydrometra Simulation for VR-Based Hysteroscopy Training <i>R. Sierra, J. Zátonyi, M. Bajka, G. Székely, M. Harders</i>	575
Brain Shift Computation Using a Fully Nonlinear Biomechanical Model <i>Adam Wittek, Ron Kikinis, Simon K. Warfield, Karol Miller</i>	583
Finite Element Model of Cornea Deformation <i>Jessica R. Crouch, John C. Merriam, Earl R. Crouch III</i>	591
Characterization of Viscoelastic Soft Tissue Properties from <i>In vivo</i> Animal Experiments and Inverse FE Parameter Estimation <i>Jung Kim, Mandayam A. Srinivasan</i>	599
A Fast-Marching Approach to Cardiac Electrophysiology Simulation for XMR Interventional Imaging <i>M. Sermesant, Y. Coudière, V. Moreau-Villéger, K.S. Rhode, D.L.G. Hill, R.S. Razavi</i>	607
An Inverse Problem Approach to the Estimation of Volume Change <i>Martin Schweiger, Oscar Camara-Rey, William R. Crum, Emma Lewis, Julia Schnabel, Simon R. Arridge, Derek L.G. Hill, Nick Fox</i>	616
A Velocity-Dependent Model for Needle Insertion in Soft Tissue <i>Jessica R. Crouch, Chad M. Schneider, Josh Wainer, Allison M. Okamura</i>	624
Material Properties Estimation of Layered Soft Tissue Based on MR Observation and Iterative FE Simulation <i>Mitsunori Tada, Noritaka Nagai, Takashi Maeno</i>	633
Simulating Vascular Systems in Arbitrary Anatomies <i>Dominik Szczerba, Gábor Székely</i>	641

Medical Image Computing - Shape II

Physiological System Identification with the Kalman Filter in Diffuse Optical Tomography <i>Solomon Gilbert Diamond, Theodore J. Huppert, Ville Kolehmainen, Maria Angela Franceschini, Jari P. Kaipio, Simon R. Arridge, David A. Boas</i>	649
Brain Surface Parameterization Using Riemann Surface Structure <i>Yalin Wang, Xianfeng Gu, Kiralee M. Hayashi, Tony F. Chan, Paul M. Thompson, Shing-Tung Yau</i>	657
Automated Surface Matching Using Mutual Information Applied to Riemann Surface Structures <i>Yalin Wang, Ming-Chang Chiang, Paul M. Thompson</i>	666
Optimization of Brain Conformal Mapping with Landmarks <i>Yalin Wang, Lok Ming Lui, Tony F. Chan, Paul M. Thompson</i>	675
A New Method for SPECT Quantification of Targeted Radiotracers Uptake in the Myocardium <i>Shimin Li, Lawrence W. Dobrucki, Albert J. Sinusas, Yi-Hwa Liu</i>	684
Tracking and Analysis of Cine-Delayed Enhancement MR <i>Thomas O'Donnell, Engin Dikici, Randolph Setser, Richard D. White</i>	692
Characterizing Vascular Connectivity from microCT Images <i>Marcel Jackowski, Xenophon Papademetris, Lawrence W. Dobrucki, Albert J. Sinusas, Lawrence H. Staib</i>	701
Shape Modeling Using Automatic Landmarking <i>Jun Xie, Pheng-Ann Heng</i>	709
A Computer-Aided Design System for Revision of Segmentation Errors <i>Marcel Jackowski, Ardeshir Goshtasby</i>	717
Statistical Modeling of Shape and Appearance Using the Continuous Medial Representation <i>Paul A. Yushkevich, Hui Zhang, James C. Gee</i>	725
Vertebral Shape: Automatic Measurement with Dynamically Sequenced Active Appearance Models <i>M.G. Roberts, T.F. Cootes, J.E. Adams</i>	733

Geodesic Active Contours with Adaptive Neighboring Influence <i>Huafeng Liu, Yunmei Chen, Hon Pong Ho, Pengcheng Shi</i>	741
A Construction of an Averaged Representation of Human Cortical Gyri Using Non-linear Principal Component Analysis <i>G. Lohmann, D.Y. von Cramon, A.C.F. Colchester</i>	749
Efficient Kernel Density Estimation of Shape and Intensity Priors for Level Set Segmentation <i>Mikael Rousson, Daniel Cremers</i>	757
Corpus Callosum Subdivision Based on a Probabilistic Model of Inter-hemispheric Connectivity <i>Martin A. Styner, Ipek Oguz, Rachel Gimpel Smith, Carissa Cascio, Matthieu Jomier</i>	765
Image Segmentation and Analysis II	
Random Walks for Interactive Organ Segmentation in Two and Three Dimensions: Implementation and Validation <i>Leo Grady, Thomas Schiwietz, Shmuel Aharon, Rüdiger Westermann</i>	773
Robust Pulmonary Nodule Segmentation in CT: Improving Performance for Juxtapleural Cases <i>K. Okada, V. Ramesh, A. Krishnan, M. Singh, U. Akdemir</i>	781
Tissue Classification of Noisy MR Brain Images Using Constrained GMM <i>Amit Ruf, Hayit Greenspan, Jacob Goldberger</i>	790
Automatic Left Atrium Segmentation by Cutting the Blood Pool at Narrowings <i>Matthias John, Norbert Rahn</i>	798
Automatic Vascular Tree Formation Using the Mahalanobis Distance <i>Julien Jomier, Vincent LeDigarcher, Stephen R. Aylward</i>	806
The Use of Unwrapped Phase in MR Image Segmentation: A Preliminary Study <i>Pierrick Bourgeat, Jurgen Fripp, Andrew Janke, Graham Galloway, Stuart Crozier, Sébastien Ourselin</i>	813

2D and 3D Shape Based Segmentation Using Deformable Models <i>Ayman El-Baz, Seniha E. Yuksel, Hongjian Shi, Aly A. Farag, Mohamed A. El-Ghar, Tarek Eldiasty, Mohamed A. Ghoneim</i>	821
CT Hepatic Venography: 3D Vascular Segmentation for Preoperative Evaluation <i>Catalin Fetita, Olivier Lucidarme, Françoise Prêteux, Philippe Grenier</i>	830
Shape-Based Averaging for Combination of Multiple Segmentations <i>T. Rohlfing, C.R. Maurer, Jr.</i>	838
Automatic Initialization Algorithm for Carotid Artery Segmentation in CTA Images <i>Martijn Sanderse, Henk A. Marquering, Emile A. Hendriks, Aad van der Lugt, Johan H.C. Reiber</i>	846
Automated Nomenclature of Bronchial Branches Extracted from CT Images and Its Application to Biopsy Path Planning in Virtual Bronchoscopy <i>Kensaku Mori, Sinya Ema, Takayuki Kitasaka, Yoshito Mekada, Ichiro Ide, Hiroshi Murase, Yasuhito Suenaga, Hirotsugu Takabatake, Masaki Mori, Hiroshi Natori</i>	854
Spectral Clustering Algorithms for Ultrasound Image Segmentation <i>Neculai Archip, Robert Rohling, Peter Cooperberg, Hamid Tahmasebpour, Simon K. Warfield</i>	862
Using the Fast Marching Method to Extract Curves with Given Global Properties <i>Xiaodong Tao, Christos Davatzikos, Jerry L. Prince</i>	870
Robust Tissue Boundary Detection for Cerebral Cortical Thickness Estimation <i>Marietta L.J. Scott, Neil A. Thacker</i>	878
Statistical Analysis of Pharmacokinetic Models in Dynamic Contrast-Enhanced Magnetic Resonance Imaging <i>Volker J. Schmid, Brandon J. Witcher, Guang-Zhong Yang, N. Jane Taylor, Anwar R. Padhani</i>	886

Image Registration III

Inter-breath-hold Registration for the Production of High Resolution Cardiac MR Volumes <i>Nicholas M.I. Noble, Redha Boubertakh, Reza S. Razavi, Derek L.G. Hill</i>	894
Consistent Estimation of Cardiac Motions by 4D Image Registration <i>Dinggang Shen, Hari Sundar, Zhong Xue, Yong Fan, Harold Litt</i>	902
Multispectral MR to X-Ray Registration of Vertebral Bodies by Generating CT-Like Data <i>Everine B. van de Kraats, Graeme P. Penney, Theo van Walsum, Wiro J. Niessen</i>	911
Articulated Rigid Registration for Serial Lower-Limb Mouse Imaging <i>Xenophon Papademetris, Donald P. Dione, Lawrence W. Dobrucki, Lawrence H. Staib, Albert J. Sinusas</i>	919
Incorporating Statistical Measures of Anatomical Variability in Atlas-to-Subject Registration for Conformal Brain Radiotherapy <i>Olivier Commowick, Radu Stefanescu, Pierre Fillard, Vincent Arsigny, Nicholas Ayache, Xavier Pennec, Grégoire Malandain</i>	927
Accurate Image Registration for Quadrature Tomographic Microscopy <i>Chia-Ling Tsai, William Warger II, Charles DiMarzio</i>	935
Riemannian Elasticity: A Statistical Regularization Framework for Non-linear Registration <i>X. Pennec, R. Stefanescu, V. Arsigny, P. Fillard, N. Ayache</i>	943
3D Model-Based Approach to Lung Registration and Prediction of Respiratory Cardiac Motion <i>Mikhail G. Danilouchkine, Jos J.M. Westenberg, Hans C. van Assen, Johan H.C. van Reiber, Boudewijn P.F. Lelieveldt</i>	951
Fast DRR Generation for 2D/3D Registration <i>W. Birkfellner, R. Seemann, M. Figl, J. Hummel, C. Ede, P. Homolka, X. Yang, P. Niederer, H. Bergmann</i>	960
Validation of PET Imaging by Alignment to Histology Slices <i>Philip J. Edwards, Ayman D. Nijmeh, Mark McGurk, Edward Odell, Michael R. Fenlon, Paul K. Marsden, David J. Hawkes</i>	968

Adaptive Subdivision for Hierarchical Non-rigid Registration of Multi-modal Images Using Mutual Information <i>Adrian Andronache, Philippe Cattin, Gábor Székely</i>	976
3-D Diffeomorphic Shape Registration on Hippocampal Data Sets <i>Hongyu Guo, Anand Rangarajan, Sarang C. Joshi</i>	984
Two-Stage Registration for Real-Time Deformable Compensation Using an Electromagnetic Tracking Device <i>Hui Zhang, Filip Banovac, Neil Glossop, Kevin Cleary</i>	992
Cadaver Validation of Intensity-Based Ultrasound to CT Registration <i>Graeme P. Penney, Dean C. Barratt, Carolyn S.K. Chan, Mike Slomczykowski, Timothy J. Carter, Phillip J. Edwards, David J.Hawkes</i>	1000
Author Index	1009

Classification of Structural Images via High-Dimensional Image Warping, Robust Feature Extraction, and SVM

Yong Fan, Dinggang Shen, and Christos Davatzikos

Section of Biomedical Image Analysis, Department of Radiology,
University of Pennsylvania, Philadelphia, PA 19104
{yong.fan, dinggang.shen, christos.davatzikos}@uphs.upenn.edu

Abstract. This paper presents a method for classification of medical images, using machine learning and deformation-based morphometry. A morphological representation of the anatomy of interest is first obtained using high-dimensional template warping, from which regions that display strong correlations between morphological measurements and the classification (clinical) variable are extracted using a watershed segmentation, taking into account the regional smoothness of the correlation map which is estimated by a cross-validation strategy in order to achieve robustness to outliers. A Support Vector Machine-Recursive Feature Elimination (SVM-RFE) technique is then used to rank computed features from the extracted regions, according to their effect on the leave-one-out error bound. Finally, SVM classification is applied using the best set of features, and it is tested using leave-one-out. The results from a group of 61 brain images of female normal controls and schizophrenia patients demonstrate not only high classification accuracy (91.8%) and steep ROC curves, but also exceptional stability with respect to the number of selected features and the SVM kernel size.

1 Introduction

Morphological analysis of medical images is performed commonly in a variety of research and clinical studies. Region of Interest volumetry (ROI) has been traditionally used to obtain regional measurement of anatomical volumes and investigate abnormal tissue structures with disease [1]. However, in practice, *a priori* knowledge about abnormal regions is not always available. Even when good *a priori* hypotheses can be made about specific ROIs, a region of abnormality might be part of an ROI, or span multiple ROIs, thereby potentially reducing statistical power of the underlying morphological analysis significantly. These limitations can be effectively overcome by methods falling under the general umbrella of High-Dimensional Morphological Analysis (we will refer to these methods as HDMA), such as voxel-based and deformation-based morphometric analysis methods, e.g. [2-5]. However, a voxel-by-voxel analysis is limited by noise, registration errors, and excessive inter-individual variability of measurements that are too localized, such as voxel-wise displacement fields, Jacobians, or residuals. Most importantly, voxel-by-voxel mass-univariate analysis of transformations or residuals does not capture multi-variate relationships in the data.

Linear methods, such as PCA [6] are not effective in capturing complex relationships in high-dimensional spaces.

In order to overcome these limitations, pattern classification methods have begun to emerge in the recent years in the field of computational anatomy [7-11], aiming at capturing nonlinear multivariate relationships among many anatomical regions, to more effectively characterize group differences. A major challenge in these methods has been the sheer dimensionality of HDMA-related measurements, which is often in the millions, coupled with the relatively small number of training samples, which is at best in the hundreds, and often just a few dozens. Accordingly, extracting a small number of most informative features from the data has been a fundamental challenge. A main emphasis of this paper is the extraction of distinctive, but also robust features from high-dimensional morphological measurements obtained from brain MR images, which are used in conjunction with nonlinear support vector machines (SVM) for classification. The proposed method is tested on classifying normal controls from schizophrenia patients in female participants.

The key elements of the proposed approach are: 1) Regional volumetric information is first extracted from a template warping transformation; herein we focus entirely on volumetric information, such as atrophy, and don't consider higher order shape characteristics. 2) The anatomy of interest (the brain, herein) is partitioned into a number of regions, via a watershed algorithm applied to the correlation map between clinical status and regional volumetric measurements; various techniques are applied to estimate the correlation map, in order to achieve robustness to outliers. 3) An SVM-RFE technique is applied to the previously rank-ordered features that are computed from the extracted regions, in order to select the most important feature set for classification. 4) A nonlinear SVM classifier is applied and tested via cross-validation and ROC analysis. 5) Group differences are visually displayed via a *discriminative direction method* [7,11]. We now detail the methodology.

2 Methods

Our classification method involves three steps: feature extraction, feature selection, and nonlinear classification, which are detailed next.

2.1 Feature Extraction

As mentioned in the Introduction, the features used for brain classification are extracted from automatically generated regions, which are determined from the training data. Several issues are taken into consideration here. First, morphological changes of brain structures resulting from pathological processes usually don't occur in isolated spots, but rather they occur in regions that can have irregular shapes and are not known *a priori*. Second, noise, registration errors and inter-individual anatomical variations necessitate the collection of morphological information from regions much larger than the voxel size, which must additionally be distinctive of the pathology of interest. Third, multivariate classification methods are most effective and generalizable when applied to a small number of reliable and discriminative features. Accordingly, features irrelevant to classification must be eliminated.

In the following, we detail the procedure for automatically generating adaptive regions from a training dataset, by first introducing the method to extract local morphological features, then defining the criteria for adaptively clustering voxels into regions, and finally extracting overall features from each region.

Construction of a morphological profile of the brain. In order to obtain the morphological profile from an individual brain image, warping the image into a template space is often a first step, leading to various morphological measurements (e.g. deformation field, Jacobian determinant, tissue density maps, or residuals) that are in the same space and therefore directly comparable across individuals. Herein we follow the framework that was proposed in [5], and which is based on a mass-preserving shape transformation framework; a similar method is used within the SPM software and is often referred to as "Jacobian modulation". We have used this approach because it is robust to registration errors, due to the mass preservation principle, in contrast to the determinant of the Jacobian that is directly affected by registration errors.

The approach in [5] uses images that are first segmented into three tissues, namely gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF)--we used the segmentation method of [12] in our experiments. Then, by using a high-dimensional image warping method [13] these segmented images are spatially transformed into a template space, by preserving the total tissue mass; this is achieved by increasing the respective density when a region is compressed, and vice versa. As a result, three tissue density maps, $f_0(u)$, $f_1(u)$, $f_2(u)$, are generated in the template space, each reflecting local volumetric measurements corresponding to GM, WM, and CSF, respectively, at location u . These three tissue density maps give a quantitative representation of the spatial tissue distribution. Regional atrophy is reflected by reduction in the respective tissue density map. Representative tissue density maps are shown in Fig.1.

Learning-based generation of adaptive regions. Brain regions are generated by spatial clustering of morphological features of similar classification power. For each morphological feature, its classification power is highly related to its discriminative power and reliability. The discriminative power of a feature can be quantitatively measured by its *relevance* to classification as well as its *generalizability* for classification. The relevance of a feature to classification can be measured by the correlation between this feature and the corresponding class label in a training dataset (e.g. normal or pathologic). In machine learning and statistics for relevance analysis, the correlation measures can be broadly divided into linear correlation and non-linear correlation. Most non-linear correlation measures are based on the information-theoretical concept of entropy, such as mutual information, computed by probability estimation. For continuous features, probability density estimation is a hard task especially when the number of available samples is limited. On the other hand, linear correlation measures are easier to compute even for continuous features and are robust to overfitting, thus they are widely used for feature selection in machine learning. Here, we used the absolute Pearson correlation coefficient, closely related to T-test [14] in the context of extracting group differences, to measure the relevance of each feature to classification. Given an image location u , the Pearson correlation between a feature $f_i(u)$ of tissue i , and class label y , is defined as

$$\rho_i(u) = \frac{\sum_j (f_{ij}(u) - \overline{f_i(u)})(y_j - \overline{y})}{\sqrt{\sum_j (f_{ij}(u) - \overline{f_i(u)})^2 \sum_j (y_j - \overline{y})^2}}, \quad (1)$$

where j denotes the j th sample in the training dataset. Thus, $f_{ij}(u)$ is a morphological feature of tissue i in the location u of j th sample (the tissue density map, here), and $\overline{f_i(u)}$ is the mean of $f_{ij}(u)$ over all samples. Similarly, y_j is a class label of the j th sample, and \overline{y} is the mean of y_j over all samples. In addition to the relevance, the *generalizability* of a feature is equally important for classification. A bagging strategy [15] is adopted to take the generalization ability into account, when measuring the discriminative power of a feature by absolute Pearson score. That is, given n training samples, a leave-one-out procedure is used to measure the discriminative power of each feature $f_i(u)$ by a conservative principle that selects the minimal absolute Pearson score from n resulted scores as this feature's discriminative power, defined as

$$P_i(u) = \min_{1 \leq k \leq n} |\rho_{ik}(u)|, \quad (2)$$

where $\rho_{ik}(u)$ is the k th leave-one-out Pearson correlation at location u of tissue map i . Maximizing the measure in (2) would select features that maximize the margin from 0, i.e., from no discrimination at all.

The *spatial consistency* of a feature is another important issue in classification, since morphological features are locally extracted and thus might not be reliable due to registration errors and inter-individual anatomical variations. A feature is spatially consistent if it is similar to other features in its spatial neighborhood, implying that small registration errors will not significantly change the value of this feature. For each feature $f_i(u)$, we measure its spatial consistency, $R_i(u)$, by an intra-class correlation coefficient that is computed from all features in its spatial neighborhood and all samples in the training dataset [16]. In our applications, the value of $R_i(u)$ is constrained to lie between 0 and 1.

For each feature, its discriminative power score $P_i(u)$ and its spatial consistency score $R_i(u)$ are both non-negative, with high score indicating better feature for classification. We combine these two measurements into one by the following equation,

$$s_i(u) = P_i(u)^p R_i(u)^r, \quad p, r > 0, \quad (3)$$

thus obtaining a single score $s_i(u)$ for each feature $f_i(u)$, which reflects the classification power of this feature for the particular classification problem. Three score maps are produced for GM, WM and CSF, respectively.

As we mentioned above, the disease-affected brain regions generally occur in the clusters of spatially contiguous voxels. Therefore, a watershed segmentation method [17] is employed to partition a brain into different regions according to the scores $s_i(u)$, and finally to obtain separate partitions for each tissue. A typical brain region partition result, with all regions generated from three tissue density maps, is shown in Fig. 1, on the right.

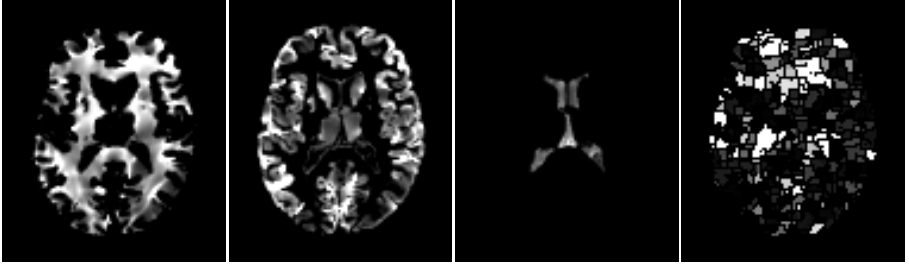


Fig. 1. Typical tissue density maps (GM, WM, CSF, from left to right) and automatically generated brain regions in which high grey-levels indicate discriminative power

Feature extraction from generated brain regions. For each region generated as described above, its corresponding volumetric measure is computed by summing up all tissue density values in this region, which effectively calculates the volumes of the corresponding regions in individual anatomies. Volumetric measures from all WM, GM, and CSF regions constitute an attribute vector to represent morphological information of the brain. Although currently we focus on local tissue volumetric information, other types of information could also be considered.

2.2 Feature Selection via SVM-RFE

Although the number of regions determined in Sec. 2.1 is dramatically smaller than the original number of brain voxels, measures obtained from many regions are less effective, irrelevant and redundant for classification. This requires a feature selection method to select a small set of the most informative features for classification. We have experimented with several feature selection methods, and determined that the SVM-RFE algorithm has the best performance. SVM-RFE is a feature subset selection method based on SVM, initially proposed for a cancer classification problem [18]. It was later extended by introducing SVM-based leave-one-out error bound criteria in [19]. The goal of SVM-RFE is to find a subset of size r among d variables ($r < d$), which optimizes the performance of the classifier. This algorithm is based on a backward sequential selection that removes one feature at a time. At each time, the removed feature makes the variation of SVM-based leave-one-out error bound smallest, compared to removing other features. In order to apply this subset selection to our problem in a reasonable time cost and to avoid local optima, we first remove the most irrelevant features by the feature ranking method [14] in which the rank score is computed by a Pearson correlation based bagging strategy as we described above, and then apply the SVM-RFE algorithm on the set of remaining features.

2.3 SVM-Based Classification

The nonlinear support vector machine is a supervised binary classification algorithm [20]. SVM constructs a maximal margin linear classifier in a high (often infinite) dimensional feature space, by mapping the original features via a kernel function. The Gaussian radial basis function kernel is used in our method.

SVM is not only empirically demonstrated to be one of the most powerful pattern classification algorithms, but also has provided many theoretic bounds on the leave-one-out error to estimate its capacity, for example, the radius/margin bound, which could be utilized in feature selection. Another reason for us to select SVM as a classifier is its inherent sample selection mechanism, i.e., only support vectors affect the decision function, which may help us find subtle differences between groups.

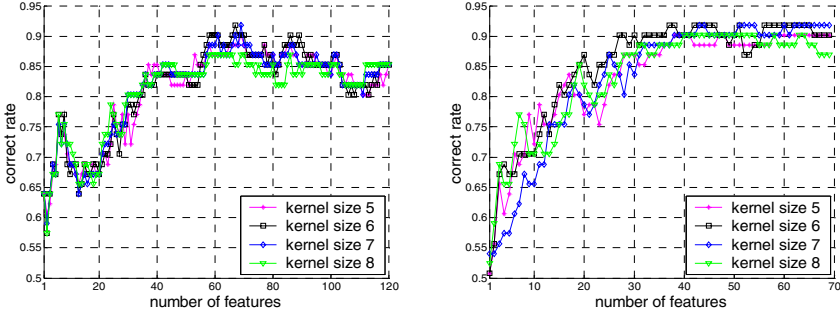


Fig. 2. Performance of ranking based feature selection (left plot) and SVM-RFE feature selection (right plot). Plotted are the average classification rates for different SVM kernel sizes and different feature numbers. Notably, the SVM-RFE algorithm starts selecting subsets of features from 69 features, which are top-ranked features, selected by ranking-based feature selection method. The SVM-RFE algorithm performs a robust selection of features and leads to stable performance.

3 Results

We tested our approach on MR T1 brain images, in order to compare the brain differences between female schizophrenia patients ($N = 23$) and normal controls ($N = 38$).

A full leave-one-out cross-validation is performed in our experiments. In each leave-one-out validation experiment, one subject was first selected as testing subject, and the remaining subjects are used for the entire adaptive regional feature extraction, feature selection and training procedure described in Section 2. Then, the classification result on the testing subject using the trained SVM classifier was compared with the ground-truth class label, to evaluate the classification performance. Absolutely *all* feature selection and training steps were cross-validated, i.e. the testing image had no influence on the construction of the classifier. By repeatedly leaving each subject out as testing subject, we obtained the average classification rate from 61 leave-one-out experiments. Finally, these experiments were repeated for different numbers of features, in order to test the stability of the results. The best average correct classification rate was 91.8% by using 37 features, selected by SVM-RFE algorithm, as shown in Fig. 2. Although a reasonably good performance was achieved just via the feature ranking method according to the scores computed by a bagging strategy [15], as

shown in Fig. 2-left, more stable performance was achieved by incorporating the SVM-RFE method (Fig. 2-right), since simple feature ranking does not consider correlations between features. Furthermore, these plots also indicate that the described algorithm is quite robust with respect to the SVM Gaussian kernel. The ROC curve of the classifier that yields the best classification result is also shown in Fig. 3, which indicates that our classifier has large area under ROC curve.

Besides using a classifier to determine the abnormality of brains, we can also use it for detecting group difference. In [7,11], the *discriminative direction method* was used to estimate the group difference from the classification function. Here, we utilized a similar method to estimate the group difference. The group differences are overlaid on the template brain, highlighting the most significant and frequently detected group differences in our leave-one-out experiments (Fig. 4).

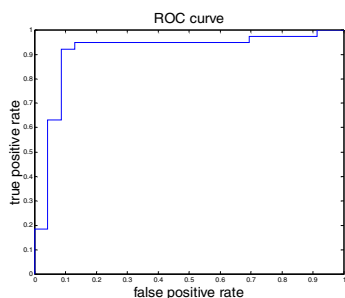


Fig. 3. ROC curve

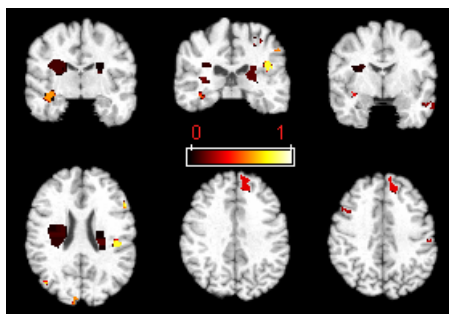


Fig. 4. Regions of most representative of the group differences, found via decision function gradient.(high light indicates more significant)

4 Discussions and Conclusions

We have presented a statistical classification method for identification of brain abnormality based on regional morphological information. The classifier is built on adaptive regional feature extraction and feature selection. In particular, brain regions are generated automatically by grouping local morphological features with similar classification power. This adaptive regional feature extraction method aims at overcoming the limitations of the traditional ROI methods that need prior knowledge of what specific regions might be affected by disease, and the limitations of the voxel based morphometry (VBM) methods that use an identical isotropic filter to collect regional morphological information in all brain locations. The robust feature selection method used in this paper further removes features that are irrelevant and redundant to classification, thus improving the classification performance. The experimental results indicate that this method can achieve high classification rate in a schizophrenic study.

References

1. Giuliania, N.R., Calhoun, V.D., Pearlson, G.D., Francis, A., Buchanan, R.W.: Voxel-based morphometry versus region of interest: a comparison of two methods for analyzing gray matter differences in schizophrenia. *Schizophrenia Research* 74 (2005) 135—147
2. Thompson, P.M.; MacDonald, D.; Mega, M.S.; Holmes, C.J.; Evans, A.; Toga, A.W.: Detection and mapping of abnormal brain structure with a probabilistic atlas of cortical. *Journal of Computer Assisted Tomography* 21 (1997) 567-581
3. Ashburner, J., Friston, K.J.: Voxel-based morphometry--the methods. *NeuroImage* 11 (2000) 805—821
4. Chung, M.K., Worsley, K.J. Paus, T., Cherif, C., Collins, D.L., Giedd, J.N., Rapoport, J.L., Evans, A.C.: A unified statistical approach to deformation-based morphometry. *NeuroImage* 14 (2001) 595-606
5. Davatzikos, C., Genc, A., Xu, D., Resnick, S.M.: Voxel-based morphometry using the ravens maps: Methods and validation using simulated longitudinal atrophy. *NeuroImage* 14 (2001) 1361—1369
6. Miller, M.; Banerjee, A.; Christensen, G.; Joshi, S.; Khaneja, N.; Grenander, U.; Matejic, L.: Statistical methods in computational anatomy. *Statistical Methods in Medical Research* 6 (1997) 267-299
7. Golland, P., Grimson, W.E.L., Shenton, M.E., Kikinis, R.: Deformation analysis for shape based classification. In: *IPMI*. (2001) 517—530
8. Gerig, G.; Styner, M.; Lieberman, J.: Shape versus Size: Improved understanding of the morphology of brain structures. In: *MICCAI* (2001) 24-32
9. Yushkevich, P.A., Joshi, S.C., Pizer, S.M., Csernansky, J.G., Wang, L.E.: Feature selection for shape-based classification of biological objects. In Taylor, C.J., Noble, J.A., eds.: *IPMI*. Volume 2732 of *Lecture Notes in Computer Science.*, Springer (2003) 114—125
10. Liu, Y., Teverovskiy, L., Carmichael, O.T., Kikinis, R., Shenton, M.E., Carter, C.S., Stenger, V.A., Davis, S., Aizenstein, H., Becker, J.T., Lopez, O.L., Meltzer, C.C.: Discriminative MR image feature analysis for automatic schizophrenia and alzheimer's disease classification. In: *MICCAI* (1). (2004) 393—401
11. Lao, Z., Shen, D., Xue, Z., Karacali, B., M.Resnick, S., Davatzikos, C.: Morphological classification of brains via high-dimensional shape transformations and machine learning methods. *NeuroImage* 21 (2004) 46—57
12. Pham, D.L., Prince, J.L.: Adaptive fuzzy segmentation of magnetic resonance images. *IEEE Transactions on Medical Imaging* 18 (1999) 737—752
13. Shen, D., Davatzikos, C.: HAMMER: Hierarchical attribute matching mechanism for elastic registration. *IEEE Transactions on Medical Imaging* 21 (2002) 1421—1439.
14. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. *Journal of Machine Learning Research* 3 (2003) 1157—1182
15. Breiman, L.: Bagging predictors. *Machine Learning* 24 (1996) 123—140
16. McGraw, K.O., Wong, S.P.: Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1 (1996) 30—46
17. Vincent, L., Soille, P.: Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13 (1991) 583—589
18. Guyon, I., Weston, J., Barnhill, S., Vapnik, V.: Gene selection for cancer classification using support vector machines. *Machine Learning* 46 (2002) 389—422
19. Rakotomamonjy, A.: Variable selection using SVM-based criteria. *Journal of Machine Learning Research* 3 (2003) 357—1370
20. Vapnik, V.N.: *Statistical Learning Theory*. Wiley (1998)

Bone Enhancement Filtering: Application to Sinus Bone Segmentation and Simulation of Pituitary Surgery

Maxime Descoteaux¹, Michel Audette², Kiyoyuki Chinzei², and Kaleem Siddiqi³

¹ Odyssee Project Team, INRIA Sophia-Antipolis / ENPC-Paris / ENS-Ulm Paris, France
Maxime.Descoteaux@sophia.inria.fr

² Surgical Assist Group, AIST, Tsukuba, Japan

³ School Of Computer Science, McGill University, Montreal, Canada

Abstract. We present a novel multi-scale bone enhancement measure that can be used to drive a geometric flow to segment bone structures. This measure has the essential properties to be incorporated in the computation of anatomical models for the simulation of pituitary surgery, enabling it to better account for the presence of sinus bones. We present synthetic examples that validate our approach and show a comparison between existing segmentation techniques of paranasal sinus CT data.

1 Introduction

Pituitary gland tumors represent the third most common primary intracranial tumors encountered in neurosurgical practice. In the majority, a surgical intervention is required. A neurosurgeon typically enters through the nose, has to break thin paranasal sinus bones and remove soft tissues while avoiding nerves and blood vessels to reach the pituitary gland (Fig.1). This requires extensive practice and precise knowledge of the anatomy, the absence of which can have serious implications on the patient [1]. Currently, the only way to train a neurosurgery resident for such an operation is by multiple observation and by elementary maneuver attempts supervised by an expert neurosurgeon. This is why for the past several years, there has been a growing interest in building a surgical simulator to provide a tool for such training. Existing surgical simulators generally involve a generic anatomical model elaborated on the basis of extensive human supervision, interacting with a fast but constitutively limited biomechanics engine.

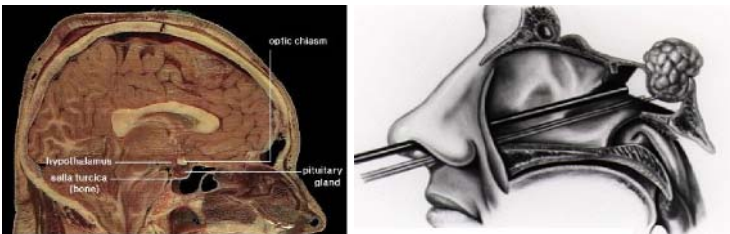


Fig. 1. Endoscopic pituitary surgery

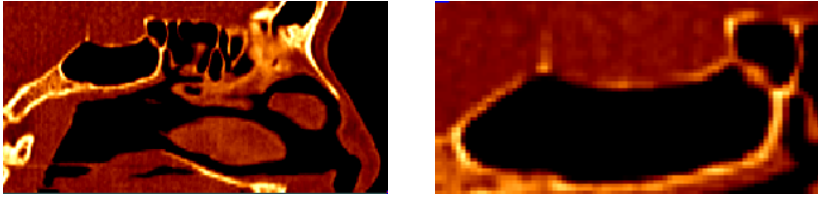


Fig. 2. Sagittal slice of CT data set in region of interest for pituitary gland surgery (see Fig. 1)

The goal of our research is to formulate a minimally supervised method for producing a set of patient-specific anatomical models, from MR and CT data sets, in a manner that can interact with a hierarchical finite-element based biomechanics engine. To do so, we need a precise 3-dimensional (3D) partition of tissue classes into bone, air, vessel, nerve and soft-tissue.

In this paper, we focus on paranasal sinus bone enhancement and segmentation from CT data. The usual methods for segmentation of bone in CT are based on thresholding followed by some image connectivity measures or manual editing which is quite tedious and prone to human error. At a coarse scale, segmentation by thresholding is quite good due to the 3-class nature of CT images and the known Hounsfield value range for bone. Air has close to no signal and bone has much higher signal than surrounding tissues (see Fig. 2). However, from Fig. 2(b), we can clearly see that thin bones can have holes and diffusive boundaries. True bones do not have these features, they are due to partial volume effects and noise present in CT data sets. For such thin bones, a simple thresholding procedure gives unsatisfying results.

In this paper, we introduce a novel algorithm for bone enhancement filtering and segmentation based on the multi-scale second order local structure of the image. We exploit the eigenvalue decomposition of the Hessian matrix which is known to give valuable local information on the blob-like, tube-like and sheet-like behavior of iso-intensity level sets in the image [5,6,3,4,2]. We propose a *sheetness* measure that can be used to drive an active surface to segment bone. This is motivated in part by Frangi’s tubular structure enhancement filtering measure [3] and Descoteaux et al [2] multi-scale geometric flow. We illustrate the power of the approach with segmentation results in regions with holes and low Hounsfield values in CT data and compare them to results from [9] based on local structure extraction with the structure tensor. We also validate the approach quantitatively on synthetic data. To our knowledge, our method is the first flow-based approach for paranasal sinus bone segmentation.

2 Using Local 3D Structure for Filtering and Segmentation

In this section, we investigate medical image filtering algorithms using the tensor matrix, \mathcal{T} and the Hessian matrix, \mathcal{H} as shape descriptors. For a 3D image \mathcal{I} , they are defined as

$$\mathcal{T} = \nabla \mathcal{I}^T \nabla \mathcal{I} = \begin{pmatrix} \mathcal{I}_x \mathcal{I}_x & \mathcal{I}_x \mathcal{I}_y & \mathcal{I}_x \mathcal{I}_z \\ \mathcal{I}_x \mathcal{I}_y & \mathcal{I}_y \mathcal{I}_y & \mathcal{I}_y \mathcal{I}_z \\ \mathcal{I}_x \mathcal{I}_z & \mathcal{I}_y \mathcal{I}_z & \mathcal{I}_z \mathcal{I}_z \end{pmatrix} \quad \mathcal{H} = \begin{pmatrix} \mathcal{I}_{xx} & \mathcal{I}_{xy} & \mathcal{I}_{xz} \\ \mathcal{I}_{xy} & \mathcal{I}_{yy} & \mathcal{I}_{yz} \\ \mathcal{I}_{xz} & \mathcal{I}_{yz} & \mathcal{I}_{zz} \end{pmatrix}.$$

In both cases, an eigenvalue analysis is performed to extract the local behavior of iso-intensity level sets in the image. Many methods have been developed for blood vessel segmentation using these models [5,6,3,4,2]. We review only a selection of the representative techniques for our bone segmentation problem.

2.1 Modeling Sheet-Like Structures with the Tensor Descriptor

As mentioned before, the usual method for bone segmentation for CT data is simple thresholding which fails on thin bone structures. Recently, Westin et al. [9] have introduced an adaptive thresholding approach using the structure tensor to segment thin bones around the eye socket and in the paranasal sinus area. The idea is to evaluate the structure tensor at every voxel of the data and determine the degree to which its shape resembles a line, a plane or a sphere. Letting $\lambda_1, \lambda_2, \lambda_3$ ($0 \leq \lambda_1 \leq \lambda_2 \leq \lambda_3$) be the eigenvalues of the structure tensor, the interest is when the structure tensor can be approximated by a plane (see [9] for details). The authors describe a planar measure $c_{plane} = \frac{\lambda_3 - \alpha\lambda_2}{\lambda_3}$ such that in theory, it has a value of 1 for plane structures and 0 for others. This measure is used to adaptively threshold the input data set, $t(x) = t_0 - \alpha c_{plane}(x)$, where t_0 is a global threshold modified locally and α is a weight factor for the planar measure.

In Section 5, we demonstrate several properties of this approach. In particular, since the tensor is based on first order variation, the c_{plane} measure is strong at boundaries (where the gradient is strong) and weak inside the bone structure. For our application, we seek a measure that is high at the center of the structure with a fall off at boundaries where a priori, our confidence in a voxel being strictly bone or strictly soft tissue is weak. Such a confidence index can guide the subsequent surface and volume meshing of bone tissue relevant to the simulation of pituitary surgery. We explore the properties of the Hessian shape operator to define such a measure.

2.2 Modeling Sheet-Like Structures Using the Hessian Operator

The Hessian matrix encodes important shape information. An eigenvalue decomposition measures the maximum changes in the normal vector (gradient vector) of the underlying intensity iso-surface in a small neighborhood. Hence, it can differentiate between tube-like, sheet-like and blob-like structures. The classification of Table 1 was first explored by Sato et al. [6], and Lorenz et al. [5] separately. In [3], Frangi defines three ratios using tube-like properties of the eigenvalues of Table 1 to separate blood vessels from other structures. Derivatives of the Hessian matrix are computed with multi-scale

Table 1. Local structure classification assuming $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$

eigenvalue conditions	local structure	examples
$\lambda_1 \approx \lambda_2 \approx 0, \lambda_3 \gg 0$	sheet-like	bone, skin
$\lambda_1 \approx 0, \lambda_2 \approx \lambda_3 \gg 0$	tube-like	vessels, nerves
$\lambda_1 \approx \lambda_2 \approx \lambda_3 \gg 0$	blob-like	nodule
$\lambda_1 \approx \lambda_2 \approx \lambda_3 \approx 0$	noise-like	background, noise

Table 2. Theoretical properties of the ratios defined to construct the sheetness measure

Ratios	sheet	tube	blob	noise
$R_{sheet} = \lambda_2 / \lambda_3 $	0	1	1	<i>undefined</i>
$R_{blob} = (2 \lambda_3 - \lambda_2 - \lambda_1)/ \lambda_3 $	2	1	0	<i>undefined</i>
$R_{noise} = \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}$	λ_3	$\sqrt{2}\lambda_3$	$\sqrt{3}\lambda_3$	0

γ -parametrized Gaussian kernels and the three quantities are integrated in a *vesselness* measure designed to be maximum when computed at the scale corresponding to the radius of the tubular objects. The vessel index is thus maximum nearby the vessel center and is zero outside. In [2] a vesselness measure is used to find putative centerlines of tubular structures along with their estimated radii and is then distributed to create a vector field which is orthogonal to vessel boundaries so that the flux maximizing flow algorithm of [8] can be applied to recover them. This method can recover low contrast and thin vessels from standard anatomical proton density weighted data sets.

Inspired by these approaches, we propose a multi-scale *sheetness* measure that enhances bone structures and then use it to drive a deformable surface that stops at bone boundaries. At every voxel, we determine whether the underlying iso-intensity surface behaves like a sheet. In this case, we know that the eigenvectors corresponding to the null eigenvalues span the plane of the plate structure and the other eigenvector is perpendicular to it. We define three ratios, R_{sheet} , R_{blob} , R_{noise} , to differentiate sheet-like structures from others. Their behavior is described in Table 2. Then, just as in [3], we can define the sheetness measure, S , as the maximum response over all scales σ at which the derivatives of the Hessian are computed,

$$S = \max_{\sigma \in \Sigma} S(\sigma) = \begin{cases} 0 & \text{if } \lambda_3 > 0 \\ \left(\exp\left(\frac{-R_{sheet}^2}{2\alpha^2}\right) \right) \left(1 - \exp\left(\frac{-R_{blob}^2}{2\beta^2}\right) \right) \left(1 - \exp\left(\frac{-R_{noise}^2}{2c^2}\right) \right), & \end{cases} \quad (1)$$

where Σ is a finite set of scales chosen by the user in the range of smallest to thickest possible bone structure ($0.5 \leq \sigma \leq 3.0$) in the data and parameters α, β, c are set to 0.5, 0.5 and half the maximum Frobenius norm (R_{noise}) respectively as suggested in [3,2].

Each term of the equation has a function and depends on the characteristics of Table 2. To avoid division by a null λ_3 in the case of noise, the *undefined* can be set to obtain the desired behavior. Breaking down the terms of Eq. 1, we have

1. $\exp\left(\frac{-R_{sheet}^2}{2\alpha^2}\right)$ is a sheet enhancement term, where the maximum occurs for sheet-like structures and the minimum for others. We set *undefined* to 1.
2. $\left(1 - \exp\left(\frac{-R_{blob}^2}{2\beta^2}\right)\right)$ is a blob and noise elimination term since it is zero for both. Moreover, the term is high for a sheet and lower for a tube. We set *undefined* to zero.
3. $\left(1 - \exp\left(\frac{-R_{noise}^2}{2c^2}\right)\right)$ is a background or noise reduction term. R_{noise} is known as the Frobenius norm and is high only in the presence of structure.

Note that we do not define a tube elimination term as the curved ends of bone structures have a behavior that is both tube-like and sheet-like. Thus, the sheetness measure

is designed to be maximum for sheet-like voxels, less important for tube-like regions and zero for other structures. The power of this approach resides in the fact that after this sheetness computation, we have a confidence sheet-like score at each voxel and in addition, for high score locations, we have the scale estimate of the radius of the sheet as well as the normal vector to the plane.

3 A Geometric Flow for Segmenting Bone Structures

There have been a few deformable model methods proposed for bone segmentation which are quite different from our approach because they are suited to 2D images from different modalities and different bones. In [7], the segmentation of carpal bones of the hand in CT images is faced with similar challenges as in our sinus bone CT data sets. A skeletally coupled curve evolution framework is proposed that combines probabilistic growth with local and global competition. Promising results are shown on 2D images with gaps and diffused edges. However, the method is based on skeletal points which would be difficult to determine in one or two voxel wide bone structures such as those in the paranasal sinuses.

In our application, we propose to use the bone enhancement measure of Eq. 1 to drive a 3D surface evolution. We construct a vector field that is both large in magnitude and orthogonal to bone structures. The key idea is to *distribute the sheetness measure, which is concentrated on the center sheet, to the bone boundaries* implied by the local scale and direction estimates coming from the multi-scale sheetness measure. At each voxel where the sheetness index is high, we consider a disc or flat ellipsoid with its two semi-minor axes aligned with the estimated plane orientation and its semi-major axis equal to the approximated radius. The sheetness measure is then distributed over every voxel on the boundary of the disc. We define the addition of the extensions carried out

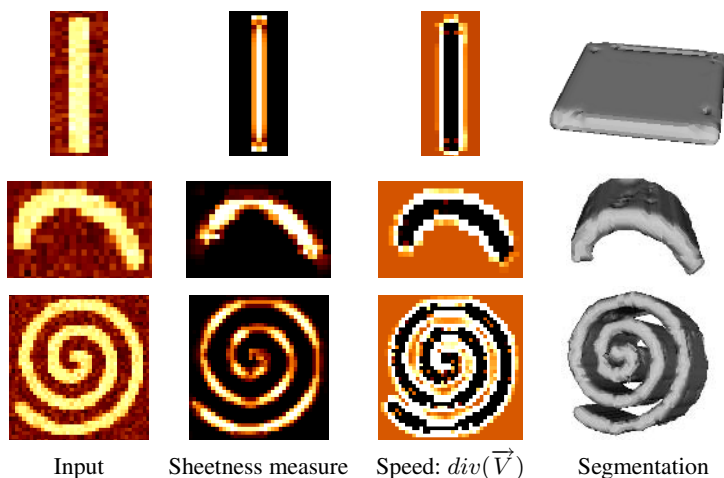


Fig. 3. Experiment on synthetic objects. We show slices of the input volumes, the sheetness measure, the speed term and a surface rendering of the segmentation

independently at all voxels to be the ϕ *distribution*. The extended vector field is now defined as the product of the normalized gradient of the *original* image with the above ϕ distribution, $\vec{V} = \phi \frac{\nabla I}{|\nabla I|}$. This vector field is an ideal candidate for the static vector field in the flux maximizing geometric flow $S_t = \text{div}(\vec{V})\vec{N}$. The flow evolves the surface S to converge to the zero-crossing of the speed term to a $\text{div}(\vec{V})$ (see Fig. 3). The flow is topologically adaptive due to its implementation with levelset methods, is computationally efficient and requires minimal user interaction. We refer the reader to [2] for further details.

4 Quantitative Validation on Synthetic Objects

In order to validate the extension of the sheetness measure to boundaries and to evaluate the effectiveness of the speed term driving the geometric flow, we constructed several binary synthetic objects of varying widths and centerline curvatures. Each volume was then smoothed using 2 iterations of mean curvature smoothing to simulate partial volume effects at their boundaries. Ground truth surface points were obtained as the 0.5 crossings of each object, obtained by linear interpolation on the voxel grid. We then added white Gaussian noise to each voxel to simulate typical noise levels in a CT acquisition process (we used a 10% signal-to-noise ratio) and followed the steps detailed above to construct the sheetness measure, obtain the ϕ distribution and then build the vector field \vec{V} , using the same parameters throughout. Empirical surface points were defined as the zero-crossings of the speed term $\text{div}(\vec{V})$ in the geometric flow.

Table 3. Agreement level between reconstruction and binary ground truth

object	average distance error (voxels)	maximum error (voxels)	ratio (%)
plate	0.22	1.00	95
rib	0.25	1.12	95
spiral	0.26	1.22	91

A visualization of the important terms is shown in Fig. 3. We evaluated the accuracy of the zero-crossings of the speed term by computing the average and maximum Euclidean distance errors between each empirical surface voxel and its closest ground truth surface voxel. We also computed the ratio between classified voxels in the segmentation and bone voxels in the binary volume. The results, shown in Table 3, indicate that the average error is typically less than 0.3 voxels and that the agreement between the reconstructed and original volume is above 90%. We have determined empirically that the maximum errors occur at the two ends of each synthetic object, which is to be expected since mean curvature smoothing causes the most shrinkage there.

5 Bone Enhancement and Segmentation on CT Data

In order to compare our algorithm with Westin et al.’s adaptive thresholding method and simple thresholding, we cropped a 53 mm x 89 mm x 98mm region of a CT data sets

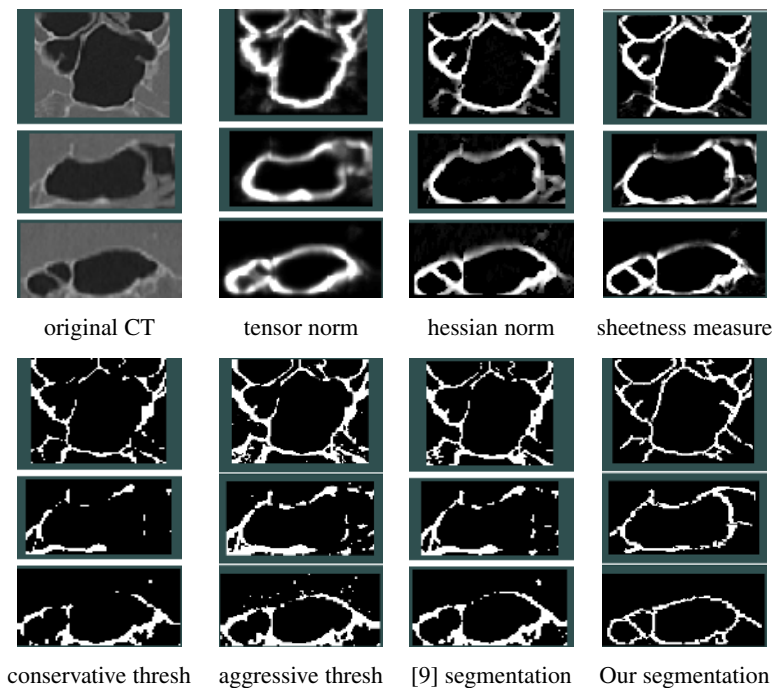


Fig. 4. Comparison between different segmentation methods. Our method connects most of the thin bone structure and reconstructs more bone than the approach in [9] and than thresholding.

around the paranasal sinuses and ran the segmentation methods. The original volume was first resampled to a $0.468mm^3$ isotropic grid.

We make several qualitative observations from Fig. 4. First, we see that both the algorithm of [9] and our method work better than simple thresholding. Second, our method appears to better exploit tissue contiguity and connect more bone structure than Westin’s segmentation. This is to be expected since our geometric flow method is designed to evolve and connect as much sheetness information as possible whereas [9] remains a thresholding method, albeit one that accounts for local structure. We have chosen adaptive thresholding (c_{plane}) parameters which we thought gave the best results and it is clear that, as the authors suggest, if incorporated in an deformable model framework a better segmentation could be obtained. In our experiments with the approach, we have found that the structure tensor picks out the direction of maximum change in intensity but spatially, it does not behave well for our application. The c_{plane} measure is strong mostly at boundaries and tends to thicken the edges as seen in the tensor norm response in Fig. 4. Our method has the advantage of extracting locations at the center of bone structures where the underlying iso-intensity level set behaves like a plane. The sheetness measure combined with a flow designed to evolve and stop at boundaries performs better as it is able propagate along low sheetness regions. Most importantly, on its own, the sheetness measure has suitable characteristics to be incorporated in a surgical simulator for pituitary intervention. The measure is high on the

center plane of bones and decreases towards the boundaries where there is an uncertainty about the tissue classification of a voxel. Hence, it can allow a surface mesh model to account for uncertainty in determining triangulated bone boundaries, to better model sinus bones in the simulation. For our application, it is thus more useful to have a confidence measure for all voxels than a binary segmentation of bones which certainly still has missing segments and no notion of distance to the boundaries.

6 Conclusions

We have presented a general multi-scale bone enhancement measure that can be used to drive a geometric flow to segment sheet-like structures. The key contribution is the introduction of a the sheetness measure based on the properties of the multi-scale Hessian shape operator which has well-founded differential geometric theory and is stable under the presence of noise. The measure gives a confidence index on the presence of bones and for voxels with high values, the scale and direction of the local bone structure is directly obtained. We are able to detect and connect very thin and diffusive bone structure boundaries. In current work, we are trying to use both the bone segmentation as well as the sheetness, vesselness and other tissue classification cues from the CT and corresponding MR data to use a global class competition levelset framework.

References

1. I. Ciric et al. Complications of transssphenoidal surgery: Results of a national survey, review of the literature, and personal experience. *Neurosurgery*, 40(2):225–236, 1997.
2. M. Descoteaux, L. Collins, and K. Siddiqi. A multi-scale geometric flow for segmenting vasculature in mri. In *MICCAI'2004*, September 2004.
3. A. Frangi, W. Niessen, K. L. Vincken, and M. A. Viergever. Multiscale vessel enhancement filtering. In *MICCAI'98*, pages 130–137, 1998.
4. K. Krissian, G. Malandain, and N. Ayache. Model-based detection of tubular structures in 3d images. *Computer Vision and Image Understanding*, 80(2):130–171, November 2000.
5. C. Lorenz, I. Carlsen, T. Buzug, C. Fassnacht, and J. Weese. Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2d and 3d medical images. In *CVRMED-MRCAS'97*, volume 1205, pages 233–242, 1997.
6. Y. Sato et al. and R. Kikinis. 3d multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. *Medical Image Analysis*, 2(2):143–168, 1998.
7. T. B. Sebastian et al. and B. B. Kimia. Segmentation of carpal bones from 3d ct images using skeletally coupled deformable models. In *MICCAI'98*, pages 1184–1194, 1998.
8. A. Vasilevskiy and K. Siddiqi. Flux maximizing geometric flows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(12):1–14, 2002.
9. C.-F. Westin, A. Bhalerao, H. Knutsson, and R. Kikinis. Using local 3D structure for segmentation of bone from computer tomography images. In *CVPR'97*, pages 794–800, 1997.

Simultaneous Registration and Segmentation of Anatomical Structures from Brain MRI

Fei Wang* and Baba C. Vemuri*

Department of Computer & Information Sciences & Engr.,
University of Florida, Gainesville, FL 32611
{fewang, vemuri}@cise.ufl.edu

Abstract. In this paper, we present a novel variational formulation of the registration assisted image segmentation problem which leads to solving a coupled set of nonlinear PDEs that are solved using efficient numerical schemes. Our work is a departure from earlier methods in that we have a unified variational principle wherein non-rigid registration and segmentation are simultaneously achieved; unlike previous methods of solution for this problem, our algorithm can accommodate for image pairs having very distinct intensity distributions. We present examples of performance of our algorithm on synthetic and real data sets along with quantitative accuracy estimates of the registration.

1 Introduction

In Medical Imaging applications, segmentation can be a daunting task due to possibly large inhomogeneities in image intensities across an image e.g., in MR images. These inhomogeneities combined with volume averaging during the imaging and possible lack of precisely defined shape boundaries for certain anatomical structures complicates the segmentation problem immensely. One possible solution for such situations is atlas-based segmentation. The atlas once constructed can be used as a template and can be registered non-rigidly to the image being segmented (henceforth called a target image) thereby achieving the desired segmentation. Many of the methods that achieve atlas-based segmentation are based on a two stage process involving, (i) estimating the non-rigid deformation field between the atlas image and the target image and then, (ii) applying the estimated deformation field to the desired shape/atlas to achieve the segmentation of the corresponding structure/s in the target image. In this paper, we develop a novel technique that will simultaneously achieve the non-rigid registration and segmentation. There is a vast body of literature for the tasks of registration and segmentation independently however, methods that combine them into one algorithm are far and few in between. In the following, we will briefly review the few existing methods that attempt to achieve simultaneous registration and segmentation.

* This research was in part funded by the NIH grants, RO1 NS046812 & NS42075. Authors thank Dr. C.M. Leonard of the UF Neuroscience Dept. for the hippocampal data sets.

In one of the earliest attempts at joint registration & segmentation, Bansal et al., [1] developed a minmax entropy framework to rigidly register & segment portal and CT data sets. In [2], Yezzi et al., present a variational principle for achieving simultaneous registration and segmentation, however, the registration part is limited to rigid motions. A similar limitation applies to the technique presented by Noble et al., in [3]. A variational principle in a level-set based formulation was presented in Pargios et. al., [4], for segmentation and registration of cardiac MRI data. Their formulation was again limited to rigid motion and the experiments were limited to 2D images. In Fischl et al., [5], a Bayesian method is presented that simultaneously estimates a linear registration and the segmentation of a novel image. Note that linear registration does not involve non-rigid deformations. The case of joint registration and segmentation with non-rigid registration has not been addressed adequately in literature with the exception of the recent work reported in Soatto *et al.*, [6] and Vemuri *et al.*, [7]. However, these methods can only work with image pairs that are necessarily from the same modality or the intensity profiles are not too disparate.

In this paper, we present a unified variational principle that will simultaneously register the atlas shape (contour/surface) to the novel brain image and segment the desired shape (contour/surface) in the novel image. In this work, the atlas serves in the segmentation process as a prior and the registration of this prior to the novel brain scan will assist in segmenting it. Another key feature/strength of our proposed registration+segmentation scheme is that *it accommodates for image pairs having very distinct intensity distributions as in multimodality data sets*. More details on this are presented in section 2.

2 Registration+Segmentation Model

We now present our formulation of joint registration & segmentation model wherein it is assumed that the image to be segmented can be modeled by piecewise constant regions, as was done in [8]. This assumption simplifies our presentation but our model itself can be easily extended to the piecewise smooth regions case. Additionally, since we are only interested in segmenting a desired anatomical shape (e.g., the hippocampus, the corpus callosum, etc.), we will only be concerned with a binary segmentation i.e., two classes namely, voxels inside the desired shape and those that are outside it. These assumptions can be easily relaxed if necessary but at the cost of making the energy functional more complicated and hence computationally more challenging. Let I_1 be the atlas image containing the atlas shape \bar{C} , I_2 the novel image that needs to be segmented and \mathbf{v} be the vector field, from I_2 to I_1 i.e., the transformation is centered in I_2 , defining the non-rigid deformation between the two images. The variational principle describing our formulation of the registration assisted segmentation problem is given by:

$$\min E(\mathbf{v}, \tilde{C}) = Seg(I_2, \tilde{C}) + dist(\mathbf{v}(\bar{C}), \tilde{C}) + Reg(I_1, I_2, \mathbf{v}). \quad (1)$$

Where, the first term denotes the segmentation functional. \tilde{C} is the boundary contour (surface in 3D) of the desired anatomical shape in I_2 . The second term

measures the distance between the transformed atlas $\mathbf{v}(\bar{C})$ and the current segmentation \tilde{C} in the novel brain image i.e., the target image and the third term denotes the non-rigid registration functional between the two images. Each of these functionals are given the following form:

$$Seg(I_2, \tilde{C}) = \int_{\Omega} (I_2 - u)^2 dx + \alpha \oint_{\tilde{C}} ds \quad (2)$$

Where, Ω is the image domain and α is a regularization parameter. $u = u_i$ if $x \in \tilde{C}_{in}$ and $u = u_o$ if $x \in \tilde{C}_{out}$. \tilde{C}_{in} and \tilde{C}_{out} denote the regions inside and outside of the curve, \tilde{C} representing the desired shape boundaries in I_2 .

For the non-rigid registration term in the energy function, we use the recently introduced information theoretic-based criteria [9] called the cross cumulative residual entropy (CCRE). In [9], CCRE was shown to outperform MI-based registration in the context of noise immunity and convergence range, motivating us to pick this criteria over the MI-based cost function. The new registration functional is defined by

$$Reg(I_1, I_2, \mathbf{v}) = - \left(\mathcal{C}(I_1(\mathbf{v}(\mathbf{x})), I_2(\mathbf{x})) + \mu \int_{\Omega} (\|\nabla \mathbf{v}(\mathbf{x})\|^2) \right) \quad (3)$$

where, cross-CRE $\mathcal{C}(I_1(), I_2()) = \mathcal{E}(I_1) - E[\mathcal{E}(I_1/I_2)]$ and $\mathcal{E}(\mathbf{I}_1) = - \int_{\mathcal{R}_+} P(|\mathbf{I}_1| > \lambda) \log P(|\mathbf{I}_1| > \lambda) d\lambda$ with $\mathcal{R}_+ = (x \in \mathcal{R}; x \geq 0)$. $\mathbf{v}(\mathbf{x})$ is as before and μ is the regularization parameter and $\|\cdot\|$ denotes Frobenius norm. Using a B-spline representation of the non-rigid deformation, one need only compute this field at the control points of the B-splines and interpolate elsewhere, thus accruing computational advantages. Using this representation, we derived analytic expressions for the gradient of the energy with respect to the registration parameters. This in turn makes our optimization more robust and efficient.

In order for the registration and the segmentation terms to “talk” to each other, we need a connection term and that is given by

$$dist(\mathbf{v}(\bar{C}), \tilde{C}) = \int_R \phi_{\mathbf{v}(\bar{C})}(\mathbf{x}) d\mathbf{x} \quad (4)$$

Where, R is the region enclosed by \tilde{C} , $\phi_{\mathbf{v}(\bar{C})}(\mathbf{x})$ is the embedding signed distance function of the contour $\mathbf{v}(\bar{C})$, which can be used to measure the distance between $\mathbf{v}(\bar{C})$ and \tilde{C} . The level-set function $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ is chosen so that its zero level-set corresponds to the transformed template curve $\mathbf{v}(\bar{C})$. Let $E_{dist} := dist(\mathbf{v}(\bar{C}), \tilde{C})$, one can show that $\frac{\partial E_{dist}}{\partial \tilde{C}} = \phi_{\mathbf{v}(\bar{C})}(\tilde{C})\mathbf{N}$ where \mathbf{N} is the normal to \tilde{C} . The corresponding curve evolution equation given by gradient descent is then

$$\frac{\partial \tilde{C}}{\partial t} = -\phi_{\mathbf{v}(\bar{C})}(\tilde{C})\mathbf{N} \quad (5)$$

Not only does the signed distance function representation make it easier for us to convert the curve evolution problem to the level-set framework (refer to section 3), it also facilitates the matching of the evolving curve \tilde{C} and the transformed

template curve $\mathbf{v}(\bar{C})$, and yet does not rely on a parametric specification of either \bar{C} or the transformed template curve. Note that since $dist(\mathbf{v}(\bar{C}), \tilde{C})$ is a function of the unknown registration \mathbf{v} and the unknown segmentation \tilde{C} , it plays the crucial role of connecting the registration and the segmentation terms.

Combining these three functionals together, we get the following variational principle for the simultaneous registration+segmentation problem:

$$\begin{aligned} min E(\tilde{C}, \mathbf{v}, u_o, u_i) = & \int_{\Omega} (I_2 - u)^2 dx + \alpha_1 \oint_{\tilde{C}} ds + \alpha_2 dist(\mathbf{v}(\bar{C}), \tilde{C}) \\ & - \alpha_3 \mathcal{C}(I_1(\mathbf{v}(\mathbf{x})), I_2(\mathbf{x})) + \alpha_4 \int_{\Omega} \|\nabla \mathbf{v}(\mathbf{x})\|^2 d\mathbf{x}. \end{aligned} \quad (6)$$

α_i are weights controlling the contribution of each term to the overall energy function and can be treated as unknown constants and either set empirically or estimated during the optimization process. This energy function is quite distinct from those used in methods existing in literature because it is achieving the Mumford-Shah type of segmentation in an active contour framework jointly with non-rigid registration and shape distance terms. We are now ready to discuss the level-set formulation of the energy function in the following section.

3 Level Set Formulation

We now present a level-set form of our formulation described earlier. For our model where the equation for the unknown curve \tilde{C} is coupled with the equations for $\mathbf{v}(\mathbf{x})$, u_o , u_i , it is convenient for us to use the level set approach as proposed in [8]. Taking the variation of $E(\cdot)$ with respect to \tilde{C} and writing down the gradient descent leads to the following curve evolution equation:

$$\frac{\partial \tilde{C}}{\partial t} = - \left[-(I_2 - u_i)^2 + (I_2 - u_o)^2 + \alpha_1 \kappa + \alpha_2 \phi_{\mathbf{v}(\bar{C})}(\tilde{C}) \right] \mathbf{N} \quad (7)$$

Note that equation (5) is used in the derivation. Equation (7) in the level-set framework is given by:

$$\frac{\partial \phi}{\partial t} = \left[-(I_2 - u_i)^2 + (I_2 - u_o)^2 + \alpha_1 \nabla \cdot \frac{\nabla \phi}{|\nabla \phi|} + \alpha_2 \phi_{\mathbf{v}(\bar{C})}(\tilde{C}) \right] |\nabla \phi| \quad (8)$$

where u_i and u_o are the mean values inside and outside of the curve \tilde{C} in the image I_2 .

As mentioned before, we use a B-spline basis to represent the displacement vector field $\mathbf{v}(\mathbf{x}, \mu)$, where μ is the transformation parameters of the B-spline basis.

$$\frac{\partial E}{\partial \mu} = \alpha_2 \frac{\partial \int_R \phi_{\mathbf{v}(\bar{C})}(\mathbf{x}) d\mathbf{x}}{\partial \mu} - \alpha_3 \frac{\partial \mathcal{C}(I_1(\mathbf{v}(\mathbf{x})), I_2(\mathbf{x}))}{\partial \mu} + \alpha_4 \frac{\partial \int_{\Omega} \|\nabla \mathbf{v}(\mathbf{x})\|^2 d\mathbf{x}}{\partial \mu} \quad (9)$$

The first term of equation(9) can be rewritten as follows:

$$\begin{aligned} \frac{\partial \int_R \phi_{\mathbf{v}(\bar{C})}(\mathbf{x}) d\mathbf{x}}{\partial \mu} &= \int_R \frac{\partial \phi_{\mathbf{v}(\bar{C})}(\mathbf{x})}{\partial \mu} d\mathbf{x} \\ &= \int_R \frac{\partial \phi_{\mathbf{v}(\bar{C})}}{\partial v} \Big|_{v=\mathbf{v}(\mathbf{x}, \mu)} \cdot \frac{\partial \mathbf{v}(\mathbf{x}, \mu)}{\partial \mu} d\mathbf{x} \end{aligned} \quad (10)$$

where $\frac{\partial \phi_{\mathbf{v}(\bar{C})}}{\partial v}$ is the directional derivative in the direction of $\mathbf{v}(\mathbf{x}, \mu)$. The second term of equation(9) is more involved and tedious. We simply state the result here without the derivations for the sake of brevity,

$$\frac{\partial \mathcal{C}(I_2, I_1 \circ \mathbf{v}(x; \mu))}{\partial \mu} = \sum_{\lambda \in I_1} \sum_{k \in I_2} \log \frac{P(i > \lambda, k; \mu)}{p_{I_2}(k)P(i > \lambda; \mu)} \cdot \frac{\partial P(i > \lambda, k; \mu)}{\partial \mu} \quad (11)$$

where $P(i > \lambda, k; \mu)$ and $P(i > \lambda; \mu)$ are the joint and marginal cumulative residual distributions respectively. $p_{I_2}(k)$ is the density function of image I_2 . The last term of equation(9) leads to,

$$\frac{\partial \int_{\Omega} \|\nabla \mathbf{v}(\mathbf{x})\|^2 d\mathbf{x}}{\partial \mu} = 2 \int_{\Omega} \nabla \mathbf{v} \cdot \frac{\partial \mathbf{v}}{\partial \mu} d\mathbf{x} \quad (12)$$

where both the matrices $\nabla \mathbf{v}$ and $\frac{\partial \mathbf{v}}{\partial \mu}$ are vectorized before the dot product is computed.

Substituting equations (10), (11) and (12) respectively back into the equation (9), we get the analytical gradient of our energy function with respect to the B-spline transformation parameters μ . We then solve for the stationary point of this nonlinear equation numerically using a quasi-Newton method.

Algorithm Summary

Given atlas image I_1 and the unknown subject's brain scan I_2 , we want the segmentation result \tilde{C} in I_2 . Initialize \tilde{C} in I_2 to be \bar{C} , set initial transformation parameters μ_0 to be zero.

1. Optimize μ_i using equation (9) with Quasi-Newton method for one step. Update the deformation field $\mathbf{v}(x; \mu_i)$.
2. Evolve ϕ in I_2 using equation (8) for one step, update \bar{C}_i as the zero level-set of ϕ .
3. Stop the registration process if the difference in consecutive iterates is less than $\epsilon = 0.01$, a pre-chosen tolerance, else go to **Step 1**.

4 Implementation Results

In this section, we present several examples results from an application of our algorithm. The results are presented for synthetic as well as real data. The first

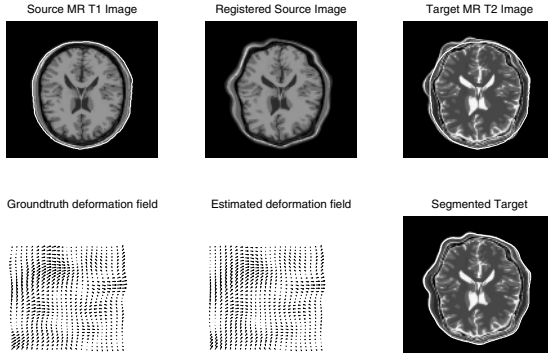


Fig. 1. Results of application of our algorithm to synthetic data (see text for details)

three experiments were performed in 2D, while the fourth one was performed in 3D. Note that the image pairs used in all these experiments have significantly different intensity profiles, which is unlike any of the previous methods, reported in literature, used for joint registration and segmentation. The synthetic data example contains a pair of MR T1 and T2 weighted images which are from the MNI brainweb site [10]. They were originally aligned with each other. We use the MR T1 image as the source image and the target image was generated from the MR T2 image by applying a known non-rigid transformation that was procedurally generated. In this case, we present the error in the estimated non-rigid deformation field, using our algorithm, as an indicator of the accuracy of estimated deformations.

Figure 1 depicts the results obtained for this image pair. With the MR T1 image as the source image, the target was obtained by applying a synthetically generated non-rigid deformation field to the MR T2 image. Notice the significant difference between the intensity profiles of the source and target images. Figure 1 is organized as follows, from left to right: the first row depicts the source image with the atlas-segmentation superposed in white, the registered source image which is obtained using our algorithm followed by the target image with the un-registered atlas-segmentation superposed to depict the amount of mis-alignment; second row depicts ground truth deformation field which we used to generate the target image from the MR T2 image, followed by the estimated non-rigid deformation field and finally the segmented target. As evident, the registration+segmentation are quite accurate from a visual inspection point of view. As a measure of accuracy of our method, we estimated the average, μ , and the standard deviation, σ , of the error in the estimated non-rigid deformation field. The error was estimated as the angle between the ground truth and estimated displacement vectors. *The average and standard deviation are 1.5139 and 4.3211 (in degrees) respectively, which is quite accurate.*

Table 1 depicts statistics of the error in estimated non-rigid deformation when compared to the ground truth. For the mean ground truth deformation

(magnitude of the displacement vector) in Column-1 of each row, 5 distinct deformation fields with this mean are generated and applied to the target image of the given source-target pair to synthesize 5 pairs of distinct data sets. These pairs (one at a time) are input to our algorithm and the mean (μ) of the mean deformation error (MDE) is computed over the five pairs and reported in Column-2 of the table. MDE is defined as $d_m = \frac{1}{\text{card}(R)} \sum_{x_i \in R} \|\mathbf{v}_0(x_i) - \mathbf{v}(x_i)\|$, where $\mathbf{v}_0(x_i)$ is the ground truth and estimated displacements respectively at voxel x_i . $\|\cdot\|$ denotes the Euclidean norm, and R is the volume of the region of interest. Column-3 depicts the standard deviation of the MDE for the five pairs of data in each row. As evident, the mean and the standard deviation of the error are reasonably small indicating the accuracy of our joint registration + segmentation algorithm. *Note that this testing was done on a total of 20 image pairs (=40) as there are 5 pairs of images per row.*

Table 1. Statistics of the error in estimated non-rigid deformation.

μ_g	μ of MDE	σ of MDE
2.4	0.5822	0.0464
3.3	0.6344	0.0923
4.5	0.7629	0.0253
5.5	0.7812	0.0714

For the first real data experiment, we selected two image slices from two different modalities of brain scans. The two slices depict considerable changes in shape of the ventricles, the region of interest in the data sets. One of the two slices was arbitrarily selected as the source and segmentation of the ventricle in the source was achieved using an active contour model. The goal was then to automatically find the ventricle in the target image using our algorithm given the input data along with the segmented ventricles in the source image. Figure 2 is organized as follows, from left to right: the first row depicts the source image with the atlas-segmentation superposed in black followed by the target image with the unregistered atlas-segmentation superposed to depict the amount of mis-alignment; second row depicts the estimated non-rigid vector field and finally the segmented target. As evident from figures 2, the accuracy of the achieved registration+segmentation visually very good. Note that the non-rigid deformation between the two images in these two examples is quite large and our method was able to simultaneously register and segment the target data sets quite accurately.

The second real data example is obtained from two brain MRIs of different subjects and modalities, the segmentation of the cerebellum in the source image is given. We selected two “corresponding” slices from these volume data sets to conduct the experiment. Note that even though the number of slices for the two data sets are the same, the slices may not correspond to each other from an anatomical point of view. However, for the purposes of illustration of our algorithm, this is not very critical. We use the corresponding slice of the 3D segmentation of the source as our atlas-segmentation. The results of an application of our algorithm are organized as before in figure 3. Once again, as evident, the visual quality of the segmentation and registration are very high.

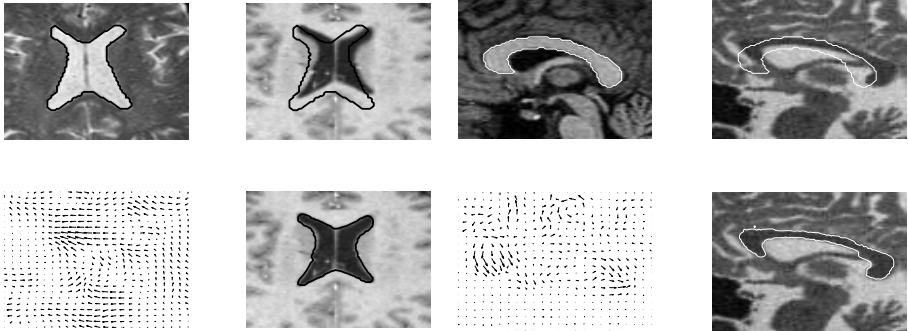


Fig. 2. Results of application of our algorithm to a pair of slices from human brain MRIs (see text for details)

Fig. 3. Corpus Callosum segmentation on a pair of corresponding slices from distinct subjects

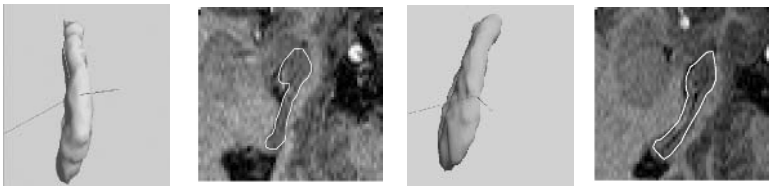


Fig. 4. Hippocampal segmentation using our algorithm on a pair of brain scans from distinct subjects. (see text for details)

Finally we present a 3D real data experiment. In this experiment, the input is a pair of 3D brain scans with the segmentation of the hippocampus in one of the two images (labeled the atlas image) being obtained using the well known PCA on the several training data sets. Each data set contains 19 slices of size 256x256. The goal was then to automatically find the hippocampus in the target image given the input. Figure 4 depicts the results obtained for this image pair. From left to right, the first image shows the given (atlas) hippocampus surface followed by one cross-section of this surface overlaid on the source image slice; the third image shows the segmented hippocampus surface from the target image using our algorithm and finally the cross-section of the segmented surface overlaid on the target image slice. To validate the accuracy of the segmentation result, we randomly sampled 120 points from the segmented surface and computed the average distance from these points to the ground truth hand segmented hippocampal surface in the target image. The hand segmentation was performed by an expert neuroanatomist. The average and standard deviation of the error in the aforementioned distance in estimated hippocampal shape are 0.8190 and 0.5121(in voxels) respectively, which is very accurate.

5 Conclusions

In this paper, we presented a novel variational formulation of the joint (non-rigid) registration and segmentation problem which requires the solution to a coupled set of nonlinear PDEs that are solved using efficient numerical schemes. Our work is a departure from earlier methods in that we have a *unified variational principle* wherein non-rigid registration and segmentation are simultaneously achieved, our algorithm can also accommodate for image pair having distinct intensity distributions. We presented several examples (twenty) on synthetic and (three) real data sets along with quantitative accuracy estimates of the registration in the synthetic data case. More extensive experimentation under different amounts of noise and varying degrees of non-rigidity needs to be performed prior to drawing conclusions on the accuracy of the proposed model. This will be the focus of our future efforts.

References

1. Bansal, R., Staib, L., Chen, Z., Rangarajan, A., Knisely, J., Nath, R., Duncan, J.: Entropy-based, multiple-portal-to-3dct registration for prostate radiotherapy using iteratively estimated segmentation. In: MICCAI. (1999) 567–578
2. Yezzi, A., Zollei, L., Kapur, T.: A variational framework for joint segmentation and registration. In: IEEE CVPR - MMBIA. (2001) 388–400
3. Wyatt, P., Noble, J.: Mrf-map joint segmentation and registration. In: MICCAI. (2002) 580–587
4. Paragios, N., Rousson, M., Ramesh, V.: Knowledge-based registration & segmentation of the left ventricle: A level set approach. In: WACV. (2002) 37–42
5. Fischl, B., Salat, D., Buena, E., et.al., M.A.: Whole brain sementation: Automated labeling of the neuroanatomical structures in the human brain. In: Neuron. Volume 33. (2002) 341–355
6. Soatto, S., Yezzi, A.J.: Deformation: Deforming motion, shape average and the joint registration and segmentation of images. In: ECCV. (2002) 32–57
7. Vemuri, B.C., Chen, Y., Wang, Z.: Registration assisted image smoothing and segmentation. In: ECCV. (2002) 546–559
8. Chan, T., Vesse, L.: An active contour model without edges. In: Intl. Conf. on Scale-space Theories in Computer Vision. (1999) 266–277
9. Wang, F., Vemuri, B.C., Rao, M., Chen, Y.: A new & robust information theoretic measure and its application to image alignment. In: IPMI. (2003) 388–400
10. : Mcconnell brain imaging centre brain database.
<http://www.bic.mni.mcgill.ca/brainweb/> (1997)

Synthetic Ground Truth for Validation of Brain Tumor MRI Segmentation*

Marcel Prastawa¹, Elizabeth Bullitt², and Guido Gerig^{1,3}

¹ Dept. of Computer Science,

² Dept. of Surgery,

³ Dept. of Psychiatry,

University of North Carolina, Chapel Hill, NC 27599, USA

{prastawa, gerig}@cs.unc.edu, bullitt@med.unc.edu

Abstract. Validation and method of comparison for segmentation of magnetic resonance images (MRI) presenting pathology is a challenging task due to the lack of reliable ground truth. We propose a new method for generating synthetic multi-modal 3D brain MRI with tumor and edema, along with the ground truth. Tumor mass effect is modeled using a biomechanical model, while tumor and edema infiltration is modeled as a reaction-diffusion process that is guided by a modified diffusion tensor MRI. We propose the use of warping and geodesic interpolation on the diffusion tensors to simulate the displacement and the destruction of the white matter fibers. We also model the process where the contrast agent tends to accumulate in cortical csf regions and active tumor regions to obtain contrast enhanced T1w MR image that appear realistic. The result is simulated multi-modal MRI with ground truth available as sets of probability maps. The system will be able to generate large sets of simulation images with tumors of varying size, shape and location, and will additionally generate infiltrated and deformed healthy tissue probabilities.

1 Introduction

The segmentation of brain tumor from magnetic resonance (MR) images is a vital process for treatment planning and for studying the differences of healthy subjects and subjects with tumor. The process of automatically extracting tumor from MR images is a challenging process, and a variety of methods have been proposed [1,2,3]. The typical standard for validation of the different segmentation methods is comparison against the results of manual raters. However, manual segmentation suffers from the lack of reliability and reproducibility, and different sites may have different methods for manually outlining tumors in MRI. The true ground truth may need to be estimated from a collection of manual segmentations [4]. Validation of the segmentation of structures other than brain tumor is typically not done since manual segmentation of edema or of the whole brain are very challenging tasks and might not represent truth very well.

Brain MRI with tumor is difficult to segment due to a combination of the following factors:

* This work was supported by NIH-NIBIB R01 EB000219.

1. The deformation of non-tumor structures due to tumor mass effect.
2. Infiltration of brain tissue by tumor and edema (swelling). Edema appears around the tumor mainly in white matter regions.
3. There is gradual transition from tumor to edema, often it is difficult to discern the boundary between the two structures.
4. The standard MR modality used to identify tumor, T1w with contrast enhancement (typically using gadolinium), is not always ideal. Blood vessels and cortical cerebrospinal fluid (csf) tend to be highlighted along with tumor, while parts of tumor that are necrotic tissue do not appear enhanced at all. It is generally impossible to segment tumor by simply thresholding the contrast enhanced T1w image.

Rexilius *et al.*[5] proposed a framework for generating digital brain phantoms with tumor. They used a biomechanical finite element model to simulate the tumor mass effect. The phantom for a healthy subject is deformed and a tumor structure from a real subject is inserted to the MRI. Their model for edema is computed from the distances to the tumor boundary and the white matter mask. This is insufficient to simulate some real infiltration properties because infiltration is not only influenced by distance to tumor. Typically, edema infiltration occurs following the white matter fibers. Their framework only considers contrast enhancement inside tumors, without enhancement of vessels and csf regions.

We propose a method for generating simulated brain tumor MRI which includes most of the difficulties encountered in real MR images. The MR images produced by the method presents the four challenges as listed above. Tumor mass effect is simulated using a biomechanical model. Infiltration of brain tissues by tumor and edema is simulated as a reaction-diffusion process that is guided by a modified diffusion tensor MR image (DT-MRI). We also simulate the process where the contrast agent accumulates in some fluid regions and outer tumor regions to generate contrast enhanced T1w MRI that reflect challenges encountered in real tumor MR images.

2 Method

The input for our method is a ground truth of a healthy subject, which is a set of spatial probabilities for white matter, gray matter, and csf as shown in

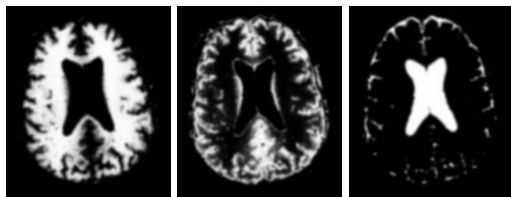


Fig. 1. The input for our phantom generation framework: ground truth for a healthy subject. From left to right: axial view of the spatial probabilities for white matter, gray matter, and csf.

Figure 1. We generate new spatial probabilities for tumor and edema, and modify the healthy probabilities to account for mass effect and infiltration. The new set of probabilities is used to simulate MR images given training data that is obtained from real brain MRI with tumors. The probability that a particular location contains contrast agent is computed to determine regions that appear highlighted in T1w MRI.

The brain tumor MRI simulation system includes the following steps:

1. Selection of a MRI of a healthy subject with probabilistic tissue segmentation. This provides the initial healthy ground truth that will be transformed into pathological ground truth.
2. Selection of a region for the initial tumor probability that describes the initial state of the tumor. Tumor growth is then simulated through deformation and infiltration.
3. Computation of a deformation field using a biomechanical model. The probabilistic maps and the DT-MRI are warped using the deformation.
4. Modification of the DT-MRI to account for destruction of white matter fibers.
5. Simulation of tumor and edema infiltration using the modified DT-MRI.
6. Simulation of the process of accumulation of contrast agent in fluid and tumor regions.
7. Generation of multi-modal MRI given the final healthy tissue, tumor, and edema probability maps.

2.1 Tumor Mass Effect

Given an initial tumor region that is obtained through user interaction, the growth of the tumor is modeled as a deformation process. The initial tumor region can also be automatically drawn at random given some prior knowledge of the spatial distribution. Meningiomas, for example, often appear near the brain periphery. The initial tumor region can have arbitrary shape and influences the resulting deformation. In the initial tumor region, the tumor probabilities are set to be one, $p_{tumor}(x) = 1$ and tissue or fluid probabilities are set to be zero. The set of spatial probabilities, with the addition of the new spatial tumor probability are deformed according to a biomechanical model.

We have chosen to use the linear elastic finite element model used in [6,5]. The external force for the system is a radial force that originates from the initial tumor region, and this force weakens by distance.

2.2 Tumor and Edema Infiltration

We use the registered and reoriented DT-MRI to simulate the tissue infiltration process, similar to the approach done by Clatz *et al.*[7,8]. However, registration and reorientation are generally insufficient to account for the mass effect. White matter fibers around tumor tend to be displaced, and as observed by Lu *et al.*[9], in regions near the tumor the mean diffusivity (MD) tends to be increased while

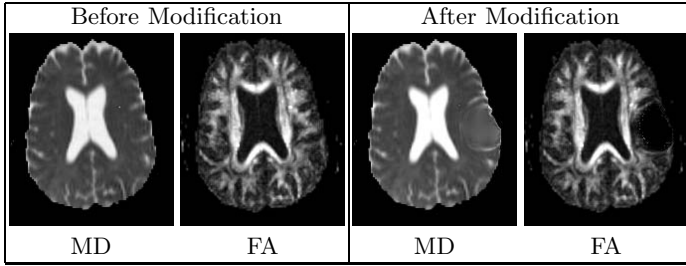


Fig. 2. Visualization of the diffusion tensor MRI through the axial views of the 3D Mean Diffusivity (MD) and Fractional Anisotropy (FA) images. The modified DT-MRI has higher MD and lower FA in the regions surrounding tumor, which models the destruction of the fibers. The MD image shows that the ventricle near the tumor is slightly deformed. The FA image shows that the white matter fibers near the tumor region are pushed away.

the fractional anisotropy (FA) tends to be decreased. This effect can be seen as a reflection of the destruction of white matter fibers due to tumor growth.

The influence of tumor mass effect on DT-MRI is modeled using a combination of image warping and non-linear interpolation. The DT-MRI is warped following the strategy described in [10], where a rigid rotation is applied to each individual tensors. The rigid rotation for each tensor is computed based on the local warping property. The destruction of white matter fibers is simulated by interpolating each individual tensor with an isotropic tensor that has higher mean diffusivity. Figure 2 shows the registered DT-MRI before and after the modification that models the influence of tumors.

The interpolation between a warped tensor D and its isotropic version E is calculated using the geodesic interpolation strategy proposed by Fletcher *et al.*[11]. The isotropic tensor E is formulated to have two times the determinant value of D , $E = (2|D|)^{\frac{1}{3}} I_{3 \times 3}$. The weight α for the interpolation between D and E is inversely proportional to the amount of deformation. We use $\alpha(x) = \exp(-\frac{K(x)}{2\sigma_K^2})$ where $K(x) = \max(1, |J(x)|) - 1$ with J being the Jacobian matrix of the spatial coordinate mapping function and σ_K is the rough estimate of the amount of deformation that destroys fibers. The modified tensor we use for computing infiltration is $D'(x) = \text{interpolate}(D(x) : \alpha, E(x) : (1 - \alpha))$. With our formulation of α , volume expansion destroys fibers while volume compression does not change the original tensors. This is done because we observed that in real tumor DT-MRI some fibers can appear condensed without being destroyed.

The spatial probability that a particular location is infiltrated by pathological cells, $p_{infiltrated} = \phi$, is evolved using the modified DT-MRI as follows:

$$\frac{\partial \phi}{\partial t} = \text{div}(a D' \nabla \phi) + b \phi(1 - \phi)$$

The first term is the DT-MRI guided diffusion, with an additional parameter a that depends on the tissue type. White matter is more likely to diffuse than

white matter, while csf is not likely to be infiltrated at all. The second term is the growth term, with b being a constant. The diffusion tensors D' are normalized so that the trace of the tensors is within the range of $[0, 1]$. The initial values for ϕ is chosen to be higher for regions with high tumor and white matter probability, $\phi(x, t = 0) = p_{white}(x) p_{tumor}(x)$. The evolution is stopped when the volume of infiltrated brain voxels is higher than a fraction of the brain tissue volume.

Tumor may not only deform tissue, but also infiltrate nearby tissue. We model this by attributing the early stages of infiltration to tumor and the later stages to edema. The probability of infiltrated tumor or edema is the probability that a location is both infiltrated and part of brain tissue. More precisely, $p_{tumor}(x) = p_{warped\ tumor} + \phi(x, t_{early}) p_{tissue}(x)$ and $p_{edema}(x) = (\phi(x, t_{final}) - \phi(x, t_{early})) p_{tissue}(x)$. The probability of observing brain tissue is $p_{tissue}(x) = p_{white}(x) + p_{gray}(x)$. The value t_{final} is the time where evolution stops, and t_{early} is a time value earlier in the process, $t_{early} < t_{final}$. The choice for the value of t_{early} depends on the type of tumor being modeled. For example, meningiomas tend to have less tumor infiltration compared to glioblastomas.

2.3 Generation of MR Image Intensities

One of the particular challenges in segmenting brain tumor from MRI is the inconsistencies in the contrast enhanced T1w image. Due to the biological process, the contrast agent is almost always accumulated in regions other than tumor, mainly in the cortical csf and the blood vessels. Additionally, the necrotic parts of the tumor do not accumulate the contrast agent at all. Tumor necrosis are typically found in the core tumor regions.

Our method models the accumulation of the contrast agent in the active tumor tissue and the cortical csf in order to generate more challenging images. The spatial probability for the accumulation of contrast agent, $p_{acc} = \eta$ is evolved using the following reaction diffusion equation:

$$\frac{\partial \eta}{\partial t} = \text{div}(u \nabla \eta) + I\{x \in X_{source}\} v \eta - I\{x \in X_{sink}\} w \eta$$

where v and w are constants and I is the indicator function. The value of u depends on the tissue type at location x , contrast agent is modeled to be more

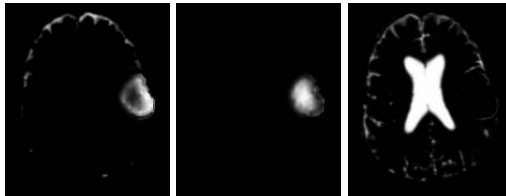


Fig. 3. Axial view of the generated probabilities related to contrast enhancement. From left to right: probability for highlighted csf or tumor, probability for non-highlighted tumor, and probability for non-highlighted csf.

likely to spread in csf than in tumor tissue. X_{source} and X_{sink} are the sets of points that act as sources or sinks respectively, the points are chosen at random. Locations with high csf probability and low distance to brain boundary are more likely to be source points. Within tumor, the outer regions are more likely to be source points while the core regions are more likely to be sink points. We initialize η so that the voxels in the source regions are equally likely to have accumulated contrast agent or to have no contrast agent at all, $\eta(x, t = 0) = I\{x \in X_{source}\} 0.5$.

The probability that a voxel would appear highlighted in contrast enhanced T1w MRI is the probability that the voxel is csf or tumor and has accumulated contrast agent, $p_{enhanced}(x) = p_{acc}(p_{csf}(x) + p_{tumor}(x))$. Figure 3 shows the generated enhancement probabilities. Our model accounts for the fact that cortical csf and active tumor regions are highlighted and that necrosis regions are not enhanced. However, it does not account for the fact that blood vessels can also appear enhanced.

Given the modified spatial probabilities of the healthy ground truth, the MR images are generated as linear combinations of a set of mean intensities for each class:

$$I^m(x) = \sum_{i=1}^{N_m} p_{C_i}(x) \mu_i$$

where m is the modality, N_m is the number of classes, C_i is one of the classes used for that modality, and μ_i is the mean intensity for class C_i . The mean class intensities are obtained from real brain tumor MRI. For T1w and T2w images, the set of classes C is composed of white matter, gray matter, csf, tumor, and edema. For the contrast enhanced T1w image, the set of classes C is composed of white matter, gray matter, non-enhancing csf, non-enhancing tumor, edema, and the class for all contrast enhanced voxels.

3 Results

The synthetic brain tumor MRI is shown together with an example of a real brain tumor MRI in Figure 4. In both cases, tumor deforms other structures and edema infiltrates brain tissue. The contrast enhanced T1w MR images also show complex highlight patterns. Figure 5 shows the ground truth for the synthetic MRI. Ground truth is presented as a set of probability maps for tissue and pathology, similar to the one provided by BrainWeb [12] for healthy subjects. This has significant advantage over binarization since validation can use probabilistic statistical analysis rather than simple volume comparison.

4 Discussion and Conclusion

We have presented a method for generating synthetic multi-modal MR images with brain tumors that present similar difficulties as real brain tumor MR images. Using sets of such images with variations of tumor size, location, extent of

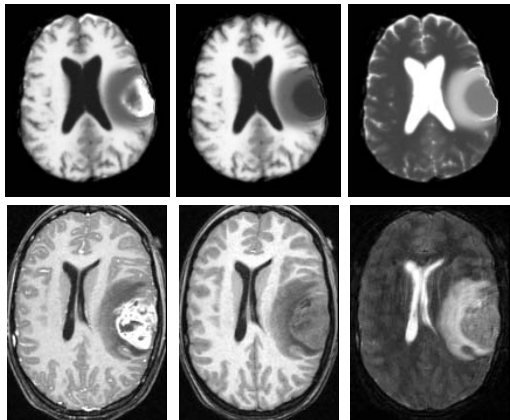


Fig. 4. The synthetic MR images compared to real MR images of a subject with meningioma and surrounding edema. Top: axial view of the synthetic MR images generated using our method. Bottom: axial view of real MR images. From left to right: contrast enhanced T1w, T1w, and T2w images.

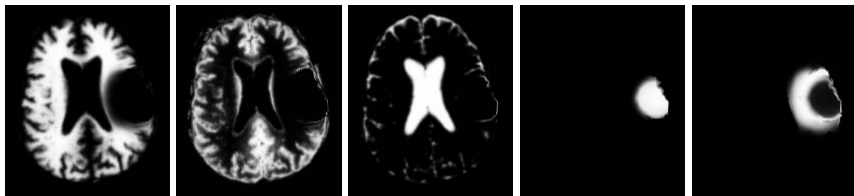


Fig. 5. Axial view of the ground truth for the synthetic MR images. From left to right: the class probabilities for white matter, gray matter, csf, tumor, and edema.

surrounding edema, and enhancing regions, segmentation methods can be tested on images that include most of the challenges for segmentation. The synthetic MRI and the ground truth allows for the validation of the segmentation of the whole brain, which includes white matter, gray matter, csf, tumor, and edema. This capability is novel as most validations done so far were focused on tumor only but not on infiltrated tissue and on deformed healthy tissue.

A possible extension to the method we proposed is the inclusion of vessel information to determine additional regions where contrast agent tend to accumulate. Blood vessel information can also be combined together with deformation and infiltration to generate more precise simulation of the tumor growth and the development of necrosis. This could lead to development of a texture model for the tumor and edema regions.

It is important to note that our goal is to generate sufficiently realistic MR images that are difficult to segment. The accurate modeling of tumor and edema growth is beyond the scope of our current work. Such an effort would require the modeling of the complex interactions between the deformation process, the

infiltration process, and the development of blood vessels [13]. Here, we focus on the generation of test images that empirically simulate pathology as seen in real images, with the main purpose to use simulated images and ground truth for validation and cross-comparison.

The method presented here may also be applied for multi-focal lesions, for example in cases of vascular strokes or multiples sclerosis. Such cases generally present multiple regions with small scale local deformation and tissue infiltration, which can be generated using our framework.

References

1. Clark, M.C., Hall, L.O., Goldgof, D.B., Velthuizen, R., Murtagh, F.R., Silbiger, M.: Automatic tumor-segmentation using knowledge-based techniques. *IEEE TMI* **117** (1998) 187–201
2. Kaus, M., Warfield, S., Nabavi, A., Black, P.M., Jolesz, F.A., Kikinis, R.: Automated segmentation of mr images of brain tumors. *Radiology* **218** (2001) 586–591
3. Prastawa, M., Bullitt, E., Ho, S., Gerig, G.: A brain tumor segmentation framework based on outlier detection. *MedIA* **8** (2004) 275–283
4. Warfield, S.K., Zou, K.H., Wells, W.M.: Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *IEEE TMI* **23** (2004) 903–921
5. Rexilius, J., Hahn, H.K., Schlüter, M., Kohle, S., Bourquain, H., Böttcher, J., Peitgen, H.O.: A framework for the generation of realistic brain tumor phantoms and applications. In: Proc. MICCAI 2004. LNCS 3217 (2004) 243–250
6. Ferrant, M., Nabavi, A., Macq, B., Jolesz, F.A., Kikinis, R., Warfield, S.K.: Registration of 3D interoperative MR images of the brain using finite element biomechanical model. *IEEE TMI* **20** (2001) 1384–1397
7. Clatz, O., Bondiau, P.Y., Hervé Delingette Grégoire Malandain, Maxme Sermesant, S.K.W., Ayache, N.: In silico tumor growth: Application to glioblastomas. In: Proc. MICCAI 2004. Volume LNCS 3217. (2004) 337–345
8. Clatz, O., Sermesant, M., Bondiau, P.Y., Delingette, H., Warfield, S.K., Malandain, G., Ayache, N.: Realistic simulation of the 3d growth of brain tumors in mr images including diffusion and mass effect. *IEEE Transactions on Medical Imaging* (2005) In print.
9. Lu, S., Ahn, D., Johnson, G., Cha, S.: Peritumoral diffusion tensor imaging of high-grade gliomas and metastatic brain tumors. *AJNR* **24** (2003) 937–941
10. Alexander, D.C., Pierpaoli, C., Basser, P.J., Gee, J.C.: Spatial transformations of diffusion tensor magnetic resonance images. *IEEE TMI* **20** (2001) 1131–1139
11. Fletcher, P.T., Joshi, S.: Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In: CVAMIA 2004. LNCS 3117 (2004) 87–98
12. Cocosco, C.A., Kollokian, V., Kwan, R.S., Evans, A.C.: BrainWeb: Online interface to a 3D MRI simulated brain database. *Neuroimage* **5** (1997)
13. Zheng, X., Wise, S.M., Cristini, V.: Nonlinear simulation of tumor necrosis, neo-vascularization and tissue invasion via and adaptive finite-element/level-set method. *Bulletin of Mathematical Biology* **67** (2005) 211–259

Automatic Cerebrovascular Segmentation by Accurate Probabilistic Modeling of TOF-MRA Images

Ayman El-Baz¹, Aly A. Farag¹, Georgy Gimel'farb², and Stephen G. Hushek³

¹ CVIP Laboratory, University of Louisville, Louisville, KY 40292, USA
{elbaz, farag}@cvip.louisville.edu
<http://www.cvip.louisville.edu>

² Department of Computer Science, Tamaki Campus,
University of Auckland, Auckland, New Zealand

³ Technical Director, iMRI Department, Norton Hospital, Louisville, KY, USA

Abstract. Accurate automatic extraction of a 3D cerebrovascular system from images obtained by time-of-flight (TOF) or phase contrast (PC) magnetic resonance angiography (MRA) is a challenging segmentation problem due to small size objects of interest (blood vessels) in each 2D MRA slice and complex surrounding anatomical structures, e.g. fat, bones, or grey and white brain matter. We show that due to a multi-modal nature of MRA data blood vessels can be accurately separated from background in each slice by a voxel-wise classification based on precisely identified probability models of voxel intensities. To identify the models, an empirical marginal probability distribution of intensities is closely approximated with a linear combination of discrete Gaussians (LCDG) with alternate signs, and we modify the conventional Expectation-Maximization (EM) algorithm to deal with the LCDG. To validate the accuracy of our algorithm, a special 3D geometrical phantom motivated by statistical analysis of the MRA-TOF data is designed. Experiments with both the phantom and 50 real data sets confirm high accuracy of the proposed approach.

1 Introduction

Accurate cerebrovascular segmentation is of prime importance for early diagnostics and timely endovascular treatment. Unless detected at early stage, serious vascular diseases like carotid stenosis, aneurysm, and vascular malformation may cause not only severe headaches but also a brain stroke or a life-threatening coma [1]. Non-invasive MRA is a valuable tool in preoperative evaluation of suspected intracranial vascular diseases. Three commonly used MRA techniques are TOF-MRA, phase contrast angiography (PCA), and contrast enhanced MRA (CE-MRA). Both TOF-MRA and PCA use flowing blood as an inherent contrast medium, while for CE-MRA a contrasting substance is injected into the circulatory system.

A variety of today's most popular techniques for segmenting blood vessels from TOF-MRA data can be roughly classified into deformable and statistical models.

The former methods iteratively deform an initial boundary surface of blood vessels in order to optimize an energy function which depends on image gradient information and surface smoothness [2]. Topologically adaptable surfaces make classical deformable models more efficient for segmenting intracranial vasculature [3]. Geodesic active contours implemented with level set techniques offer flexible topological adaptability to segment MRA images [4] including more efficient adaptation to local geometric structures represented e.g. by tensor eigenvalues [5]. Fast segmentation of blood vessel surfaces is obtained by inflating a 3D balloon with fast marching methods [6]. Two-step segmentation of a 3D vascular tree from CTA data sets in [7] is first carried out locally in a small volume of interest. Then a global topology is estimated to initialize a new volume of interest. A multi-scale geometrical flow is proposed in [8] to segment vascular tree from MRI images.

The statistical approach extracts the vascular tree automatically, but its accuracy depends on underlying probability data models. The TOF-MRA image is multi-modal in that signals in each region-of-interest (e.g. blood vessels, brain tissues, etc) are associated with a particular mode of the total marginal probability distribution of signals. To the best of our knowledge, up-to-now there exists only one adaptive statistical approach for extracting blood vessels from the TOF-MRA data proposed by Wilson and Noble [9]. They model the marginal data distribution with a mixture of two Gaussian and one uniform components for the stationary CSF, brain tissues, and arteries, respectively. To identify the mixture (i.e. estimate all its parameters), they use a conventional EM algorithm. Furthermore, a region-based deformable contour for segmenting tubular structures was derived in [10] by combining signal statistics and shape information.

2 LCDG-Model of a Multi-modal TOF-MRA Image

Let $q; q \in \mathbf{Q} = \{0, 1, \dots, Q-1\}$, denote the Q -ary gray level. The discrete Gaussian (DG) is defined as the probability distribution $\Psi_\theta = (\psi(q|\theta) : q \in \mathbf{Q})$ on \mathbf{Q} such that $\psi(q|\theta) = \Phi_\theta(q+0.5) - \Phi_\theta(q-0.5)$ for $q = 1, \dots, Q-2$, $\psi(0|\theta) = \Phi_\theta(0.5)$, $\psi(Q-1|\theta) = 1 - \Phi_\theta(Q-1.5)$ where $\Phi_\theta(q)$ is the cumulative Gaussian probability function with a shorthand notation $\theta = (\mu, \sigma^2)$ for its mean, μ , and variance, σ^2 .

We assume the number K of dominant modes, i.e. regions, objects, or classes of interest in a given TOF-MRA image, is already known. In contrast to a conventional mixture of Gaussians and/or other simple distributions, one per region, we closely approximate the empirical gray level distribution for a TOF-MRA image with an LCDG having C_p positive and C_n negative components such that $C_p \geq K$:

$$p_{\mathbf{w}, \Theta}(q) = \sum_{r=1}^{C_p} w_{p,r} \psi(q|\theta_{p,r}) - \sum_{l=1}^{C_n} w_{n,l} \psi(q|\theta_{n,l}) \quad (1)$$

under the obvious restrictions on the weights $\mathbf{w} = [w_{p,\cdot}, w_{n,\cdot}]$: all the weights are non-negative and

$$\sum_{r=1}^{C_p} w_{p,r} - \sum_{l=1}^{C_n} w_{n,l} = 1 \quad (2)$$

To identify the LCDG-model including the numbers of its positive and negative components, we modify the conventional Expectation-Maximization (EM) algorithm to deal with the LCDG.

First the numbers $C_p - K$, C_n and parameters \mathbf{w} , Θ (weights, means, and variances) of the positive and negative DG components are estimated with a sequential EM-based initializing algorithm. The goal is to produce a close initial LCDG-approximation of the empirical distribution. Then under the fixed C_p and C_n , all other model parameters are refined with an EM algorithm that modifies the conventional one in [11] to account for the components with alternating signs.

2.1 Sequential EM-Based Initialization

Sequential EM-based initialization forms an LCDG-approximation of a given empirical marginal gray level distribution using the conventional EM-algorithm [11] adapted to the DGs. At the first stage, the empirical distribution is represented with a mixture of K positive DGs, each dominant mode being roughly approximated with a single DG. At the second stage, deviations of the empirical distribution from the dominant K -component mixture are modeled with other, “subordinate” components of the LCDG. The resulting initial LCDG has K dominant weights, say, $w_{p,1}, \dots, w_{p,K}$ such that $\sum_{r=1}^K w_{p,r} = 1$, and a number of subordinate weights of smaller values such that $\sum_{r=K+1}^{C_p} w_{p,r} - \sum_{l=1}^{C_n} w_{n,l} = 0$.

The subordinate components are determined as follows. The positive and negative deviations of the empirical distribution from the dominant mixture are separated and scaled up to form two new “empirical distributions”. The same conventional EM algorithm is iteratively exploited to find the subordinate mixtures of positive or negative DGs that approximate best the scaled-up positive or negative deviations, respectively. The sizes $C_p - K$ and C_n of these mixtures are found by sequential minimization of the total absolute error between each scaled-up deviation and its mixture model by the number of the components. Then the obtained positive and negative subordinate models are scaled down and then added to the dominant mixture yielding the initial LCDG model of the size $C = C_p + C_n$.

2.2 Modified EM Algorithm for LCDG

Modified EM algorithm for LCDG maximizes the log-likelihood of the empirical data by the model parameters assuming statistically independent signals:

$$L(\mathbf{w}, \Theta) = \sum_{q \in \mathbf{Q}} f(q) \log p_{\mathbf{w}, \Theta}(q) \quad (3)$$

A local maximum of the log-likelihood in Eq. (3) is given with the EM process extending the one in [11] onto alternating signs of the components. Let $p_{\mathbf{w}, \Theta}^{[m]}(q) = \sum_{r=1}^{C_p} w_{p,r}^{[m]} \psi(q|\theta_{p,r}^{[m]}) - \sum_{l=1}^{C_n} w_{n,l}^{[m]} \psi(q|\theta_{n,l}^{[m]})$ denote the current LCDG at iteration m . Relative contributions of each signal $q \in \mathbf{Q}$ to each positive and negative DG at iteration m are specified by the respective conditional weights

$$\pi_p^{[m]}(r|q) = \frac{w_{p,r}^{[m]} \psi(q|\theta_{p,r}^{[m]})}{p_{\mathbf{w},\Theta}^{[m]}(q)}; \quad \pi_n^{[m]}(l|q) = \frac{w_{n,l}^{[m]} \psi(q|\theta_{n,l}^{[m]})}{p_{\mathbf{w},\Theta}^{[m]}(q)} \quad (4)$$

such that the following constraints hold:

$$\sum_{r=1}^{C_p} \pi_p^{[m]}(r|q) - \sum_{l=1}^{C_n} \pi_n^{[m]}(l|q) = 1; \quad q = 0, \dots, Q-1 \quad (5)$$

The following two steps iterate until the log-likelihood changes become small:

- E-step**^[m+1]: Find the weights of Eq. (4) under the fixed parameters $\mathbf{w}^{[m]}$, $\Theta^{[m]}$ from the previous iteration m , and
- M-step**^[m+1]: Find conditional MLEs $\mathbf{w}^{[m+1]}$, $\Theta^{[m+1]}$ by maximizing $L(\mathbf{w}, \Theta)$ under the fixed weights of Eq. (4).

Considerations closely similar to those in [11] show this process converges to a local log-likelihood maximum. Let the log-likelihood of Eq. (3) be rewritten in the equivalent form with the constraints of Eq. (5) as unit factors:

$$L(\mathbf{w}^{[m]}, \Theta^{[m]}) = \sum_{q=0}^Q f(q) \left[\sum_{r=1}^{C_p} \pi_p^{[m]}(r|q) \log p^{[m]}(q) - \sum_{l=1}^{C_n} \pi_n^{[m]}(l|q) \log p^{[m]}(q) \right] \quad (6)$$

Let the terms $\log p^{[m]}(q)$ in the first and second brackets be replaced with the equal terms $\log w_{p,r}^{[m]} + \log \psi(q|\theta_{p,r}^{[m]}) - \log \pi_p^{[m]}(r|q)$ and $\log w_{n,l}^{[m]} + \log \psi(q|\theta_{n,l}^{[m]}) - \log \pi_n^{[m]}(l|q)$, respectively, which follow from Eq. (4). At the E-step, the conditional Lagrange maximization of the log-likelihood of Eq. (6) under the Q restrictions of Eq. (5) results just in the weights $\pi_p^{[m+1]}(r|q)$ and $\pi_n^{[m+1]}(l|q)$ of Eq. (4) for all $r = 1, \dots, C_p$; $l = 1, \dots, C_n$ and $q \in \mathbf{Q}$. At the M-step, the DG weights $w_{p,r}^{[m+1]} = \sum_{q \in \mathbf{Q}} f(q) \pi_p^{[m+1]}(r|q)$ and $w_{n,l}^{[m+1]} = \sum_{q \in \mathbf{Q}} f(q) \pi_n^{[m+1]}(l|q)$ follow from the conditional Lagrange maximization of the log-likelihood in Eq. (6) under the restriction of Eq. (2) and the fixed conditional weights of Eq. (4). Under these latter, the conventional MLEs of the parameters of each DG stem from maximizing the log-likelihood after each difference of the cumulative Gaussians is replaced with its close approximation with the Gaussian density (below ‘‘c’’ stands for ‘‘p’’ or ‘‘n’’, respectively):

$$\mu_{c,r}^{[m+1]} = \frac{1}{w_{c,r}^{[m+1]}} \sum_{q \in \mathbf{Q}} q \cdot f(q) \pi_c^{[m+1]}(r|q)$$

$$(\sigma_{c,r}^{[m+1]})^2 = \frac{1}{w_{c,r}^{[m+1]}} \sum_{q \in \mathbf{Q}} (q - \mu_{c,i}^{[m+1]})^2 \cdot f(q) \pi_c^{[m+1]}(r|q)$$

This modified EM-algorithm is valid until the weights \mathbf{w} are strictly positive. The iterations should be terminated when the log-likelihood of Eq. (3) does not change or begins to decrease due to accumulation of rounding errors.

The final mixed LCDG-model $p_C(q)$ is partitioned into the K LCDG-submodels $P_{[k]} = [p(q|k) : q \in \mathbf{Q}]$, one per class $k = 1, \dots, K$, by associating the subordinate DGs with the dominant terms so that the misclassification rate is minimal.

3 Experimental Results

Experiments were conducted with the TOF-MRA images acquired with the Picker 1.5T Edge MRI scanner having spatial resolution of $0.43 \times 0.43 \times 1.0$ mm. The size of each 3D data set is $512 \times 512 \times 93$. The TOF-MRA images contain three classes ($K = 3$), namely, darker bones and fat, brain tissues, and brighter blood vessels. A typical TOF-MRA slice, its empirical marginal gray level distribution $f(q)$, and the initial 3-component Gaussian dominant mixture $p_3(q)$ are shown in Fig. 1. Figure 2 illustrates basic stages of our sequential EM-

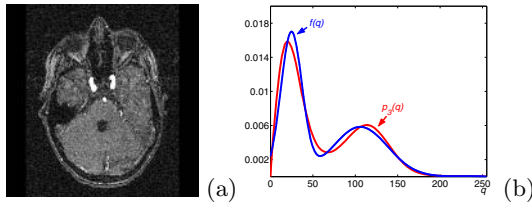


Fig. 1. Typical TOF-MRA scan slice (a) and deviations between the empirical distribution $f(q)$ and the dominant 3-component mixture $p_3(q)$ (b)

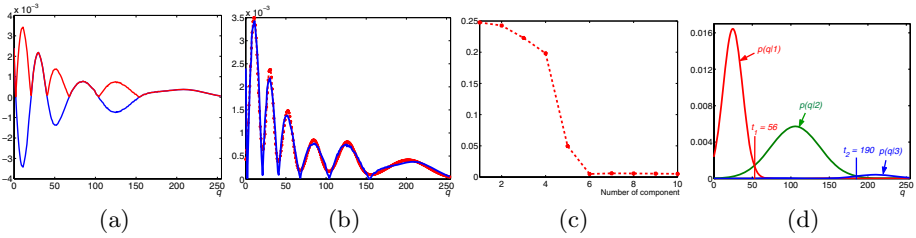


Fig. 2. Deviations and absolute deviations between $f(q)$ and $p_3(q)$ (a), the mixture model (b) of the absolute deviations in (a), the absolute error (c) as a function of the number of Gaussians approximating the scaled-up absolute deviations in (a), and the initial estimated LCDG-models for each class (d)

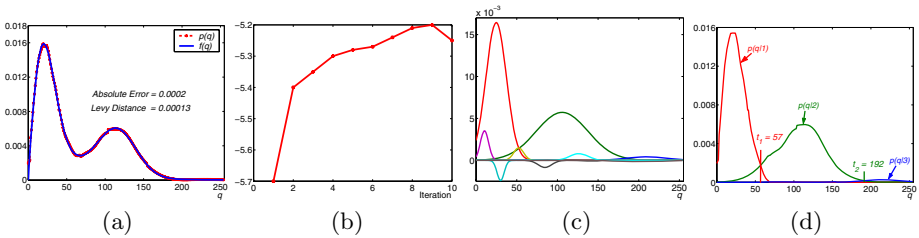


Fig. 3. Final 3-class LCDG-model overlaying the empirical density (a), the log-likelihood dynamics (b) for the refining EM-iterations, the refined model components (c), and the class LCDG-models (d)

based initialization by showing the scaled-up alternating and absolute deviations $f(q) - p_3(q)$, the best mixture model estimated for the absolute deviations (these six Gaussian components give the minimum approximation error), and the initial LCDG-models for each class. The scaling makes the sums of the positive or absolute negative deviations for $q = 0, \dots, Q - 1$ equal to one. Figure 3

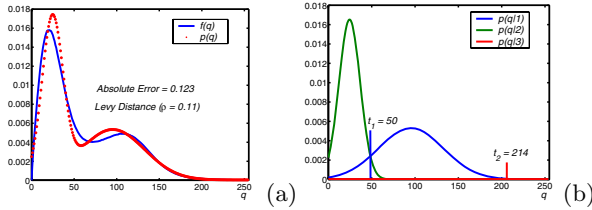


Fig. 4. Wilson-Noble's model [9]: the estimated distribution (a) and the class models (b)

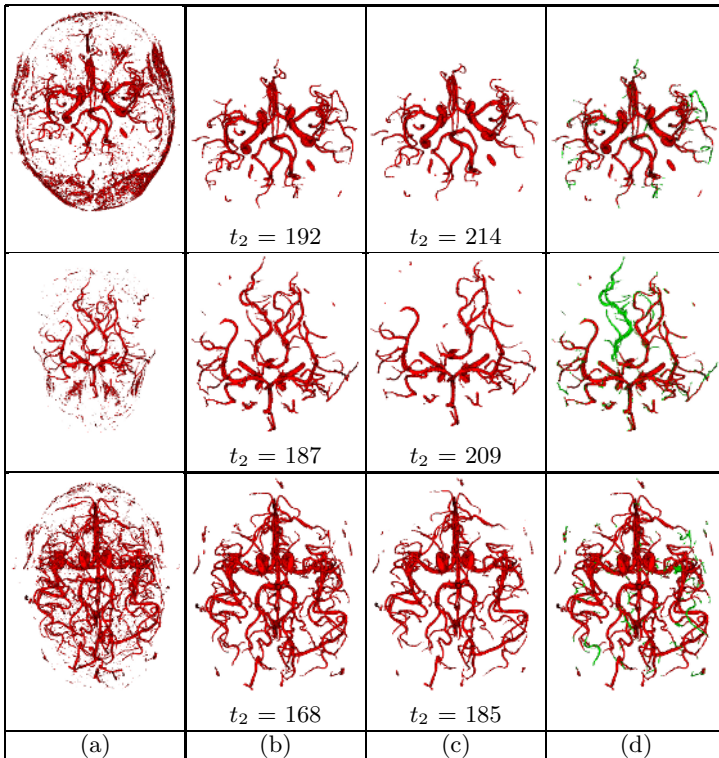


Fig. 5. Each row relates to one patient: our segmentation before (a) and after (b) noise and small fat voxels are eliminated with the connectivity filter, the Wilson-Noble's segmentation (c) after the connectivity filter, and the differences between (b) and (c); the green voxels are missed by the Wilson-Noble's approach and the red ones are detected by the both approaches

presents the final LCDG-model after refining the initial one with the modified EM-algorithm and shows successive changes of the log-likelihood at the refinement iterations. The final LCDG-models of each class are obtained with the best separation thresholds $t_1 = 57$ and $t_2 = 192$. First nine refining iterations increase the log-likelihood from -5.7 to -5.2 .

To highlight the advantages of our approach over the existing one, Fig. 4 shows results obtained with the model of Wilson and Noble [9]. To measure the estimation quality, we use the Levy distance between two distributions [12] and the absolute error. The Levy distance between the empirical distribution and its estimated model is 0.11 and 0.00013 and the absolute errors are 0.123 and 0.0002 for the Wilson-Noble's and our approach, respectively. The larger Levy distance and absolute error indicate the notably worse approximation which strongly affects the accuracy of separating the blood vessels from the background. Because of a typically higher separation threshold, e.g. $t_2 = 214$ versus our $t_2 = 192$ in this particular example, the Wilson-Noble's approach misses some blood vessels, as shown in Fig. 5.

Both the approaches have been compared on 50 data sets. Results of the three tests are depicted in Fig. 5. As the first column, (a), suggests, TOF-MRA is sensitive to tissues like subcutaneous fat with a short T1 response that may obscure the blood vessels in the segmented volume. To eliminate them, the volume is processed with an automatic connectivity filter which selects the largest connected tree structures using a 3D volume growing algorithm [13]. The results after applying such a filter to our and Wilson-Noble's segmentation in Fig. 5 show that the latter approach fails to detect sizeable fractions of the vascular trees which are validated by the expert (radiologist) that the green parts which are detected by our approaches follow the topology of the brain vascular tree.

4 Validation

It is very difficult to get accurate manually segmented complete vasculatures to validate our algorithm. Thus to evaluate its performance, we have created a wooden phantom shown in Fig. 6(a) with topology similar to the blood vessels. Furthermore, the phantom mimics bifurcations, zero and high curvature that exist in any vasculature system, and it has a varying radius to simulate both large and small blood vessels. The phantom was scanned by CT and then manually segmented to obtain the ground truth. The blood vessel and non-vessel signals for the phantom are generated according to the distribution $p(q|3)$ and $p(q|1)$, $p(q|2)$, respectively, in Fig. 3(d) using the inverse mapping methods. The resulting phantom's histogram was similar to that in Fig. 3(a).

Let the total segmentation error be a percentage of erroneous voxels with respect to the overall number of voxels in the manually segmented 3D phantom. Figure 6 shows our approach is 15 times more accurate than the Wilson-Noble's one (the total errors 0.31% and 4.64%, respectively). The error constituents per each 2D slice for both the approaches are also plotted in Fig. 6.

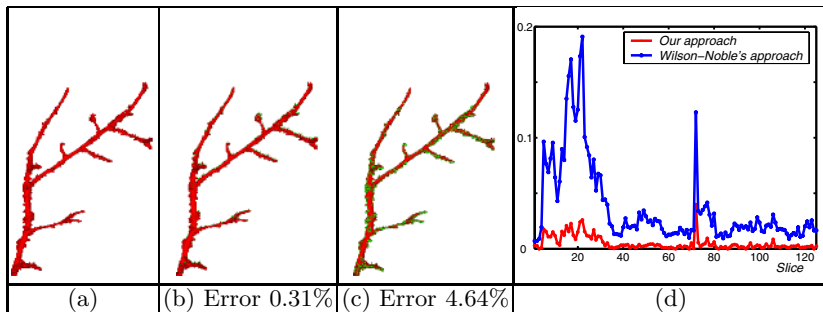


Fig. 6. The 3D geometrical phantom (a), our (b) and Wilson-Noble's (c) segmentation, and total errors per each phantom's slice for both the approaches (d)

5 Conclusions

We presented a new statistical approach to find blood vessels in multi-modal TOF-MRA images. The LCDG-model accurately approximates the empirical marginal gray level distribution yielding the high quality segmentation. The accuracy of our approach is validated using a specially designed 3D geometrical phantom.

Our present implementation on C++ programming language using a single 2.4 GHZ Pentium 4 CPU with 512 MB RAM takes about 49 *sec* for 93 TOF-MRA slices of size 512x512 pixels each.

The LCDG-model is also suitable for segmenting PC-MRA and CTA images which are not presented due to space limitations.

References

1. Health Resources. Patient Resources, *Disorder of the Month. Cerebrovascular Disease. Neurosurgery: On-Call* [serial online], July 1999.
2. V. Caselles, R. Kimmel, and G. Sapiro., "Geodesic active contours," *Int. J. Computer Vision*, vol. 22, pp. 61–79, 1997.
3. T. McInerney and D. Terzopoulos, "Medical image segmentation using topologically adaptable surface", *Proc. CVRMed-MRCAS'97*, pp. 23–32, 1997
4. L. M. Lorigo, O. D. Faugeras, W. E. L. Grimson, and R. Keriven, "Curves: Curve evolution for vessel segmentation", *Medical Image Analysis*, vol. 5, pp. 195–206, 2001.
5. O. Wink, W. J. Niessen, and M. A. Viergever, "Fast delineation and visualization of vessels in 3-D angiographic images," *IEEE Trans. Med. Imaging*, vol. 19, pp. 337–346, 2000.
6. T. Deschamps and L. D. Cohen, "Fast extraction of tubular and tree 3D surfaces with front propagation methods", *Proc. 16th ICPR*, pp. 731–734, 2002.
7. R. Manniesing and W. Niessen, "Local speed functions in level set based vessel segmentation", *Proc. MICCAI'04*, pp. 475–482, 2004.

8. M. Descotesux, L. Collins, and K. Siddiqi, "Geometric flows for segmenting vasculature in MRI: Theory and validation", *Proc. MICCAI'04*, pp. 500–507, 2004.
9. D. L. Wilson and J. A. Noble, "An adaptive segmentation algorithm for time-of-flight MRA data", *IEEE Trans. Med. Imaging*, vol. 18, pp. 938–945, 1999.
10. D. Nain, A. Yezzi, and G. Turk, "Vessels segmentation using a shape driven flow", *Proc. MICCAI'04*, pp. 51–59, 2004.
11. M. I. Schlesinger and V. Hlavac, *Ten Lectures on Statistical and Structural Pattern Recognition*, Kluwer Academic, Dordrecht, 2002.
12. J. W. Lamperti, *Probability*, J. Wiley & Sons, New York, 1996.
13. A. A. Farag, A. El-Baz, and G. Gimel'farb, "Density estimation using modified Expectation Maximization for a linear combination of Gaussians", *Proc. IEEE Int. Conf. Image Processing, Oct. 24–27, 2004, Singapore*, IEEE CS Press: Los Alamitos, vol. 1, pp. 194–197, 2004.

A Segmentation and Reconstruction Technique for 3D Vascular Structures

Vincent Luboz¹, Xunlei Wu¹, Karl Krissian², Carl-Fredrik Westin², Ron Kikinis², Stéphane Cotin¹, and Steve Dawson¹

¹ The SIM Group – CIMIT/ MGH, 65 Landsdowne Street, Cambridge, MA 02139, USA
{vluboz, xwu4, scotin, sdawson}@partners.org
<http://www.medicalsim.org/index.htm>

² SPL, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115, USA
{karl, westin, kikinis}@bwh.harvard.edu
<http://splweb.bwh.harvard.edu:8000/>

Abstract. In the context of stroke therapy simulation, a method for the segmentation and reconstruction of human vasculature is presented and evaluated. Based on CTA scans, semi-automatic tools have been developed to reduce dataset noise, to segment using active contours, to extract the skeleton, to estimate the vessel radii and to reconstruct the associated surface. The robustness and accuracy of our technique are evaluated on a vascular phantom scanned in different orientations. The reconstructed surface is compared to a surface generated by marching cubes followed by decimation and smoothing. Experiments show that the proposed technique reaches a good balance in terms of smoothness, number of triangles, and distance error. The reconstructed surface is suitable for real-time simulation, interactive navigation and visualization.

1 Introduction

Stroke is a leading cause of death. Our team is developing a real-time neuro-interventional radiology simulation system where physicians will be able to learn and practice without putting patients at risk. This kind of application requires a streamlined data processing from a patient's computer tomography angiogram (CTA) to a computer representation of the vasculature. Therefore, the reconstructed vascular network has to be *smooth* for visualization, *structured* for blood flow computation, and *efficient* for real time collision detection/collision response between interventional tools and vessel wall. The method presented in this paper generates virtual vasculature through segmentation and surface reconstruction.

Current techniques for processing vascular images can be divided in two main approaches: techniques for centerline enhancement, including multi-scale approaches, usually based on the Hessian matrix; and techniques for contour extraction, including statistical approaches: Expectation Maximization [1], random Markov fields, and geometrical approaches: region growing, adaptive thresholding, active contours that can be explicit, like snakes, or implicit, like level sets [2, 3]. These techniques usually perform better after noise reduction. A topological representation of the vascular network can be obtained from both approaches either by computing ridges or by applying a thinning technique like homotopic skeletonization.

The result of the segmentation is then processed to generate an efficient and structured representation of vascular structure for our purpose. Bühler et al. [4] presents a comprehensive and up-to-date survey on surface reconstruction techniques. Our surface reconstruction algorithm improves upon [5] in that the coarse base mesh generation is followed by smoothing through surface subdivision. The base mesh and multi-scale subdivision approach, as shown in subsequent sections, is proved to be a robust framework for real-time physics-based flow computation, smooth tissue/tool interaction, as well as high-fidelity anatomical visualization.

Section 2 will describe our semi-automatic tools which reduce dataset noise, segment using active contours, compute the skeleton, estimate the vessel radii and reconstruct the associated surface. In section 3, we present different tests performed on the phantom and used to evaluate the robustness and the accuracy of our method. Finally, discussion and conclusion are presented in the last section.

2 Materials and Method

Our process consists in the following pipeline: anisotropic diffusion, level set evolution, skeletonization, pruning, connection and smoothing, radius estimation, reconstruction. This approach preserves the topology at junctions and gives semi-automatically a skeleton requiring little interaction before surface reconstruction.

2.1 Segmentation

The first step of the segmentation is to apply an anisotropic diffusion filter based on [6]. This filter reduces the noise while preserving small vascular structures enabling better segmentation which is important in the region of the brain. Next, we remove the skull bones, the sinuses and the skin, who the similar intensity as the vessels and might disturb the segmentation process, using morphological operations.

We then segment the vessel contours by the means of a level set evolution. For efficiency, we initialize the active contour using a threshold on the image intensity. The level set equation [7, 2], evolves a surface according to three different forces: an advection force that pushes the surface towards the edges of the image, a smoothing term that keeps the surface smooth and a balloon force that allows expansion of the contours within vascular structures. The smoothing term is proportional to the minimal curvature of the surface [8]. The balloon force relies on the intensity statistics to either expand or shrink the evolving contour. This force is expressed as $\exp(-(I-m)^2/\sigma^2)-\tau$ where I is the intensity, m stands for the mean intensity of the vessels, σ is their standard deviation, and τ is a threshold (0.2 by default) that allows shrinking the contour when the position is unlikely to belong to a vessel. A 3D model of the vessels is obtained as the iso-surface of intensity zero from the result of the level set evolution, using the Marching-Cubes algorithm [9]. From the result of the level set, a skeletonization technique is applied to obtain a simple topological representation of the vascular network. It is based on homotopic thinning where voxels are removed in the order of the Euclidean distance to the segmented surface. Voxels are iteratively removed if they are simple [10] and if they are not end-points, such that they have more than one neighbor in a 3x3x3 neighborhood.

After the skeletonization, small vessels can still have connectivity discrepancies along the centerlines near small branches. Consequently, a pruning is applied to remove small leaves (lines with at least an extremity which is not a junction). We then connect the lines that are close but disconnected because of the resolution of the medical dataset. This connection is made only if the lines are close to each other and their directions are matching within a small difference, which could be induced by the vessel curvature. At this step, some manual work is often needed. This work consists in connecting lines that are too far to be connected automatically or in removing lines that are too long to be deleted.

Once the connected skeleton is finished, the radius of the lines is extracted from the dataset. This is done using the intensity gradient in the binary image obtained from the level set evolution. Starting from the centerline, we grow a circle in the plan of the cross-section and stop when there is a relevant local maximum of the intensity gradient, thus giving estimated radii along the centerlines.

2.2 Surface Reconstruction

The goal of our surface reconstruction is to generate a smooth surface that can be easily refined to suit the needs of efficient collision detection/response, stable vessel deformation, real-time flow simulation, also multi-scale anatomical visualization. Our algorithm improves over [5] in four main areas:

1. Handling directed graphs with loops and multiple roots.

One branch is allowed to have multiple parents and children. Artery vessels can form loops, e.g. circle of Willis. *One branch can connect to a single branch forming 1-furcation.* This is useful to construct a unified directed graph for both artery and venous sides. *Multiple trees can be reconstructed at the same time.*

2. Trunk branch selection based on angle and radii variance.



Fig. 1. *Left:* Trunk branch selection: using both vessel average radii and branching angle to determine the continuation trunk branch. Although $\theta_j < \theta_i$, B_1^in is chosen as the trunk branch of B_0^in , due to the similarity of their average radii. *Middle:* Cross section distribution gets denser at thinner regions of a vessel. *Right:* the density is higher where a vessel turns or twists.

To patch the surface at vessel joints, both algorithms define at a trunk branch with respect to the current branch and form polygons to connect the trunk surface and other joint branches base meshes. Since n_i , the cross section normal at the beginning or end of branch B_i , is computed by differentiating neighboring sampling points, the approximation can be misleading when centerlines are under sampled. Our scheme considers both branching angle and vessel radii to reduce under-sampling artifacts which improves the reconstruction robustness. First, n_i^in where $i > 0$ are reversed. Then, we com-

pute the disparity $\Omega_i \equiv \lambda \theta_{i+1} + (1-\lambda) |r_i - r^{in}_o|$, where $\lambda \in [0, 1]$ is the weight balancing the influence of branching angle and of the average radii variance. The algorithm picks the branch with minimal Ω as the trunk branch. In the left half of Fig. 1, although $\theta_j < \theta_s$, B_j^{in} is chosen due to the similarity of their average radii.

3. Adaptive cross sections distribution.

Our cross section distribution scheme (2) considers radii and centerline curvature:

$$x_{i+1} - x_i = \alpha \left(\frac{r_{i+1}}{1 + \beta \kappa_{i+1}} + \frac{r_i}{1 + \beta \kappa_i} \right) \quad i \in [0, N_{segment} - 1] \quad (2)$$

where x_i is the curvilinear coordinate of the cross section center. r_i and κ_i [11] are the corresponding radius and Gaussian curvature, respectively, obtained by linear interpolation between two adjacent raw skeleton samples where $\alpha > 0$ is the desired spacing scalar and $\beta > 0$ is the weight on curvature influence. Eq. (2) states that after filtering, the centers of two adjacent cross sections are placed closer if the vessel is thin or turns. A straight branch does not need many cross sections to resemble its original geometry. Assembling (2) for all i yields $(N_{seg}-1)$ nonlinear algebraic equations with $(N_{seg}-1)$ unknowns, since x_0 and x_N are set to be the curvilinear coordinates of the vessel end nodes. Broyden's method [12] is used to solve for all x_i .

4. Robust joint tiling: *end-segment-grouping* and *adjacent-quadrant-grouping*.

We connect every branch to its trunk using both end segments regardless the branching angles so that a single recursive joint tiling is needed. *End-segment-grouping*, unifies all the outgoing branches together such that the *connecting patches* connect the bottom of the outgoing branch's base mesh with both end segments of the trunk branches, i.e. $Seg(N-1)$ and $Seg(0)$, demonstrated in the left half of Figure 2.

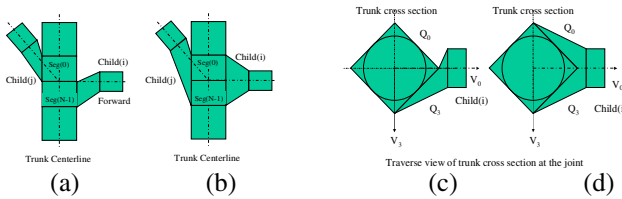


Fig. 2. (a) Felkel's method connects forward $Child(j)$ to $Seg(0)$ and backward $Child(i)$ to $Seg(N-1)$. (b) *End-segment-grouping* connects $Child(i)$ and $Child(j)$ both $Seg(N-1)$ and $Seg(0)$. The bottle-neck effect is reduced. When $Child(i)$ lies close to the boundary of Q_3 , (d) *adjacent-quadrant-grouping* uses both Q_0 and Q_3 eliminates twisting artifact by using only Q_3 in [5] (c).

When the outgoing centerline forms a small angle with the trunk centerline, using single end segment produces bottle-neck effect. The artifact is reduced when both end segments are deployed for the joint tiling. When the outgoing centerline lies in or close to the bisection plane of two trunk centerlines, using a single end segment loses the symmetry. This symmetry is nicely preserved by connecting the mesh of $Child(i)$ to the same sides of $Seg(N-1)$ and $Seg(0)$. *End-segment-grouping* not only reduces the patching artifacts in both extreme cases, but yields smoother trunk-to-branch transition under all branching configuration.

We improve the joint tiling not just in the trunk centerline direction. *Adjacent-quadrant-grouping* is designed to use 2 adjacent sides of the end hexahedron seg-

ments. When a child centerline lies close to the boundary of 2 quadrants, tiling with only one quadrant introduces twists. This artifact is eliminated by adding the neighboring quadrant into the tiling, e.g. Q_0 and Q_3 are grouped together as a whole when tiling $Child(i)$ to the trunk mesh. When $Child(i)$ lies close to a quadrant center, our approach uses only current quadrant for the tiling as in [5].

With these improvements, the proposed reconstruction scheme is able to handle more general directed graph. It is less prone to artifacts due to initial data sampling. It is also more robust to present full range of bifurcation configuration. The reconstructed smooth vascular surface is suitable for the purpose of efficient and stable physics modeling, and smooth visualization.

3 Tests and Evaluation

Our vascular phantom, in Fig. 3(a), is composed of a Plexiglas box filled with silicon gel and nylon tubing forming a simplified vasculature. Vessel radii range from $0.78mm$ (simulating small brain vessels) to $2.34mm$ (simulating the middle cerebral artery). After CTA scans of the phantom, the segmentation leads to a skeleton from which the reconstruction module generates a smooth surface as shown in Fig. 3(b).

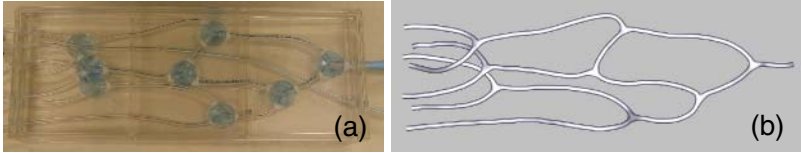


Fig. 3. (a) The silicon phantom with nylon tubing. (b) Reconstruction of the 3D surface

To evaluate the rotational invariance and robustness of our method, we scanned this phantom in 12 different orientations. The phantom orientations are obtained via a 45° or 90° rotation on one or more axes. The CTA scan resolution is $0.6 \times 0.6 \times 1.25$ mm.

3.1 Evaluation of the Robustness of the Vessel Lengths and Radii

The segmentation method, described in section 2, was applied to those datasets with following parameters: 1000 level set iterations; using intensity threshold of 2300 with standard deviation (SD) of 750; 5mm pruning; 2mm maximal distance for gap connection; radius estimation with gradient computed from the derivatives of a Gaussian kernel with SD $0.4mm$. Line orientation was manually corrected before the final surface reconstruction (Fig. 3(b)). To evaluate the results, the lengths and the radii have been analyzed through the Brand-Altman method [13]. Fig. 4 shows that length variation stays within $1.0mm$, while 2σ length = $3.5mm$. In only 6 out of 204 cases (17tubes x 12scans), the radius variation is out of $[-2 \sigma$ radius, 2σ radius], where 2σ radius = $0.2mm$.

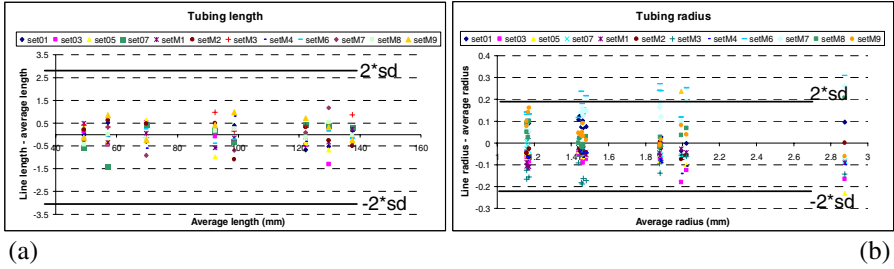


Fig. 4. Brand-Altman plots for (a) the tube lengths and (b) the average radii. The upper and the lower limits represent 2σ . Very few length and radius values are away from their average.

3.2 Evaluation of the Mesh Accuracy and Smoothness

We measure surface smoothness and the distance between two surfaces. *Hausdorff* distance is computed using MESH software¹. Smoothness is measured as the RMS of the minimal and the maximal surface curvatures, κ_1 and κ_2 respectively. They are computed by fitting a 2nd polynomial to each vertex and its direct neighbors: using this small region considers the surface noise in the smoothness measure. The lower the value, the smoother the surface is. Fig. 5(a) depicts the distance between our reconstructed model at 3 subdivision levels, L_0, L_1, L_2 and the surface, S_0 , obtained from the Marching Cubes algorithm applied to the result of the level set segmentation. The RMS is always less than one voxel ($<0.6mm$) and lower than $0.4mm$ on L_1 .

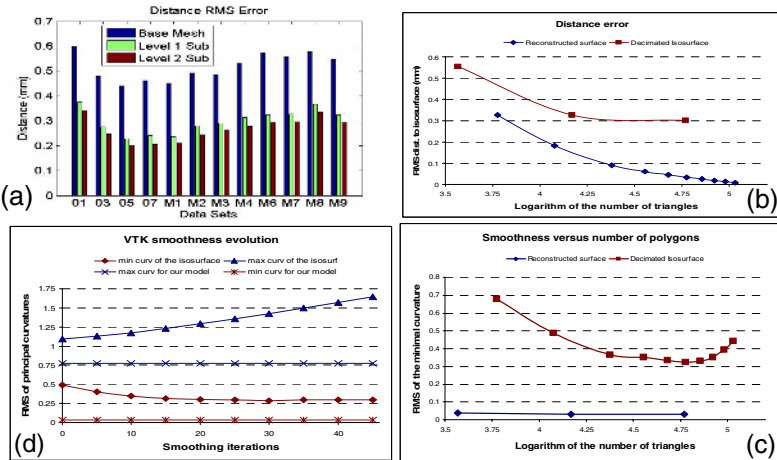


Fig. 5. (a) Root Mean Square (RMS) distance error for 12 data sets at 3 subdivision. (b) RMS distance error on M7 versus the number of triangles, after different decimations of the original iso-surface. (c) RMS of the min curvature, versus the number of triangles. (d) Smoothness evolution for different smoothing levels compared to the smoothness our model at level 1.

¹ <http://mesh.berlios.de>

We compared the smoothness and the distance error on M7 obtained from our reconstruction to the ones obtained using the VTK². Fig. 5(b) shows the distance between S_0 and L_0, L_1, L_2 compared to the distance between S_0 and its decimations $S_0^d, d \in [0, 9]$ using *vtkDecimatePro*. The RMS after decimation is always smaller, because most errors occur at vessel extremities or junctions. However, our model allows simpler mesh, with reasonable error ($RMS < 0.6mm$) and good smoothness shown in Fig. 5(c). This figure displays the RMS of κ_1 on L_0, L_1, L_2 and on S_0^d . It is almost constant (0.03) for all levels and much better than any S_0^d (between 0.3 and 0.7). For completeness, we also depict the RMS of κ_1 and κ_2 for L_1 and S_0^8 having similar number of triangles. Fig. 5(d) plots the evolution of these smoothness measures according to the number of smoothing iterations, where we smooth S_0^8 by applying the *vtkSmoothPolyDataFilter*. The RMS of κ_2 always increases with the smoothing due to vessel shrinkage, and the RMS of κ_1 decreases to 0.25 while L_1 have a value of 0.03. This shows our model smoothness superiority over VTK smoothing.

3.3 Results on a Clinical Dataset

Before evaluating on a full clinical dataset, we apply our method to a portion of a patient vascular network in Fig. 6(a). The testing data contains the end of the vertebral arteries, joining into the basilar artery which then split into the posterior cerebral arteries. A local level set, ignoring the small vessels, followed by an iso-surface reconstruction allowed getting those arteries shown in Fig. 6(b). After skeletonization and surface reconstruction shown in Fig. 6(c), both surfaces are compared. Fig. 6(b) shows the color code of the distance between the two surfaces. The RMS distance error is lower than 0.4mm with 5% of iso-surface triangles. The result of our streamlined process, on a full CTA dataset, is depicted in Fig. 6(d).

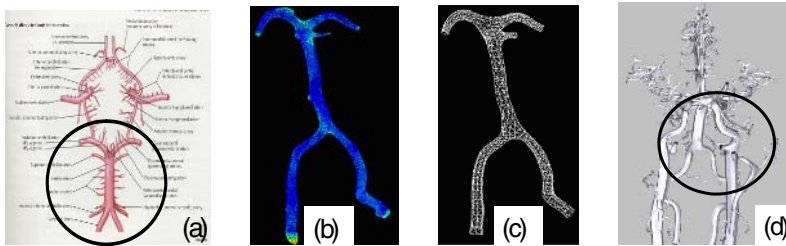


Fig. 6. (a) Anatomy of the circle of Willis; (b) segmented iso-surface; (c) reconstructed surface. The color code in (b) ranges from blue (0.0mm) to red (3.0mm), (d) reconstructed arterial side.

4 Discussion and Conclusion

The method presented in this article deals with the segmentation and the reconstruction of the vascular network. The final reconstructed vascular surface is aimed to be integrated in a neuro vascular training and procedural planning simulator. The goal of our method is to streamline the process from the CTA scan of a patient to a structured,

² Visual Tool Kit library available at <http://www.vtk.org>

smooth, and efficient vascular model with minimum manual interactions. It has shown interesting results in term of accuracy and robustness. Indeed, its evaluation on the phantom in 12 different orientations produces homogeneous skeletons and radii. The generated surfaces are close to the reference ones and are much smoother. The main drawback of our method is that it is not fully automatic and it estimates circular vessel cross sections.

In future work, the main effort will focus on reducing the amount of manual work. On the segmentation side, one main difficulty is to separate tangent vessels, which are merged at the current imaging resolution. Another difficulty is to fully detect small vessels. Both tasks would benefit from an a priori knowledge based on an anatomical atlas. As a perspective, we would like to integrate a labeling tool in the skeletonization step. This feature could give the name of the arteries and veins and consequently help in the training/learning process of our simulator. An automatic correction of centerline orientation is also under investigation based on graph theory. For the estimation of the cross-sections, fitting an ellipse instead of a circle would help to match their real geometry without sacrificing the smoothness and the low complexity of the mesh. Finally, testing the whole method on more patients would help to validate it on large scale networks before integrating it in the neuro-vascular intervention training system mentioned above.

References

1. Wells, W. M. and Grimson, W. E. L. and Kikinis, Ron and Jolesz, Ferenc A. Adaptive Segmentation of MRI Data Springer-Verlag, 1995 , 905 , 59-69.
2. Sethian, J.A. Level Set Methods and Fast Marching Methods: Evolving Interfaces in Comp. Geom., Fluid Mech., Comp. Vision and Materials Sci. Cambridge Univ. Press, 1999.
3. Suri, J.S., Liu, K., Singh, S., Laxminarayan, S.N., Zeng, X., Reden, L.: Shape recovery algorithms using level sets in 2D/3D medical imagery: a state-of-the-art review IEEE Transactions on Information Technology in Biomedicine, 2002 , 6 , 8-28.
4. Bühler, K., Felkel, P., La Cruz, A.: Geometric Methods for Vessel Visualization and Quantification - A Survey. VRVis Research Center, Austria, 2002, Technical Report, pp. 24-48.
5. Felkel, P., Wegenkittl, R., Bühler, K.: Surface Models of Tube Trees. In: Computer Graphics International (CGI'04), pp. 70-77.
6. Krissian, Karl Flux-based anisotropic diffusion applied to enhancement of 3-D angiogram. IEEE Trans Med Imaging, 2002, 21 , 1440-2.
7. Osher, S., Sethian, J.A. Fronts propagating with curvature dependent speed: algorithms based on the Hamilton-Jacobi formalism J. Comput. Physics, 1988 , 79 , 12-49.
8. Lorigo, L.M., Faugeras, O.D., Grimson, W.E.L., Keriven, R., Kikinis, R., Nabavi, A., Westin, C.-F. CURVES: Curve Evolution for Vessel Segm. MedIA, 2001 , 5 , 195-206.
9. Lorensen, W.E., Cline, H.E. Marching Cubes: A high resolution 3-D surface construction algorithm Computer Graphics, 1987, 21, 163-169.
10. Malandain, G., Bertrand, G., Ayache, N. Topological Segmentation of Discrete Surfaces IJCV, 1993 , 10 , 183-197.
11. Calabi, E., Olver, P.J., Shakiban, C., Tannenbaum, A., Haker, S.: Differential and Numerically Invariant Signature Curves Applied to Object Recognition. IJCV, 1998 (26):107-35.
12. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical Recipes in C, University Press, Cambridge, 1992.
13. Bland, J.M., Altman, D.G.: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, i: 307-310, 1986.

MRA Image Segmentation with Capillary Active Contour

Pingkun Yan and Ashraf A. Kassim

Department of Electrical & Computer Engineering, National University of Singapore
{pingkun, ashraf}@nus.edu.sg

Abstract. Precise segmentation of three-dimensional (3D) magnetic resonance angiography (MRA) image can be a very useful computer aided diagnosis (CAD) tool in clinical routines. Our objective is to develop a specific segmentation scheme for accurately extracting vasculature from MRA images. Our proposed algorithm, called the capillary active contour (CAC), models capillary action where liquid can climb along the boundaries of thin tubes. The CAC, which is implemented based on level sets, is able to segment thin vessels and has been applied for verification on synthetic volumetric images and real 3D MRA images. Compared with other state-of-the-art MRA segmentation algorithms, our experiments show that the introduced capillary force can facilitate more accurate segmentation of blood vessels.

1 Introduction

Magnetic resonance angiography (MRA) is a noninvasive medical imaging modality that produces three-dimensional (3D) images of vessels and cavities. Accurate extraction of 3D vascular structures from MRA images has become increasingly important for diagnosis and quantification of vascular diseases.

Existing MRA segmentation techniques can be broadly divided into two categories: skeleton-based and nonskeleton-based. Skeleton-based techniques are those indirect methods which segment and reconstruct the vessels by first detecting the centerlines of the vessels. Several methods have been developed based on this principle and multiscale schemes to allow for the diversity of vessel sizes [1, 2, 3]. In these approaches, the centerline models can be generated explicitly, implicitly or via postprocessing by vessel modeling methods. Contrary to above methods, nonskeleton-based techniques are those that compute the vessels in 3D directly. In this category, deformable model based methods have received considerable attention and success.

Klein *et al.* [4] proposed to reconstruct 2D vessel boundaries or 3D vessel walls using deformable surface models represented by B-spline surfaces. However, it is not possible to employ parameterized deformable models to effectively deal with whole vessel trees, as the models would be required to change topology during evolution. Yim *et al.* [5] proposed a deformable surface model based on triangulated meshes for vessel construction in 3D. Nevertheless, it may be problematic to apply these methods [4, 5] for segmentation of vessels from low contrast MRA

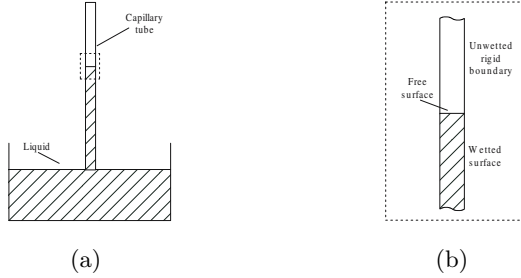


Fig. 1. Capillary action. (a) Capillary tube, (b) Surfaces of a three-phase system.

images. Geometric deformable models based on level set theory [6, 7], which can freely adapt into complex topologies of objects, were then proposed and applied on MRA images segmentation. Chen and Amini [8] employed a hybrid model using both parametric and geometric deformable models for segmentation of an entire vascular tree. However, these methods may have difficulty in extracting tiny vessels from 3D images. Small vessels and their branches, which exhibit much variability, are very important in planning and performing neurosurgical procedures. Greater details can provide more precise navigation and localization information for computer guided procedures. Lorigo *et al.* [9] proposed the “CURVES” algorithm to extract thin vessels, which uses geodesic active contour [7] based on the co-dimension two level set method [10]. However, to make the evolving curve stop at the object boundaries, the evolution speed needs to be multiplied by a heuristic factor.

Our work aims to develop an image segmentation methodology for automatically extracting the whole vasculature from 3D angiography. Inspired by the capillary action associated with thin tubes as shown in Fig. 1(a), in which liquid climbs up to some height without external pulling force, an algorithm for vasculature extraction is proposed. In the capillary action, the thinner the tube, the higher the liquid level in the tube. The situation is very similar to the segmentation of blood vessels if we imagine thin blood vessels as capillary tubes. Hence, it may be useful to employ this mechanism for segmenting thin vessels in low contrast situations. By modeling this phenomenon and fitting it into image segmentation problem, the *capillary active contour* (CAC) is obtained for MRA image segmentation.

2 Capillary Geodesic Active Contour

The capillary action can be understood as an energy minimization process, which involves surface tensions. We first introduce the energy functional associated with the *free surface* and *wetted surface*. Following that, the volume constraint of the fluid is considered. To minimize the overall energy functional, the Euler-Lagrange equations are computed for each part and these derived parts are

integrated together to get the final evolution equation of the fluid surface. Since the capillary force is incorporated, the derived method is called *capillary active contour*.

2.1 Free Surface Energy

Free surface is the part of the liquid surface that is not in contact with the solid boundaries. Capillary action can only be observed when the adhesion tension between molecules of fluid and molecules of solid is larger than the cohesion tension between fluid molecules. Therefore, molecules on the free surface tend to be attracted toward the solid boundaries. Fluid molecules will have lower potential energy when they are nearer to solid boundaries and this energy will be minimized when they are on the solid boundaries.

The gradient magnitude $|\nabla I|$ is calculated to indicate possible boundaries, where $I : [0, a] \times [0, b] \times [0, c] \rightarrow \mathbb{R}^+$ denotes an image. A uniform decreasing function $g : [0, +\infty[\rightarrow \mathbb{R}^+$ is defined, where $g(r) \rightarrow 0$ as $r \rightarrow +\infty$. Function $g(|\nabla I(S(q))|)$ is used to describe the energy coefficient associated with the surface $S(q) : [0, 1] \rightarrow \mathbb{R}^3$, which is the surface tension coefficient in physics. Let $S_f(t)$ denote the free surface of the liquid at time t . The capillary energy of free surface can be expressed as

$$E(S_f(t)) = \int_q g(S(q)) \left| \frac{\partial S(t, q)}{\partial q} \right| dq \quad (1)$$

by integrating over q , where $S(t, q) \in S_f(t)$. By computing the Euler-Lagrange of (1), the evolution equation of the free surface is obtained as

$$S_t = g\kappa\mathcal{N} - (\nabla g \cdot \mathcal{N})\mathcal{N}, \quad (2)$$

where κ is the Euclidean curvature, and \mathcal{N} is the unit normal vector of the surface.

2.2 Wetting Surface Energy

Since the adhesion force is larger than the cohesion force, the fluid molecules will be attracted to the solid boundaries and the unwetted surface is then converted into wetted surface. The total energy involved in this dynamic process can be modeled as

$$E(S_w(t)) = \beta S_w + \beta^* S_w^*, \quad (3)$$

where S_w is the area wetted by the fluid and S_w^* is the area in contact with the outer medium (unwetted surface). Since S_w can be extended in an arbitrarily continuous way into S_w^* , we have

$$E(S_w(t)) = \hat{\beta} S_w^* \quad (4)$$

where $\hat{\beta} = \beta^* - \beta$. Thus, minimizing the wetted surface energy is equivalent to minimizing the surface S_w^* through evolving the contact line. However, solving

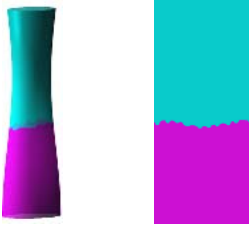


Fig. 2. 3D tubular surface is stretched to get 2D surface

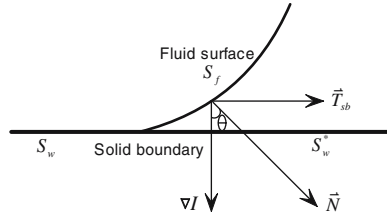


Fig. 3. The evolving direction of the contact line

this problem directly in 3D is problematic, because the surface to be minimized is not enclosed by the contact line. Since we are only interested in an immediate neighborhood of the contact line, it is always possible to find some view point that all the surface area is under the contact line (see Fig. 2). In addition, the 3D surface can be “cut” along a line and then stretched to becoming a 2D plane as shown in Fig. 2.

Let $\mathcal{C}(t, x) : [0, 1] \rightarrow \mathbb{R}^2$ denote the contact line between the fluid surface and the unwetted surface. The length of the curve $\int \left| \frac{\partial \mathcal{C}(t, x)}{\partial x} \right| dx$ is considered as a regularization term to keep the curve smooth. Thus, the energy associated with the wetted surface can be written as

$$E(S_w(t)) = \int \mathcal{C}(t, x) dx + \lambda \int \left| \frac{\partial \mathcal{C}(t, x)}{\partial x} \right| dx. \tag{5}$$

where λ is a real positive constant parameter. Then the 2D evolving equation of contact line \mathcal{C} is obtained as

$$\mathcal{C}_t = (1 + \lambda \hat{\kappa}) \mathcal{N}_\Gamma, \tag{6}$$

where $\hat{\kappa}$ and \mathcal{N}_Γ are the Euclidean curvature and the unit normal vector of the contact line in 2D, respectively. However, when working in 3D, the curvature $\hat{\kappa}$ of the 2D contact line becomes a co-dimension two curvature $\hat{\kappa}_2$ [10, 9]. Furthermore, as we are considering the evolution of liquid along vessel walls, the direction of evolution should be parallel to the wall at each point, *i.e.*, the tangential direction \mathcal{T}_{sb} as shown in Fig. 3. Then (6) is changed into

$$\mathcal{C}_t = (1 + \lambda \hat{\kappa}_2) \mathcal{T}_{sb}. \tag{7}$$

The tangential subvector can be calculated as

$$\mathcal{T}_{sb} = \mathcal{N} - \frac{\nabla g}{|\nabla g|} \cos \theta, \tag{8}$$

where $\cos \theta = \mathcal{N} \cdot \frac{\nabla g}{|\nabla g|}$.

Thus, the overall evolution equation describing the energy of wetting surfaces can be written as

$$\mathbf{S}_t = (1 + \lambda\hat{\kappa}_2) \left(\mathcal{N} - \frac{\nabla g}{|\nabla g|} \cos \theta \right). \quad (9)$$

2.3 Volume Constraint

The volume constraint is considered as $E(V) = cV$, which can be expanded as

$$E(V) = c \int S(t, q) dq,$$

where c is a Lagrange multiplier. By applying the gradient descent minimization, we have evolution equation $\mathbf{S}_t = c\mathcal{N}$, which is a constant velocity for minimizing the volume enclosed by the surface. Since liquid is bounded by solid, the fluid surface will snap to the boundaries. Thus, the constant velocity is multiplied by function g and then the evolution equation is modified into

$$\mathbf{S}_t = g(S(q))c\mathcal{N}. \quad (10)$$

2.4 Evolution Equation

Based on the results obtained separately in the above sections, the final evolution equation is obtained by integrating these terms

$$\mathbf{S}_t = g(\kappa + c)\mathcal{N} - (\nabla g \cdot \mathcal{N})\mathcal{N} + \alpha(1 + \lambda\hat{\kappa}_2) \left(\mathcal{N} - \frac{\nabla g}{|\nabla g|} \cos \theta \right), \quad (11)$$

where parameter α is a real positive constant. The constant term c in (11) acts like balloon force in [11], which facilitates the evolving surface snapping to solid boundaries. Comparing the new speed function (11) with those of the geodesic active contour [7], the third term is new which comes from the capillary action. The new term makes the fluid surface move along the solid boundaries. The capillary action term is expected to facilitate the evolving surface adapting into thin parts of objects, *e.g.* thin vessels.

2.5 Level Set Evolution Equation

According to the level set theory and noting the fact that

$$\mathcal{N} = -\frac{\nabla \Psi}{|\nabla \Psi|},$$

evolving a surface S under the speed function (11) is equivalent to updating a volumetric map Ψ with

$$\Psi_t = g(\kappa + c)|\nabla \Psi| + \nabla g \cdot \nabla \Psi + \alpha(1 + \lambda\hat{\kappa}_2)|\nabla \Psi|(1 - \cos^2 \theta) \quad (12)$$

where S is the zero level set embedded in Ψ and

$$\cos \theta = \frac{\nabla \Psi \cdot \nabla g}{|\nabla \Psi| |\nabla g|}. \quad (13)$$

3 Results and Discussion

3.1 Capillary Force

Fig. 4 shows the application of the CAC in a computer generated cylinder, which illustrates the effects of the capillary force. Starting from the initialization visualized in Fig. 4(b), evolution is done according to the level set speed function in (12). In our experiment, all the parameters are fixed except that α is varying. The evolution results with different parameter settings after 50 iterations are visualized in Fig. 4.

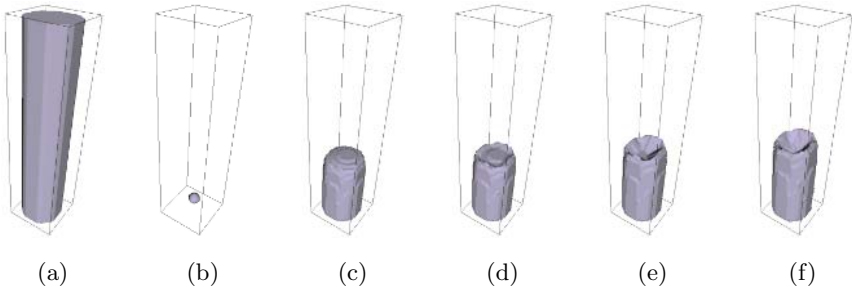


Fig. 4. Illustration of the effects on varying capillary force coefficient α . (a) Generated cylinder. (b) Initialization of the algorithm. (c) $\alpha = 0$. (d) $\alpha = 0.25$. (e) $\alpha = 0.5$. (f) $\alpha = 0.75$.

When α is set to zero, the CAC in fact evolves exactly like a geodesic active contour [7]. The evolution of the free surface in the axial direction of the tubular object is driven by the mean curvature κ and constant speed c in (12). Since this term is multiplied by the uniquely decreasing function g , its contribution will be small when the surface is near the object boundary, and becomes zero when the surface is on the edges. Hence, we can see the free surface is convex as shown in Fig. 4(c). When the tubular object becomes much thinner, the evolution may stop because of the small value of g at everywhere. After incorporating the capillary force, the free surface is attracted by the unwetted surface when near the solid boundaries and *meniscus* [12] can be observed as shown in Fig. 4(d)(e)(f). With larger value of α , the surface evolution along the object boundary will be stronger and faster. Therefore, the capillary force can facilitate the evolution of the free surface even when the vessels are very thin.

3.2 MRA Segmentation

The CAC method is also applied on 3D MRA images of cerebral vasculature. The image size is $512 \times 512 \times 120$ voxels with spacing $0.43\text{mm} \times 0.43\text{mm} \times 1.2\text{mm}$. The initial surface is obtained by thresholding the raw dataset. The MIP of MRA data set used in our experiments is shown in Fig. 5. To make the results

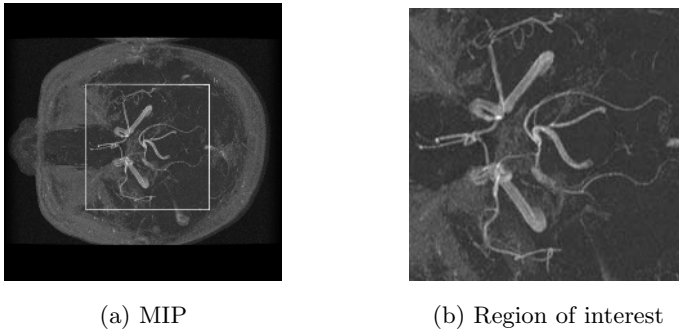


Fig. 5. MIP of the cerebral MRA data set

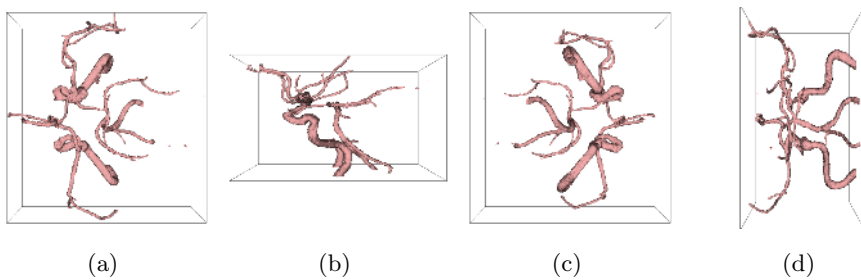


Fig. 6. MRA segmentation results using CURVES

clearer, a region of interest (ROI) is extracted (Fig. 5(b)) and only vessels inside the ROI are segmented and visualized.

We have compared our proposed CAC method with the state-of-the-art CURVES [9]. Identical parameter settings in the evolution equations are used for both methods except for the capillary force coefficient, which is specific to the CAC. The segmentation results of CURVES and CAC are shown in Fig. 6

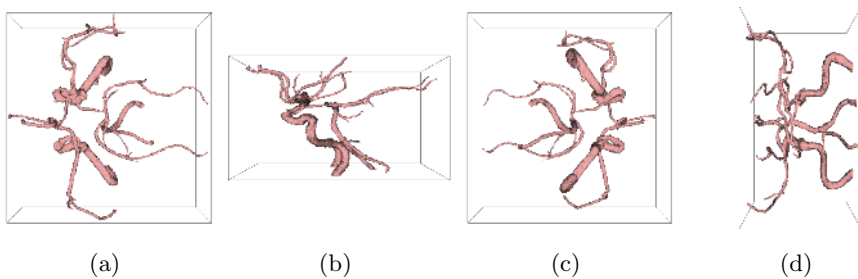


Fig. 7. MRA segmentation results using CAC

and Fig. 7 with different points of view. The CURVES algorithm successfully extracts much of the vasculature but fails to identify some parts when compared with the CAC.

4 Conclusions

In this paper, we propose the capillary active contour and apply it to both 3D synthetic and cerebral MRA images segmentation. The incorporated capillary force adapts the evolving surface into very thin branches of blood vessels and obtains more accurate segmentation as demonstrated in our experiments. Compared with other techniques, the CAC can achieve more details of vasculature. Our approach is geometric in nature and topology free due to that implicit representation of the evolving surface is used.

References

1. Frangi, A.F., Niessen, W.J., Hoogeveen, R.M., van Walsum, T., Viergever, M.A.: Model-based quantitation of 3-D magnetic resonance angiographic images. *IEEE Trans. Med. Imag.* **18** (1999) 946–956
2. Krissian, K., Malandain, G., Ayache, N., Vaillant, R., Troussset, Y.: Model based detection of tubular structures in 3D images. *Computer Vision and Image Understanding* **80** (2000) 130–171
3. Aylward, S.R., Bullitt, E.: Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction. *IEEE Trans. Med. Imag.* **21** (2002) 61–75
4. Klein, A.K., Lee, F., Amini, A.A.: Quantitative coronary angiography with deformable spline models. *IEEE Trans. Med. Imag.* **16** (1997) 468–482
5. Yim, P.J., Boudewijn, G., Vasbinder, C., Ho, V.B., Choyke, P.L.: Isosurfaces as deformable models for magnetic resonance angiography. *IEEE Trans. Med. Imag.* **22** (2003) 875–881
6. Malladi, R., Sethian, J.A., Vermuri, B.C.: Shape modeling with front propagation: A level set approach. *IEEE Trans. Pattern Anal. and Machine Intell.* **17** (1995) 158–174
7. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. *Int. J. Computer Vision* **22** (1997) 61–79
8. Chen, J., Amini, A.A.: Quantifying 3-D vascular structures in MRA images using hybrid PDE and geometric deformable models. *IEEE Trans. Med. Imag.* **23** (2004) 1251–1262
9. Lorigo, L.M., Faugeras, O., Grimson, W.E.L., Keriven, R., Kikinis, R., Nabavi, A., Westin, C.F.: CURVES: Curve evolution for vessel segmentation. *Medical Image Analysis* **5** (2001) 195–206
10. Ambrosio, L., Soner, H.M.: Level set approach to mean curvature flow in arbitrary codimension. *J. Differential Geometry* **43** (1996) 693–737
11. Cohen, L.D.: On active contour models and balloons. *CVGIP: Image Understanding* **53** (1991) 211–218
12. Finn, R.: *Equilibrium Capillary Surfaces*. Springer-Verlag, New York (1986)

Spatial Graphs for Intra-cranial Vascular Network Characterization, Generation, and Discrimination

Stephen R. Aylward¹, Julien Jomier¹, Christelle Vivert¹,
Vincent LeDigarcher¹, and Elizabeth Bullitt²

¹ CADDLab, Department of Radiology
aylward@unc.edu

² Department of Surgery, University of North Carolina, USA

Abstract. Graph methods that summarize vasculature by its branching topology are not sufficient for the statistical characterization of a population of intra-cranial vascular networks. Intra-cranial vascular networks are typified by topological variations and long, wandering paths between branch points.

We present a graph-based representation, called spatial graphs, that captures both the branching patterns and the spatial locations of vascular networks. Furthermore, we present companion methods that allow spatial graphs to (1) statistically characterize populations of vascular networks, (2) generate the central vascular network of a population of vascular networks, and (3) distinguish between populations of vascular networks. We evaluate spatial graphs by using them to distinguish the gender and handedness of individuals based on their intra-cranial vascular networks.

1 Introduction

Intra-cranial vasculature varies across individuals to such an extent that only a few of the largest intra-cranial vessels are sufficiently consistent to be depicted and labeled in anatomical atlases [13]. Variations in the location and connectivity of even those named vessels are common in a healthy population [2]. Furthermore, between branchpoints, intra-cranial vessels will often follow long and wandering walks. These walks must be preserved to truly capture the form and function (filling regions, etc.) of intra-cranial vasculature. Our interest in characterizing intra-cranial vasculature arises from the mounting evidence that a genetic relationship exists between mental disorders and vascular network formation [10][15].

We have developed a novel, graph-theoretic method for representing and analyzing the intra-cranial vascular networks of individuals and populations. Our "spatial graphs" capture the anatomical locations, branching patterns, and tortuous paths of intra-cranial vascular networks. Spatial graph formation begins with a centroid voronoi tessellation (CVT) [11] of intra-cranial space. Those CVT regions are potential nodes of a spatial graph. A vascular network, overlaid on a voronoi tessellation of space, defines a spatial graph by recording the branching sequence of CVT nodes visited by the network. These branching sequences can

be equivalently recorded as weighted, asymmetric adjacency matrices. Adjacency matrix and node-specific statistics, such as the centrality [3] and branch probability at each node, can be collected for a population of vascular networks. We have defined a method whereby those statistics can be used to probabilistically generate vascular networks that are central to a population. We have also developed a method with which those statistics can be used to compute membership scores for an individual’s vascular network; an individual’s membership scores from different population-specific graphs and be used to classify that individual.

Forming graph representations of vascular networks is not a new concept. In 1993 Dr. Gerig presented [9] an intra-cranial vessel segmentation technique and illustrated the reduction of those segmentations to graphs that represent their topology. Such branch-based graphs are the basis of much of the vessel-based liver lobe and heart vessel segmentation work being conducted in a variety of labs [8][14]. In computer-vision, graph theory has been applied to the characterization of the medial structure of objects [12]. To the best of our knowledge, a method has not been previously demonstrated that applies graph theory to the characterization, generation, and discrimination of populations of intra-cranial vascular networks.

In the next section we present our methods. The subsequent section describes three evaluations of our methods: (1) visually assessing the central vascular network generated from the spatial graph statistics from right-handed males, (2) using male and female-specific spatial graphs of intra-cranial vasculature to determine the gender of testing individuals, and (3) using a right-handed-specific population graph to determine the handedness of testing individuals.

2 Methods

This section is divided into three subsections: forming spatial graphs, generating vascular networks that are central to population of vascular networks, and computing graph membership measures. While these method descriptions focus on intra-cranial vasculature, these methods are applicable to any branching, directed, cyclic, space-occupying structure.

2.1 Forming Spatial Graphs

Spatial graph formation involves vessel extraction, inter-subject registration, space partitioning, and recording graph statistics. These steps are detailed next.

1) *Extract the vascular networks from the individuals’ MRA data.* These encodings must capture the location of vessels, their directions of blood flow, and their branchpoints to accurately encode a spatial graph as a directed graph. We have developed a vessel network that has been shown to produce accurate localizations [1] and accurate flow direction and branch point information [4]. Our method works for contrast CT and time-of-flight MR angiograms (MRA).

2) *Align the individuals’ vascular networks to a common coordinate frame.* We achieve this by (a) rigidly registering, using mutual-information, the individuals’ MRA to their T1 images; (b) performing an affine, mutual information

registration between the T1 images and the BrainWeb T1 atlas [7]; (c) composing the corresponding MRA-to-T1 and T1-to-atlas transforms; and (d) applying the composed transform to the extracted vessels. [6]

3) *Partition intra-cranial space.* The goal of this step is to define a partitioning that captures vascular locations, branches, and paths in intra-cranial space. This information is best captured when each partition has equal likelihood of containing a vessel; otherwise under-utilized partitions would offer less discriminatory information. We therefore partition space using expected vascular density as follows.

Expected vascular density is computed using the vascular distance maps for a population [6]. Specifically, for each vascular network, the squared distance to the nearest vessel is computed throughout intra-cranial space. Those distance maps are then inverted so that points nearer the vessels have larger values; thus, intensity is related to likelihood of finding a vessel at that point. A mean distance map is computed as a voxel-by-voxel average of those maps.

An equi-probable partitioning of expected vascular density is computed using centroid voronoi tessellation (CVT) [11]. For this paper, 1000 CVT regions are resolved using a probabilistic version of Lloyd’s method [11]. This algorithm is computed for 1000 iterations using 100,000 new, randomly generated samples per iteration. Random samples are generated from the mean density map using the rejection method [11], so, the likelihood of selecting a voxel as a sample is proportional to the image intensity at that voxel. Each iteration, the mass centroids of the voronoi regions are updated using a new set of samples. Results are illustrated in Fig. 1.

4) *Recording graph statistics.* The adjacency matrix and node-specific statistics summarize how vascular networks map onto the CVT regions.

For adjacency matrix computation, as each vessel’s directed traversal crosses from CVT region i to the next region j , the corresponding position A_{ij} in the

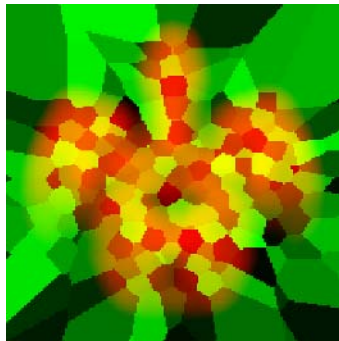


Fig. 1. An axial slice (at the level of the Circle of Willis) through the mean vessel distance map and CVT partitioning formed from 18 individuals. The fuzzy (red) object in the background is the expected vascular density. The crisp (green) overlay uses random intensities to depict the computed CVT regions.

population’s adjacency matrix is increased by $1/(\text{total number of vessels in that individual’s network})$. This weighting compensates for level-of-detail variations in the vascular networks.

Node statistics summarize local and connectivity information. At each node we record (a) the number of times it contains a branch point, (b) the mean radius of the vessel points that are contained within it, and (c) the number of times a root artery, e.g., a carotid or basilar artery, originates from within it. Additionally, in this paper, we compute node centrality as a local summary of connectivity. Our node centrality computation is an adaptation of a method developed for determining the “popular people” in a social network, i.e., the most “central person” in directed communications [3]. To compute such centrality, a value e_i represents the externally defined status of person i in the network. A parameter α is set to weight the influence of that external status on internal exchanges. The centrality c_i of a node/person is then computed as

$$c = (I - \alpha A^t)^{-1} e \quad (1)$$

where A is the asymmetric adjacency matrix and I is the identity matrix. We applied this measure to our vascular networks by giving every CVT node equal external status, i.e., $e_i = 1$ for all i , and setting $\alpha = 0.01$ to reduce the influence of external status. In the future, it may be useful to relate external status to the volume of a vessel’s fill region or its radius. We also removed the transpose of A so that the directed links in our networks are inverted; all vessels are made to flow to their roots, e.g., the carotids. After applying Equ. 1, the node that most often contains the root of the largest tree receives the largest centrality value. The next distal node is the adjacent node with the highest centrality value, and so forth. Using adjacency and node statistics, vascular graphs can be generated and classified, as described next.

2.2 Generating Central Vascular Networks

Generating the most central vascular network from a population’s spatial graph involves a greedy traversal of the population’s adjacency matrix and the consideration of the branch probabilities at its nodes. Since each node correspond to a volume in space, a branching sequence of nodes defines a branching network passing through those volumes in space. Our generation process involves two user specified threshold values. The first threshold θ_A indicates the minimum level of adjacency required for a vessel traversal to continue. The second threshold θ_b indicates the minimum branch-support required for a branch point to be taken. Branch-support from a node i to node j is defined by the product of the probability of branching at node i and the adjacency weight from i to j .

Generation begins at the most likely root node, as reported by the recorded node statistics. The largest weighted adjacency from that node is then chosen as the next transition in the generated graph. Adjacent nodes, not selected for continued generation, are searched for branch-support values above θ_b . If a branch node is found, it is queued for generation after the current vessel’s generation

terminates. A vessel’s generation terminates when it encounters a node from which none of the adjacent nodes have an adjacency value above θ_A . Future research should consider statistics regarding radius, mean curvature, etc. when generating vessels.

2.3 Graph Membership Measures

Membership scores for a new vascular network, represented by the spatial graph g , are conditional on a population’s spatial graph G . We evaluate two membership scores in this paper: $\Sigma_A(g|G)$ is the sum of the adjacency weights in G for all transitions that also exist in g , and $\Sigma_c(g|G)$ is the sum of the centrality values for the CVT nodes in G that are also touched by g . Both of these values should be maximized when a vascular network is compared with a graph that represents the population to which that vascular networks belongs. The next section describes experiments which test this theory.

3 Results

Three experiments were conducted to evaluate spatial graphs. Each of these is an unsolved problem in medical image analysis.

For these experiments, MRA and T1-weighted MR images were collected on a Siemens Allegra head-only 3T MR system. MRA data was acquired using a time-of-flight sequence at $0.5 \times 0.5 \times 0.8 \text{mm}^3$ voxel spacing. T1-weighted MR data was acquired at $1 \times 1 \times 1 \text{mm}^3$.

The images from an individual were either used for training or testing, but never for both. Individuals were labeled by gender and handedness.

All of the experiments began with a common CVT partitioning of intra-cranial space. Specifically, vascular networks from 9 training-data, right-handed males (ages 22-55, mean age 34.4) and 9 training-data, right-handed females (ages 20-54, mean age 35.3) were combined to produce a common mean density map. The CVT of that map was computed once. Adjacency matrices, node centrality, and other measures were computed using the designated individuals’ vascular networks combined with this common CVT partitioning. Using a common CVT partitioning may have reduced sensitivity and specificity, but it allowed node statistics to be compared across the populations’ spatial graphs. This CVT partitioning is depicted in Fig. 1.

3.1 Experiment 1: Vascular Graph Generation

The first experiment demonstrates the generation of the central vascular network for right-handed males. Using the 9 training-data, right-handed males, a right-handed, male population spatial graph was computed. The vascular network generation method described above was then applied. For visualization, the mid-points on the lines between the centroids of each generated sequence of nodes were used as the control points for a b-spline curve through space. A constant radius of 0.5 mm was used for visualizing the vessels as tubes.

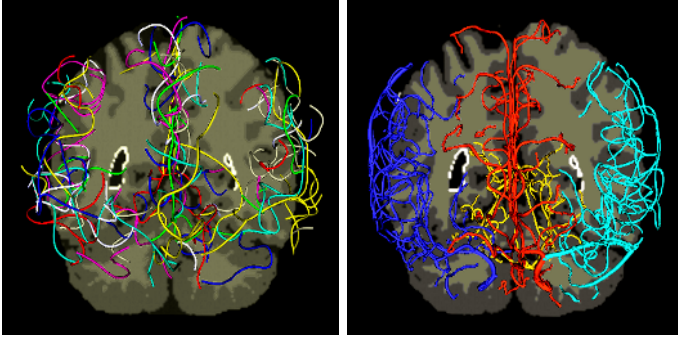


Fig. 2. Left: A vascular network generated from the right-handed male graph. Right: A vascular network from one of the training right-handed males. Carotids, circle of willis, vascular groups, other vascular patterns are evident in the generated network.

The generated central vascular network is visualized in Fig. 2. Subjectively, there are striking similarities between the generated and actual intra-cranial vascular networks. The generation method’s threshold values can be manipulated to produce various levels of branching detail. The image shown used a low branching threshold, $\theta_b = 0.001$, so as to simulate extensive branching detail. Major vascular groups, the carotids, circle of willis, basilar artery, communicating arteries, and other major vessels are evident. Using b-splines to visualize the node adjacency sequence may have exaggerated intra-node vessel curvature.

3.2 Experiment 2: Distinguishing Gender

The second experiment involved distinguishing males from females based on intra-cranial vasculature. Two population-specific spatial graphs were computed: one using 9 training-data, right-handed males and the other using 9 training-data, right-handed females. For each population-specific graph, the membership

Table 1. (Upper Table) Mean difference for each membership measure for testing individuals given gender-specific population graphs. (Lower Table) Mean differences for each membership measure for testing individuals given right-handed and mirrored-right-handed population graphs. Items in bold highlight population-specific graphs that correctly classified every individual in the corresponding testing data.

	$\Sigma_A(g G_{male}) - \Sigma_A(g G_{female})$	$\Sigma_C(g G_{male}) - \Sigma_C(g G_{female})$
Testing Males	$\mu = \mathbf{76.6}$ $\sigma = \mathbf{5.08}$	$\mu = -7.4$ $\sigma = 0.89$
Testing Females	$\mu = \mathbf{-18.6}$ $\sigma = \mathbf{5.02}$	$\mu = \mathbf{-15.1}$ $\sigma = \mathbf{0.75}$
	$\Sigma_A(g G_{right}) - \Sigma_A(g G_{rightmirrored})$	$\Sigma_C(g G_{right}) - \Sigma_C(g G_{rightmirrored})$
Testing Right	$\mu = \mathbf{5.83}$ $\sigma = \mathbf{3.40}$	$\mu = -4.01$ $\sigma = 18.5$
Testing Left	$\mu = \mathbf{-82.5}$ $\sigma = \mathbf{12.62}$	$\mu = 9.06$ $\sigma = 21.8$

measures $\Sigma_A(g|G)$ and $\Sigma_c(g|G)$ were computed for 3 testing-data, right-handed males (ages 23, 26, 57) and 3 testing-data, right-handed females (ages 33, 43, and 57).

Gender classification results are shown in Tab. 1. The Σ_A measure produced correct classifications for every testing-data male and every testing-data female. Results for Σ_c contained 50% classification error: mislabeling every testing-data male as female. The success of Σ_A and failure of Σ_c implies that gender differences are not based on the location or number of vessels, but it is the directed vascular paths through space that distinguish males from females. Further research is focusing on visualizing these differences.

3.3 Experiment 3: Determining Handedness

Our third experiment involved analyzing brain lateralization to determine the self-reported handedness of individuals. To study brain lateralization, instead of creating a population-specific spatial graph for left-handed males, we mirrored the right-handed graph along its x-axis. This mirroring approximates brain lateralization since the BrainWeb atlas' mid-sagittal plane is at the center of and normal to the x-axis. Using these right-handed male and mirrored-right-handed-male population-specific spatial graphs, membership measures $\Sigma_A(g|G)$ and $\Sigma_c(g|G)$ were computed for the 3 testing-data right-handed males and for 3 testing-data left-handed males (ages 19, 26, and 29).

Results from handedness classification using intra-cranial vasculature are shown in Tab. 1. Again the adjacency membership measure $\Sigma_A(g|G)$ lead to correct labels for all testing-data individuals. The centrality measure $\Sigma_c(g|G)$ had a 66% classification error; mislabeling two right-handed individuals as left-handed and two left-handed individuals as right-handed. Again, the success of $\Sigma_A(g|G)$ and the failure of $\Sigma_c(g|G)$ suggests that the direct paths of intra-cranial vasculature is key to distinguishing these populations.

4 Discussion

We have developed spatial graphs as a technique which is well suited for characterizing intra-cranial vasculature. Initial results indicate that spatial graphs can be used to classify the gender and handedness of individuals. These successes suggests that spatial graphs can be used to study the genetic relationship that may exists between mental disorders and vascular network formation [10][15].

One particularly interesting finding is that when various membership measures are compared, classification experiments suggest that it is the paths taken by the vessels in a network (their adjacency matrices), not their centrality, that distinguishes genetic vascular variations.

While these results are encouraging, we are very cautious about drawing any conclusions beyond the limited data involved. It is critical that these methods be applied to larger datasets before suggesting that any general "truths" have been revealed. Future work will also pursue more advanced statistical analysis techniques.

References

1. Aylward, S and Bullitt, E, *Initialization, Noise, Singularities and Scale in Height Ridge Traversal for Tubular Object Centerline Extraction*, IEEE TMI 21(2):61-75, 2002
2. Baskaya, MF; et al. *Surgical and Angiographic Anatomy of the Posterior Communicating and Anterior Communicating Arteries*, Neuro-anatomy 3:38-42, 2004
3. Bonacich, P and Lloyd, P, *Eigenvector-Like Measures of Centrality for Asymmetric Relations*, Social Networks 23:191-201, 2001
4. Bullitt, E; Aylward, S; Liu, A; Stone, J; Mukherji, S; Coffey, C; Gerig, G; Pizer, SM, *3D Graph Description of the Intracerebral Vasculature from Segmented MRA and Tests of Accuracy by Comparison with X-Ray Angiograms*, IPMI:308-321, 1999
5. Bullitt, E; Aylward, S; Smith, K; Mukherji, S; Jiroutek, M; and Muller, K, *Symbolic Description of Intracerebral Vessels Segmented From MRA and Evaluation by Comparison with X-Ray Angiograms*, MedIA 5:157-169, 2001
6. Chillet, D; Jomier, J; Cool, D; Aylward, S, *Vascular Atlas Formation Using a Vessel-to-Image Affine Registration Method*, MICCAI :335-342, 2003
7. Cocosco, CA; Kollokian, V; Kwan, RK-S; Evans, AC, *BrainWeb: Online Interface to a 3D MRI Simulated Brain Database* NeuroImage 5(4), 1997
8. Ezquerro, N; Capell, S; Klein, L; Duijves, P, *Model-Guided Labeling of Coronary Structure*, IEEE TMI 17(3):429-441, 1989
9. Gerig, G; Koller, T; Székely, G; rechbühler, C; übler, O, *Symbolic Description of 3-D Structures Applied to Cerebral Vessel Tree Obtained From MR Angiography Volume Data*, IPMI :94-111, 1993
10. Hesser, BA; et. al, *Down Syndrome Critical Region Protein 1 (DSCR1), a Novel VEGF Target Gene That Regulates Expression of Inflammatory Markers on Activated Endothelial Cells*, Blood 104(1):149-158, 2004
11. Ju, L; Du, Q; Gunzburger, M, *Probabilistic Methods for Centroidal Voronoi Tessellations and Their Parallel implementations*, Parallel Computing 28:1477-1500 2002
12. Macrini, D.; Shokoufandeh, A; Dickinson SJ; Siddiqi, K; Zucker; SW, *View-Based 3-D Object Recognition Using Shock Graphs*, ICPR:24-32, 2002
13. Netter, F, *Atlas of Human Anatomy*, Novartis, East Hanover, new Jersey, pages 525, 1989
14. Selle, D; Preim, B; Schenk, A; Peitgen, H.-O., *Analysis of Vasculature for Liver Surgical Planning* IEEE TMI 21(11):1344-1357, 2002
15. Stalmans, I; et al., *VEGF: A Modifier of the del22q11 (DiGeorge) Syndrome?* Nature Medicine 9:173-182, 2003

Surface Alignment of 3D Spherical Harmonic Models: Application to Cardiac MRI Analysis

Heng Huang¹, Li Shen³, Rong Zhang¹, Fillia Makedon¹,
Bruce Hettleman², and Justin Pearlman^{1,2}

¹ Department of Computer Science, Dartmouth College, Hanover, NH 03755

² Department of Cardiology, Dartmouth Medical School, Lebanon, NH 03756

³ Computer and Information Science Department, Univ. of Mass. Dartmouth, MA 02747
hh@cs.dartmouth.edu

Abstract. The spherical harmonic (SPHARM) description is a powerful surface modeling technique that can model arbitrarily shaped but simply connected 3D objects and has been used in many applications in medical imaging. Previous SPHARM techniques use the first order ellipsoid for establishing surface correspondence and aligning objects. However, this first order information may not be sufficient in many cases; a more general method for establishing surface correspondence would be to minimize the mean squared distance between two corresponding surfaces. In this paper, a new surface matching algorithm is proposed for 3D SPHARM models to achieve this goal. This algorithm employs a useful rotational property of spherical harmonic basis functions for a fast implementation. Applications of medical image analysis (*e.g.*, spatio-temporal modeling of heart shape changes) are used to demonstrate this approach. Theoretical proofs and experimental results show that our approach is an accurate and flexible surface correspondence alignment method.

1 Introduction

Surface representation and shape modeling play increasingly prominent roles in many computer vision and image processing applications. Medical image analysis is one of the most important applications. Many techniques have been developed for modeling and inspecting anatomic structures in the diagnosis and treatment of disease. The spherical harmonics approach has been used for the representation of shapes in many types of biomedical image data to help perform functional information analysis or classify different pathological symptoms.

Many spherical harmonic based shape descriptions have been developed for medical image analysis. Chen *et al.* [1] use this method to model and analyze left ventricular shape and motion. Matheny *et al.* [2] and Burel *et al.* [3] use 3D and 4D surface harmonics to reconstruct rigid and nonrigid shapes. Since they start from an initial radial surface function $r(\theta, \phi)$, their method is capable of representing only star-shaped or convex objects without holes. Brechbühler *et al.* [4] present the SPHARM description that is an extended spherical harmonic method for modeling any simply connected 3D object. The object surface is represented as $\mathbf{v}(\theta, \phi) = (x(\theta, \phi), y(\theta, \phi), z(\theta, \phi))^T$ and spherical harmonics expansion is used for all three coordinates. Gerig and Styner have

applied SPHARM in many medical imaging applications (*e.g.*, shape analysis of brain structures [5,6,7]). It has also been used for shape modeling and functional analysis for cardiac MRI [8].

In order to compare different SPHARM models, a shape registration step is often necessary for aligning these models together and extracting their shape descriptors (*i.e.*, excluding translation, rotation, and scaling). Like shape registration using ICP [9], two important substeps are involved in aligning SPHARM models: (1) creating surface correspondence, and (2) minimizing the distance between the corresponding surface parts. Once the surface correspondence is established, the distance minimization becomes relatively easy. Thus, the focus of this paper is on creating surface correspondence for two 3D SPHARM models.

Previous studies [5,6] used the first order ellipsoid for shape registration. The parameter net on this ellipsoid is rotated to a canonical position such that the north pole is at one end of the longest main axis, and the crossing point of the zero meridian and the equator is at one end of the shortest main axis. The aligned parameter space creates surface correspondence between two models: two points with the same parameter pair (θ, ϕ) on two surfaces are defined to be a corresponding pair. This alignment technique works only if the first order ellipsoid is a real ellipsoid, as in the case of hippocampal data [6], but not if it is an ellipsoid of revolution or a sphere. There are also other cases in which first order ellipsoid alignment may not work. One example is given in Fig. 1 for the heart ventricle case.

In this paper, instead of aligning the first order ellipsoid, we employ a more general metric for establishing surface correspondence: minimizing the mean squared distance between two SPHARM surfaces. A fast surface alignment algorithm is proposed to achieve this. Based on the rotational properties of harmonics analysis, we prove that a new set of SPHARM coefficients after a rotated parametrization can be directly generated from the original set. Thus we can easily obtain a new SPHARM model for a re-parameterized object by rotating its parametrization along the surface. This process is faster than a standard recalculation (*e.g.*, solving a linear equation) of SPHARM coefficients for a re-parameterized object. This work is motivated by the need for better shape modeling and analysis in current medical applications. Some of these applications are used to demonstrate our algorithm in this paper.

2 Methods

2.1 Surface Description Using SPHARM

The SPHARM technique [4] can be used to model arbitrarily shaped, simply connected 3D objects. The object surface is represented by using spherical harmonics expansion for all three coordinates,

$$\mathbf{v}(\theta, \phi) = (x(\theta, \phi), y(\theta, \phi), z(\theta, \phi))^T = \sum_{l=0}^{\infty} \sum_{m=-l}^l \mathbf{c}_l^m Y_l^m(\theta, \phi). \quad (1)$$

The coefficients $\mathbf{c}_l^m = (c_{lx}^m, c_{ly}^m, c_{lz}^m)^T$ are 3D vectors. Their components, c_{lx}^m , c_{ly}^m , and c_{lz}^m are usually complex numbers. The coefficients up to a user-desired degree can be

estimated by solving a set of linear equations in a least square fashion. The object surface can be reconstructed using these coefficients, and using more coefficients leads to a more detailed reconstruction. Thus, a set of coefficients actually form an object surface description.

2.2 Fast Rotation Theorem for Spherical Harmonic Parametrization

Theorem (Parametrization Rotation). The parametrization spatial rotation on the surface can be decomposed into three rotations of mapping parameter meshes onto the x -sphere, y -sphere, and z -sphere. Let $\mathbf{v}(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l c_l^m Y_l^m(\theta, \phi)$ be a SPHARM parametric surface. After rotating the parameter net on the surface in Euler angles (α, β, γ) , the new coefficients $c_l^m(\alpha\beta\gamma)$ is

$$c_l^m(\alpha\beta\gamma) = \sum_{m'=-l}^l c_l^{m'} D_{mm'}^l(\alpha\beta\gamma). \quad (2)$$

Proof. According to Euler's rotation theorem, any rotation of the coordinate system (e_1, e_2, e_3) can be decomposed into three elementary rotations $R(\alpha, \beta, \gamma)$. The $SO(3)$ harmonics provide the tool to express the rotated version of a function on the sphere extended by spherical harmonics [10]. The effect of such a rotation on the spherical harmonic basis functions is [11]

$$R_{ZYZ}(\alpha\beta\gamma) Y_l^m(\theta, \phi) = \sum_{m'=-l}^l Y_l^{m'}(\theta, \phi) D_{m'm}^l(\alpha\beta\gamma), \quad (3)$$

where $R_{ZYZ}(\alpha\beta\gamma)$ represents the rotation operator dependent on the Euler angles; the rotation matrices $D_{m'm}^l(\alpha\beta\gamma)$ (also called the $SO(3)$ matrix elements) are calculated by

$$D_{m'm}^l(\alpha\beta\gamma) = e^{-im'\alpha} d_{m'm}^l(\beta) e^{-im\gamma},$$

where

$$\begin{aligned} d_{m'm}^l(\beta) &= \sum_{t=\max(0, m-m')}^{\min(l+m, l-m')} (-1)^t \times \frac{\sqrt{(l+m)!(l-m)!(l+m')!(l-m')!}}{(l+m-t)!(l-m'-t)!(t+m'-m)!t!} \\ &\times \left(\cos \frac{\beta}{2} \right)^{(2l+m-m'-2t)} \left(\sin \frac{\beta}{2} \right)^{(2t+m'-m)}. \end{aligned}$$

Since we employ the SPHARM surface modeling technique, the surface coordinate information of a 3D object is coded onto three unit spheres: an x -sphere, a y -sphere, and a z -sphere. These three spherical functions are expanded using spherical harmonics and represented by $f(\theta, \phi)$ ($f \in \{x, y, z\}$). We denote $f'(\theta, \phi)$ as the new function after applying a rotation operator $R_{ZYZ}(\alpha\beta\gamma)$ to $f(\theta, \phi)$ on the f -sphere:

$$f'(\theta, \phi) = R_{ZYZ}(\alpha\beta\gamma) f(\theta, \phi), \quad (4)$$

thus

$$\mathbf{v}'(\theta, \phi) = R_{obj}(\alpha\beta\gamma) (x(\theta, \phi), y(\theta, \phi), z(\theta, \phi))^T \quad (5)$$

where $R_{obj}(\alpha\beta\gamma) = [R_{ZY Z}(\alpha\beta\gamma) \ 0 \ 0; 0 \ R_{ZY Z}(\alpha\beta\gamma) \ 0; 0 \ 0 \ R_{ZY Z}(\alpha\beta\gamma)]$.

Then $\mathbf{v}'(\theta, \phi) = R_{obj}(\alpha\beta\gamma)\mathbf{v}(\theta, \phi)$ represents the new parametrization on the surface, which can be generated by rotating the original parametrization along the object's surface about Euler angles (α, β, γ) . In other word, the result of applying the rotation matrix $R_{ZY Z}(\alpha\beta\gamma)$ on the mapping meshes of x, y, z -sphere is to rotate the parameter mesh on the object's surface at the same orientation. Because of the distortions introduced by spherical parameterization, the result of rotation is not identical to the result of applying Euler angles on the sphere, but both will have nearly the same orientation. Thus we only use R_{obj} , which we refer to as the parametric rotation matrix, for rotating the parameter mesh along the surface of an object. Substituting Eq. (1) and Eq. (3) into Eq. (4) gives

$$\begin{aligned} \sum_{l=0}^L \sum_{m=-l}^l c_{lf}^m(\alpha\beta\gamma) Y_l^m(\theta, \phi) &= f'(\theta, \phi) = R_{ZY Z}(\alpha\beta\gamma) \sum_{l'=0}^L \sum_{m'=-l'}^{l'} c_{l'f}^{m'} Y_{l'}^{m'}(\theta, \phi) = \\ \sum_{l'=0}^L \sum_{m'=-l'}^{l'} c_{l'f}^{m'} R_{ZY Z}(\alpha\beta\gamma) Y_{l'}^{m'}(\theta, \phi) &= \sum_{l'=0}^L \sum_{m'=-l'}^{l'} c_{l'f}^{m'} \sum_{n=-l'}^{l'} Y_l^n(\theta, \phi) D_{nm'}^{l'}(\alpha\beta\gamma) \end{aligned}$$

and multiplying $\bar{Y}_k^j(\theta, \phi)$ on both sides (adjusting the k from 0 to L and j from $-k$ to k) and integrating on the sphere. Since all Kronecker delta values are zero except at $k = l = l'$ and $j = m = n$, we get the following:

$$c_{lf}^m(\alpha\beta\gamma) = \sum_{m'=-l'}^{l'} c_{l'f}^{m'} D_{mm'}^{l'}(\alpha\beta\gamma) = \sum_{m'=-l}^l c_{l'f}^{m'} D_{mm'}^l(\alpha\beta\gamma). \quad (6)$$

According to the above derivation, the harmonics expansion coefficients transform among themselves during rotation. Each new spherical harmonic coefficient $c_{lf}^m(\alpha\beta\gamma)$ after applying a rotated function $R_{ZY Z}(\alpha\beta\gamma)$ is a linear combination of the coefficients $c_{l'f}^{m'}$ of the original function $f(\theta, \phi)$ ($f \in \{x, y, z\}$). We can use this property to calculate the new SPHARM model $\mathbf{v}'(\theta, \phi)$ for the object surface after a rotated parametrization, and we only need the old coefficients $\{c_{lx}^m, c_{ly}^m, c_{lz}^m\}$ and rotation matrices $D_{mm'}^l(\alpha\beta\gamma)$.

2.3 Surface Correspondence Difference Measure

The surface correspondence alignment problem is generally formulated in terms of the optimal parameters, such as (α, β, γ) , that minimize some surface distance function. In this paper, we adopt the Euclidean distance as the distance function between surfaces. Formally, for two surfaces given by $\mathbf{v}_1(s)$ and $\mathbf{v}_2(s)$, their distance $D(\mathbf{v}_1, \mathbf{v}_2)$ is defined as [6]

$$D(\mathbf{v}_1, \mathbf{v}_2) = \left(\sum_{f \in \{x, y, z\}} \sum_{l=0}^L \sum_{m=-l}^l (c_{lf_1}^m - c_{lf_2}^m)^2 \right)^{1/2}. \quad (7)$$

2.4 Fast Surface Alignment Algorithm

We use a sampling-based search algorithm, which fixes one parameter mesh and rotates the other to carry out a greedy search on its surface with a small step size, to align the surface correspondence by minimizing the surface distance defined in Eq. (7). First, we uniformly distribute sample points on the surface as the new north poles locations, and the new SPHARM coefficients can be calculated by using Eq. (2) with new Euler angles set (α, β, γ) . For each candidate north pole, the parameter mesh is counterclockwise rotated along the new north-south axes (the rotation angle ω ranges from 0 to 2π). In order to calculate the coefficients of the new rotated parameter mesh using Eq. (2), we must transform the rotation angle ω into the Euler angles (α, β, γ) .

The original north and south poles of a surface's parameter mesh are mapped onto the axis $e_3 = (0\ 0\ 1)$ and $-e_3 = (0\ 0\ -1)$ in the x -, y -, and z -sphere. After rotation using Euler angles $(\alpha_p, \beta_p, \gamma_p)$, the north and south pole coordinates switch from $v(\theta, \phi)$ to $v'(\theta, \phi) = R_{obj}(\alpha_p, \beta_p, \gamma_p)v(\theta, \phi)$, ($\theta = 0$ or π). Simultaneously the axis e_3 in the coordinate systems of the three mapping spheres is changed to $\hat{e}_3 = R_{ZZZ}(\alpha_p, \beta_p, \gamma_p)(0\ 0\ 1)^T$. Because \hat{e}_3 also contains the origin and has unit length direction, we apply the Rodrigues' rotation formula [13] for computing the rotation matrix $R_{\hat{e}_3} \in SO(3)$ corresponding to a rotation by an angle ω about the fixed axis \hat{e}_3

$$R_{\hat{e}_3}(\omega) = I + S \sin\theta + S^2(1 - \cos\theta), \quad S = \begin{bmatrix} 0 & -\hat{e}_{3x} & \hat{e}_{3y} \\ \hat{e}_{3z} & 0 & -\hat{e}_{3x} \\ -\hat{e}_{3y} & \hat{e}_{3x} & 0 \end{bmatrix},$$

where I is the identity matrix. We can obtain the Euler angles (α, β, γ) by solving the equation $R_{ZZZ}(\alpha\beta\gamma) = R_{\hat{e}_3}(\omega)$. These Euler angles can then be used to calculate the coefficients of new parameter mesh using Eq. (2).

In the second step, we use the BFGS algorithm [14] to locally minimize Eq. (7) starting from the result of the first step. Because the result of the first step is already close to the target, this step generally needs only a few iterations. Although the dimension of the Jacobian matrix is large, the matrix is quite sparse. The computational time of this step is very low.

3 Experiments and Discussions in Medical Image Analysis

The fast alignment algorithm for surface correspondence described above was used for shape analysis in selected medical image analysis applications. Based on segmented MRI data of heart, we use the SPHARM method to do surface reconstruction and apply the surface alignment algorithm presented in this paper to determine a correspondence between shapes. This aligned correspondence allows researchers to access more functional details.

3.1 Comparison of Methods

In previous shape analysis study using the SPHARM description [5], researchers choose to use the three major axis of the first order ellipsoid (which is computed from the

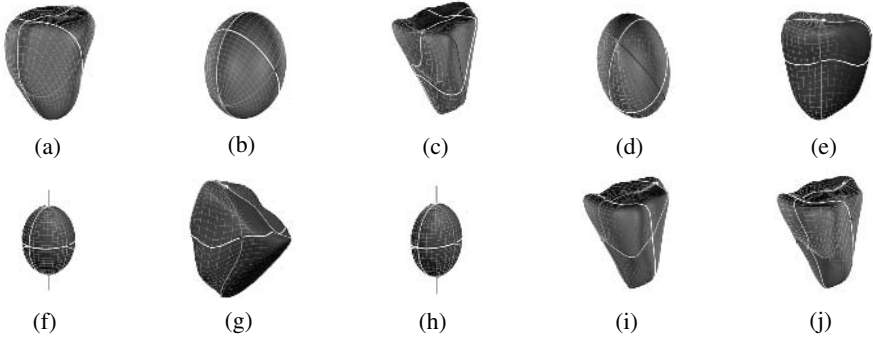


Fig. 1. Comparison of methods: (a) shows the reconstructed SPHARM surface of left ventricle, (b) is the first order ellipsoid of surface (a); (c) shows the reconstructed SPHARM surface of right ventricle, (d) is the first order ellipsoid of surface (c). By using the previous method, the first order ellipsoids and parametrizations are rotated to the positions in (f) and (h), and the SPHARM surfaces and parametrizations are rotated as (e) and (g). By using our algorithm, (i) shows the result of poles alignment. North and south poles are aligned close to the poles of (a). And the parameter mesh is rotated along the north pole. After using the BFGS algorithm [14] in the second step, the last alignment result is shown in (j).

first order SPHARM coefficients) as the intrinsic coordinate system. Parametrization is rotated in the parameter space for normalization so that three main ridges of the first order ellipsoid are moved to the equator [5,6]. Their method works well if two or more objects have a similar orientation (e.g., aligning hippocampal shapes). However this method may not work in some cases.

Fig. 1.(a) and Fig. 1.(c) show the reconstructed surface of two ventricles of the heart (left ventricle and right ventricle). We separate the parametrization on the surface into eight regions using five lines ($\theta = \pi/2$ in middle line, in north pole $\theta = 0$, in south pole $\theta = \pi$, the other four lines separately represent $\phi = 0, \pi/2, \pi, 3\pi/2, 2\pi$). The correspondences between surfaces in Fig. 1.(a) and Fig. 1.(c) are unordered as the visualizations. Fig. 1.(b) and Fig. 1.(d) show their first order ellipsoids. By using the previous method, the first order ellipsoids and parametrizations are rotated to the positions in Fig. 1.(f) and Fig. 1.(h). Three main directions of the ellipsoids are moved to the equator. The surface correspondence is created when the first order ellipsoid is aligned. As the result, the SPHARM surfaces and parametrizations should be rotated as Fig. 1.(e) and Fig. 1.(g). A limitation of this approach is that it may not represent the real surface correspondence between two surfaces. The reason for this is that the left ventricle and right ventricle have two very different orientations of their first order ellipsoid that are obvious in Fig. 1.(b) and Fig. 1.(d). Thus, although the first order ellipsoids are rotated to the normalized positions, the surfaces are rotated to the opposite orientations.

Our new alignment algorithm produces a correct alignment in these cases, because it is a general surface alignment method that does not depend on any orientation information. Fig. 1.(i) and Fig. 1.(j) show the results generated by our algorithm. Fig. 1.(a) is the fixed surface and the parametrization in Fig. 1.(c) is rotated to Fig. 1.(j).

The effectiveness of our algorithm can also be demonstrated by computing the surface correspondence distance defined in Eq. (7). The surface correspondence distance

between surfaces in Fig. 1.(e) and Fig. 1.(g) is $258.6536mm$, but the surface correspondence distance between surfaces in Fig. 1.(a) and Fig. 1.(j) is $62.4798mm$. Our surface alignment algorithm derives a better result.

3.2 Alignment for Temporal Heart Sequences

This new surface alignment algorithm also provides a promising method for studying spatio-temporal structures. In the previous research [15], surface tracking techniques (tracking points on 3D shape using 2D images) are used to create temporal sequence descriptions for points on the left ventricle inner surface through each heart cycle. Such temporal sequence descriptions can quantify the ventricular mechanical asynchrony or synchrony, which has important diagnostic and prognostic values, and can help determine optimal treatment in heart failures where a heart has a highly asynchronous contraction. Because the points are tracked on 2D images and mapped to a 3D surface, this method can only describe the heart contraction and dilation along the plane direction, and is not accurate for the perpendicular direction.

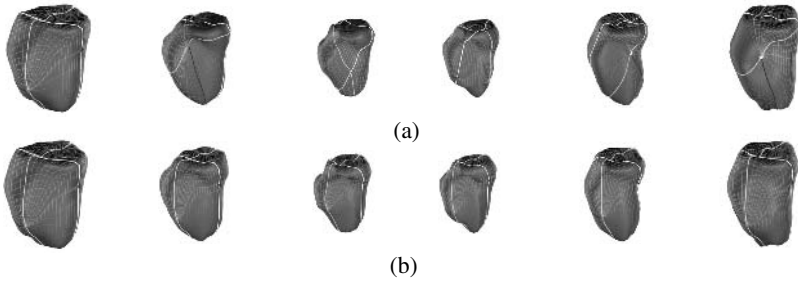


Fig. 2. (a) shows a shape sequence of a left ventricular inner surface during one heart cycle before surface alignment. The shape sequence in (b) is the result after surface alignment.

Combining the SPHARM description and our surface alignment methods offers a set of spatio-temporal surface correspondences for medical image analysis research. Our new algorithm generates more reasonable surface correspondences for the left ventricle sequence, and these surface correspondences describe the heart contraction and dilation in every direction of 3D space. Based on this new model, more valuable diagnostic and prognostic information can be derived for helping make clinical determinations.

Fig. 2.(a) is a shape sequence of a left ventricular inner surface during one heart cycle. Before surface alignment, the parametrization of every surface is unordered. The shape sequence in Fig. 2.(b) is the result after surface alignment. During the alignment procedure, every shape is aligned with its anterior shape.

4 Conclusions

This paper addresses the problem of finding surface correspondences between SPHARM parametric surfaces. We propose a theorem based on the SPHARM rotational property

to easily calculate SPHARM coefficients of the new parameter mesh along the surface. The distance between surfaces is defined as the objective function. Its efficacy is demonstrated in experiments based on several medical research problems, where we observe a significant improvement in robustness relative to existing shape modeling and analysis techniques.

There are several future directions. 1) The current algorithm can be integrated into an ICP-like framework for registration of 3D parametric surfaces. After finding the corresponding points on the surfaces, we can use an ICP-like algorithm to minimize the value of the correspondence. 2) Considerable research has focused on generating an equal area parametric mesh for a given surface. Our alignment algorithm can be combined with such parametrization methods to produce a better surface correspondence. These results are very useful for the shape and functional analysis in medical imaging.

References

1. Chen Chang Wen, Huang, T.S., Arrott, M.: Modeling, analysis, and visualization of left ventricle shape and motion by hierarchical decomposition. *IEEE PAMI* **16(4)** (1994) 342–356.
2. Matheny, A., Goldgof, D.B.: The use of three- and four-dimensional surface harmonics for rigid and nonrigid shape recovery and representation. *IEEE PAMI* **17(10)** (1995) 967–981.
3. Burel, G., Hennocq, H.: Determination of the Orientation of 3D Objects Using Spherical Harmonics. *Graphical Models and Image Processing* **57(5)** (1995) 400–408.
4. Brechbühler, Ch., Gerig, G., and Kübler, O.: Parametrization of closed surfaces for 3D shape description. *Computer Vision and Image Understanding* **61(2)** (1995) 154–170.
5. Styner, M., Gerig, G.: Three-Dimensional Medial Shape Representation Incorporating Object Variability. *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (2002) 651–656.
6. Gerig, G., Styner, M.: Shape versus Size: Improved Understanding of the Morphology of Brain Structures. *International Conference on Medical Image Computing and Computer Assisted Intervention LNCS 2208* (2001) 24–32.
7. Styner, M., Lieberman, A.J., Pantazis, D., and Gerig, G.: Boundary and Medial Shape Analysis of the Hippocampus in Schizophrenia. *Med. Im. Ana. MEDIA*. **8(3)** (2004) 197–203.
8. Huang, H., Shen, L., Ford, J., Makedon, F., Zhang, R., *et al.*: Functional analysis of cardiac MR images using SPHARM modeling. *Proceedings of the SPIE* **5747** (2005) 1384–1391.
9. Besl, P. J., McKay, N. D.: A method for registration of 3-D shapes. *IEEE Trans. on Pattern Analysis and Machine Intelligence* **14(2)** (1992) 239–256.
10. Chirikjian, G.S., Kyatkin, A.B.: *Engineering Applications of Noncommutative Harmonic Analysis: With Emphasis on Rotation and Motion Groups*. CRC Press (2000).
11. Biedenharn, L.C., Louck, J.C.: *Angular Momentum in Quantum Physics*. Addison-Wesley: Reading MA (1981).
12. Ritchie, D.W., Kemp, G.J.L.: Fast computation, rotation, and comparison of low resolution spherical harmonic molecular surfaces. *J. Comp. Chem.* **20(4)** (1999) 383–395.
13. <http://mathworld.wolfram.com>
14. Fletcher, R.: *Practical Methods of Optimization*. Princeton University Press, John Wiley and Sons, 2nd edition (1987).
15. Huang, H., Shen, L., Makedon, F., Zhang, S., Greenberg, M., Gao, L., Pearlman, J.: A clustering-based approach for prediction of cardiac resynchronization therapy. *ACM Symposium on Applied Computing*. (2005) 260–266.

Unified Point Selection and Surface-Based Registration Using a Particle Filter

Burton Ma and Randy E. Ellis

School of Computing, Queen's University at Kingston, Canada K7L 3N6
mab@cs.queensu.ca, ellis@cs.queensu.ca

Abstract. We propose an algorithm for jointly performing registration point selection and interactive, rigid, surface-based registration. The registration is computed using a particle filter that outputs a sampled representation of the distribution of the registration parameters. The distribution is propagated through a point selection algorithm derived from a stiffness model of surface-based registration, allowing the selection algorithm to incorporate knowledge of the uncertainties in the registration parameters. We show that the behavior of target registration error improves as the quality measure of the registration points increases.

1 Introduction

One method of registering a patient to preoperative 3D medical images for use in image-guided surgery is to digitize anatomical registration points and match them to surface models derived from the images. A problem that has gone largely unexplored in this framework is how to intraoperatively guide the surgeon to regions of the anatomy that might contain good registration points.

A few selection algorithms have been described that could be applied to preoperatively choosing registration point sets. Simon [13] proposed to choose surface model points that maximized the noise amplification index [11] of a 6×6 scatter matrix. Gelfand et al. [2] used the condition number of Simon's scatter matrix to optimize ICP registration stability when aligning pairs of surfaces with significant overlap. Uniform sampling of the sphere of normal vectors was proposed by Rusinkiewicz and Levoy [12] as a method of point selection for registering pairs of range images. We [9] showed that Simon's scatter matrix was equivalent to the well known spatial-stiffness matrix of an unactuated, kinematically unconstrained elastic mechanism; we proposed a point-selection algorithm based on the coordinate frame invariant analysis of Lin et al. [8].

The problem with applying these kinds of selection algorithms to online guidance is that they all attempt to optimize some criteria based on points and surface normals defined in the model coordinate frame, but these features are uncertain during the course of a registration. If the true registration was known, then the model features could be inferred by finding the model points that corresponded to the patient registration points. One possible solution is to use the current set of N patient points and a registration algorithm to estimate the corresponding model points, and then use the selection algorithm to suggest the next best model registration point. A possibly superior approach is to use a registration algorithm that estimates the uncertainties of the registration parameters

and propagates the uncertainties through the selection process. If the algorithm represents the uncertainties as a covariance matrix, for instance Lavallée et al. [7], then the uncertainties can be propagated as a finite set of estimates of the registration parameters using a technique such as the unscented transform of Julier and Uhlmann [6].

In this article we unify our point selection algorithm [9] with our sampling-based registration algorithm [10] to produce a filter-selection algorithm that estimates the distribution of the registration parameters (instead of only the covariance) and computes which regions of the model are likely to contain good registration points.

2 Stiffness-Based Point Selection

We [9] proposed an algorithm for sequentially constructing a set of model registration points by greedily maximizing a quality measure derived from a stiffness model of shape-based registration. Our approach considers the N registration points \mathbf{p}_i to be the attachment locations for unloaded linear springs with orientations given by the surface normal vectors \mathbf{n}_i . This configuration of springs leads to a stiffness relationship

$$\mathbf{w} = \mathbf{K}\mathbf{t} \tag{1}$$

$$\begin{array}{l} \text{force} \\ \text{torque} \end{array} \begin{bmatrix} \mathbf{f} \\ \boldsymbol{\tau} \end{bmatrix} = \left(\sum_{i=1}^N \begin{bmatrix} \mathbf{n}_i \\ \mathbf{p}_i \times \mathbf{n}_i \end{bmatrix} \begin{bmatrix} \mathbf{n}_i \\ \mathbf{p}_i \times \mathbf{n}_i \end{bmatrix}^T \right) \begin{bmatrix} \boldsymbol{\nu} \\ \boldsymbol{\omega} \end{bmatrix} \begin{array}{l} \text{linear displacement} \\ \text{rotational displacement} \end{array}$$

where \mathbf{w} is the wrench representing the reaction force and torque, and \mathbf{t} is the twist representing the applied linear and rotational displacement.

Simon [13] analyzed \mathbf{K} by examining its singular values, which are not frame invariant and are not directly comparable to one another because rotational and translational displacements have differing units. Frame-invariant quantities, called the principal translational and rotational stiffnesses, can be calculated from \mathbf{K} (Lin et al. [8]); moreover, by considering a task-specific point target, the rotational stiffnesses can be scaled so that they are directly comparable to the translational stiffnesses. The scaled stiffnesses are called the equivalent rotational stiffnesses, and a stiffness quality measure is $Q = \min(\mu_{\text{eq},1}, \mu_{\text{eq},2}, \mu_{\text{eq},3}, \sigma_1, \sigma_2, \sigma_3)$ where the $\mu_{\text{eq},i}$ are the equivalent rotational stiffnesses and the σ_i are the translational stiffnesses. Q characterizes the least constrained displacement of the mechanism; maximizing Q will minimize the worst-case displacement of the mechanism. Our algorithm [9] takes as input a set of N model registration points \mathbf{p}_i with normal vectors \mathbf{n}_i and a surface model from which to select points. The quality measure Q is calculated and heuristics are used to find the model point \mathbf{p}_{N+1} that maximizes the increase in Q . At least six point/normal vector pairs are required by the selection algorithm because \mathbf{K} does not have full rank otherwise.

3 Distribution of Registration Parameters

We [10] proposed an algorithm that used a particle filter described by van der Merwe et al. [15] to estimate the distribution of the registration parameters. The registration

parameters are represented by a state vector $\mathbf{x}_t = [\theta \ \phi \ \delta \ dx \ dy \ dz]_t^T$ where θ , ϕ , and δ are the three rotation parameters, dx , dy , and dz are the three translation parameters, and t is the time index equal to the current number of registration points processed. The patient registration point acquired at time t is treated as a control input \mathbf{u}_t . The state space model we used was

$$\begin{aligned} \mathbf{x}_{t+1} &= \mathbf{x}_t + \mathcal{N}(\mathbf{0}, \mathbf{Q}_t) \\ \mathbf{y}_t &= \begin{bmatrix} \mathbf{r}(\theta_t, \phi_t, \delta_t)(\mathbf{u}_1 + [dx_t \ dy_t \ dz_t]^T) \\ \vdots \\ \mathbf{r}(\theta_t, \phi_t, \delta_t)(\mathbf{u}_t + [dx_t \ dy_t \ dz_t]^T) \end{bmatrix} + \mathcal{N}(\mathbf{0}, \mathbf{R}_t) \end{aligned} \quad (2) \quad (3)$$

Equation 2 is called the process model and it describes how the state evolves in one time step; the model has a time-invariant state, except for the additive process noise, because the registration state is constant. The zero-mean Gaussian process noise with covariance \mathbf{Q}_t allows the filter to move from initial estimates of the state to successively better estimates. \mathbf{Q}_t is annealed towards $\mathbf{0}$ over time as the estimates become better.

The observation model is given by Equation 3 where $\mathbf{r}(\theta_t, \phi_t, \delta_t)$ is a rotation about a remote center and $[dx_t \ dy_t \ dz_t]^T$ is a translation vector. We use rotational parameters that surgeons are most familiar with, measured as order-independent angles of rotation in the coronal, sagittal, and transverse planes. The form of the rotation matrix $\mathbf{r}(\theta, \phi, \delta)$ can be found in Iyun et al. [5]. In this article, we consider $[dx_t \ dy_t \ dz_t]^T$ to be the translation to a remote center of rotation. The model is simply the estimated registration transformation applied to the patient registration points concatenated into a single vector; the length of the vector at time t is $3t$. We assume additive, zero-mean Gaussian noise with covariance \mathbf{R}_t ; the noise is the displacement of each transformed registration point to the surface of the model. Like the process noise covariance \mathbf{Q}_t , the measurement noise covariance \mathbf{R}_t must be annealed.

The output of the particle filter is a set of P equally weighted samples (particles) representing the posterior distribution of the registration parameters. To conserve space in this article we refer the reader to the article by van der Merwe et al. [15] for details on the mechanics of the specific particle filter we use. Note that the particle filter is not restricted to using Gaussian noises.

3.1 Establishing an Initial Distribution

A prior distribution \mathbf{x}_0 needs to be specified for the particle filter. Strictly speaking, the prior is supposed to be independent of the observations. A non-informative prior (a uniform distribution over the 6-dimensional state space) is impractical because an unwieldy number of particles would be required to adequately sample the space. We take a pragmatic approach and estimate an initial distribution using the first four registration points. The first registration point is used to estimate the translation to a remote center of rotation and the remaining three points are used to estimate the rotation. We assume that each of the first four points comes from a predefined region of the accessible surface of the bone. A small number of samples s_1 , s_2 , s_3 , and s_4 are drawn from the four

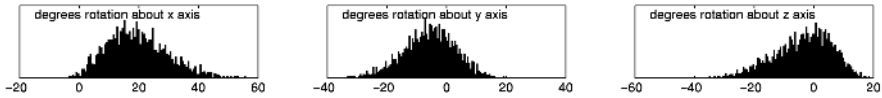


Fig. 1. Distributions of rotation parameters (θ , ϕ , and δ from left to right) obtained using the first four registration points

regions and the patient registration points are registered to all $s_1 \cdot s_2 \cdot s_3 \cdot s_4$ configurations of model registration points. This sets the number of particles as $P = s_1 \cdot s_2 \cdot s_3 \cdot s_4$ and yields an initial distribution from which we can start the filter. The resulting distributions (see example in Figure 1) are surprisingly wide in the range of rotation parameters.

4 A Unified Filter-Selection Algorithm

Our filter-selection algorithm uses the output of a particle filter registration algorithm to propagate the distribution of registration parameters through a point selection process. Each of the P registration estimates (particles) is applied to the set of patient registration points and the corresponding model points are inferred by performing a nearest neighbor calculation; correspondences weighted by distance (for instance, [1], [3], or [4]) could also be used. Each of the P sets of model correspondences is used as input to the point selection algorithm that calculates the increase δQ in quality measure Q for each of the M model points. This calculation requires $O(PM)$ time and is the major drawback of this algorithm. If we stored the distribution of δQ for each model point then we would also require $O(PM)$ storage; we compute only the mean increase ΔQ for each model point which requires $O(M)$ storage.

If we examine the model points with mean quality measure higher than a certain value—for example, the 95th percentile—then what we find is that those model points tend to come from one or more continuous regions of the model. This presents the surgeon with the task of locating any point inside a region, which is much easier than trying to locate a single point. Examples of point-selection regions are shown in Figure 2. The unified filter-selection algorithm is shown in Figure 3.



Fig. 2. Regions (in black) with mean quality measure greater than the 95th percentile after 7, 8, 9, and 10 (from left to right) registration points have been processed using the algorithm described in Figure 3

- Digitize 6 registration points \mathbf{U} so that the stiffness matrix \mathbf{K} has full rank.
- For $i = 6, 7, \dots, N$ (main loop)
 - Update the particle filter to obtain P registration samples $\{\mathbf{x}_1, \dots, \mathbf{x}_P\}$.
 - Initialize $\Delta Q[1..M] = 0$ where M is the number of model points. $\Delta Q[m]$ is the mean increase in the quality measure for the m^{th} model point, where the mean is calculated over the P particles.
 - For $j = 1, 2, \dots, P$
 - * Apply the estimated registration transformation to \mathbf{U} : $\mathbf{Y} = \mathbf{T}(\mathbf{x}_j)\mathbf{U}$.
 - * Find the model points \mathbf{P} closest to \mathbf{Y} .
 - * Compute the quality measure Q of \mathbf{P} .
 - * For $k = 1, 2, \dots, M$
 - $\mathbf{p}_k = k^{\text{th}}$ model point
 - $\mathbf{n}_k =$ normal vector at \mathbf{p}_k
 - $\Delta Q[k] = \Delta Q[k] + \delta Q(\mathbf{p}_k, \mathbf{n}_k)/P$ where $\delta Q(\mathbf{p}_k, \mathbf{n}_k)$ is an approximation of the increase in quality measure if point \mathbf{p}_k with normal \mathbf{n}_k is added to the current set of estimated model registration points \mathbf{P} ; see [9] for details.
 - Find the model points with ΔQ in the range $[lo, hi]$ (lo and hi might be the 95th and 100th percentiles, for example) defining regions on the model that contain good registration points. Digitize a patient point \mathbf{u}_{i+1} corresponding to a point inside these regions.
 - $\mathbf{U} = \{\mathbf{U}, \mathbf{u}_{i+1}\}$.

Fig. 3. A filter-selection algorithm

5 Materials and Methods

Bone-surface models of a distal radius, proximal femur, and proximal tibia were used to validate the proposed filter-selection algorithm. The models were derived from CT scans of patient volunteers. The surgically accessible area and a target were defined on each model. These features are shown in Figure 4. Four regions were defined on each phantom from where the initial four points needed to initialize the particle filter were drawn from; an additional two large regions were defined on each phantom to yield the requisite six points for the selection phase of the algorithm. The number of particles used was $7^4 = 2401$ for the radius and tibia, and $8 \times 7^3 = 2744$ for the femur. The simulations were run on a PC with an AMD Athlon 2400 CPU and 256MB of RAM.

Registration simulations were run using each model with 200 trials, each trial executing the steps shown in Figure 5. The total number of registration points N was

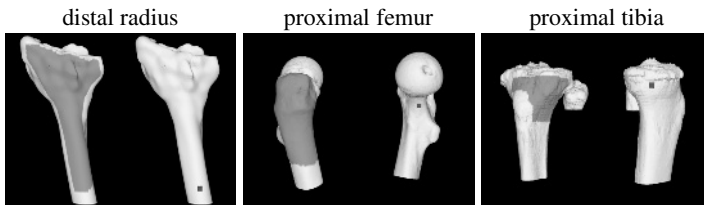


Fig. 4. Features used in the experimental validation of integrated point selection and registration. Accessible regions are shown in gray and targets are shown as cubes.

1. Randomly select registration points from the six predefined regions; $\mathbf{U} = \{\mathbf{u}_1, \mathbf{u}_1, \dots, \mathbf{u}_6\}$.
2. Add Gaussian noise $\mathcal{N}(0, (0.35\text{mm})^2)$ to the x , y , and z components of the points in \mathbf{U} .
3. Displace \mathbf{U} by the inverse of the randomly selected registration transformation.
4. Calculate the initial state distribution (Section 3.1).
5. For $i = 6..N$
 - (a) Apply one iteration of the main loop of the filter-selection algorithm (Figure 3).
 - (b) Add Gaussian noise drawn from $\mathcal{N}(0, (0.35\text{mm})^2)$ to the new registration point \mathbf{u}_{i+1}

Fig. 5. Simulation steps for the experiment described in Figure 3

20. We took \mathbf{Q}_t to be uncorrelated with initial variances of $(3^\circ)^2$ and $(3\text{mm})^2$ for the rotational and translational components. We took \mathbf{R}_t to be uncorrelated with initial variances of $(2\text{mm})^2$ in each of the x , y , and z components of the observation vector. Because we used a relatively small number of registration points, the annealing factors were both chosen to be the relatively small value of 0.7 and we annealed for a total of 10 time steps. Three different ranges of $[l_o, h_i]$ were used to determine if point selection had an effect on registration precision and accuracy; the ranges used, in percentiles of δQ , were $[50, 55]$, $[75, 80]$, and $[95, 100]$. The component-wise median of the filter distribution was used as the estimate of the registration state.

6 Results

The mean of the 200 estimated registration states was within 0.2° in each of the true rotation states, and within 0.1mm of the true translation states; this was true regardless of the model used or the range of $[l_o, h_i]$. This strongly suggests that there is no bias in the filter estimates of registration. The standard deviations of the 200 estimated registration states are shown in Table 1 for each model and range of $[l_o, h_i]$. There was a trend towards decreasing variance as points that lead to increasing ΔQ values were used. This is more clearly seen in Figure 6 which shows the distributions of target registration error (TRE). There was almost no difference in the distribution of TRE between point sets with ΔQ values of $[50, 55]$ and $[75, 80]$. There was clear improvement in TRE

Table 1. Standard deviations of the registration state parameters

Model	$[l_o, h_i]$	θ_x	θ_y	θ_z	t_x	t_y	t_z
radius	$[50, 55]$	0.57°	0.58°	1.57°	0.47mm	0.27mm	0.45mm
	$[75, 80]$	0.50°	0.55°	1.55°	0.48mm	0.21mm	0.40mm
	$[95, 100]$	0.38°	0.38°	1.22°	0.47mm	0.18mm	0.35mm
femur	$[50, 55]$	0.95°	0.56°	1.32°	0.23mm	0.48mm	0.38mm
	$[75, 80]$	0.70°	0.52°	1.22°	0.27mm	0.49mm	0.38mm
	$[95, 100]$	0.87°	0.55°	0.91°	0.19mm	0.32mm	0.41mm
tibia	$[50, 55]$	1.59°	0.82°	1.05°	0.45mm	0.36mm	0.79mm
	$[75, 80]$	1.49°	0.87°	0.97°	0.47mm	0.40mm	0.64mm
	$[95, 100]$	1.09°	0.71°	0.73°	0.34mm	0.32mm	0.51mm

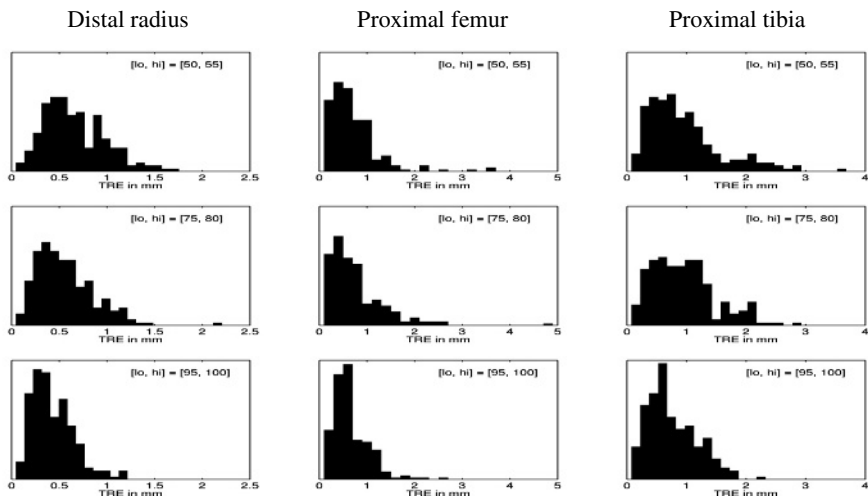


Fig. 6. Distribution of target registration errors over 200 trials for each model and range of ΔQ values $[l_o, h_i]$

when using the highest range, $[95, 100]$, of ΔQ values; however, the actual decrease in expected TRE was very small. The computation time to process 20 registration points was between two and three minutes on a relatively low-cost PC.

7 Discussion

Estimating the translation to a remote center of rotation appears to be an effective state-space model for rigid registration. This produces a better initial distribution than the random sampling approach taken in [10]; we used fewer particles and obtained more accurate registrations. Granger and Pennec [3] noted that their EM-ICP registration algorithm was most sensitive to errors in the translational component of the initial transformation estimate. This appears to be true for the particle-filter registration method as well. Figure 1 shows that the range of the rotation prior is at least 60° in each component for the tibia model. If this observation is true in general, then the initial distribution could be established using a single region to estimate the translation to the remote center of rotation; the rotation parameters could be estimated by having the surgeon hold a tracked target to approximately indicate the directions of the anatomic axes.

One feature of online point selection is that the surgeon is given a clear visualization of where good registration points are likely located. This may be an advantage when the surgeon has little experience performing registration, such as when attempting a particular procedure for the first time, or during adoption of image-guided surgery.

Both the particle filter and the point-selection algorithm can exploit parallel computation, and this would be one way to provide integrated point selection at an acceptable speed. A better approach might be to replace the particle filter with a more efficient type of filter. In particular, the family of Gaussian mixture sigma-point particle (GMSPP) filters described by van der Merwe and Wan [14] are reportedly much more efficient in

terms of computation than particle filters. These filters represent the posterior using a Gaussian mixture model. Because there is a deterministic way to draw a small number of samples from a Gaussian distribution (for example, see Julier and Uhlmann [6]) that preserves all of the relevant information, only a small number of particles might need to be considered during point selection if a GMSPP filter is used.

This work shows that it is possible to unify surface-based registration, optimal selection of the next point from the patient's anatomy, and estimation of the uncertainty in the resulting registration. Optimization of the algorithm will bring it into the timeframe required in orthopedic surgery. Extensions to this work might include application to deformable atlas-based registration and to volumetric registration of one image to another.

References

- [1] H. Chui and A. Rangarajan. A feature registration framework using mixture models. In *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, June 2000.
- [2] N. Gelfand, L. Ikemoto, S. Rusinkiewicz, and M. Levoy. Geometrically stable sampling for the ICP algorithm. In *3DIM 2003*, Banff, Canada, Oct 2003. IEEE.
- [3] S. Granger and X. Pennec. Multi-scale EM-ICP: A fast and robust approach for surface registration. In *ECCV 2002*, pages 418–432, Copenhagen, Denmark, May 2002. Springer.
- [4] W. E. L. Grimson, G. J. Ettinger, S. J. White, T. Lozano-Pérez, W. M. Wells III, and R. Kikinis. An automatic registration method for frameless stereotaxy, image guided surgery, and enhanced reality visualization. *IEEE T Med Imaging*, 15(2):129–140, 1996.
- [5] O. Iyun, D. P. Borschneck, and R. E. Ellis. Computer-assisted correction of bone deformities. In *MICCAI 2002*, pages 232–240, Tokyo, Japan, 2002. Springer.
- [6] S. J. Julier and J. K. Uhlmann. A general method for approximating nonlinear transformations of probability distributions. Technical report, Oxford University, 1996.
- [7] S. Lavallée, R. Szeliski, and L. Brunie. Anatomy-based registration of three-dimensional medical images, range images, x-ray projections, and three-dimensional models using octree-splines. In *Computer-Integrated Surgery*, pages 115–143. The MIT Press, 1996.
- [8] Q. Lin, J. Burdick, and E. Rimon. A stiffness-based quality measure for compliant grasps and fixtures. *IEEE T Robotic Autom*, 16(6):675–688, 2000.
- [9] B. Ma and R. E. Ellis. Spatial-stiffness analysis of surface-based registration. In *MICCAI 2004*, pages 623–630, St. Malo, France, 2004. Springer.
- [10] B. Ma and R. E. Ellis. Surface-based registration with a particle filter. In *MICCAI 2004*, pages 566–573, St. Malo, France, 2004. Springer.
- [11] A. Nahvi and J. M. Hollerbach. The noise amplification index for optimal pose selection in robot calibration. In *IEEE ICRA*, Minneapolis, Minnesota, Apr 1996.
- [12] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *3DIM 2001*, pages 145–152, Québec City, Canada, May 2001. IEEE.
- [13] D. A. Simon. *Fast and Accurate Shape-Based Registration*. PhD thesis, Carnegie Mellon University, Pittsburgh, Pennsylvania, Dec 1996.
- [14] R. van der Merwe and E. Wan. Gaussian mixture sigma-point particle filters for sequential probabilistic inference in dynamic state-space models. In *ICASSP*, 2003.
- [15] R. van der Merwe, A. Doucet, N. de Freitas, and E. Wan. The unscented particle filter. Technical Report CUED/F-INFENG/TR 380, Cambridge Univ. Dept. of Eng., 2000.

Elastic Registration of 3D Ultrasound Images

Pezhman Foroughi¹ and Purang Abolmaesumi^{1,2}

¹ Department of Electrical and Computer Engineering, Queen's University, Canada

² School of Computing, Queen's University, Canada
purang@cs.queensu.ca

Abstract. 3D registration of ultrasound images is an important and fast-growing research area with various medical applications, such as image-guided radiotherapy and surgery. However, this registration process is extremely challenging due to the deformation of soft tissue and the existence of speckles in these images. This paper presents a novel intra-modality elastic registration technique for 3D ultrasound images. It uses the general concept of attribute vectors to find the corresponding voxels in the fixed and moving images. The method does not require any pre-segmentation and does not employ any numerical optimization procedure. Therefore, the computational requirements are very low and it has the potential to be used for real-time applications. The technique is implemented and tested for 3D ultrasound images of liver, captured by a 3D ultrasound transducer. The results show that the method is sufficiently accurate and robust and does not easily get trapped with local minima.

1 Introduction

Ultrasound is an easy-to-use, inexpensive and real-time imaging tool with no ionizing radiation which is widely used for diagnosis applications. Recently, there has been major interest in the employment of this imaging technique for intra-operative registration in radiotherapy and image-guided surgery. Furthermore, ultrasound registration could be used for measuring tissue mechanics properties. However, due to speckles and shadowing effects, the registration of this imaging modality is a challenging process. Previous attempts include the development of techniques for registration of ultrasound to the other image modalities such as CT and MRI [1–5], as well as intra-modality registration of these images [6–9].

Several methods have been used for elastic registration of ultrasound images. Shekhar *et al.* [6] investigated the registration of ultrasound volumes based on the mutual information measure. Median filter was found to be the proper smoothing filter for this method. The results of different transformation modes were presented. Krucker *et al.* [7] developed a subvolume-based algorithm for elastic ultrasound registration. The image was divided into subvolumes interactively and local rigid registrations were computed. Connectivity of the entire volume was ensured by global interpolation using thin-plate splines after each iteration. Xiao *et al.* [8] used a correlation-based approach for registration of 3D free-hand ultrasound images of the breast. Deformable registration of cardiac images has been investigated more specifically in [9, 10].

In this paper, we propose a novel technique for 3D elastic registration of ultrasound images. In our approach, the concept of attribute vector, introduced in [11], is employed to find corresponding voxels in the ultrasound volumes. The elements of the attribute vector are revised to meet the requirements of ultrasound images. Using the attribute vectors, some leading points are automatically selected in the moving image, and their corresponding points in the fixed image are found. Finally, the moving image is warped into the fixed image.

This paper is organized as follows. Section 2 talks about the methodology and details of the implementation. Experimental results on 3D liver images are presented in Section 3. Section 4 concludes the paper and presents the future work.

2 Method

The proposed algorithm is demonstrated in Figure 1. At the first step, both fixed and moving images are passed through a Gaussian filter. This filter is used to generate more robust attribute vectors by reducing the effect of speckle, while providing different scales of images by adjusting its variance. In the next step, the elements of attribute vectors are calculated in two different scales for each voxel in both fixed and moving images. After this, the leading points with the highest magnitude attribute vectors are selected and registered, and finally the moving image is warped to the fixed image. The following paragraphs explain these steps in details.

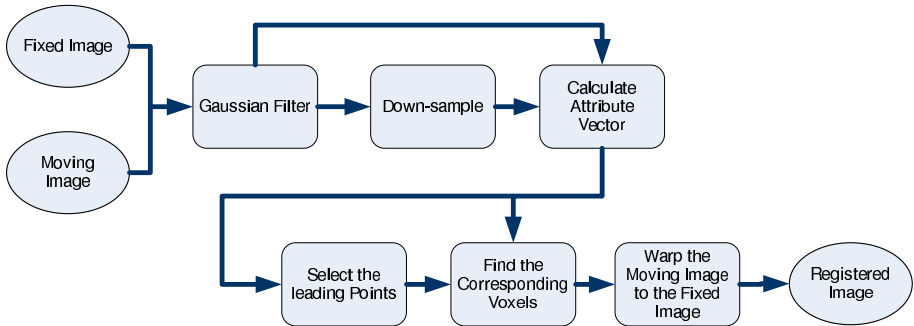


Fig. 1. The registration process

2.1 Attribute Vectors

Inspired by [12], we have used the concept of attribute vectors to extract corresponding feature points from the two ultrasound images. In [12] GMIs (geometric moment invariants) calculated from local spatial intensity histograms, form the attribute vector, which is assigned to each voxel in the image. Each attribute vector reflects the underlying anatomy and distinguishes between different features in an image.

Due to the presence of speckles in ultrasound images, our experiments show that GMIs do not generate robust attribute vectors, as they mainly characterize the speckle pattern. Depending on the properties of the imaging modality,

other elements could be selected for the attribute vector. Our proposed method replaces GMIs with a set of new elements explained in the following paragraph to consider the properties of ultrasound images.

The first element of our attribute vector is the intensity of the voxel, which guarantees that the bright voxels will not be registered to the dark ones and vice versa. The other two elements are the magnitude of the gradient and the second derivative (Laplacian of the Gaussian), which are selected to extract the boundary information of the features in the images. All the elements are calculated in two scales. The scale with higher variance reveals the more general specifications, and the one with lower variance depicts the more detailed features. The generated attribute vector is able to robustly find the correspondence across the images taken from the same anatomy at different times. All the elements are normalized between zero and one in order to fit the similarity function, described later. The variance of each element could also be used in the normalization process. The next step is to select the leading points which is described in the following section.

2.2 Leading Point Selection

One of the major advantages of our proposed technique is its fast computational speed. To this end, instead of searching for corresponding voxel points for all the voxels within the moving and fixed image data sets, the algorithm automatically selects distinct major features of the moving image and registers them to the fixed image. The displacement of the rest of the voxel points are simply calculated based on the major feature points' displacements. This selection process includes the calculation of an importance function for each voxel. The importance of voxel v is simply defined as:

$$im(v) = \sum_{i=1}^n (w_i \times e_i(v)) \quad (1)$$

where $e_i(v)$ is the i 'th element of attribute vector at point v , and n is the number of elements of attribute vector. w_i is the weight selected for i 'th element. The points that their importance function is greater than a threshold are selected as leading points. Our experiments show that the number of leading points is not very sensitive to the selected threshold. These leading points are processed in a descending order based on the values assigned from Equation (1) to each of them. Once a leading point is selected, no more points within a certain radius to that point are selected in order to decrease the number of leading points, and hence, increase the speed of registration. In the current implementation the radius size is set to 1.2 mm. This is done by setting the importance function around the selected point to zero, after its selection.

Registration of the selected points has lower risk of local minima, since they are usually located on the edge points in images. These boundary voxels have characteristics that are clearly distinguishable from neighbor areas using our proposed attribute vectors.

2.3 Matching

Let $e_i(v)$ be the i 'th element of attribute vector assigned to voxel v . Then, the similarity of two voxels v_1 and v_2 is measured with:

$$Sim(v_1, v_2) = \prod_{i=1}^n (1 - \|e_i(v_1) - e_i(v_2)\|)^{w_i} \quad (2)$$

where $Sim(s_1, s_2)$ shows the similarity of attribute vectors at points s_1 and s_2 , which is one for maximum similarity and zero for no similarity. w_i depicts the contribution of i 'th element of attribute vector in similarity function. If w_i 's are all set to one, all the elements of the attribute vector will have the same contribution in determining the similarity.

To find the corresponding voxels in the fixed and moving images, the minimum of an energy function is found in a search area around the coordinates of the voxel in the fixed image. The energy function uses the described similarity function for several points around the two center points. The size of the search area is application dependent, and is directly proportional to the amount of deformation between the two data sets being registered. The advantage of using a limited search area is to reduce the computational complexity of the algorithm, as well as reducing the possibility of getting trapped in local minima. Our energy function is similar to but not exactly the same as the ones defined in [8, 11]: Let p_i be the i 'th voxel in the neighborhood of the center point p , and q_j be the j 'th voxel in the neighborhood of the center point q . The energy function, when examining the correspondence of p and q , will be:

$$Eng(p, q) = \frac{\sum_i w(i)(1 - Sim(p_i, q_i))}{\sum_i w(i)} + \alpha d(p, q) \quad (3)$$

where w_i is a weighting function which is maximum in the center of search area, $d(p, q)$ is the distance of p from q , and α is a small coefficient. When the two sets of pixels p_i 's and q_i 's are similar, the energy function is minimized. The second term causes the energy function to be minimum at closer points to the center, when there is no absolute minimum. The reason for adding this term is that when there are two candidates for the corresponding voxel, the probability that the closer voxel is the right choice is higher.

Due to shadowing effects and the nature of ultrasound images, it is possible that a feature disappears from one image to the other. To face this phenomenon, when a match is found, the energy function at the matching point should be less than a threshold. Otherwise, that point is removed from the list of leading points. In all of our experiments for different data sets and deformations in the volumes, a single threshold value was used.

2.4 Warping

The final step of the registration process is to calculate the movement of all the voxels of the moving image with respect to the fixed image, based on the

calculated displacement of the leading feature points. The displacement of each voxel is mainly determined by the leading voxels that are close to that point. Displacement of each leading voxel is distributed to the points around it by a Gaussian function, which means that the displacement is faded through the neighbors. The variance of the function is set according to the volume scale factors. The faded displacement vectors of leading points could not be simply added together. Otherwise, since the selected leading voxels might be anywhere in the image, the displacement of the points between two corresponding leading voxels in the moving and the fixed images might get much larger than the original displacement. Therefore, the leading voxels contribute to the displacement of each point, inversely proportional to their distance to the point. The following equation is used in the warping process for voxel v :

$$Disp(v) = \frac{\sum_{i=1}^{n_{lead}} f_v(d_i)G(d_i)Disp(l_i)}{\sum_{i=1}^{n_{lead}} f_v(d_i)} \quad (4)$$

where l_i 's are the leading voxels, and n_{lead} is the number of them. d_i is the distance of v to l_i , and $f_v(d_i)$ is the inverse proportional function defined as $\frac{1}{d_i^2+\epsilon}$, where ϵ is a small positive number. G is the Gaussian function that fades the effect of far leading voxels. The inverse proportional function causes the leading points to be warped exactly to their corresponding voxels, found in previous step.

3 Results

Ultrasound images were captured by a GE Voluson 730 3D/4D ultrasound machine. The RAB4-8P transducer was used to collect 3D images from two volunteers' livers, and the images were saved in Cartesian format. These images were captured at multiple locations under different forces applied to the abdomen by various operator's hand pressures on the transducer. The deformation in ultrasound volume data sets was also caused by respiration, heartbeat and changes in the physical position of the volunteers. For each volunteer, five sets of images were captured. Each set of images, consisted of two volumes at two different moments of the breathing cycle, were taken from one part of the volunteer's liver. For each set, one of the volumes was selected as the fixed image, and the other one (moving image) was registered to it. The average displacement between the images in each set was estimated to be 8.7 voxels with a maximum displacement of 20.8 voxels (the scale factor was set to 0.6 mm/voxel).

The code has been implemented using MATLABTM. The matching and warping parts are written in C, and compiled with the C compiler of MATLABTM to reduce the computation time. Figures 2(a,b,c) demonstrate one slice of the fixed, moving and registered data sets. The difference images shown in Figures 2(d,e) depict the accuracy of the alignment of features after registration. It should be noted that the registration is performed in 3D, therefore some out-of-plane features, which may not exist in the demonstrated moving and fixed images,

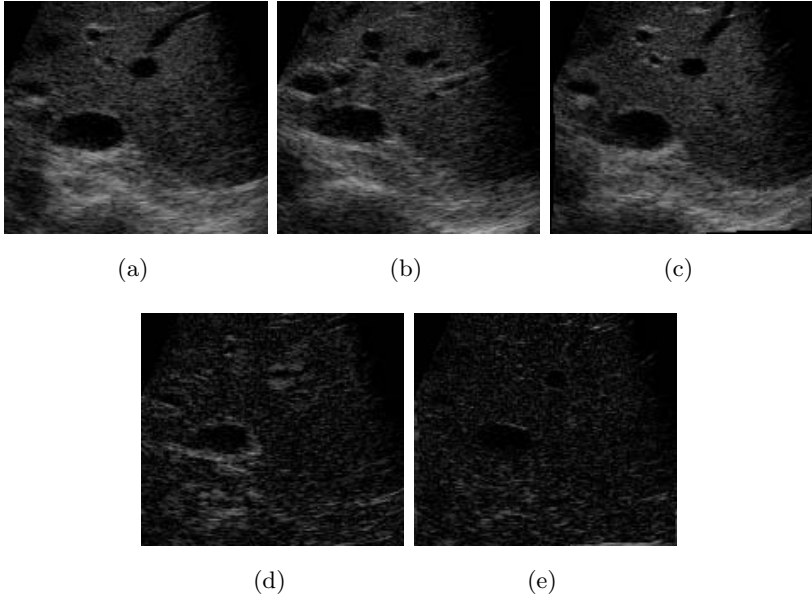


Fig. 2. (a) is a middle slice of the fixed volume, (b) is the same slice in the moving volume, and (c) is the same slice in the registered volume. (d) shows the absolute of difference between the fixed and the moving volumes, and (e) is the absolute of difference between the fixed and the registered volumes.

have been warped to the shown registered slice. Figure 3 shows the results of another experiment. Three slices in three orthogonal planes are demonstrated. The experiments show that the algorithm is robust, accurate and not sensitive to local minima, when the maximum displacement caused by deformation is less than the radius of the search area.

To evaluate the accuracy of the technique, two experiments were carried out. First, eight voxels on distinct features of the fixed and the registered volumes were manually selected. The distance between the corresponding voxels were then considered as the registration error. Using this method, the average error of registered images was about two voxels on average considering the uncertainties in the images and poor quality of them. In the second experiment, the matrices of displacement vectors generated in the previous experiments were used to simulate natural deformation. The original volumes were then registered to the deformed volumes, and the results were compared to the original displacement matrices. The average of mean error for 20 volumes deformed with this method was 1.50 voxels.

For two 185x113x199 voxels volumes, the program took about 436 seconds to finalize registration on a Pentium 4 2.4GHz with 512MB RAM machine. Pre-filtering and calculating the attribute vectors took about 74 seconds; The time required for selection of leading points and registering them was about 203 seconds, and warping took the rest of time.

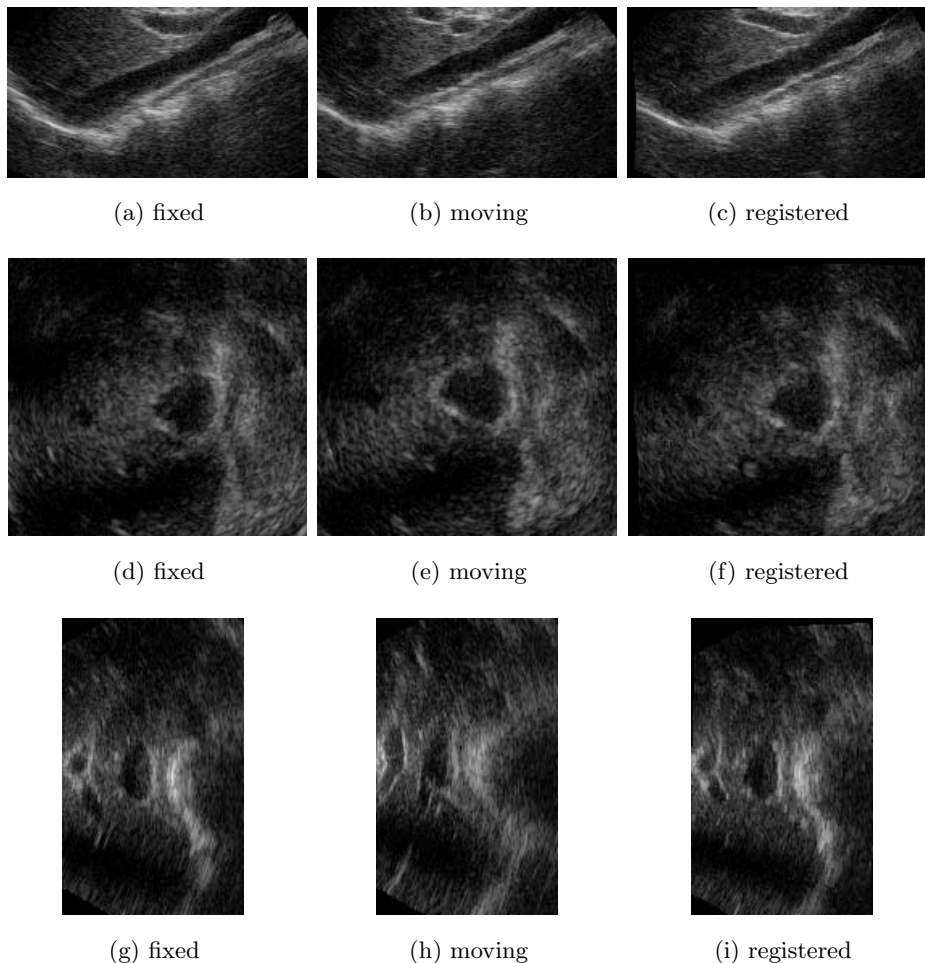


Fig. 3. (a),(b), and (c) are the slices of fixed, moving, and registered volumes, parallel to ZY plane; (d) to (f) are the slices parallel to ZX plane, and (g) to (i) are the slices parallel to XY plane

4 Conclusion and Future Work

We have developed a new method for 3D elastic registration of ultrasound images, which is robust and flexible and is intended eventually for intra-operative surgical applications. The algorithm does not require any pre-segmentation or numerical optimization, and as a result, the running time is decreased. There are a number of parameters, such as the size of the search area, that should be set correctly for the algorithm to function properly. However, as long as the search area size is larger than the deformation between the two image data sets, the performance of the technique is not very sensitive to the changes in these

parameters. Future work will include the optimization of the code to reduce the running time of the algorithm to within an acceptable range for intra-operative applications, as well as improving the normalization method of the elements of the attribute vector. The algorithm should also be tested for the registration of ultrasound images of other anatomical parts in addition to the liver.

References

1. Aylward, S.R., Jomier, J., Guyon, J.P., Weeks, S.: Intra-operative 3d ultrasound augmentation. In: *IEEE International Symposium on Biomedical Imaging*. (2002) 421–424
2. Pagoulatos, N., Edwards, W.S., Haynor, D.R., Kim, Y.: Interactive 3d registration of ultrasound and magnetic resonance images based on a magnetic position sensor. *IEEE Transactions on Information Technology in Biomedicine* **3** (1999) 278–288
3. King, A.P., Blackall, J.M., Penney, G.P., Hawkes, D.J.: Tracking liver motion using 3-d ultrasound and a surface based statistical shape model. In: *Mathematical Methods in Biomedical Image Analysis*. (2001) 145–152
4. Arbel, T., Morandi, X., Comeau, R.M., Collins, D.L.: Automatic nonlinear mri-ultrasound registration for the correction of intra-operative brain deformations. In: *MICCAI*. (2001) 565572
5. King, A.P., Blackall, J.M., Penney, G.P., Edwards, P.J., Hill, D.L.G., Hawkes, D.J.: Bayesian estimation of intra-operative deformation for image-guided surgery using 3-d ultrasound. In: *MICCAI*. (2000) 588597
6. Shekhar, R., Zagrodsky, V.: Mutual information-based rigid and nonrigid registration of ultrasound volumes. *IEEE Transactions on Medical Imaging* **21** (2002) 9–22
7. Krucker, J.F., LeCarpentier, G.L., Fowlkes, J.B., Carson, P.L.: Rapid elastic image registration for 3-d ultrasound. *IEEE Transactions on Medical Imaging* **21** (2002) 1384–1394
8. Xiao, G., Brady, M., Noble, J.A., Burcher, M., English, R.: Nonrigid registration of 3-d free-hand ultrasound images of the breast. *IEEE Transactions on Medical Imaging* **21** (2002) 405–412
9. Shekhar, R., Zagrodsky, V., Garcia, M.J., Thomas, J.D.: Registration of real-time 3-d ultrasound images of the heart for novel 3-d stress echocardiography. *IEEE Transactions on Medical Imaging* **23** (2004) 1141–1149
10. Makela, T., Clarysse, P., Sipila, O., Pauna, N., Pham, Q.C., Katila, T., Magnin, I.E.: A review of cardiac image registration methods. *IEEE Transactions on Medical Imaging* **21** (2002) 1011–1021
11. Shen, D., Davatzikos, C.: Hammer: hierarchical attribute matching mechanism for elastic registration. *IEEE Transactions on Medical Imaging* **21** (2002) 1421–1439
12. Shen, D.: Image registration by hierarchical matching of local spatial intensity histograms. In: *MICCAI*. (2004) 582–590

Tracer Kinetic Model-Driven Registration for Dynamic Contrast Enhanced MRI Time Series

Giovanni A. Buonaccorsi¹, Caleb Roberts¹, Sue Cheung¹, Yvonne Watson¹, Karen Davies¹, Alan Jackson¹, Gordon C. Jayson², and Geoff J.M. Parker¹

¹ Imaging Science and Biomedical Engineering,
University of Manchester, Manchester, UK
`giob@manchester.ac.uk`

² Cancer Research UK Dept of Medical Oncology,
Christie Hospital, Manchester, UK

Abstract. Motion during time-series data acquisition causes model-fitting errors in quantitative dynamic contrast-enhanced (DCE) MRI studies. Motion correction techniques using conventional registration cost functions may produce biased results because they were not designed to deal with the time-varying information content due to contrast enhancement. We present a locally-controlled, 3D translational registration process driven by tracer kinetic modeling that successfully registers abdominal DCE-MRI data at high temporal resolution and compare this method to a similar approach based on registration to the time series mean image in data from 8 patients. When the registration is driven by an appropriate model, we find significant improvements in model-fitting. Also, model-driven registration influences parameter estimates and reduces repeat study variability in measurements of blood volume.

1 Introduction

The quantitative, model-based, analysis of contrast agent uptake kinetics in dynamic contrast enhanced MRI (DCE-MRI) at high temporal resolution allows us to estimate the magnitude and spatial distribution of physiological kinetic parameters such as K^{trans} (the volume transfer coefficient of contrast agent between capillaries and the extravascular extracellular space — EES), v_e (the volume of the EES per unit volume of tissue) and v_p (the blood plasma volume per unit volume of tissue). These parameters provide useful information on microvascular status for studies of tumors and inflammatory conditions [1,2,3].

These parameters are usually estimated by deriving a contrast agent concentration time course and fitting a tracer kinetic model (e.g. [4,5]) within each voxel in a volume of interest (VOI). Model fitting errors arise because patient and physiological motion during data acquisition alters the voxel-to-tissue mapping. Conventional motion correction (e.g. registering each image volume in the time series to the time series mean [6]) has limited success because the passage of the contrast agent introduces new image features. The resulting time-varying information content is problematic even for inter-modal cost functions (e.g. mutual

information [7,8], correlation ratio [9]), which are most effective with a consistent information content between the input and reference images.

Other investigators have registered pre-contrast to post-contrast images [10] and incorporated a simplified tracer kinetic model directly into a registration cost function [11], but it has not yet been practicable to register an entire DCE-MRI time-series at high temporal resolution using a more detailed physiological model of enhancement in abdominal tumors. We propose a novel, tracer kinetic model-driven registration method that uses a computationally efficient, locally-controlled, 3D translational registration focused on the tumor VOI.

2 Method

2.1 Modelling

We used the standard [4] and extended [5] Kety models as applied to DCE-MRI data. The former assumes that the blood plasma volume, v_p , is small by comparison with the EES volume, v_e , and estimates K^{trans} and v_e only. This approximation fails in “rim-enhancing” tumors where angiogenesis may be pronounced [12], so in tumor studies it is usually more appropriate to use the extended Kety model, which includes a vascular term, v_p , and has the following integral form (note that the standard Kety model lacks the $v_p C_p(t)$ term):

$$C_t(t) = v_p C_p(t) + K^{trans} \int_{t'=0}^t C_p(t') \exp \left[\frac{-K^{trans} (t - t')}{v_e} \right] dt'. \quad (1)$$

In Equation 1, $C_t(t)$ is the tracer concentration in the tissue (mMol), which varies with time, t ; $C_p(t)$ is the tracer concentration in blood plasma (mMol) and K^{trans} is expressed in min^{-1} .

Each model requires the contrast agent concentration in the arterial supply, $C_p(t)$. This arterial input function (AIF) must be measured or a functional form must be assumed. The standard Kety model [4] used a bi-exponential AIF based on contrast agent concentrations measured from blood samples taken at intervals after injection of Gd-DTPA in healthy volunteers. However, this simple AIF may cause significant errors in the estimation of tracer kinetic parameters [13] and where possible we derive a measured AIF from the DCE-MRI time series [14].

The models are expressed in terms of a contrast agent concentration time course, $C_t(t)$. We compute this from the MR signal intensities using the standard relations for a spoiled gradient echo acquisition [15]. We estimate the parameters K^{trans} , v_e and v_p by fitting the above models to the measured data using a voxel-by-voxel simplex optimization [16].

2.2 Data Acquisition

We acquired abdominal images on a Philips 1.5 T Intera MRI scanner using the whole body coil for transmission and reception. The conversion to $C_t(t)$ used a baseline T_1 measurement comprising 3 axial spoiled gradient echo (Fast

Field Echo) volumes with 2° , 10° and 20° flip angles and 4 signal averages. The dynamic series comprised 75 consecutively-acquired axial volumes with a 20° flip angle, 1 average and a 4.97 s temporal resolution. All studies used 25 slices, $375 \times 375 \text{ mm}^2$ field of view, 128×128 matrix, 4.0 ms TR , and 0.82 ms TE . Slice thickness was 4 or 8 mm. Tumor VOIs were defined independently of the registration study by an expert health professional, based on co-localized T_2 -weighted and T_1 -weighted pre- and post-contrast image volumes.

Motion was graded (also independently of the registration study) on a 4 point scale. For registration, we selected 10 data sets from 8 patients with abdominal tumors who scored “moderate” or “severe” on this scale. In 5 patients we were also able to assess the reproducibility of fitted parameters via scan-rescan tests.

2.3 Registration

We assumed the tumors to be small relative to the abdomen, and to be more rigid than their surrounding tissues [17]. This allowed us to use a 3D translational registration procedure that calculated the registration cost function only within the tumor VOI but applied the resulting translation to the whole image volume. All registrations used FLIRT [18] (from the FSL package <http://www.fmrib.ox.ac.uk/fsl>) with the following iterative scheme.

Step 1: Fit the chosen model to the original, motion-corrupted, time series data within the tumor VOI to obtain initial estimates of K^{trans} , v_e and v_p .

Step 2: Output 3D maps (volumes) of the signal intensities generated by the model fit (back-converted from the modelled concentration values) for each time point in the dynamic series. These 3D maps are synthetic image volumes for which the intensity in each voxel is the output of a model fit that depends on the whole original time series. The synthetic maps therefore display no motion from time point to time point and can be used as reference volumes to drive the registration of the original time point image volumes.

Step 3: Perform a 3D translational registration using a standard FLIRT cost function, matching each original time point volume to its corresponding synthetic reference volume. To achieve local control, we calculate the registration cost function only within the tumor VOI. For robustness, we reject registrations where the translation magnitude (extracted from the translation matrix output by FLIRT) exceeds visually-defined extremes of motion for the given data set.

Step 4: Re-fit the chosen model to the registered time series.

Step 5: Repeat Steps 2–4 until a minimum is found in the median model fit sum of squared errors within the tumor VOI, at each stage fitting the model to the “last-registered” time series to generate new synthetic reference images then registering the “last-registered” time series to the new synthetic images.

We tested the registration using synthetic reference images from the standard Kety model with the bi-exponential AIF (labelled STD-EXP) and the extended Kety model with a measured AIF (EXT-MSD). For comparison with a more standard registration procedure, we also used the time-series mean image as the synthetic reference (we may view the time series mean as the simplest model of time-varying signal). In this case the EXT-MSD model was applied to the data

after each time-series mean registration loop and the SSE calculated to compare results with the model-driven registration.

3 Results

Figure 1 compares the sum of squared errors (SSE) on the model fits after model-driven registration and registration to the time series mean. As localized failures in the voxel-by-voxel fitting process may give outliers in the individual voxel SSE, for robustness we quote the median SSE over the tumor VOI. Registration significantly reduced the SSE (i.e. reduced model fitting errors) for 8 of the 10 data sets in each group. The SSEs were generally lower for registration to the time series mean than for model-driven registration, implying that the time series mean image was a more effective registration target. However, this global statistic hides some important details that we will discuss later. Note that the magnitudes of the recovered translations were always less than 15 mm in the through-plane (i.e. superior-inferior) direction and less than 8 mm in the in-plane directions, which is reasonable for these data sets.

Figure 2 quantifies the changes from the pre-registration median tumor VOI estimates of K^{trans} , v_e and v_p after EXT-MSD model-driven registration and registration to the time-series mean. Model-driven registration resulted in greater changes for all three parameters.

Figure 3 plots the model fits and data for a small VOI within the enhancing rim of a tumor in one data set, before and after registration. The data show the expected features of a “first pass” peak in contrast agent concentration, followed by a prolonged washout. In each case the post-registration data (circles) cluster more tightly around the fitted curve than the pre-registration data (crosses), consistent with a reduced SSE. However, STD-EXP model-driven registration and registration to the time series mean distorted the data behavior around the

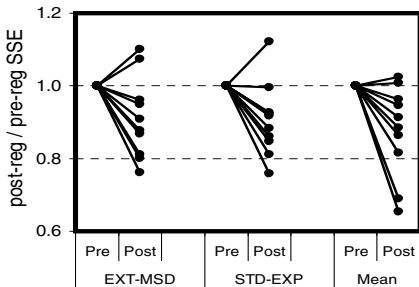


Fig. 1. Post-registration SSE relative to its pre-registration value for model-driven registration and for registration to the time series mean image — a downward slope indicates reduced SSE

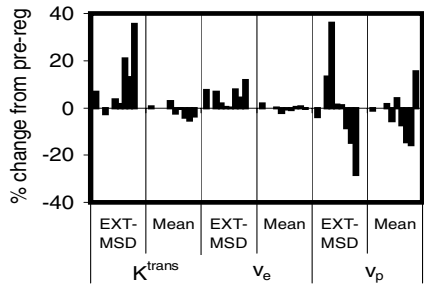


Fig. 2. Changes in the median tumor VOI estimates of the model parameters after model-driven and time series mean image registration, expressed as a percentage change from the pre-registration estimates

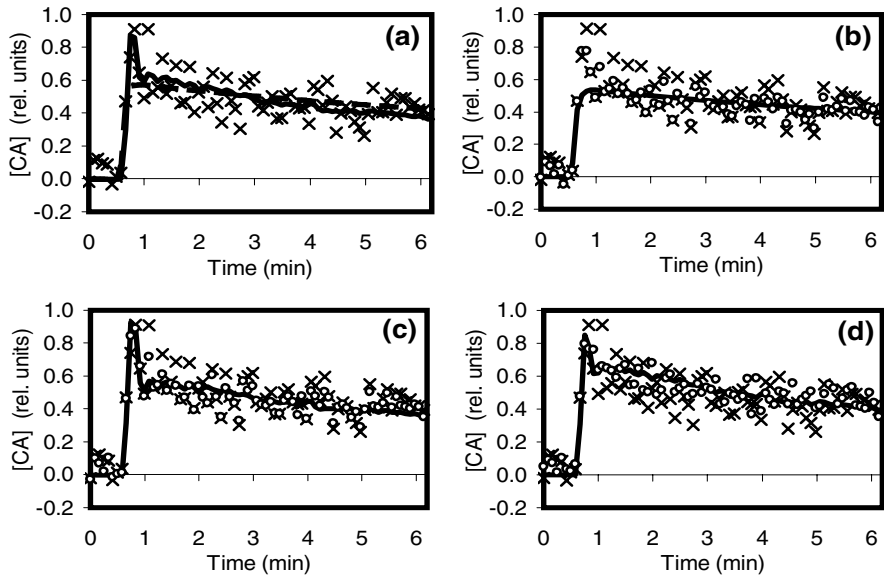


Fig. 3. Model fit curves with data before and after registration. The lines are the fitted models, the crosses pre- and the circles post-registration data. (a) Pre-registration fits for EXT-MSD (solid) and STD-EXP (dashed). Post-registration fits for (b) STD-EXP and (c) EXT-MSD model-driven registration and for (d) an EXT-MSD fit after registration to the time-series mean.

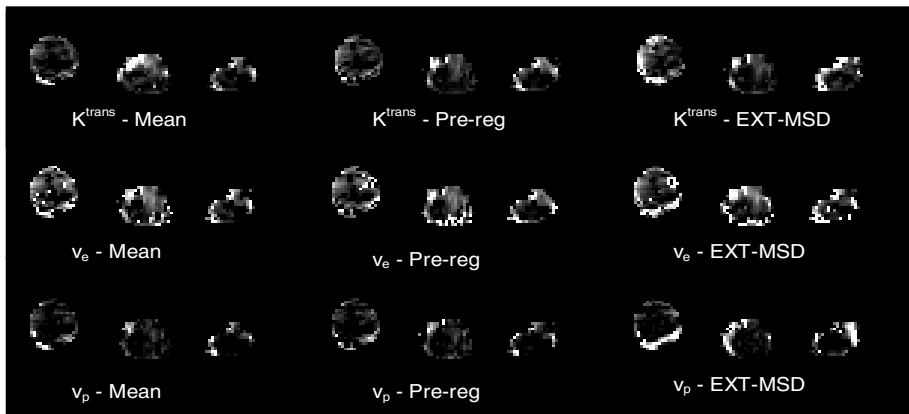


Fig. 4. Maps of the estimated parameters before registration (center column) and after registration to the time series mean (left) and EXT-MSD model-driven registration (right). In each set, the left-hand map is the axial view, the center map the sagittal view and the right-hand map the coronal view. Gray scale values are consistent for each parameter and lighter shading corresponds to higher parameter values. Gray scale ranges: K^{trans} 0–0.4 min^{-1} ; v_e 0–0.8; v_p 0–0.2.

first-pass peak, while EXT-MSD model-driven registration did not (we discuss this point more fully later in the paper).

Figure 4 shows the effect of registration on the spatial localization of parameter estimates. There are clear differences when comparing post-registration parameter maps to pre-registration maps for EXT-MSD model-driven registration — in particular, the maps of v_p and K^{trans} show a more clearly-defined enhancing rim with higher parameter values. However, the differences after registration to the time series mean are less evident. These observations are compatible with the post-registration changes in the median VOI parameter values, which were always of greater magnitude after model-driven registration (Figure 2).

For 5 patients, we were able to perform scan-rescan tests to establish whether registration could reduce inter-visit variability in parameter estimates. We found only one significant reduction after EXT-MSD model-driven registration (in v_p , $P < 0.05$ for a paired one-tailed t-test on the absolute percent changes in parameter estimates between 2 visits, before and after registration) and no significant reductions after time series mean registration.

4 Discussion

Conventional motion correction using standard registration cost functions does not cope well with features that appear and disappear between images, as happens when using a contrast agent. Investigators have observed biased output when using mutual information in a non-rigid registration algorithm for pre- vs. post-contrast MR images of focal breast lesions [19], and addressed this by adding a constraint to the cost function to penalize the artefactual deformations arising in the enhancing region of post-contrast images [20]. Another group successfully incorporated a simplified analytical form of a kinetic model into a non-rigid registration cost function [11].

Neither of these approaches can be directly applied when fitting more comprehensive kinetic models to DCE-MRI time series data from abdominal tumors at high temporal resolution. We have therefore developed a locally-controlled, 3D translational registration method that employs synthetic reference images generated from tracer kinetic model fits or from the time series mean and evaluated its performance using the standard and extended Kety models. We significantly reduced the SSE for 8 / 10 data sets in each group, indicating that in general registration to a synthetic reference image reduces motion corruption in the time series. Registration to the time series mean gave the greatest SSE reduction, implying that the mean image provides the best reference, but the SSE is a global statistic that hides some important details that contradict this conclusion.

We can see these details by looking at the data and model fits for a small VOI within the enhancing rim of a tumor (Figure 3). Both registration to the time series mean and standard Kety model-driven registration distorted the time-series data in a manner that is inconsistent with the expected and experimentally-observed behavior of DCE-MRI data. Registration to the time

series mean reduced the height of the first-pass peak in contrast agent concentration and gave raised concentrations during the washout phase, showing that the averaging process reduces the scatter in the data at the expense of important genuine structure. Standard Kety model-driven registration also reduced the height of the first-pass peak — it is significant that this model does not account for the blood plasma pool and so cannot accurately model this peak. Such distortions were not observed with the extended Kety model, which *does* account for the blood plasma pool. Therefore, if kinetic model-driven registration is to be successful, we must use a model that adequately characterizes the time course.

Extended Kety model-driven registration gave greater changes in median tumor VOI estimates of K^{trans} , v_e and v_p than registration to the time-series mean, but the interpretation of these changes is unclear, partly because these are median values from a heterogeneous volume. In the 3D parameter maps, extended Kety model-driven registration gave higher localized values of K^{trans} and v_p , with a more defined enhancing rim pattern consistent with typical tumor structure. However, without ground truth data we cannot determine whether these changes are “correct” (future studies will address this problem using synthetic images for which ground truth *is* available). It is nevertheless significant that K^{trans} and v_p depend strongly on the first-pass peak, which we have seen to be distorted by time series mean registration but maintained by extended Kety model-driven registration.

Only extended Kety model-driven registration gave a statistically significant reduction in the variability of parameter estimates across 2 DCE-MRI studies in 5 patients, and only with v_p . Nevertheless, this suggests that model-driven registration may reduce variability and thereby improve the sensitivity of quantitative DCE-MRI to genuine parameter changes — a factor of significant interest to clinical trials of novel anti-vascular and anti-angiogenic drugs (see e.g. [3]).

It is encouraging to have obtained such positive results from a simplified registration procedure. Our 3D translational registration is clearly not sufficient to fully recover tumor motion, which will include rotations and shape changes, requiring a non-linear or warping registration, and we hope to address this in future work. Despite its limitations, our registration process has reduced the adverse effects of motion in quantitative DCE-MRI studies, and shown that the kinetic model must be sufficiently detailed to avoid distorting the time-course data. Our implementation is fast enough for routine use (c. 2 hours to register 75 time point image volumes on a standard PC).

Based on the evidence in this paper, we propose that kinetic model-driven registration is a useful additional tool for the quantitative analysis of DCE-MRI data that addresses the problem of changing image information content (due to contrast enhancement) that can confound conventional registration methods.

Acknowledgements. We would like to acknowledge the financial support received from Cancer Research UK.

References

1. Tofts, P., Brix, G., Buckley, D., et al.: Estimating kinetic parameters from dynamic contrast-enhanced T_1 -weighted MRI of a diffusable tracer: standardized quantities and symbols. *J. Magn. Reson. Imag.* **10** (1999) 223–32
2. Parker, G., Padhani, A.: T_1 -weighted dynamic contrast-enhanced MRI. In Tofts, PS, ed.: *Quantitative MRI of the Brain*. 1st edn. Wiley, New York (2003)
3. Jackson, A and Buckley, DL and Parker, GJM, ed.: *Dynamic contrast-enhanced magnetic resonance imaging in oncology*. 1st edn. Springer, Berlin (2005)
4. Tofts, P., Kermode, A.: Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. I. Fundamental concepts. *Magn. Reson. Med.* **17** (1991) 357–67
5. Tofts, P.: Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *J. Magn. Reson. Imag.* **7** (1997) 91–101
6. Jenkinson, M., Bannister, P., Brady, J., et al.: Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* **17** (2002) 825–41
7. Wells III, W., Viola, P., Atsumi, H., et al.: Multi-modal volume registration by maximisation of mutual information. *Med. Imag. Anal.* **1** (1997) 35–51
8. Maes, F., Collignon, D., Vandermeulen, D., et al.: Multimodality image registration by maximisation of mutual information. *IEEE Trans. Med. Imag.* **16** (1997) 187–98
9. Roche, A., Malandain, G., Pennec, X., et al.: The correlation ratio as a new similarity measure for multimodal image registration. *Lect. Notes Comput. Sc.* **1496** (1998) 1115–24
10. Rueckert, D., Sonoda, L., Hayes, C., et al.: Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans. Med. Imag.* **19** (1999) 712–21
11. Hayton, P., Brady, M., Tarassenko, L., et al.: Analysis of dynamic MR breast images using a model of contrast enhancement. *Med. Imag. Anal.* **1** (1997) 207–24
12. Folkman, J.: Role of angiogenesis in tumor growth and metastasis. *Semin. Oncol.* **29** (2002) 15–18
13. Parker, G., Tanner, S., Leach, M.: Pitfalls in the measurement of tissue permeability over short time-scales using multi-compartment models with a low temporal resolution blood input function. *Proc. ISMRM* **4** (1996) 1582
14. Parker, G., Jackson, A., Waterton, J., et al.: Automated arterial input function extraction for T_1 -weighted DCE-MRI. *Proc. ISMRM* **11** (2003) 1264
15. Haase, A.: Snapshot FLASH MRI. Application to T_1 , T_2 and chemical shift imaging. *Magn. Reson. Med.* **13** (1990) 77–89
16. Press, WH and Teukolsky, SA and Vetterling, WT and et al.: *Numerical Recipes in C*. 2nd edn. Cambridge University Press, Cambridge, UK (1992)
17. Sarvazyan, A., Goukassian, D., Maevsky, E., et al.: Elastic imaging as a new modality of medical imaging for cancer detection. *Proc. Int. Workshop on Interaction of Ultrasound with Biological Media*, (1994) 69–81
18. Jenkinson, M., Smith, S.: A global optimisation method for robust affine registration of brain images. *Med. Imag. Anal.* **5** (2001) 143–56
19. Tanner, C., Schnabel, J., Chung, D., et al.: Volume and shape preservation of enhancing lesions when applying non-rigid registration to a time series of contrast enhancing breast images. *Lect. Notes Comput. Sc.* **1935** (2000) 327–37
20. Tanner, C., Schnabel, J., Degenhard, A., et al.: Validation of volume-preserving non-rigid registration: application to contrast-enhanced MR mammography. *Lect. Notes Comput. Sc.* **2488** (2002) 307–14

Generalised Overlap Measures for Assessment of Pairwise and Groupwise Image Registration and Segmentation

William R. Crum¹, Oscar Camara¹, Daniel Rueckert², Kanwal K. Bhatia²,
Mark Jenkinson³, and Derek L.G. Hill¹

¹ Centre for Medical Image Computing, Dept. of Medical Physics,
University College London, WC1E 6BT, UK

{b.crum, o.camara-rey, Derek.hill}@ucl.ac.uk

² Visual Information Processing, Department of Computing,
Imperial College London, SW7 2BZ, UK

{dr, bhatia98}@doc.ic.ac.uk

³ FMRIB Centre, John Radcliffe Hospital, University of Oxford, OX3 9DU, UK
mark@fmrib.ox.ac.uk

Abstract. Effective validation techniques are an essential pre-requisite for segmentation and non-rigid registration techniques to enter clinical use. These algorithms can be evaluated by calculating the overlap of corresponding test and gold-standard regions. Common overlap measures compare pairs of binary labels but it is now common for multiple labels to exist and for fractional (partial volume) labels to be used to describe multiple tissue types contributing to a single voxel. Evaluation studies may involve multiple image pairs. In this paper we use results from fuzzy set theory and fuzzy morphology to extend the definitions of existing overlap measures to accommodate multiple fractional labels. Simple formulas are provided which define single figures of merit to quantify the total overlap for ensembles of pairwise or groupwise label comparisons. A quantitative link between overlap and registration error is established by defining the overlap tolerance. Experiments are performed on publicly available labeled brain data to demonstrate the new measures in a comparison of pairwise and groupwise registration.

1 Introduction

Effective validation techniques are an essential pre-requisite for segmentation and non-rigid registration techniques to enter clinical use. Registration for medical applications seeks a mapping from one image (or set of images) to another such that structural or functional correspondence is achieved i.e. identifiable features or regions are correctly mapped between images. In medical image segmentation, the objective is to identify regions that have some functional or structural significance. If a pre-labeled image can be correctly registered to another image then the labeling problem is solved for that image. Conversely, if a pair of images contains corresponding labeled regions then the registration problem is at least partially solved by constructing a mapping between corresponding labels. Two common scenarios are that automatic image regis-

tration has been performed on the basis of voxel intensity information or that a novel segmentation technique has been applied to an image. The problem with validating these cases at present is the paucity of metrics of quality, especially where the segmentation or registration technique generates fuzzy (i.e. fractional or probabilistic) labels or is evaluated over multiple test images. To date most evaluation has relied on the use of simple measures of regional overlap, defined for single labels, which assume that each voxel is either labeled or not labeled. In this paper we use results from fuzzy set theory and fuzzy morphology to extend existing definitions of overlap to (i) measure overlap of partial volume labels (ii) compute overlap measures for multiple labels defined on multiple image pairs (iii) compute overlap measures for groupwise registration and (iv) establish a link between measures of overlap and estimates of target registration error. Experiments are performed to establish the behavior of the new overlap measures and to compare pairwise and groupwise registration performed on publicly available data.

We consider an existing labeling (E) and a test labeling (T) that may have been obtained by a new segmentation technique or by using the result of image registration to map a label set from one image to another. The most obvious quantitative comparison of regions is by volume [1] however two labelings may have similar volumes but very different shapes, or even locations. The Hausdorff-Chebyshev metric defines the largest difference between two contours or surfaces but can be computationally expensive to compute and is not symmetric between E and T (although it can be made so). The Modified Williams Index has been developed for comparison of multiple expert observers boundaries against computer generated boundaries and is the ratio between the average computer-to-observer agreement and the average inter-observer agreement [2].

For a comparison of voxel-wise binary labelling the number of true and false positives and negatives can be determined and measures of region overlap can be computed. These are generally of the form:

$$O = \frac{N(T \cap E)}{N(T \cup E)}, \text{ or } \frac{N(T \cap E)}{N(E)} \text{ or } \frac{2N(T \cap E)}{N(T) + N(E)} \quad (1)$$

In equation 1 $N(E)$ indicates the number of voxels which belong to the label E etc. In this work we will concentrate on the first (left-most) expression (the Tanimoto coefficient) although a similar development could be made with the other forms. These and other overlap measures are reviewed from the perspective of the so-called “overlapping area matrix” by Beauchemin and Thomson [3]. Measures of correspondence based on information theory have also been proposed [4]. For labels defined in a probabilistic fashion, Gerig [5] suggests a probabilistic overlap and has provided software to compute this and other overlap measures.

The work in this paper extends previous overlap definitions to cope with contemporary applications. The motivation is to develop regional overlap measures that can be intuitively pooled across labels and subjects to provide single figures of merit. Potential applications include assessment of registration of ensembles of subjects, particularly in group-wise (“target-less”) applications. We also take the opportunity to formalize a previously suggested link between overlap measures and target registration errors in registration via the definition of the overlap tolerance, τ .

2 Methods

We first redefine the overlap measure for fuzzy labels. In equation 1, the overlap is obtained by summing the number of distinct voxels in the label intersection and union. However, both binary labels which have been interpolated following registration and those which model partial occupancy at a voxel can be characterized by a number $[0, 1]$ at each voxel that defines the fraction of voxel that is labelled. Results from fuzzy set theory for the intersection and union of two fuzzy sets can immediately be applied to rewrite equation 1 to give an overlap, O_F which is a function of the label values at each voxel summed over all voxels, i , in the image.

$$O_F = \frac{\sum_{\text{voxels},i} \text{MIN}(T_i, E_i)}{\sum_{\text{voxels},i} \text{MAX}(T_i, E_i)} \quad (2)$$

Equation 2 computes the overlap of a single pair of fuzzy labels defined on a pair of images. The numerator and denominator can be accumulated across multiple labels and multiple image pairs to compute a single overlap figure, O_{PMF} , which describes the total overlap of a set of fuzzy labels defined on a set of image pairs. The overlap is the ratio of the total fuzzy intersection to the total fuzzy union.

$$O_{PMF} = \frac{\sum_{\text{pairs},k} \beta_k \sum_{\text{labels},l} \alpha_l \sum_{\text{voxels},i} \text{MIN}(T_{kli}, E_{kli})}{\sum_{\text{pairs},k} \beta_k \sum_{\text{labels},l} \alpha_l \sum_{\text{voxels},i} \text{MAX}(T_{kli}, E_{kli})} \quad (3)$$

In equation 3, α_l is a label-specific weighting factor that affects how much each label contributes to the accumulated overlap and β_k is a pair-specific weighting factor that affects the relative contribution of each image pair to the accumulated overlap. We defer a discussion of the possible values of α_l and β_k to section 2.2. There are now emerging techniques where a single groupwise registration is performed instead of multiple pairwise registrations. A groupwise overlap measure, O_{GMF} , can be constructed by considering permutations of image pairs within the group and applying equation 3. Note that there is no simple relationship between multiple overlaps evaluated separately using equation 2 and the ensemble overlap evaluated using equation 3.

2.1 Overlap and Target Registration Error: Overlap Tolerance

The overlaps described above do not consider the proximity of the non-overlapping label portions that may also provide important information. One previously proposed method of determining this proximity is the tolerance, τ [6]. The standard overlap measures can be considered to be $\tau=0$ since pairs of label voxels have to occupy the same space to be considered overlapping. However, $\tau > 0$ allows labels to be considered overlapping if they lie within τ mm of each other. Therefore as the tolerance increases, the fractional overlap $\rightarrow 1$ as the condition for overlapping voxels is relaxed. Previously this has been described for binary labels and integer values of the tolerance. We now define overlap for fractional labels and non-integer tolerances. Starting from the definition of overlap for fractional labels, O_F , (equation 2) the definition of fuzzy overlap to a tolerance, τ , can be written:

$$O_f(\tau) = \frac{\sum_{\text{voxels},i} \text{MAX}(\text{MIN}((D_\tau T)_i, E_i), \text{MIN}(T_i, (D_\tau E)_i))}{\sum_{\text{voxels},i} \text{MAX}(T_i, E_i)} \quad (4)$$

In equation 4, D_τ is a fuzzy dilation operator that can be represented as a voxel mask of dilation coefficients centered on each voxel of interest with τ specifying the extent of the operator. In 1D, where the voxel dimension is 1mm for example, $D_0=\{1\}$, $D_1=\{1, 1, 1\}$, $D_2=\{1, 1, 1, 1, 1\}$ etc. When considering fractional tolerances then $D_\gamma=\{\gamma, 1, \gamma\}$, $D_{1+\gamma}=\{\gamma, 1, 1, 1, \gamma\}$ etc where $0 \leq \gamma \leq 1$. Then the fuzzy dilation applied at a single voxel located at the origin, $L(0)$, in 1D can be written as $L^*(0) = \text{MAX}(D_\tau(i)L(i))$ where i is in the range $[-k, +k]$ and $k = (\text{int})(1+\tau)$. This definition is consistent with that of [7]: $D_v(\mu)(x) = \sup\{t\{v(y-x), \mu(y)\}, y \in S\}$ where S is the image-space, v is a structuring element, μ is a fuzzy set and x, y are both elements of S . t is a t-norm which in our case is simply defined as $t(a, b) = ab$.

Note that $O_f(\tau)$ is an increasing function of τ for $\tau \geq 0$. The maximum possible overlap given by equation 4 can be established by assuming that both D and T have fractional labels in the range $[0, 1]$. Then when $\tau \gg 1$ the numerator reduces to $\text{MAX}(T_i, E_i)$ and the maximum overlap is therefore 1 as expected. Now consider a pair of misregistered images where every voxel is independently labeled and the same set of labels exists in each image but are not necessarily coincident. Then the overlap of any pair of labels can be computed as described above. For each labeled voxel in the target image the smallest tolerance, τ_1 , for which the overlap with its partner in the source image is 1, can be computed. Then the map of tolerances is a map of target registration error. For labels spanning multiple voxels, τ_1 estimates the maximum displacement between non-overlapping voxels belonging to corresponding labels.

2.2 Parameter Choices

In equation 3, weights α and β were introduced to respectively define the relative contribution of labels and subjects to the overlap measures. With $\alpha=1$, all labels are implicitly weighted by their volume. This may not be desirable as smaller labels may represent a greater registration or segmentation challenge. Two alternative choices are to set α either to (i) the inverse mean volume of the current label pair to give all labels equal weighting or (ii) the inverse mean volume squared of the current label pair to weight by the inverse volume. We examine the effect of these different α values below. In the experiments reported in this paper we have set $\beta=1$ but β could be used to weight inversely with the variance of labeling accuracy.

3 Experiments A, B and C

Nine T1-weighted MR-brain images with labels from the Internet Brain Segmentation Repository¹ were used. Each image had ten binary anatomical labels, one for each of the following structures: amygdala, caudate, cerebellum, cortex, hippocampus, lateral ventricle, pallidum, putamen, thalamus and white matter.

¹ <http://www.cma.mgh.harvard.edu/ibsr>

Experiment A investigates the decrease in overlap measures in response to forced misregistration. Each of the nine images (the targets) was paired with a copy of itself (the sources) and the accumulated overlap measure was computed for each pair. Then a non-rigid misregistration algorithm, which acted to *reduce* an image similarity measure, was applied to each source image and associated labels using tri-linear interpolation. The mean misregistration displacement over each pair was computed together with the new accumulated overlaps. The overlaps were also computed for a single label, S constructed from the union of the ten labels above.

Experiment B examines the relationship between the overlap tolerance and the misregistration error. For each of the misregistered labels in experiment A including the combined label, S, the tolerance was found for which the overlap O_F was 0.99 and compared with the mean misregistration error computed for each label from the applied transformations.

Experiment C compares a group of pairwise registrations and a groupwise registration. The 9 images from experiment A were (i) each registered independently to a tenth image in a pairwise fashion using a B-spline approach [8] and (ii) registered in a groupwise fashion to a common reference frame representing the average shape of the population, also using a B-spline approach [9]. Both techniques overlay a mesh of uniformly spaced control points onto each of the images; deforming the control points deforms the underlying images. The control points are manipulated until the normalised mutual information is maximised. The groupwise technique does not use an explicit anatomical reference; instead an average shape is calculated implicitly by constraining the sum of all the deformations to be equal to zero using a Gradient Projection Method. The cases were compared by assuming that both had been performed in a groupwise fashion and computing the groupwise overlap measure, O_{GMF} by permuting and accumulating all possible pairwise overlaps as in equation 3.

4 Results

Experiment A. Fig. 1 plots the accumulated overlap against the mean applied misregistration for the three different label weightings, α , for each pair in experiment A.

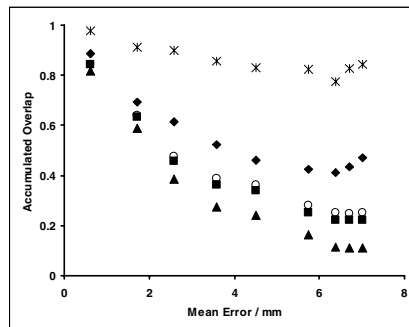


Fig. 1. The results of experiment A. The overlap accumulated over 10 labels plotted against the mean applied misregistration error for 9 image pairs. ♦=Volume-weighting, ■=Equal-weighting, ▲=Inverse Volume-weighting, *= Union Label S and ○ = Simple Average.

Also plotted are the overlap for the single union label S and the average overlap computed for each pair. It can be seen the trend is for the measured overlap to decrease as a function of increasing applied misregistration. There is a distinction between the 3 different weighting schemes with the inverse volume weighted overlaps decreasing fastest. This is to be expected, as the overlap of smaller regions will be more sensitive to misregistration. The average overlap for each pair is nearly coincident with the equally weighted accumulated overlap as expected. The overlap of large structures (as deduced from the volume-weighted and union plots) does not always strictly decrease with increasing misregistration. This result is considered in section 4.

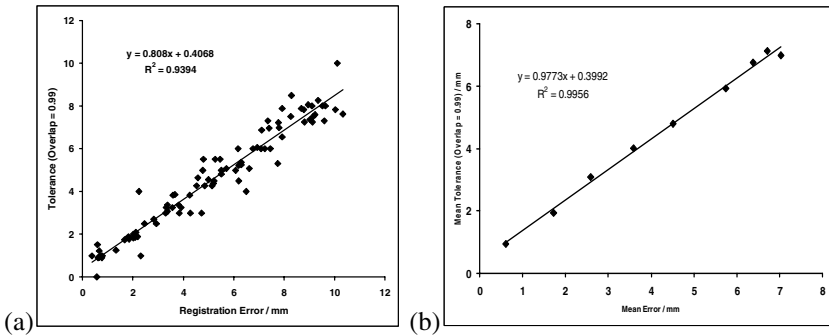


Fig. 2. Results for experiment B. (a) The smallest tolerance that gives an overlap of 0.99 against the applied misregistration error for each of 10 label-pairs on 9 subjects. (b) The results of (a) averaged over each subject.

Experiment B. Fig. 2a plots the tolerance computed iteratively using equation 4 for each label on each pair against the mean applied misregistration for that label. Figure 2b plots the tolerance averaged over all the labels on each pair against the mean applied misregistration calculated over the aggregated label S. In both cases there is a strong linear relationship between the tolerance and the applied misregistration.

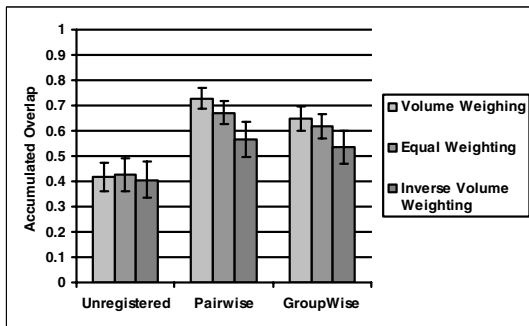


Fig. 3. Results for experiment C. The total accumulated overlap computed in a groupwise fashion for the unregistered, pairwise registered and groupwise registered images. Results are shown for volume-weighting, equal weighting and inverse volume-weighting between labels. Error bars represent the standard deviation of the overlap computed over all subject pairs.

Experiment C. Fig. 3 shows the groupwise overlap, O_{GMF} , for the unregistered, pairwise and groupwise registered images for volume weighting, equal weighting and inverse volume weighting between structures. The overlap consistently ranks the pairwise overlaps above the groupwise overlaps. While the pairwise registration was refined to a control point spacing of 2.5mm, the groupwise was only refined to a spacing of 5mm for computational efficiency. Therefore the scale of non-rigid deformations is restricted in the groupwise registration compared to the pairwise. Also the minimization of the groupwise cost function is more complex and therefore more likely to find local minima; this is an area of continuing research. A more interesting observation is that the inverse volume weighted overlaps are far more similar than the equal and volume weighted overlaps indicating that small structures are being registered more consistently. This is probably because the small, deep brain structures have relatively consistent anatomy whereas the larger structures such as cortex and white matter are known to vary significantly between individuals.

5 Discussion and Conclusion

We have developed overlap measures to allow comparison of multiple fuzzy labels defined on multiple subjects. The specific case of pairwise and groupwise registration has been considered here but these measures could also be applied to related problems of segmentation. We have re-introduced the idea of overlap tolerance and used it to relate registration error to overlap. We have demonstrated a linear relationship between the overlap tolerance and the applied misregistration in one experiment but this cannot be considered a completely general result. The misregistration acted normally to edge features so preferentially displaced high-contrast boundaries of structures rather than rearranging low-contrast features. The tolerance would be an insensitive indicator of registration error if the misregistrations were occurring within labels; this is a property of labels rather than these overlap measures. Another application for the tolerance might be to initialize registration problems where labels exist on some parts of the image.

The framework presented in this paper allows single overlap measures that encompass multiple labels defined on multiple image pairs to be generated in a natural way. Weighting can be applied to prefer smaller labels and/or to accommodate other prior information about the images. Such evaluation tools are necessary for the clinical adoption of new registration and segmentation techniques.

Acknowledgements

We acknowledge the following support: Medical Images and Signals IRC (EPSRC GR/N14248/01 and MRC D2025/31) (WRC), Modelling, Understanding and Predicting Structural Brain Change (EPSRC GR/S48844/01), (OC) and Integrated Brain Image Modelling (EPSRC GR/S82503/01), (MJ). We thank David Kennedy and the Centre for Morphometric Analysis at MGH for use of the IBSR data.

References

1. Collins, D., Dai, W., Peters, T. and Evans, A. Automatic 3D model-based neuroanatomical segmentation, *Human Brain Mapping* 3 (1995) 190-208.
2. Chalana, V. and Kim, Y. A methodology for evaluation of boundary detection algorithms on medical images, *IEEE Transactions on Medical Imaging*, 16(5) (1997) 642-652.
3. Beauchemin, M., Thomson, K.P.B. The evaluation of segmentation results and the overlapping area matrix, *International Journal of Remote Sensing* 18(18) (1997) 3895-3899.
4. Bello F., Colchester A.C.F., Measuring global and local spatial correspondence using information theory, *Proceedings of MICCAI 1998*, LNCS 1496: 964-973.
5. Gerig, G., Jomier, M. and Chakos, M. Valmet: A new validation tool for assessing and improving 3D object segmentation, *Proceedings of MICCAI 2001* 516-528.
6. Crum, W.R., Griffin, L.D., Hill, D.L.G. and Hawkes, D.J., Zen and the Art of Medical Image Registration: Correspondence, Homology and Quality, *NeuroImage* 20 (2003) 1425-1437.
7. Bloch, I., Fuzzy spatial relationships for image processing and interpretation: a review, *Image and Vision Computing* 23(2) (2005) 89-110.
8. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O. and Hawkes, D.J. Non-rigid registration using free-form deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8) (1999) 712-721.
9. Bhatia, K.K., Hajnal, J.V., Puri, B.K., Edwards A.D. and Rueckert, D., Consistent group-wise non-rigid registration for atlas construction, *IEEE Symposium on Biomedical Imaging (ISBI)*, pp 908-911, Arlington 2004.

Uncertainty in White Matter Fiber Tractography

Ola Friman and Carl-Fredrik Westin

Laboratory of Mathematics in Imaging,
Department of Radiology Brigham and Women's Hospital, Harvard Medical School

Abstract. In this work we address the uncertainty associated with fiber paths obtained in white matter fiber tractography. This uncertainty, which arises for example from noise and partial volume effects, is quantified using a Bayesian modeling framework. The theory for estimating the probability of a connection between two areas in the brain is presented, and a new model of the local water diffusion profile is introduced. We also provide a theorem that facilitates the estimation of the parameters in this diffusion model, making the presented method simple to implement.

1 Introduction

When estimating a quantity based on some available data, it is natural to ask what uncertainty is associated with this estimate. In white matter fiber tractography, connections in the brain are estimated by tracing the direction of maximal water diffusion in diffusion weighted MRI images. In this process, uncertainty arises due to noise and partial volume effects, i.e., crossing, merging and splitting fiber tracts. Uncertainty is naturally quantified by means of probability, and in this work we are interested in the probability of a fiber connection between two points A and B in the brain, given the measured diffusion data \mathcal{D} . We write this probability $p(A \rightarrow B|\mathcal{D})$. There is some previous work in the literature on this subject [1–6], but most of these approaches account for uncertainty in heuristic ways. Exceptions are Tuch [3] and Behrens et al. [6], who were the first to present approaches with stronger theoretical support. In this paper we build on the work by Tuch and Behrens et al., and present a more detailed theory as well as a computationally tractable solution.

2 Global Modeling and Estimation

A fiber can be modeled as a finite length path, described by a train of vectors $\mathbf{v}_{1:n} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. Denote by Ω_A^n the set of all possible paths of length n that originate in a point A , and assume that we can assign a probability $p(\mathbf{v}_{1:n})$ to each path in this space. Hence,

$$\int_{\Omega_A^n} p(\mathbf{v}_{1:n}) = 1 \quad \text{and} \quad \sum_{n=1}^{\infty} \int_{\Omega_A^n} p(n)p(\mathbf{v}_{1:n}) = 1, \quad (1)$$

where $p(n)$ is the probability of length n path. If we have any prior information about the expected path lengths this can be encoded in $p(n)$. Without prior knowledge, $p(n)$

is chosen to be a uniform distribution over a range of reasonable lengths. Integrals over path spaces, like the ones in Eq. 1, also arise for example in quantum mechanics [7] and computer graphics [8]. Now, let Ω_{AB}^n be the set of all paths of length n between A and another point B . We can then find the probability of a fiber going from A to B , given the diffusion data \mathcal{D} , by summing the probabilities for all paths of all lengths between these points:

$$p(A \rightarrow B|\mathcal{D}) = \sum_{n=1}^{\infty} \int_{\Omega_{AB}^n} p(n)p(\mathbf{v}_{1:n}|\mathcal{D}). \quad (2)$$

In the above equation we have suppressed the dependence of the path length n on \mathcal{D} . The integrations in Eq. 2 are performed over $2n$ -dimensional path spaces Ω_{AB}^n , and we cannot hope to find analytical solutions. Hence, we must resort to numerical integration, and it is only by applying Monte Carlo methods we can hope to estimate such high-dimensional integrals. Since $\Omega_{AB}^n \subset \Omega_A^n$, we can approximate the integrals by drawing a large number of random paths $\mathbf{v}_{1:n}^k$, $k = 1, \dots, N_n$, from $p(\mathbf{v}_{1:n}|\mathcal{D})$ over the domain Ω_A^n , i.e., with starting point in A , and calculate the fraction of paths that reach B . This technique is known as Rejection sampling. Formally, by defining an indicator function

$$\mathbb{I}(\mathbf{v}_{1:n}^k) = \begin{cases} 1 & \mathbf{v}_{1:n}^k \in \Omega_{AB}^n \\ 0 & \text{Otherwise} \end{cases}, \text{ we have } p(A \rightarrow B|\mathcal{D}) \approx \sum_{n=1}^{\infty} \sum_{k=1}^{N_n} p(n) \frac{\mathbb{I}(\mathbf{v}_{1:n}^k)}{N_n}.$$

Instead of calculating $p(A \rightarrow B|\mathcal{D})$, Tuch [3] consider the path of maximal probability between A and B . As a measure of connectivity between A and B , the maximal probability path has the interesting property of being symmetric. In general, this does not hold for $p(A \rightarrow B|\mathcal{D})$; the probability will depend on whether we start sampling paths in A or in B . The reason is that $p(A \rightarrow B|\mathcal{D})$ pertains to the probability for a *single* fiber, which cannot split or merge. Furthermore, note that in order to calculate $p(A \rightarrow B|\mathcal{D})$, there is no need to explicitly evaluate the path probabilities $p(\mathbf{v}_{1:n}|\mathcal{D})$; it is sufficient if we can draw sample paths. In contrast, to find the maximal probability path, the path probabilities must be computed. In the approach described in the following sections, both alternatives are possible.

2.1 Sampling Paths

To implement the above scheme we need a method for drawing random paths $\mathbf{v}_{1:n} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ from a probability density function (pdf) $p(\mathbf{v}_{1:n})$. Since the sampled paths are models of actual fibers, a certain regularity must be imposed upon them. For simplicity and notational convenience we assume that the vector \mathbf{v}_k only depends on the previous vector \mathbf{v}_{k-1} , and not on $\mathbf{v}_{k-2}, \dots, \mathbf{v}_1$. This assumption is trivially relaxed. We also assume a pre-determined vector length, i.e., a step length, so that we can work with normalized vectors $\hat{\mathbf{v}}_k$. The probability for a path of length n then factorizes into

$$p(\mathbf{v}_{1:n}|\mathcal{D}) = p_1(\hat{\mathbf{v}}_1|\mathcal{D}) \prod_{k=2}^n p_k(\hat{\mathbf{v}}_k|\hat{\mathbf{v}}_{k-1}, \mathcal{D}). \quad (3)$$

Hence, random paths can be built sequentially by first drawing a random direction $\hat{\mathbf{v}}_1$, then a random direction $\hat{\mathbf{v}}_2$ given $\hat{\mathbf{v}}_1$, and so on. This procedure is known as Sequential

Importance Sampling [1]. An interpolation problem arises because we only have the diffusion data \mathcal{D} on a discrete grid, whereas the sampled paths do not have such a restriction. Due to the probabilistic framework we can use the method suggested in [6], which uses the diffusion data at a grid point chosen at random based on the distance to the current sample point. The sequential sampling is terminated when the path reaches an area where the uncertainty regarding the next step is too high.

3 Local Modeling and Estimation

The goals in this section are to find a pdf of the local fiber orientation $p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1}, \mathcal{D})$ and to draw random samples from this pdf. Examples of pdf's are shown in Fig. 1a. The local pdf's should take the uncertainty arising from partial volume effects and measurement noise into account. There are a few different ways to approach this problem. For example, with q -ball imaging [3], it is possible to resolve complex fiber neighborhoods and to obtain a detailed description of the local diffusion profile. The drawback with this technique is that it does not account for the uncertainty stemming from noise. Another approach is to acquire several diffusion data sets and generate a pdf by bootstrapping [9], with the obvious disadvantage that several data set are required. The most commonly used approach in diffusion weighted imaging is to use an explicit model for the water diffusion profile. As pioneered by Behrens et al. [6], the uncertainty in fiber direction can then be inferred via Bayes' theorem, as described below.

3.1 Local Bayesian Modeling

Assume that we have a model that relates the diffusion measurements \mathcal{D} with the underlying tissue properties and architecture. Such a model necessarily contains at least one fiber direction $\hat{\mathbf{v}}_k$ and a set of nuisance parameters collectively denoted by $\boldsymbol{\theta}$. By applying Bayes' theorem we have

$$p(\hat{\mathbf{v}}_k, \boldsymbol{\theta} | \hat{\mathbf{v}}_{k-1}, \mathcal{D}) = \frac{p(\mathcal{D} | \hat{\mathbf{v}}_k, \boldsymbol{\theta}) p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1}) p(\boldsymbol{\theta})}{p(\mathcal{D})}, \quad (4)$$

where we have assumed that the prior distribution can be factorized $p(\hat{\mathbf{v}}_k, \boldsymbol{\theta} | \hat{\mathbf{v}}_{k-1}) = p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1}) p(\boldsymbol{\theta})$. To find $p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1}, \mathcal{D})$ we need to marginalize Eq. 4 over the nuisance parameters $\boldsymbol{\theta}$. Furthermore, the normalizing factor

$$p(\mathcal{D}) = \int_{\hat{\mathbf{v}}_k, \boldsymbol{\theta}} p(\mathcal{D} | \hat{\mathbf{v}}_k, \boldsymbol{\theta}) p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1}) p(\boldsymbol{\theta}) \quad (5)$$

is in general difficult to evaluate because of the high-dimensional integral and an intractable integrand. The posterior in Eq. 4 has to be calculated in every step in the sequential sampling of the fiber paths, and unless an approximation for the integral in Eq. 5 is found the cost of the Bayesian approach will be prohibitive. Behrens et al. [6] avoid evaluating the integral by applying Markov Chain Monte Carlo (MCMC) methods for drawing samples from the posterior. In the following sections we describe an alternative method, and discuss different observation models $p(\mathcal{D} | \hat{\mathbf{v}}_k, \boldsymbol{\theta})$ and priors, $p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1})$ and $p(\boldsymbol{\theta})$, that allow fast computation of the posterior.

Observation Model. The true underlying voxel intensities μ_i in the diffusion weighted images depend on the local water diffusion profile, and the fiber orientation is generally assumed to coincide with the direction of largest diffusivity. In Table 1, three models of how the voxel intensity is modulated by the water diffusion are listed. The gradient directions $\hat{\mathbf{g}}_i$ and the b -values b_i are known experimental parameters. In these models,

Table 1. Three models of how the voxel intensity μ_i depends on the water diffusion profile

Gaussian model	$\mu_i = \mu_0 e^{-b_i \hat{\mathbf{g}}_i^T \mathbf{D} \hat{\mathbf{g}}_i}$
Constrained model	$\mu_i = \mu_0 e^{-\alpha b_i} e^{-\beta b_i (\hat{\mathbf{g}}_i^T \hat{\mathbf{v}})^2}$
Compartment model	$\mu_i = \mu_0 [(1-f)e^{-b_i d} + f e^{-b_i d (\hat{\mathbf{g}}_i^T \hat{\mathbf{v}})^2}]$

the diffusion tensor \mathbf{D} or the local fiber direction $\hat{\mathbf{v}}$ are parameters of interest. The remaining parameters are nuisance parameters. The Gaussian model is widely used in diffusion imaging [10], and the Compartment model was proposed by Behrens et al. [6] (although here it is given in a reparameterized version). We propose the Constrained model, which is obtained if we assume that the two smallest eigenvalues of \mathbf{D} are equal, i.e., $\lambda_2 = \lambda_3 = \alpha$. This gives

$$\mathbf{D} = \lambda_1 \hat{\mathbf{v}}_1 \hat{\mathbf{v}}_1^T + \alpha (\hat{\mathbf{v}}_2 \hat{\mathbf{v}}_2^T + \hat{\mathbf{v}}_3 \hat{\mathbf{v}}_3^T) = (\lambda_1 - \alpha) \hat{\mathbf{v}}_1 \hat{\mathbf{v}}_1^T + \alpha \mathbf{I} = \beta \hat{\mathbf{v}}_1 \hat{\mathbf{v}}_1^T + \alpha \mathbf{I}, \quad (6)$$

which after substitution into the expression for the Gaussian model yields the Constrained model. The Constrained model describes cigar-shaped diffusion profiles, with a sphere and a line as extremes. While all three models in Table 1 describe diffusion profiles, the Constrained model and the Compartment model can also be viewed as modeling the effect of an underlying fiber. They can also be extended to handle multiple fiber directions. In this work, however, we assume that there is only one fiber direction in each voxel, and any deviations from this model will be captured as uncertainty in this direction via the posterior in Eq. 4. The reason we prefer the Constrained model over the Compartment model is out of mathematical tractability, as will become clear later.

The voxel intensity y_i in a diffusion weighted image is a noisy observation of μ_i . Moreover, the intensity decays exponentially with the water diffusivity, as implied by the models in Table 1. Hence, by taking the logarithm of the observations, $z_i = \ln y_i$, we obtain a more linear relationship between observations and model parameters. In references [10, 11] it is shown that $z_i = \ln \mu_i + \epsilon$, where $\epsilon \in N(0, \sigma^2 / \mu_i^2)$, is a good model for the observation noise. That is, after taking the logarithm the noise can be modeled as additive Gaussian with a variance that depends on the voxel intensity. The joint distribution for the observed (log-)data $\mathcal{D} = [z_1, \dots, z_N]$ is then

$$p(\mathcal{D} | \hat{\mathbf{v}}, \boldsymbol{\theta}) = \prod_{i=1}^N \frac{\mu_i}{\sqrt{2\pi\sigma^2}} e^{-\frac{\mu_i^2}{2\sigma^2} (z_i - \ln \mu_i)^2}, \quad (7)$$

where μ_i is to be replaced by the expression for any of the models in Table 1, and $\boldsymbol{\theta}$ denotes the nuisance parameters in this model. This expression is used in the equations for the posterior, Eq. 4 and Eq. 5.

Priors. Via the probability functions $p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1})$ and $p(\boldsymbol{\theta})$ we encode our prior knowledge about fiber regularity and nuisance parameters. Starting with the fiber regularity, a simple family of distributions is given by

$$p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1}) \propto \begin{cases} (\hat{\mathbf{v}}_k^T \hat{\mathbf{v}}_{k-1})^\gamma & \hat{\mathbf{v}}_k^T \hat{\mathbf{v}}_{k-1} \geq 0 \\ 0 & \hat{\mathbf{v}}_k^T \hat{\mathbf{v}}_{k-1} < 0 \end{cases} \quad \text{with } \gamma \geq 0. \quad (8)$$

This prior gives preference to continue in the previous step-direction, with a decreasing probability for sharper turns until it reaches a zero probability for turns of 90 degrees and above. The γ -parameter controls the sharpness of the distribution, i.e., the regularity of the path. In our experiments we use $\gamma = 1$. Turning to the nuisance parameters in the models in Table 1, e.g. the Constrained model for which $\boldsymbol{\theta} = \{\mu_0, \alpha, \beta, \sigma^2\}$, we generally do not have any detailed prior information, except that these parameters should be non-negative. Considering the computational effort required to evaluate the integral in Eq. 5 we choose dirac impulses as priors for the nuisance parameters, for example $p(\boldsymbol{\theta}) = \delta(\mu_0 - \hat{\mu}_0) \delta(\alpha - \hat{\alpha}) \delta(\beta - \hat{\beta}) \delta(\sigma^2 - \hat{\sigma}^2)$. That is, we fix the nuisance parameters to some values $\hat{\mu}_0, \hat{\alpha}, \hat{\beta}$ and $\hat{\sigma}^2$. The integral in Eq. 5, as well as the posterior pdf in Eq. 4, is then evaluated over the unit sphere. Hence, they are easily calculated numerically and significant computational savings are made. Also, note that the Constrained model and the Compartment model in Table 1 differ only in the way the nuisance parameters enter, and once these parameters are fixed the models are essentially equivalent.

3.2 Point Estimates of Local Model Parameters

To define the priors above, or to initialize an MCMC process for drawing samples from the posterior, it is important to have access to good point estimators of the parameters in the observation models in Table 1. For the Gaussian model, point estimates are readily obtained by means of linear least squares estimation [10, 11]. The parameters in the Constrained model can be found through the following non-trivial theorem:

Theorem. *Let \mathbf{D} be a symmetric 3×3 matrix with eigenvalue factorization $\mathbf{D} = \lambda_1 \hat{\mathbf{v}}_1 \hat{\mathbf{v}}_1^T + \lambda_2 \hat{\mathbf{v}}_2 \hat{\mathbf{v}}_2^T + \lambda_3 \hat{\mathbf{v}}_3 \hat{\mathbf{v}}_3^T$, $\lambda_1 \geq \lambda_2 \geq \lambda_3$. The closest, in terms of the Frobenius norm, symmetric matrix \mathbf{C} with the two smallest eigenvalues equal is given by*

$$\mathbf{C} = \lambda_1 \hat{\mathbf{v}}_1 \hat{\mathbf{v}}_1^T + \frac{\lambda_2 + \lambda_3}{2} (\hat{\mathbf{v}}_2 \hat{\mathbf{v}}_2^T + \hat{\mathbf{v}}_3 \hat{\mathbf{v}}_3^T). \quad (9)$$

A proof of this theorem is left out due to space restrictions. Hence, we can find the parameters of the Constrained model by first solving for the full diffusion tensor \mathbf{D} in the Gaussian model and then setting $\alpha = (\lambda_2 + \lambda_3)/2$, $\beta = \lambda_1 - \alpha$ and $\hat{\mathbf{v}} = \hat{\mathbf{e}}_1$. The Compartment model is more difficult to handle mathematically and we need to apply computationally demanding non-linear optimization techniques to find accurate estimates of its parameters.

3.3 Drawing Samples from the Posterior

Finally, to perform the sequential sampling of fiber paths, we need to draw random samples of the fiber direction from the posterior in Eq. 4. For drawing samples from

complicated and high dimensional pdf's one can always resort to MCMC techniques. However, while MCMC sampling is conceptually simple there are implementational issues, such as convergence and so-called mixing of the chain, that complicate practical use. In contrast, with the dirac priors suggested above we need only draw samples from a pdf defined on the unit sphere. This can be accomplished by evaluating the pdf at a sufficiently large number of points evenly spread over the unit sphere, effectively approximating the continuous pdf with a discrete pdf, from which it is straightforward to draw random samples. We must make take care to sample the continuous pdf densely enough so that high probability areas are covered well. In our experiments we evaluate the posterior at 2,562 predefined unit length vectors obtained by a fourfold tessellation of an icosahedron.

4 Results

Here we show results obtained by applying the methodology described in the above sections to a diffusion weighted data set with 31 different gradient directions and a b -value of 1000 s/mm^2 . Figure 1a shows example pdf's of the fiber orientation in three different voxels in this data set. To produce sample fiber paths we sequentially draw fiber orientations from such pdf's. A Matlab implementation typically produces a few sample paths per second on a high-end PC, and the sampling process is trivially parallelized. Figure 1b shows 3,000 sample fibers originating in one point in the splenium of Corpus callosum (marked in Fig. 1c). The fibers have been colored according to how the probability evolves along the paths (Eq. 3). From the sampled fibers in Fig. 1b, the probability map in Fig. 1c was produced as described in Section 2. This map indicates the probability, or our belief, of a single fiber going from the starting point to any other point in the image slice. Finally, Fig. 1d shows more examples of sampled fiber paths, started in four different points in the axial slice. Note how the stochastic framework handles splitting fiber bundles and partial volume effects. For example, a single sample fiber in the Cingulum bundles (red and yellow fibers) easily slips into the underlying Corpus callosum bundle, but the massive sampling approach gives a robust delineation.

5 Discussion

To conclude this paper, we have presented a method for assessing uncertainty in white matter fiber tractography. The main contribution is a novel way of obtaining pdf's of the local fiber direction, including a new model of the water diffusion suitable for a stochastic tracking framework, and a theorem that allows fast computation. Also, while the procedure for calculating the probability of a connection has been used before [2, 4, 5, 6], the theoretical justification in Section 2 has not been described previously. Among the cited works, the approach presented by Behrens et al. is closest to the one presented here. The essential difference lies in how the posterior pdf of the fiber orientation is approximated in order to obtain a computationally feasible solution. We use simplified priors for the nuisance parameters, which result in an algorithm that is computationally fast and easy to implement. Since some of the priors are estimated from the data, the approach can be classified as empirical Bayesian. In contrast, the approach presented

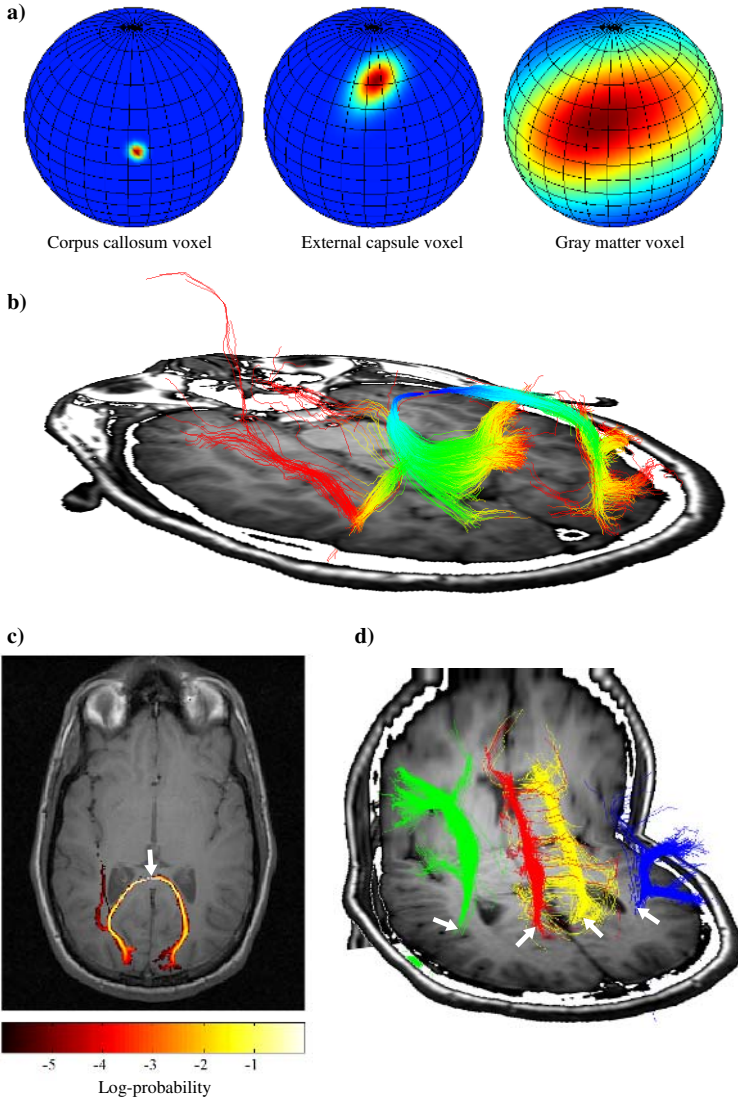


Fig. 1. **a)** Example pdf's of the local fiber orientation taken at three different locations. A uniform prior for the step direction was used when generating these pdf's, i.e, the previous step direction has no influence. **b)** 3,000 fiber samples originating in a point located at the splenium of Corpus callosum (marked in c)). Each fiber is colored according to how the probability evolves along its path. **c)** Probability map of the existence of a fiber going from the marked point to all other voxels in the image slice. The map was generated using the fiber samples in b). **d)** More examples of fiber samples. The sampling was started in the four different points marked by the arrows, and the fibers are colored accordingly.

by Behrens et al. employs full Bayesian modeling of the priors combined with MCMC sampling, which is slower but has the theoretical advantage of taking uncertainty in nuisance parameter estimates into account. However, to exploit this advantage we need to assess convergence and adaptively control the mixing of a large number of Markov chains automatically, which we have found difficult. Further work is required to investigate if there are any interesting differences in result between the faster empirical Bayesian and the full Bayesian approaches.

Acknowledgments

This work was funded by NIH grant P41-RR13218. The authors thank Dr. Raul San-Jose Estepar, Dr. Gunnar Farnebäck and Dr. Gordon Kindlmann for valuable input.

References

1. Björnemo, M., Brun, A., Kikinis, R., Westin, C.F.: Regularized stochastic white matter tractography using diffusion tensor MRI. In: Proceedings of the 5th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'02), Tokyo, Japan (2002)
2. Koch, M., Norris, D., Hund-Georgiadis, M.: An investigation of functional and anatomical connectivity using magnetic resonance imaging. *NeuroImage* **16** (2002) 241–250
3. Tuch, D.: Diffusion MRI of complex tissue structure. PhD thesis, Harvard-MIT Division of Health Sciences and Technology (2002)
4. Parker, G., Haroon, H., Wheeler-Kingshott, C.: A framework for a streamline-based probabilistic index of connectivity (PICO) using a structural interpretation of MRI diffusion measurements. *Journal of Magnetic Resonance Imaging* **18** (2003) 242–254
5. Hagmann, P., Thiran, J., Jonasson, L., Vandergheynst, P., Clarke, S., Maeder, P., Meuli, R.: DTI mapping of human brain connectivity: Statistical fibre tracking and virtual dissection. *NeuroImage* **19** (2003) 545–554
6. Behrens, T., Woolrich, M., Jenkinson, M., Johansen-Berg, H., Nunes, R., Clare, S., Matthews, P., Brady, J., Smith, S.: Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine* **50** (2003) 1077–1088
7. Kleinert, H.: Path integrals in quantum mechanics, statistics, and polymer physics. World Scientific, Singapore (1995)
8. Veach, E.: Robust Monte Carlo methods for light transport simulation. PhD thesis, Stanford University (1997)
9. Lazar, M., Alexander, A.: Bootstrap white matter tractography (BOOT-TRAC). *NeuroImage* **24** (2005) 524–532
10. Basser, P., Mattiello, J., LeBihan, D.: Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance, Series B* **103** (1994) 247–254
11. Salvador, R., Pena, A., Menon, D., Carpenter, T., Pickard, J., Bullmore, E.: Formal characterization and extension of the linearized diffusion tensor model. *Human Brain Mapping* **24** (2005) 144–155

Fast and Simple Calculus on Tensors in the Log-Euclidean Framework

Vincent Arsigny¹, Pierre Fillard¹, Xavier Pennec¹, and Nicholas Ayache¹

INRIA Sophia - Projet Epidaure, BP 93, 06902 Sophia Antipolis Cedex, France
{Vincent.Arsigny, Pierre.Fillard, Xavier.Pennec,
Nicholas.Ayache}@Sophia.Inria.fr

Abstract. Computations on tensors have become common with the use of DT-MRI. But the classical Euclidean framework has many defects, and affine-invariant Riemannian metrics have been proposed to correct them. These metrics have excellent theoretical properties but lead to complex and slow algorithms. To remedy this limitation, we propose new metrics called Log-Euclidean. They also have excellent theoretical properties and yield similar results in practice, but with much simpler and faster computations. Indeed, Log-Euclidean computations are Euclidean computations in the domain of matrix logarithms. Theoretical aspects are presented and experimental results for multilinear interpolation and regularization of tensor fields are shown on synthetic and real DTI data.

1 Introduction: Calculus on Tensors

Tensors, i.e. symmetric positive-definite matrices in medical imaging, appear in many contexts: Diffusion Tensor MRI (DT-MRI or DTI) [2], modeling of anatomical variability [7], etc. They are also a general tool in image analysis, especially for segmentation, motion and texture analysis (see [1] for references on this subject). Many approaches have been proposed in the literature to process tensors [13, 15, 14, 4, 5]. But in order to carry out general computations on these objects, one needs a consistent operational framework. This is necessary to completely generalize to tensors statistical tools and Partial Differential Equations (PDEs). The framework of Riemannian metrics has recently emerged as particularly adapted to this task [11].

One can directly use a Euclidean structure on square matrices to define a metric on the tensor space, for instance with the following distance: $\text{dist}^2(S_1, S_2) = (\text{Trace}((S_1 - S_2)^2))$. This is straightforward and leads *a priori* to simple computations. But this framework is practically and theoretically unsatisfactory for three main reasons. First, symmetric matrices with null or negative eigenvalues appear during Euclidean computations. And from a physical point of view, in DTI, a diffusion exactly equal to zero is impossible: above 0°Kelvin, water molecules *will* move in *all* directions. Even worse, a negative diffusion is meaningless. This occurs during iterated Euclidean computations, for instance during the estimation of tensors from diffusion-weighted images, the regularization of tensors fields, etc. To avoid going out of the tensor space, it has been proposed

to regularize only *features* extracted from tensors, like first eigenvectors [5] or orientations [4]. The regularization is propagated to tensors in a second step. This is not completely satisfactory, since it would be preferable to regularize tensors *directly* in order to take into account *all* the information they carry.

Second, a tensor corresponds typically to a *covariance matrix*. The value of its determinant is a direct measure of the dispersion of the associated multivariate Gaussian. The reason is that the volume of associated confidence regions are proportional to the square root of the covariance determinant. But the Euclidean averaging of tensors leads very often to a *tensor swelling effect*: the determinant (and thus the dispersion) of the Euclidean mean can be larger than the original determinants! In DTI, diffusion tensors are assumed to be covariance matrices of the local Brownian motion of water molecules. Introducing more dispersion in computations amounts to introducing more diffusion, which is physically unacceptable. For illustrations of this effect, see [4].

Third, the Euclidean metric is unsatisfactory in terms of *symmetry* with respect to matrix inversion. The Euclidean mean for tensors is an *arithmetic* mean which does not yield the identity for a tensor and its matrix inverse. When tensors model variability, one would rather have in many cases a *geometric* mean.

To fully circumvent these difficulties, affine-invariant Riemannian metrics have been recently proposed for tensors in [12, 8, 9, 10]. With them, negative and null eigenvalues are at an infinite distance, the swelling effect disappears and the symmetry with respect to inversion is respected. The price paid for this success is a high computational burden, essentially due to the curvature induced on the tensor space. Practically, this yields slow and hard to implement algorithms.

We propose a new Riemannian framework to fully overcome these computational limitations while preserving excellent theoretical properties. It is based on new metrics named *Log-Euclidean*, which are particularly simple to use. They result in classical Euclidean computations in the domain of matrix logarithms. In Section 2, we give an overview of the theory of Log-Euclidean metrics, detailed in [1]. In particular, we briefly compare these new metrics to affine-invariant metrics. In Section 3, we highlight the differences between the three frameworks with experimental results on synthetic and real DT-MRI data in interpolation and regularization. Computations are very simple and experimentally much faster in the Log-Euclidean than in the affine-invariant framework.

2 Presentation of the Log-Euclidean Framework

Complete proofs for all the results presented in this Section are given in [1].

Existence and Uniqueness of the Logarithm. A tensor S has a unique *symmetric matrix logarithm* $L = \log(S)$. It verifies $S = \exp(L)$ where \exp is the matrix exponential. Conversely, each symmetric matrix is associated to a tensor by the exponential. L is obtained from S by changing its eigenvalues into their natural logarithms, which can be done easily in an orthonormal basis in which S (and L) is diagonal.

A Vector Space Structure on Tensors. Since there is a one-to-one mapping between the tensor space and the *vector* space of symmetric matrices, one can transfer to tensors the addition “+” and the scalar multiplication “.” with the matrix exponential. This defines on tensors the *logarithmic multiplication* \odot and the *logarithmic scalar multiplication* \otimes , given by:

$$\begin{cases} S_1 \odot S_2 \stackrel{def}{=} \exp(\log(S_1) + \log(S_2)) \\ \lambda \otimes S \stackrel{def}{=} \exp(\lambda \cdot \log(S)) = S^\lambda. \end{cases} \quad (1)$$

The logarithmic multiplication is *commutative* and coincides with matrix multiplication whenever the two tensors S_1 and S_2 commute in the matrix sense. With \odot and \otimes , the tensor space has by construction a *vector space structure*, which is *not* the usual structure directly inherited from square matrices.

Log-Euclidean Metrics. When one considers only the multiplication \odot on the tensor space, one has a *Lie group* structure [11], i.e. a space which is both a smooth manifold and a group in which algebraic operations are *smooth mappings*. Among Riemannian metrics in Lie groups, the most convenient in practice, when they exist, are *bi-invariant metrics*, i.e. distances that are invariant by multiplication and inversion. For the tensor Lie group, bi-invariant metrics exist and are particularly simple. We have named such metrics *Log-Euclidean* metrics, since they correspond to Euclidean metrics in the domain of logarithms. From a Euclidean norm $\|\cdot\|$ on symmetric matrices, they can be written:

$$\text{dist}(S_1, S_2) = \|\log(S_1) - \log(S_2)\|. \quad (2)$$

Boundary Problems and Symmetry. Contrary to the classical Euclidean framework on tensors, one can clearly see from Eq. (2) that matrices with null or negative eigenvalues are at an infinite distance from tensors and will not appear in practical computations. Moreover, distances are not changed by inversion.

Invariance by Similarity. Log-Euclidean metrics are not affine-invariant. However, some of them are invariant by similarity (orthogonal transformation and scaling). This means that if tensors are covariance matrices, computations on tensors using these metrics will be invariant with respect to a change of coordinates obtained by a similarity. The similarity-invariant Log-Euclidean metric used throughout this article is given by:

$$\text{dist}(S_1, S_2) = (\text{Trace}(\{\log(S_1) - \log(S_2)\}^2))^{\frac{1}{2}}. \quad (3)$$

Euclidean Calculus in the Logarithmic Domain. The tensor vector space with a Log-Euclidean metric is in fact *isomorphic* (the algebraic structure of vector space is conserved) and *isometric* (distances are conserved) with the corresponding Euclidean space of symmetric matrices. As a consequence, the Riemannian framework for statistics and analysis is extremely simplified [1]. In particular, the Log-Euclidean mean of N tensors with arbitrary positive weights

$(w_i)_{i=1}^N$ such that $\sum_{i=1}^N w_i = 1$ is a direct generalization of the geometric mean of positive numbers and is given explicitly by:

$$\mathbb{E}_{LE}(S_1, \dots, S_N) = \exp \left(\sum_{i=1}^N w_i \log(S_i) \right). \quad (4)$$

This is remarkable: *in this framework, the processing of tensors is simply Euclidean in the logarithmic domain. Final results obtained on logarithms are mapped back to the tensor domain with the exponential.* Hence, statistical tools or PDEs are readily generalized to tensors in this framework.

Comparison with Affine-Invariant Metrics. As shown experimentally in Section 3, Log-Euclidean computations provide results similar to their affine-invariant equivalent, presented in [12]. The reason is the two families of metrics provide two generalizations of the geometric mean of positive numbers on tensors. Contrary to the Log-Euclidean mean, there is in general no closed form for the affine-invariant mean but rather a *barycentric equation*. Nevertheless, the determinants of the two means are both equal to the scalar geometric mean of the determinants of the averaged tensors [1]. This explains their likeness and the absence of swelling effect in both cases. This resemblance between the two means propagates to general computations which involve averaging, such as interpolation, extrapolation and regularization. The two means are even identical in a number of cases, in particular when averaged tensors commute. Yet they are *not* equal in general: Log-Euclidean means are slightly more anisotropic.

3 Experimental Results

3.1 Bilinear and Trilinear Interpolation

Often, voxels in clinical DT images are quite anisotropic. But algorithms tracking white matter fascicles are more efficient with isotropic voxels [3]. An adequate interpolation method is therefore important for such algorithms.

Fig. 1 shows the results obtained for the bilinear interpolation of four tensors with three methods: Euclidean (linear interpolation of coefficients), affine-invariant and Log-Euclidean bilinear interpolations. There is a pronounced swelling effect in the Euclidean case, which is not physically acceptable. On the contrary, both Riemannian interpolations provide the same geometric interpolation of determinants. There is a slightly larger anisotropy in Log-Euclidean means, which is a general effect discussed in [1]. The computation of the affine-invariant mean is iterative (we use the Gauss-Newton method described in [12]), whereas the closed form given by Eq. (4) is used directly in the Log-Euclidean case. This has a large impact on computation times: 0.003s (Euclidean), 0.009s (Log-Euclidean) and 1s (affine-invariant) for a 5×5 grid on a Pentium M 2 GHz.

To compare the Euclidean and Riemannian bilinear interpolations on real data, we have reconstructed by bilinear interpolation a down-sampled DTI slice. One column out of two and one line out of two were removed. The slice was

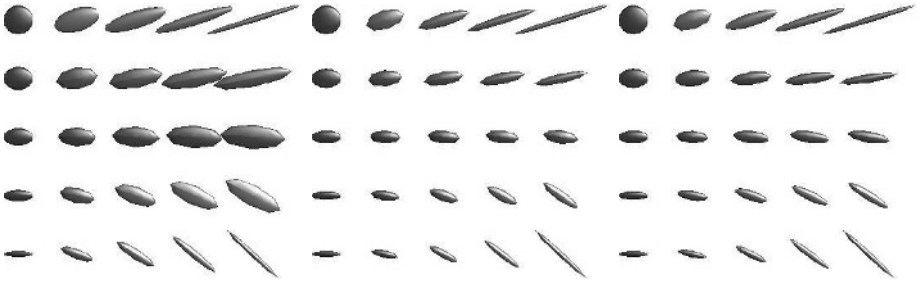


Fig. 1. Bilinear interpolation of 4 tensors at the corners of a grid. Left: Euclidean interpolation. **Middle:** affine-invariant interpolation. **Right:** Log-Euclidean interpolation. Note the characteristic swelling effect observed in the Euclidean case, which is not present in both Riemannian frameworks. Note also that Log-Euclidean means are slightly more anisotropic than their affine-invariant counterparts. The coloring of ellipsoids is based on the direction of dominant eigenvectors.

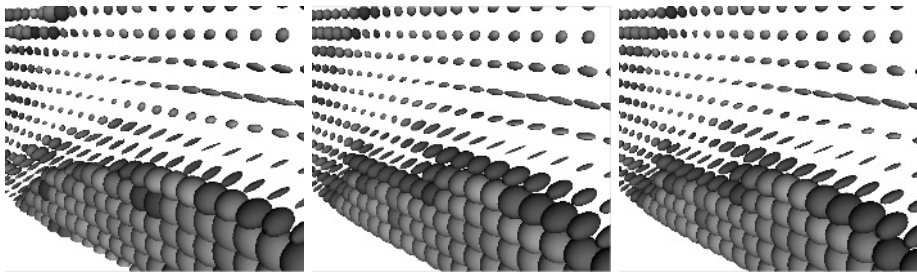


Fig. 2. Bilinear interpolation in a real DTI slice. Left: Original DTI slice, before down-sampling. **Middle:** Euclidean interpolation. **Right:** Log-Euclidean interpolation. Half the columns and lines of the original DTI slice were removed before reconstruction with a bilinear interpolation. The slice is taken in the mid-sagittal plane and displayed in perspective. Note how the tensors corresponding to the corpus callosum (in red, above the large and round tensors corresponding to a part of the ventricles) are better reconstructed (more anisotropic) in the Log-Euclidean case.

chosen in the mid-sagittal plane where strong variations are present in the DT image. The results in Fig. 2 show that the tensors corresponding to the corpus callosum are better reconstructed in the Log-Euclidean case. Affine-invariant results are very close to Log-Euclidean results and not shown here.

3.2 Regularization of Tensor Fields

DT images are corrupted by noise, and regularizing them can be a crucial preliminary step for DTI-based algorithms that reconstruct the white matter connectivity. As shown in [12], Riemannian metrics provide a general framework to provide such a regularization. We focus here on a typical Riemannian crite-

rior for the regularization of tensor fields. An anisotropic regularization can be obtained by the minimization of a Φ -functional [4] $Reg(S)$ given by:

$$Reg(S) = \int_{\text{Image}} \Phi(\|\nabla S\|_{S(x)}(x)) dx. \quad (5)$$

$\Phi(s)$ is a function penalizing large values of the norm of the *spatial gradient* ∇S of the image. Here, we use $\Phi(s) = \sqrt{1 + s^2/\kappa^2}$. Contrary to the Euclidean case, the norm of ∇S *depends explicitly on the current point* $S(x)$ and is given by:

$$\|\nabla S\|_{S(x)}^2 = \sum_{i=1}^d \left\| \frac{\partial S}{\partial x_i}(x) \right\|_{S(x)}^2. \quad (6)$$

In general and in particular in the affine-invariant case, this dependence on the current point leads to complex resolution methods. Practically, this implies in the affine-invariant case an intensive use of matrix inverses, square roots, exponentials and logarithms [12]. But in the Log-Euclidean framework, the general Riemannian formulation is extremely simplified. The reason is that the dependence on the current tensor disappears on logarithms [1]:

$$\|\nabla S\|_{S(x)} = \|\nabla \log(S)\|. \quad (7)$$

Consequently, *the energy functional can be minimized directly on the vector field of logarithms*. The regularized tensor field is given in a final step by the matrix exponential of regularized logarithms. Interestingly, mathematical issues such as existence and uniqueness of PDEs on tensors in the Log-Euclidean framework are simply particular cases of the classical theory of PDEs on vector-valued images.

In the following experiments, the minimization method used is a first-order gradient descent with a fixed time step dt . We use an explicit finite difference scheme on logarithms in the Log-Euclidean case and the geodesic marching scheme described in [12] in the affine-invariant case. In the Euclidean framework, we also use affine-invariant geodesic marching rather than a classical explicit scheme to limit the appearance of non-positive eigenvalues, proceeding similarly as in [4]. Homogeneous Neumann boundary conditions are used, $\kappa = 0.05$, $dt = 0.1$ and 100 iterations are performed.

As a first example, we restore a noisy synthetic image of tensors. Results are shown in Fig. 3: the negative impact of the Euclidean swelling effect is clearly visible. On the contrary, both Riemannian frameworks yield proper results, the only (small) difference being slightly more anisotropy for Log-Euclidean results.

Let us now turn to a real DTI volume of the brain with $128 \times 128 \times 30$ voxels with spatial dimensions of $1.875 \times 1.875 \times 4 \text{ mm}^3$. The b-value is 1000 s.mm^{-2} . As shown in Fig. 4, both Riemannian results are qualitatively very satisfactory: the smoothing is done without blurring the edges. They are also very similar to each other, with only slightly more anisotropy in the Log-Euclidean case. As before, the Euclidean results are marred by a pronounced swelling effect. Computations are much faster in the Log-Euclidean case: 30 minutes instead of 122 minutes for affine-invariant results on a Pentium Xeon 2.8 GHz with 1 Go of RAM.

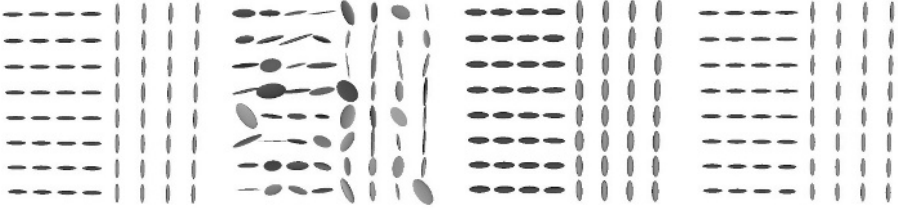


Fig. 3. Regularization of a synthetic DTI slice. **Left:** original synthetic data. **Middle Left:** noisy data. **Middle Right:** Euclidean regularization. **Right:** Log-Euclidean regularization. The original data is properly reconstructed in the Log-Euclidean case, as opposed to the Euclidean case where the result is marred by the swelling effect.

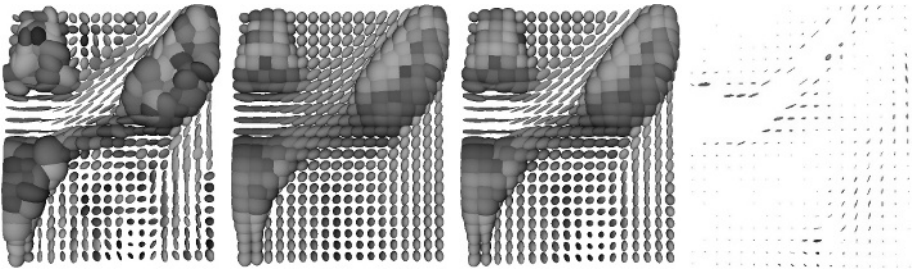


Fig. 4. Regularization of a real DTI volume. **Left:** close-up on the top right ventricle and nearby. **Middle Left:** Euclidean regularization. **Middle Right:** Log-Euclidean regularization. **Right:** highly magnified view ($\times 100$) of the absolute value (the absolute value of eigenvalues is taken) of the difference between Log-Euclidean and affine-invariant results. Note that there is *no* tensor swelling in the Riemannian cases, contrary to the Euclidean case. Log-Euclidean and affine-invariant results are very similar, the only difference being slightly more anisotropy in Log-Euclidean results.

4 Discussion and Perspectives

In this work, we have presented a particularly simple and efficient Riemannian framework for tensor calculus, called Log-Euclidean. As in the affine-invariant case, the defects of the Euclidean framework are corrected with Log-Euclidean metrics, but without any unnecessary technicality. Indeed, Riemannian computations on tensors are converted into Euclidean computations on vectors in this novel framework. In practice, classical statistical tools and PDEs for vectors can be directly used on the matrix logarithms of tensors, which are simple vectors. Moreover, all usual operations on tensors can be efficiently carried out in this framework, like the joint estimation and smoothing of DTI from diffusion-weighted images, as shown in [6].

In future work, we will study in further details the restoration of noisy DT images. In particular, we plan to quantify the impact of the regularization on the tracking of fibers in the white matter of the human nervous system. We also

intend to use this new framework to better model and reconstruct the anatomical variability of the human brain with tensors [7].

Acknowledgments and Pending Patent

The authors thank Denis Ducreux, MD, Kremlin-Bicêtre Hospital (France), for the DT-MRI data he kindly let us use in this work.

A patent is pending for the Log-Euclidean processing framework of tensors (Deposit Number 0503483, April 7th, 2005, property of INRIA, France).

References

1. V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Fast and simple computations on tensors with Log-Euclidean metrics. Research Report 5584, INRIA, May 2005.
2. P. Basser, J. Mattiello, and D. Le Bihan. MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66:259–267, 1994.
3. P. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi. In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine*, 44:625–632, 2000.
4. C. Chefd’hotel, D. Tschumperlé, R. Deriche, and O. Faugeras. Regularizing flows for constrained matrix-valued images. *J. Math. Im. Vis.*, 20(1-2):147–162, 2004.
5. O. Coulon, D. Alexander, and S. Arridge. Diffusion tensor magnetic resonance image regularization. *Medical Image Analysis*, 8(1):47–67, 2004.
6. P. Fillard, V. Arsigny, X. Pennec, and N. Ayache. Joint estimation and smoothing of clinical DT-MRI with a Log-Euclidean metric. Research Report RR-5607, INRIA, Sophia-Antipolis, France, June 2005.
7. P. Fillard, V. Arsigny, X. Pennec, P. Thompson, and N. Ayache. Extrapolation of sparse tensor fields: Application to the modeling of brain variability. In G. Christensen and M. Sonka, editors, *Proc. of IPMI’05, LNCS*, Glenwood springs, Colorado, USA, July 2005. Springer. To appear.
8. P.T. Fletcher and S.C. Joshi. Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In *Proc. of CVAMIA and MMBIA Workshops, Prague, Czech Republic, May 15, 2004*, LNCS 3117, pages 87–98. Springer, 2004.
9. C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras. Statistics on multivariate normal distributions: A geometric approach and its application to diffusion tensor MRI. Research Report 5242, INRIA, June 2004.
10. M. Moakher. A differential geometry approach to the geometric mean of symmetric positive-definite matrices. *SIAM Jour. on Mat. Anal. and Appl.*, 2004. To appear.
11. Xavier Pennec. Probabilities and statistics on Riemannian manifolds: a geometric approach. Research Report 5093, INRIA, January 2004.
12. Xavier Pennec, Pierre Fillard, and Nicholas Ayache. A Riemannian framework for tensor computing. Research Report 5255, INRIA, July 2004.
13. C. Poupon, C. A. Clark, V. Frouin, J. Regis, I. Bloch, D. Le Bihan, and J.-F. Mangin. Regularization of diffusion-based direction maps for the tracking of brain white matter fascicles. *Neuroimage*, 12(2):184–95, August 2000.
14. Z. Wang, B. Vemuri, Y. Chen, and T. H. Mareci. A constrained variational principle for simultaneous smoothing and estimation of the diffusion tensors from complex DWI data. *IEEE TMI*, 23(8):930–939, 2004.
15. C.-F. Westin, S. E. Maier, H. Mamata, A. Nabavi, F. A. Jolesz, and R. Kikinis. Processing and visualization of diffusion tensor MRI. *Media*, 6:93–108, 2002.

3D Curve Inference for Diffusion MRI Regularization

Peter Savadjiev¹, Jennifer S.W. Campbell²,
G. Bruce Pike², and Kaleem Siddiqi¹

McGill University, Montréal, QC, Canada

¹ School of Computer Science & Centre For Intelligent Machines

² McConnell Brain Imaging Centre, Montréal Neurological Institute

Abstract. We develop a differential geometric framework for regularizing diffusion MRI data. The key idea is to model white matter fibers as 3D space curves and to then extend Parent and Zucker’s 2D curve inference approach [8] by using a notion of *co-helicity* to indicate compatibility between fibre orientation estimates at each voxel with those in a local neighborhood. We argue that this provides several advantages over earlier regularization methods. We validate the approach quantitatively on a biological phantom and on synthetic data, and qualitatively on data acquired *in vivo* from a human brain.

1 Introduction and Related Work

In this article, we consider the problem of regularizing orientation fields obtained using either Diffusion Tensor Imaging (DTI) [1] or High Angular Resolution Diffusion Imaging (HARDI) [12] measurements. We view both cases in a differential geometric setting where white matter fibers are modeled as 3D space curves. We extend Parent and Zucker’s 2D curve inference approach [8] by using a notion of *co-helicity* to model the compatibility between orientation estimates in 3D. In this context, curvature and torsion play key roles in the interpretation of a *tangent bundle*, where average local support is maximized using relaxation labeling techniques. We demonstrate that using 3D curve inference for the regularization of such data has advantages over earlier methods, including: 1) the possibility of representing multiple orientations at the same voxel, 2) applicability to both DTI and HARDI data, 3) numerical robustness in the vicinity of sparse measurements and 4) estimates of curvature and torsion at each voxel, which can be useful to guide fiber tracking algorithms.

We begin by briefly mentioning some of the key approaches to diffusion MRI regularization. Poupon *et al.* [9] use a “spaghetti plate” model for the regularization of DTI data. Westin *et al.* [7] regularize the individual elements of each DTI tensor using a six-dimensional, multivariate Gaussian Markov Random Field. Vemuri *et al.* [13] propose a constrained variational principle for the simultaneous estimation and regularization of DTI tensors directly from diffusion weighted images. Coulon *et al.* [5] combine variational and anisotropic diffusion methods to separately regularize orientation fields and diffusivities. Tschumperlé and

Deriche [11] also propose variational methods for regularizing orientation fields in DTI data while preserving discontinuities, with the constraint that orthonormality between the eigenvectors at each location is preserved.

All of the above algorithms have the inherent limitations that: 1) they were designed explicitly for DTI data and cannot be trivially extended to handle HARDI data and 2) they assume (explicitly or implicitly) a single fiber orientation at each voxel and thus cannot handle branchings or crossings. To our knowledge the only algorithms which address these concerns, at least in part, are those proposed in [10, 3, 4]. Whereas some of these latter methods have been validated on synthetic data, few (or none) have been demonstrated on a biological phantom with known ground truth fiber orientations.

2 3D Curve Inference

We assume as input a 3D curve orientation distribution function sampled on a regular (typically rectangular) 3D lattice. Following Parent and Zucker's 2D curve inference methodology [8], local curve orientation estimates can be interpreted as an initial *tangent bundle* at each location in the lattice. Our goal is to obtain estimates of the trace, tangent, curvature and torsion fields of curves in the 3D volume. This is done by using a notion of *co-helicity* between three tangents, which is the natural extension to 3D of Parent and Zucker's *co-circularity* constraint between a pair of tangents in 2D [8]. In this framework, an *osculating helix* (which has constant curvature and torsion) is used to approximate a curve passing through a given point.

2.1 Properties of a Helix and Co-Helicity

A circular helix is a curve inscribed on the surface of a cylinder, such that at all points on the curve, the associated tangent vector forms a constant angle with the cylinder's axis. Consider such a helix, parametrized by t , with its axis coinciding with the z -axis. Its equations and those of its unit tangent and unit normal are given by:

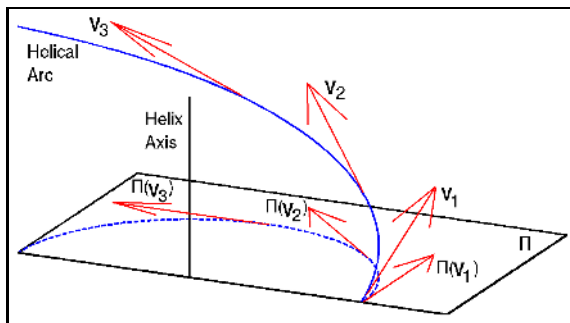


Fig. 1. An illustration of co-helicity between three vectors \mathbf{v}_1 , \mathbf{v}_2 and \mathbf{v}_3

$$\mathbf{x}(t) = (x(t), y(t), z(t)) = (r \cos(t), r \sin(t), ct), \quad (1)$$

$$\frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|} = \frac{(x'(t), y'(t), z'(t))}{\|(x'(t), y'(t), z'(t))\|} = \frac{(-r \sin(t), r \cos(t), c)}{\sqrt{r^2 + c^2}}, \quad (2)$$

$$\frac{\mathbf{x}''(t)}{\|\mathbf{x}''(t)\|} = \frac{(x''(t), y''(t), z''(t))}{\|(x''(t), y''(t), z''(t))\|} = (-\cos(t), -\sin(t), 0). \quad (3)$$

Here r is the radius of the helix and c is a constant defining the vertical separation of the helical loops (measured along the helix axis). The orthographic projection of a helix onto the xy plane is a circle with radius r .

Definition 1. *Three orientations or vectors in 3D are co-helical if there is a helix to which all three are tangent.*

The concept of co-helicity is illustrated in Fig. 1, which shows a helical arc, three of its tangent vectors labeled \mathbf{v}_1 , \mathbf{v}_2 and \mathbf{v}_3 , along with the helix axis and its orthogonal plane Π . The orthographic projection of the helix onto Π is a circle (shown with a dashed curve), and the projections of \mathbf{v}_1 , \mathbf{v}_2 and \mathbf{v}_3 onto Π are labeled $\Pi(\mathbf{v}_1)$, $\Pi(\mathbf{v}_2)$ and $\Pi(\mathbf{v}_3)$. These three projections are co-circular [8], since they are all tangent to the same circle in plane Π .

Theorem 1. *Given three unit vectors \mathbf{v}_1 , \mathbf{v}_2 and \mathbf{v}_3 in \mathcal{R}^3 specified at three locations, it is possible to determine whether or not they are co-helical. Furthermore, if they are co-helical, it is possible to recover the parameters of the helix passing through these locations and having \mathbf{v}_1 , \mathbf{v}_2 and \mathbf{v}_3 as tangent vectors.*

Owing to space restrictions we omit the proof, but present an algorithm (Algorithm 1) for determining whether or not \mathbf{v}_1 , \mathbf{v}_2 and \mathbf{v}_3 in \mathcal{R}^3 are co-helical, and if so, for recovering the parameters of the helix.

2.2 Relaxation Labeling

In Parent and Zucker's 2D curve inference framework [8], tangent fields are estimated over a 2D image by maximizing the *average local support*, defined by $A(p) = \sum_{i=1}^n s_i(\lambda) p_i(\lambda)$. Expressions of this type can be maximized using the relaxation labeling algorithm of [6]. Here $p_i(\lambda)$ indicates the confidence in orientation λ at pixel i and its (local) neighborhood support given by $s_i(\lambda) = \sum_{j=1}^n \sum_{\lambda'=1}^m r_{ij}(\lambda, \lambda') p_j(\lambda')$. Each j denotes a node (i.e. pixel) in the neighborhood of i , λ' is an orientation at that node, $r_{ij}(\lambda, \lambda')$ represents the compatibility between orientation λ at node i and orientation λ' at node j , and there are a total of m orientations and n nodes. The compatibility coefficients are designed so that *co-circular* tangents lend one another support, while other configurations are suppressed. This imposes a constraint on the variation of curvature, while providing as a bi-product local estimates of curvature as well. The compatibility coefficients are further refined to include a partitioning into curvature classes [8].

The key to extending the above 2D curve inference framework to 3D is to replace the notion of *co-circularity* between two tangents with one of *co-helicity*

Algorithm 1: Determining co-helicity between three orientation estimates

Data : Three 3D vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$

Result : 1 if there is a helix to which $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are tangent, 0 otherwise.

1. (testing for a special case) If the three vectors are collinear, return 1.
2. Define a vector \mathbf{n} as the cross-product of the difference vectors $\mathbf{v}_3 - \mathbf{v}_2$ and $\mathbf{v}_2 - \mathbf{v}_1$. This is the axis of the putative helix.
3. Compute the projections $\Pi(\mathbf{v}_1), \Pi(\mathbf{v}_2)$ and $\Pi(\mathbf{v}_3)$ of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ onto the plane that has \mathbf{n} as its normal vector. That fixes the value of the parameter r in the helix equation (1). That is, r is the radius of the circle, lying in the plane orthogonal to \mathbf{n} , to which $\Pi(\mathbf{v}_1), \Pi(\mathbf{v}_2)$ and $\Pi(\mathbf{v}_3)$ are all tangent.
4. Without loss of generality we can set the parameter t (see (1)) of \mathbf{v}_1 to $t_1 = 0$, the parameter t of \mathbf{v}_2 is then simply the angle $t_2 = \arccos\left(\left\langle \frac{\Pi(\mathbf{v}_1)}{\|\Pi(\mathbf{v}_1)\|}, \frac{\Pi(\mathbf{v}_2)}{\|\Pi(\mathbf{v}_2)\|} \right\rangle\right)$.
5. Using (1) and (2), it is easy to show that

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \frac{r^2 \cos(t_2 - t_1) + c^2}{r^2 + c^2} \quad (4)$$

(assuming that \mathbf{v}_1 and \mathbf{v}_2 have been normalized to unit length). With the values for r, t_1 and t_2 obtained in steps 1 and 2, we can derive c using (4).

6. Given c , we can calculate (using (2)) the angle that the tangent vectors should make with the helix axis, as well as the displacement in the direction of the helix axis of the position of \mathbf{v}_1 and \mathbf{v}_2 ((1)). It is possible to verify whether \mathbf{v}_1 and \mathbf{v}_2 possess indeed these properties. If so, return 1. If not, return 0.
 7. If any of the above steps fail (e.g. there is no one circle to which $\Pi(\mathbf{v}_1), \Pi(\mathbf{v}_2)$ and $\Pi(\mathbf{v}_3)$ are all tangent in step 3) return 0.
-

between three tangents. The compatibility condition obviously has to be generalized to a notion of higher order compatibility. Following [6], this is done by replacing the support function by

$$s_i(\lambda) = \sum_{j, \lambda'} \sum_{k, \lambda''} r_{ijk}(\lambda, \lambda', \lambda'') p_j(\lambda') p_k(\lambda''). \quad (5)$$

Here $r_{ijk}(\lambda, \lambda', \lambda'')$ represents the compatibility (co-helicity) between orientation λ at node i , orientation λ' at node j and orientation λ'' at node k , as determined using Algorithm 1.

2.3 Implementation

The 3D curve inference algorithm can use an arbitrarily large label set. In our implementation, we use a label set of 90 unit direction vectors distributed isotropically over a hemisphere, obtained using an electrostatic charge repulsion algorithm. The relaxation labeling technique is implemented according to Algorithm 8.2 in [6], together with the radial projection method (Appendix A in [8]). The

support coefficients are calculated in a spherical neighborhood of each voxel. We obtain a significant increase in efficiency by precomputing the co-helical configurations of tangent triplets and storing these in a look-up table. The initial values for $p_i(\lambda)$ are obtained from the value of the orientation distribution function (ODF) along orientation λ at voxel i , and a discretization of the allowed range of variation of curvature and torsion (the 3D extension of the curvature classes in [8]) is implemented. Note that the output of our algorithm is not an ODF any longer, but a regularized tangent bundle, which is an indication of confidence in the underlying curve (i.e. white matter fibre) orientations at each voxel.

3 Validation

Quantitative Validation On a Biological Phantom: A biological phantom was created from two excised Sprague-Dawley rat spinal cords embedded in 2% agar. Two diffusion-weighted datasets were acquired using this phantom, with 90 diffusion encoding directions, with b values of 1300 s/mm^2 and 3000 s/mm^2 , respectively. The first was used for diffusion tensor reconstruction of the diffusion ODF and the second for high angular resolution reconstruction, using the q-ball technique [12]. A T1-weighted image of this phantom is shown in Fig. 2 (top left). The ground truth orientations were determined by extracting the centerlines of each cord using the technique of [2] and then smoothly extending the orientation estimates in the center to the boundary of the cord, for each cord (Fig. 2, top

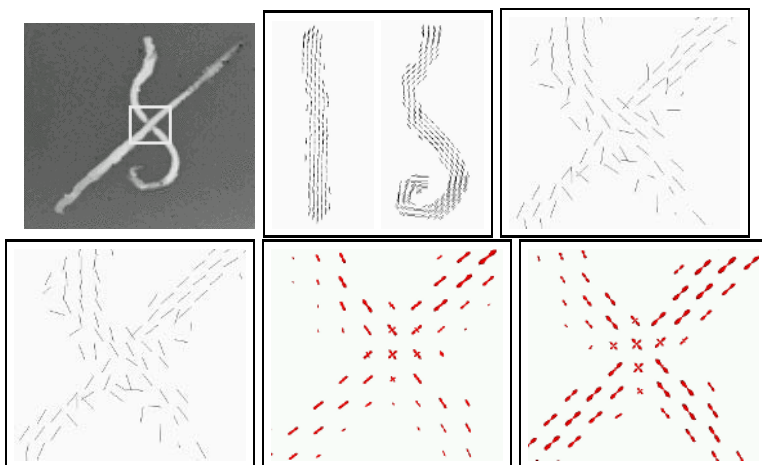


Fig. 2. Top Row: A biological phantom created by overlaying two rat cord spinal cords (left), ground truth orientation estimates shown separately for the two cords (middle), and the principal eigenvector orientations of the DTI image in the vicinity of the crossing (right). Bottom Row: The regularized orientation estimates obtained using the technique of [11] (left), those obtained using 3D curve inference on the DTI reconstruction (middle), and the regularized HARDI reconstruction using 3D curve inference (right).

middle). Fig. 2 (top right) shows the principal components of the diffusion tensors in the region of the crossing, indicated by a rectangle in Fig. 2 (top left). The bottom row of Fig. 2 shows the results obtained using Tschumperlé and Deriche’s orthonormal vector regularization [11] (OVR-DTI) (left), curve inference on the DTI dataset (CI-DTI) (middle) and curve inference on the HARDI dataset (CI-HARDI) (right). The average orientation errors in degrees (± 1 std. deviation) between the ground truth dataset and the unregularized as well as regularized datasets for each case are shown in Fig. 3 (left). Observe that the results obtained using curve inference give significantly lower errors. Furthermore, the lowest error is obtained by applying curve inference to HARDI data. The phantom data is challenging because of eddy current induced artefacts in many boundary voxels, where the measured principal direction of diffusion is perpendicular to the ground truth orientation. In the case of the curve inference experiments, the error is measured between the maximally supported orientation at each voxel in the regularized data, and the corresponding ground truth orientation.

Quantitative Validation On Synthetic Data: A synthetic DTI dataset in a $100 \times 50 \times 100$ voxel grid was created by placing anisotropic diffusion tensors with their principal direction vector aligned with one of three curves: a planar sine wave, and two helices with different curvature and torsion. Partial volume averaging effects were simulated in voxels where the helices intersect the sine wave, and background voxels were filled with isotropic (spherical) tensors. All tensors had a mean eigenvalue of 3. In anisotropic regions the principal eigenvalue was set to 7 and the others to 1. In voxels with crossings, the two principal directions had eigenvalues of 4, and the other eigenvalue was set to 1. The two angles describing the orientation of each tensor were perturbed (independently) by adding Gaussian noise, with mean 0 and a standard deviation of 22.9° (0.4 rad). The original (noiseless) dataset was treated as the ground truth. One view of the noisy dataset is shown in Fig. 3 (right); it is important to note that the helices are non-planar curves. Validation results are shown in the right column of the table in Fig. 3 (left). Observe that once again curve inference (CI-DTI) achieves a significant reduction in orientation error, compared to both the noisy unregularized data, as well as the result obtained with OVR-DTI [11].

Qualitative Validation On Human Brain Data: We conclude with regularization results using *in vivo* human brain data. Fig. 4 (top left) shows a slice through the data, which consists of a DTI reconstruction of the diffusion ODF. The data covers an area with multiple fibre directions, including voxels with partial volume averaging effects, as well as voxels with cerebro-spinal fluid and grey matter that do not exhibit high diffusion anisotropy. Fig. 4 (top right) shows one region of interest (ROI) of the regularization result using 3D curve inference, and Fig. 4 (bottom left) shows a slice through the corresponding 3D fractional anisotropy image of the brain, with a white rectangle indicating the ROI. A zoom-in on the bottom-left part of the regularization result is given in Fig. 4 (bottom right), showing the recovery of multiple fibre directions in voxels with partial volume averaging effects. Qualitatively, intra-voxel crossings and the apparent variation

	Phantom	Synthetic
Unreg. DTI	$27.1 \pm 29.0^\circ$	$24.9 \pm 14.5^\circ$
Unreg. HARDI	$21.0 \pm 21.9^\circ$	N/A
CI-DTI	$11.0 \pm 16.3^\circ$	$11.2 \pm 8.2^\circ$
CI-HARDI	$10.3 \pm 17.3^\circ$	N/A
OVR-DTI	$24.0 \pm 22.2^\circ$	$21.2 \pm 12.9^\circ$




Fig. 3. LEFT: Table of validation results showing average orientation errors in degrees for the biological phantom data set and the synthetic data set. RIGHT: A snapshot of the noisy synthetic data set, prior to regularization.

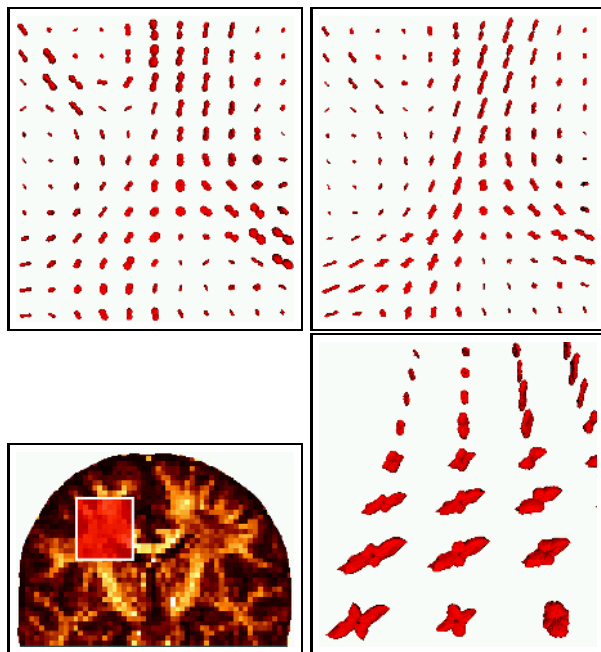


Fig. 4. A ROI through the brain DTI data (top left) with the regularization results using curve inference (top right). A slice through the associated fractional anisotropy image (bottom left) with a white rectangle enclosing the ROI. A zoom-in on the result from a different viewpoint in a region of partial volume averaging effects (bottom right).

of curvatures are recovered well from the DTI data. Quantitative evaluation of these results is difficult, due to the lack of ground truth.

4 Conclusion

We have presented a differential geometric framework for regularizing diffusion MRI data, where a notion of *co-helicity* is used to compute support for orienta-

tions given neighbouring orientation estimates. Since multiple orientations can receive support in the same voxel, the algorithm is applicable to configurations of crossings or branchings, and it handles DTI and HARDI data in an identical way. It can also be applied to the regularization of any set of ODFs distributed over a discrete 3D lattice. As a bi-product of the algorithm, we obtain discrete estimates of the curvature and torsion of the likely curves at each voxel, which could be used to guide fiber tracking algorithms. We have validated the algorithm quantitatively on a biological phantom and on a synthetic data set, and qualitatively on *in vivo* human brain data.

Acknowledgments. We thank Sylvain Bouix for his centerline extraction algorithm and Vladimir V. Rymar for his help with preparing the biological phantom. This work was supported by grants from FQRNT, NSERC and CIHR.

References

1. P.-J. Basser, J. Mattiello, and D. Le Bihan. MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66:259–267, 1994.
2. S. Bouix, K. Siddiqi, and A. Tannenbaum. Flux driven automatic centerline extraction. *Medical Image Analysis*, 9(3):209–221, 2005.
3. Y. Chen, W. Guo, Q. Zeng, X. Yan, F. Huang, H. Zhang, G. He, B.-C. Vemuri, and Y. Liu. Estimation, smoothing, and characterization of apparent diffusion coefficient profiles from high angular resolution DWI. In *Proc. IEEE CVPR*, pages 588–593, 2004.
4. Y. Cointepas, C. Poupon, D. Le Bihan, and J.-F. Mangin. A spin glass framework to untangle fiber crossing in MR diffusion based tracking. In *Proc. MICCAI 2002, LNCS 2488*, pages 475–482, 2002.
5. O. Coulon, D. Alexander, and S. Arridge. A regularization scheme for diffusion tensor magnetic resonance images. In *Proc. IPMI 2001, LNCS 2082*, pages 92–105, 2001.
6. R. Hummel and S. Zucker. On the foundations of relaxation labeling processes. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 5:267–287, 1983.
7. M. Martín-Fernández, C.-F. Westin, and C. Alberola-Lopez. 3D Bayesian regularization of diffusion tensor MRI using multivariate gaussian markov random fields. In *Proc. MICCAI 2004, LNCS 3216*, pages 351–359, 2004.
8. P. Parent and S. Zucker. Trace inference, curvature consistency, and curve detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11:823–839, 1989.
9. C. Poupon, C. Clark, V. Frouin, J. Régis, I. Bloch, D. Le Bihan, and J.-F. Mangin. Regularization of diffusion-based direction maps for the tracking of brain white matter fascicles. *NeuroImage*, 12:184–195, 2000.
10. A. Ramírez-Manzanares and M. Rivera. Brain nerve bundles estimation by restoring and filtering intra-voxel information in diffusion tensor MRI. In *Proc. IEEE VLISM*, 2003.
11. D. Tschumperlé and R. Deriche. Orthonormal vector sets regularization with PDE’s and applications. *Int. Journal of Computer Vision*, 50:237–252, 2002.
12. D. Tuch, T. Reese, M. Wiegell, and V. Wedeen. Diffusion MRI of complex neural architecture. *Neuron*, 40:885–895, 2003.
13. Z. Wang, B. Vemuri, Y. Chen, and T. Mareci. A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI. *IEEE Trans. on Medical Imaging*, 23:930–939, 2004.

Fiber Tract-Oriented Statistics for Quantitative Diffusion Tensor MRI Analysis^{*}

Isabelle Corouge^{1,2}, P. Thomas Fletcher³, Sarang Joshi^{1,4},
John H. Gilmore², and Guido Gerig^{1,2}

¹ Departments of Computer Science, ² Psychiatry, ⁴ Radiation Oncology,
University of North Carolina, Chapel Hill, USA

³ Scientific Computing and Imaging Institute, University of Utah, USA
corouge@unc.edu, fletcher@sci.utah.edu, sjoshi@unc.edu,
John.Gilmore@med.unc.edu, gerig@cs.unc.edu

Abstract. Diffusion tensor imaging (DTI) has become the major modality to study properties of white matter and the geometry of fiber tracts of the human brain. Clinical studies mostly focus on regional statistics of fractional anisotropy (FA) and mean diffusivity (MD) derived from tensors. Existing analysis techniques do not sufficiently take into account that the measurements are tensors, and thus require proper interpolation and statistics based on tensors, and that regions of interest are fiber tracts with complex spatial geometry. We propose a new framework for quantitative tract-oriented DTI analysis that includes tensor interpolation and averaging, using nonlinear Riemannian symmetric space. As a result, tracts of interest are represented by the geometry of the medial spine attributed with tensor statistics calculated within cross-sections. Examples from a clinical neuroimaging study of the early developing brain illustrate the potential of this new method to assess white matter fiber maturation and integrity.

Keywords: Diffusion tensor interpolation, diffusion tensor statistics, DTI analysis, fiber tract modeling.

1 Introduction

Diffusion tensor imaging of brain structures measures diffusion properties by the local probability of self-motion of water molecules. A tensor field is calculated from directional gradient images and characterizes amount and locally preferred directions of local diffusivity. While diffusion can be considered isotropic in fluid it appears highly anisotropic along neural fiber tracts due to inhibition of free diffusion of intra- and extracellular fluid. DTI has become the preferred modality to explore white matter properties

^{*} This research is supported by the NIH NIBIB grant P01 EB002779, the NIMH Silvio Conte Center for Neuroscience of Mental Disorders MH064065, and the UNC Neurodevelopmental Disorders Research Center HD 03110. The work is also funded by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant U54 EB005149-01, project NAMIC. We acknowledge the Insight Toolkit community for providing the software framework. Dr. Weili Lin, UNC Radiology, is acknowledged for active support of developing an improved neonatal DT MRI acquisition technique.

associated with brain connectivity *in vivo*. The literature proposes a variety of DTI processing techniques, ranging from tensor field computation to quantitative analysis, and including visualization, regularization, registration, tractography and population statistics. Few of these methods make use of the full tensor information though most would benefit from an appropriate mathematical framework for tensor operations and tensor statistics calculation. For instance, tensor interpolation is required in regularization, registration and spatial normalization. Tensor statistics calculation becomes necessary for statistical DTI analysis in population studies. So far, analysis schemes have mostly focused on measuring properties in regions of interest and to a lesser extent along fiber bundles [1], [2], and they have not made use of the full tensor information. Consequently, clinical studies have mostly been limited to statistics of FA or MD maps on a voxel-by-voxel basis [3].

As opposed to voxel-based analysis, we propose an object-oriented approach in which the fiber tracts act as coordinate systems for quantitative DTI analysis. Our concept provides a complete representation of each individual bundle, describing both geometry and diffusion properties. The representation includes model of the geometry of individual bundles and statistics of diffusion tensors to be associated with the geometric model. In this paper, we focus on the computation of DTI tensor interpolation and DTI tensor statistics. Tensor information is integrated across cross-sections and represented along bundles. We thus supplement our previous model of tract geometry [4] with statistics of diffusion tensors, from which we derive diffusion properties.

2 Theoretical Framework

We denote the space of all diffusion tensors, i.e., the space of all 3×3 symmetric, positive-definite matrices, as $PD(3)$. Averaging and interpolation of diffusion tensors can be formulated as a least-squares minimization problem in this space. This definition depends on the choice of metric, or distance, on the space $PD(3)$. Treating diffusion tensors as vectors in the space \mathbb{R}^9 , one can define a linear average of N diffusion tensors p_1, \dots, p_N as $\mu = \frac{1}{N} \sum_{i=1}^N p_i$. This definition minimizes the Euclidean metric on \mathbb{R}^9 . However, linear averages suffer from a “swelling” effect where diffusion tensors with the same determinant will have an average with a larger determinant. Linear interpolation of diffusion tensors suffers from this same effect. We adopt a more natural metric for averaging and interpolation by treating $PD(3)$ as a curved manifold, or more specifically, a Riemannian symmetric space. We use tensor averaging and interpolation methods, first presented in [5], [6], that are based on the notion of geodesic distance within this space. In a similar approach [7], interpolation is limited to only two tensors, whereas our work applies to averaging and interpolation of an arbitrary number of tensors and thus allows a full 3D interpolation of DTI data. The symmetric space metric does not suffer from the swelling effect of the linear metric, that is, diffusion tensors with the same determinant will have an average with the same determinant.

Symmetric spaces [8] arise from transformation groups on manifolds. The Riemannian metric is chosen to be invariant under the group transformations. The symmetric space structure of $PD(3)$ arises from transformations by $GL(3)$, the group of positive-determinant matrices. The transformation of a diffusion tensor $p \in PD(3)$ by a matrix

$g \in GL^+(3)$ is given by $p \mapsto gpg^T$. Because of the algebraic nature of the symmetric space structure, distance and geodesic computations on $PD(3)$ are also algebraic in nature. For instance, the geodesic distance between two tensors $p_1, p_2 \in PD(3)$ can be computed using singular-value decomposition (SVD) as follows:

- Let $p_1 = U\Lambda U^T$ be the SVD of p_1 , set $g = U\sqrt{\Lambda}$.
- Compute the action of g^{-1} on p_2 : $y = g^{-1}p_2(g^{-1})^T$.
- Again using SVD, compute the eigenvalues σ_i of y .
- The geodesic distance is $d(p_1, p_2) = \left(\sum_{i=1}^3 \log(\sigma_i)^2\right)^{\frac{1}{2}}$.

2.1 Statistics of Diffusion Tensors

We now define the mean and variance of diffusion tensors respecting the geometry of the space. Following Fréchet [9], we define the average as the minimum mean squared error estimator under the natural Riemannian metric defined above. Given a set of diffusion tensors $p_1, \dots, p_N \in PD(3)$ the mean is defined as

$$\mu = \arg \min_{p \in PD(3)} \sum_{i=1}^N d(p, p_i)^2. \tag{1}$$

This minimization problem can be solved using a gradient descent method as described in [5]. This is analogous to the algorithm for computing the intrinsic mean given by Pennec [10]. Having defined the mean, we define the sample variance of the data as the expected value of the squared geodesic distances from the mean:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N d(\mu, p_i)^2. \tag{2}$$

2.2 Interpolation of Tensors

For developing consistent interpolation between diffusion tensors we extend the above definition of the mean to weighted averaging. Using a least-squares criterion, we define the weighted average of diffusion tensors $p_1, \dots, p_N \in PD(3)$ as

$$\text{Ave}(\{w_i\}, \{p_i\}) = \arg \min_{p \in PD(3)} \sum_{i=1}^N w_i d(p, p_i)^2, \tag{3}$$

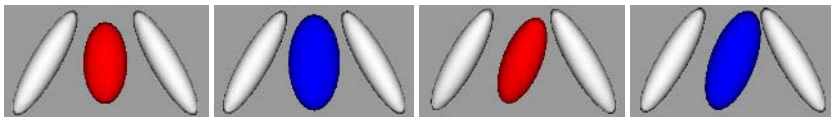


Fig. 1. Synthetic examples of weighted averages of tensors. The white ellipsoids average to the red ellipsoid with the geodesic method and to the blue ellipsoid with the linear method. Left: weights = {0.5,0.5}. Right: weights = {0.75,0.25}. It can be observed that the linear method does not preserve the determinant.

where w_1, \dots, w_N are positive real weights that sum to 1. Figure 1 provides synthetic examples of weighted averages of tensors.

For interpolating tensors within a voxel, trilinear weights may be used for the w_i . In this paper, we only focus on trilinear weights although higher order interpolation may be defined using the same concept. This interpolation is a natural generalization of trilinear interpolation of scalar values, i.e., if we replaced the diffusion tensors in the above definitions with real numbers, we would arrive at trilinear interpolation. It follows easily from the use of trilinear weights that the interpolation function does indeed interpolate the corner points. It can also be shown that the interpolation function is continuous on $[0, 1]^3$ (see [6] for a proof).

3 Fiber Tract Modeling and Analysis Methodology

Interpolation and averaging of tensors is applied for quantitative fiber tract-oriented analysis of DTI. The geometry of an individual fiber tract is modeled, basically with what is commonly called a point distribution model (PDM) [11]. Diffusion tensor statistics are computed across fiber tract sections and are associated with the mean geometric model, resulting in a compact description of diffusion properties along the fiber tract. An overview of our framework is illustrated in Fig. 2.

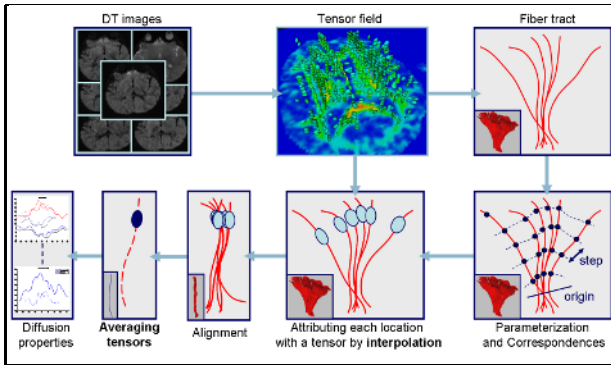


Fig. 2. Overview of the DTI analysis framework

3.1 Preprocessing: Tensor Field Computation and Fiber Extraction

The tensor field is computed from DTI data by solving the Stejskal-Tanner's diffusion equation system [12]. A tractography algorithm [13] extracts streamlines following the principal diffusion tensor directions between source and target regions of interest, with sub-voxel precision. Our latest version of the tractography tool includes tensor interpolation as described in Sect. 2.2. Except at branching or crossing points, the extracted 3D curves are assumed to represent the most likely pathways through the tensor field.

3.2 Geometric Modeling

An individual fiber tract, described by a set of streamlines, acts as a training set from which we estimate a template shape, the mean shape, and statistical deviations by learning its inherent shape variability. A brief review of the geometric modeling of one individual fiber tract is presented below (see [4] for a more comprehensive description).

First, fibers represented as polylines are reparametrized by cubic B-spline curves. This ensures an equidistant sampling along each fiber as well as a consistent sampling for all fibers. Second, an origin, which can be reliably identified across subjects, is defined for each fiber tract. This is either a geometric criterion, e.g., a cross-section with minimal area, or anatomical information like intersection with the midsagittal plane. Points with the same arc-length along the fiber tract are defined as homologous. This explicit point to point matching has been proven relevant in [2] where we demonstrated that it properly aligns local shape features across all curves in a fiber bundle. Given this correspondence, the alignment of all curves in the training set is achieved by Procrustes analysis [14]. Only estimated translations and rotations are applied to fibers. Indeed, a size normalization is not desirable since the training fibers belong to the same individual fiber tract. Given the set of aligned shapes, the mean shape is estimated by averaging the spatial coordinates at each corresponding location over the tract. Additionally, statistical shape deviations from this template shape along the tract can be characterized by extracting the principal modes of deformation via a principal component analysis.

3.3 Attributing the Geometric Model with Diffusion Tensor Statistics

The estimated mean shape models the geometry of the fiber tract. A complete representation of the tract, describing both geometry and diffusion properties, is obtained by attributing each location along the mean curve with statistics of diffusion tensors calculated over cross-sections.

Computing the Mean Tensors along the Fiber Tract. First, each sample point x from the set of reparameterized fibers is assigned a tensor p . Since the tensor field is defined on the discrete voxel grid while x lies on a continuous curve, a geodesic interpolation (see Sect. 2.2) is required to compute the tensor p at the location x . The tensor p is given by the weighted average of the eight voxel tensor values in the nearest $2 \times 2 \times 2$ neighborhood of x , the weights being defined by trilinear interpolation (see (3)). Let \mathcal{P} be the set of obtained tensors, $\mathcal{P} = \{p_{f,i}\}$ with f indexing the reparameterized fibers across the tract and i the location along each reparameterized fiber. Then, the tensor set \mathcal{P} is aligned by rotation. Let $R_f \in SO(3)$ be the rotation estimated by Procrustes analysis for the reparameterized fiber f . Each tensor $p_{f,i}$ lying on f is rotated to the tensor $p'_{f,i}$ by the group action $R_f: p'_{f,i} = R_f p_{f,i} R_f^T, \forall i$. Last, at each corresponding location i along the tract, the mean tensor μ_i is computed from the set of aligned tensors, $\{p'_{f,i}\}$, as defined in (1). In addition, tensor diffusion variability can be assessed at each location of the average curve by computing the geodesic standard deviation according to (2). For visualization purposes, each average tensor is translated to its corresponding average location on the average curve. Since diffusion tensors are invariant to translation, this does not affect any diffusion property.

Deriving Diffusion Properties. At each location along the template curve, diffusion properties are derived from the average tensor. We consider the measures: *i*) the three eigenvalues of the average diffusion tensor, $\lambda_1, \lambda_2, \lambda_3$, which represent the diffusivities along the three principal directions of the tensor, *ii*) the mean diffusivity, MD, defined by the first moment of the diffusion tensor eigenvalues, *iii*) the fractional anisotropy, FA, which is a normalized measure of the tensor shape and defines a distance to isotropy.

4 Experiments and Results

We have applied our new fiber tract-oriented DTI analysis technique to image data of a prospective study of neonatal brain structure in children at high risk for schizophrenia [15]. This study includes 3-Tesla MRI and DTI of neonates at 2 weeks of age with follow-up at 1 year. Local diffusion properties in white matter as measured by DTI have been implicated to be associated with axon density, degree of myelination and density of fluid. DTI of neonates in general present a decrease of FA and an increase of MD from central to peripheral regions, reflecting the typical pattern of early structuring of white matter. Over the first few months till age 1 and further, there is a rapid development of myelination of white matter. This is expected to be demonstrated by a thinning of local tensors and therefore by an increase of FA and a decrease of MD as a function of age. We will focus on the commissural bundles of the corpus callosum, specifically on the genu and splenium fiber tracts which connect prefrontal cortices and parts of the temporal, parietal and occipital lobes, respectively. It is known that the splenium is myelinating earlier than the genu, and that the presence of myelin sheaths has a significant effect on ability of water to diffuse [16].

Diffusion Tensor Image Data. Neonates and follow-ups were scanned on a Siemens head-only 3T scanner (Allegra). A single shot echo planar (EPI) diffusion tensor sequence with total scan time of approximately 4 minutes was used. The imaging parameters were: TR/TE/TH=4219ms/92.2ms, isotropic voxels with 2mm slice distance and inplane resolution = 2×2 mm, 5 averages, and number of slices ranging from 44 to 65. Seven images were acquired for each slice, one without diffusion gradient ($b=0$) while the remaining six with $b=1000$ s/mm² and diffusion gradients along the standard orientations [17]. We selected 8 cases from our large image database: 4 neonates at 2 weeks and 4 infants at age 1 year. For all cases, the genu and splenium tracts were extracted by tractography (see Fig. 3), the regions of interest being manually defined on the FA image using our SNAP tool.

Average of Diffusion Tensors in Cross-Sections Along Tracts. The geometric model and associated diffusion tensor statistics are computed for both tracts of each subject as described in Sect. 3. The Procrustes alignment for clustering of dispersed bundles has been skipped in these experiments, because tracking resulted in coherent, compact bundles. Figure 4 shows the mean tensors along the mean curve for each selected fiber tract for typical neonate and one year-old cases. As expected, the one year-old case presents much sharper and more elongated mean tensors than the 2 weeks-old case, which is explained by myelination of these white matter tracts.

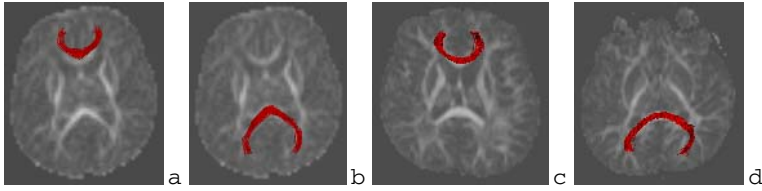


Fig. 3. DTI data with fiber tracts overlaid on axial sections of FA images. Genu and splenium of: a), b) one typical neonate case, c), d) one typical one year-old case.

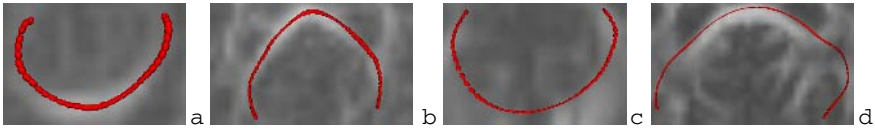


Fig. 4. Average tensors calculated in cross-sections displayed along central spine of each bundle. Genu and splenium of: a), b) one typical neonate case, c), d) one typical one year-old case.

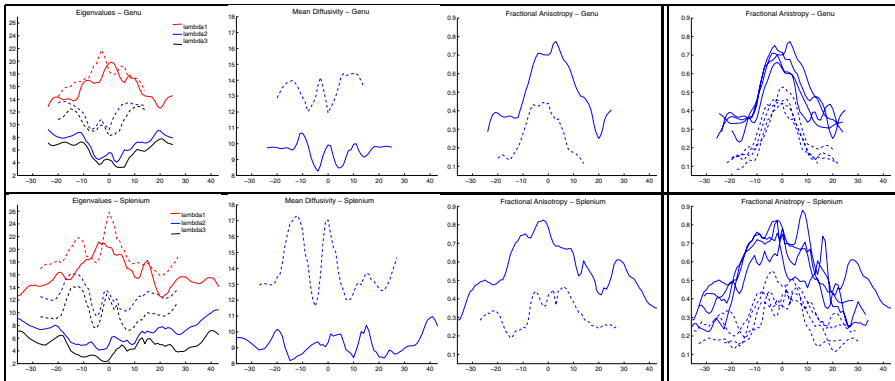


Fig. 5. Diffusion properties derived from average tensors plotted as a function of arc-length. From left to right: Eigenvalues, MD and FA for the genu (top row) and the splenium (bottom row) of one typical neonate (dashed lines) and one typical one year-old (solid lines). Last column displays FA for all 8 subjects: 4 neonates (dashed lines) and 4 one year-old (solid lines).

Diffusion Properties Along Fiber Tracts. The diffusion properties computed from the mean tensors, confirm the observation of increase of tensor elongation (see Fig. 5). For both tracts, the maximum diffusivity, λ_1 , appears quite similar for all cases, whereas the median and minimum diffusivities, λ_2 and λ_3 , are definitely higher in the two weeks-old cases (see dashed versus solid lines). Similarly, the mean diffusivity is much higher in the two weeks-old cases compared to the one year-old cases. This is consistent with the FA plots that, on the contrary, show higher values for the one year-old cases. These

observed diffusion properties complies with the fact that white matter becomes more and more dense and structured with age. The overall pattern of each diffusion property for all subjects is similar. This indicates an early local structure of unmyelinated fiber tracts at birth, possibly explained by high axonal density, that develops into a more organized structure due to myelination with increasing age.

Validation. A preliminary validation of our quantitative DTI analysis framework has been performed on the basis of 6 repeated DTI scans of the same subject with slight change of head position [18]. The analysis is applied to 3 callosal bundles and the uncinate fasciculus from each single scan and an average DTI computed after rigid body alignment. Diffusion properties along tracts are compared between scans. Standard deviation from the set of 6 scans lies with 5% for FA and MD and 2% for λ_1 . This test/re-test validation demonstrated a good reproducibility of the methodology.

5 Discussion

We have presented a new framework for fiber tract-oriented quantitative analysis of DTI data. It combines a geometric model of fiber tracts with diffusion tensor statistics. Unlike most other statistical analysis of DTI data, we do not compute statistics on scalar measurements derived from tensors but we compute statistics on diffusion tensors followed by calculation of tensor properties. We use non linear statistics for tensor interpolation and averaging. The extended set of features provided by the new methodology which includes integrated measurements across bundles and along fiber tracts seems to be suitable to study white matter fiber tract properties in cross-sectional and longitudinal studies. The new quantitative DTI analysis technique describes properties of tracts and is therefore superior to conventional region-of-interest measurements and might lead to an improved understanding of MRI/DTI findings and its association to normal/abnormal brain development at early age. Our analysis confirms earlier findings in regard to decrease of FA and increase of MD towards peripheral regions. The expected change over the first year of development is clearly demonstrated by FA and MD plotted as a function of arc-length. Analysis of the three associated eigenvalues reveals more insight, namely that the FA increase is due to decrease of the second and third eigenvalue whereas the diffusion along the major tensor direction remains similar. Further analysis on a larger population with extended age range might reveal more insight into the trajectory of growth as measured by DTI and measured as a function of anatomical location. We presently have 60 neonates and will increase this number to a total of 125. Ultimately, the proposed compact representation of the geometry of a tract and of associated diffusion properties aims at being used for inter-subject comparison and statistical analysis. This implies correspondence issues that are currently investigated by arc-length parameterization and could include local shape features of curves which have been shown to yield typical patterns along major fiber tracts [2]. Group comparison in clinical studies would require more advanced statistical techniques, for instance for comparison of probability distributions of tensors and hypothesis testing.

References

1. Ding, Z., Gore, J., Anderson, A.: Classification and quantification of neuronal fiber pathways using diffusion tensor MRI. *Magnetic Resonance in Medicine* **49** (2003) 716–721
2. Corouge, I., Gouttard, S., Gerig, G.: Towards a shape model of white matter fiber bundles using Diffusion Tensor MRI. In: Proc. IEEE ISBI. (2004) 344–347
3. Lim, K., Helpert, J.: Neuropsychiatric applications of DTI - a review. *NMR in Biomedicine* **15** (2002) 587–593
4. Corouge, I., Gouttard, S., Gerig, G.: A statistical shape model of individual fiber tracts extracted from Diffusion Tensor MRI. In: MICCAI. Vol. 3217 of LNCS. (2004) 671–679
5. Fletcher, P.T., Joshi, S.: Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In: ECCV 2004, Workshop CVAMIA, Vol. 3117 of LNCS. (2004) 87–98
6. Fletcher, P.T.: Statistical variability in nonlinear spaces: Application to shape analysis and DT-MRI. PhD thesis, University of North Carolina (2004)
7. Batchelor, P.G., Moakher, M., Atkinson, D., Calamante, F., Connelly, A. A rigorous framework for diffusion tensor calculus. *Magnetic Resonance in Medicine* **53** (2005) 221–225
8. Helgason, S.: *Differential Geometry, Lie Groups, and Symmetric Spaces*. Academic Press (1978)
9. Fréchet, M.: Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. H. Poincaré* (1948) 215–310
10. Pennec, X.: Probabilities and statistics on Riemannian manifolds: basic tools for geometric measurements. In: IEEE Workshop on Nonlinear Signal and Image Processing. (1999)
11. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models - their training and application. *Computer Vision and Image Understanding* **61** (1995) 38–59
12. Westin, C.F., Maier, S., Mamata, H., Nabavi, A., Jolesz, F., Kikinis, R.: Processing and visualization for diffusion tensor MRI. *Medical Image Analysis* **6** (2002) 93–108
13. Fillard, P., Gerig, G.: Analysis tool for diffusion tensor MRI. In: MICCAI. Vol. 2879 of LNCS. (2003) 967–968
14. Goodall, C.: Procrustes methods in the statistical analysis of shape. *J.R. Statist. Soc. B* **53** (1991) 285–239
15. Gilmore, J., Zhai, G., Wilber, K., Smith, J., Lin, W., Gerig, G.: 3T magnetic resonance imaging of the brain in newborns. *Psychiatry Research Neuroimaging* **132** (2004) 81–85
16. Rutherford, M.: *MRI of the Neonatal Brain*. WB Saunders Ltd. (2002) ISBN: 0702025348.
17. Basser, P., Pierpaoli, C.: Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *J. Magn. Reson. B* **3** (1996) 209–219
18. Gerig, G., Corouge, I., Vachet, C., Krishnan, K.R., MacFall, J.R.: Quantitative analysis of diffusion properties of white matter fiber tracts: A validation study. In: ISMRM, May 2005.

White Matter Tract Clustering and Correspondence in Populations

Lauren O'Donnell^{1,2} and Carl-Fredrik Westin^{1,3}

¹ MIT Computer Science and Artificial Intelligence Lab

² Harvard-MIT Division of Health Sciences and Technology, Cambridge MA, USA
lauren@csail.mit.edu

³ Laboratory of Mathematics in Imaging,
Brigham and Women's Hospital, Harvard Medical School, Boston MA, USA
westin@bwh.harvard.edu

Abstract. We present a novel method for finding white matter fiber correspondences and clusters across a population of brains. Our input is a collection of paths from tractography in every brain. Using spectral methods we embed each path as a vector in a high dimensional space. We create the embedding space so that it is common across all brains, consequently similar paths in all brains will map to points near each other in the space. By performing clustering in this space we are able to find matching fiber tract clusters in all brains. In addition, we automatically obtain correspondence of tractographic paths across brains: by selecting one or several paths of interest in one brain, the most similar paths in all brains are obtained as the nearest points in the high-dimensional space.

1 Introduction

Diffusion MRI measures water diffusion in tissue. Where many cells have similar orientations, such as in muscle or nervous tissue, the MRI measurements describe that local cellular pattern. Starting from these measurements of diffusion (represented as diffusion tensors) in the brains of several subjects, we produce an approximation of the neural fiber tract anatomy of the population.

Due to the complexity of the data, manual exploration of diffusion tensor images for the purposes of segmentation, registration, diagnosis, surgical planning, etc. is currently very difficult. One may choose to visualize tensors directly using small glyphs, or attempt to reconstruct anatomical structure by interpolating paths that follow the principal direction of diffusion. Detailed three-dimensional neuroanatomical knowledge is necessary to pinpoint the expected location of a white matter fiber, so it is not easy to choose a region of interest for display of the glyphs or paths. By creating an automatic segmentation of all possible paths, one could reduce this data interaction problem to a simple choice of which group of paths to display. Not only would this aid in visualization of the data, but also in the automatic quantification of properties of interest such as anisotropy measures, and in finding cross-patient anatomical correspondences.

Early work in grouping of tractographic paths by Brun et. al has used color to enhance visual perception of connectivity using information from spectral

embeddings [2, 1]. Related work in segmentation of tractographic paths has employed clustering approaches which divide the problem into two parts: the decision of how to quantify distance (or similarity) between paths, and the choice of clustering method. Several distance measures have been proposed in the literature. In one of the earliest approaches, Ding et al. calculate the mean distance separating paths using pointwise correspondences between path segments [4]. Their method is specific to paths which have been seeded in one image slice, but in later approaches more general distances have been defined. Brun et al. introduced a 9-D vector for tract shape approximation which they define as the mean and lower triangular part of the covariance matrix of the points on a path [1]. They compute the distance between paths as the Euclidean distance between the corresponding 9-D vectors. Gerig et al. and Corouge et al. propose distances that do pointwise comparison of tract shapes: they define three measures related to the Hausdorff distance [6, 3]. In more recent work by Jonasson et al. (who use paths through high angular resolution diffusion data) a path similarity measure is calculated based on the number of times two paths share the same voxel [7]. Two general types of clustering methods have been employed in the literature, hierarchical clustering [3, 6, 10], and the spectral clustering approach [1, 7].

To our knowledge, there is little prior work on the topic of matching tractographic paths across patients. In work by Zhang et al., a two-step process first performs clustering on each patient, then describes the clusters with 9-D vectors (the average start point, end point, and “middle point” of all paths in the cluster). These feature vectors are used to match clusters across patients [10].

2 Methods

Our method has several steps: tractography, estimation of distances/similarities between paths, spectral embedding, and finally clustering and/or anatomical region of interest (ROI) selection in the population. Figure 1 illustrates this process using data from the population of brains described further in Section 3. In the rest of this section we explain each step in our algorithm and its parameters.

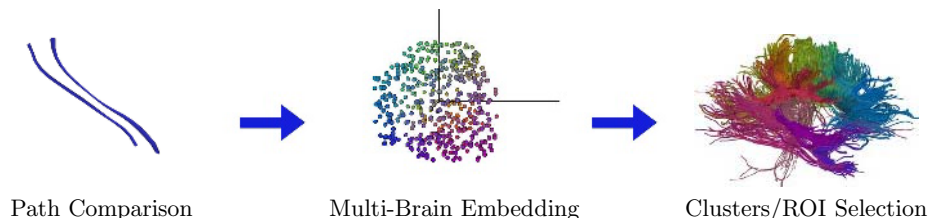


Fig. 1. Visual overview of the method using data from the population: two paths, a random sample of 500 embedding vectors, and population tract clusters in one brain

2.1 Tractography

We perform tractography using open-source software (www.slicer.org), which does second-order Runge-Kutta integration. We produce paths starting at all

points in a region of interest that covers most of the brain. Paths stop when the FA (fractional anisotropy) or curvature become too low or high, respectively. For the experiments in this paper, the region of interest was the area with FA greater than 0.3, and the settings used for the thresholds were FA of 0.1 and curvature of 1.1. In addition, to limit the number of paths, paths shorter than a length cutoff of approximately 30mm were discarded. To give an idea of the numbers of paths involved, there are about $256 \times 256 \times 50$, or 3 million, voxels in a dataset, of which we select approximately 300,000 for path seeding. After pruning the number of paths is between 5,000 and 25,000 for one brain.

2.2 Similarity Measure for Tract Paths

We use a symmetrized Hausdorff distance, as suggested in [6], and we compare paths in all brains. The Hausdorff distance H_{ij} is defined as the maximum of pointwise minimum distances between a pair of curves i and j . We symmetrize the distance by averaging H_{ij} with H_{ji} . In the common case where a short and long path are aligned, the average distance is a more forgiving measure than the maximum distance, but it preserves some information about the length mismatch, unlike the minimum distance. We convert the distance measure to a similarity measure W_{ij} using a Gaussian kernel where σ controls the distance over which paths can influence each other. For the experiments in this paper, σ was set to 10.

$$W_{ij} = e^{-H_{ij}/\sigma^2} \quad (1)$$

2.3 Spectral Clustering

Spectral clustering methods group data using eigenvectors of a data affinity matrix. The method we employ is described in [8] and is identical to the k-way normalized cuts procedure from [5], except we do not need to postprocess the clusters to correct for oversegmentation. First the symmetric tract similarity matrix W is scaled using a diagonal matrix D whose entries are the row (or column) sums of W . This produces L :

$$L = D^{-\frac{1}{2}} W D^{-\frac{1}{2}} \quad (2)$$

Embedding vectors are then calculated from the eigensystem of L ($L = U\Lambda U^T$) by applying a scaling to the rows of U . The scaling converts the eigenvectors to the solution one would get by solving the generalized eigensystem $(D - W)y = D\lambda y$, which minimizes the Normalized Cut [9]. From [5], the embedding vector for the j th tract is

$$E_j = \frac{1}{\sqrt{D_{jj}}}(U_{j,2}, U_{j,3}, \dots, U_{j,n}) \quad (3)$$

where n is the number of eigenvectors chosen for the embedding, and the column index of U starts at 2 in order to skip the first (major and smoothest) eigenvector. Care is needed in choosing n because the eigenvectors become noisy as the

eigenvalues decrease. For the experiments in this paper we have set n to 20. In spectral clustering, the next step is to cluster the embedding vectors, usually by k-means.

2.4 The Nystrom Method

In our application it is important to avoid computing the full W matrix, which for 5 brains would be at least $25,000 \times 25,000$. The Nystrom method is a way to interpolate values of a function, using known values and interpolation weights. We use this method to approximate the eigenvectors of W , as described in [5].

Instead of computing W directly, a random subset of paths is selected for comparison, and the similarity measure is calculated between each path and that subset. The matrix A contains similarities between paths in the subset, and the matrix B contains similarities between other paths and the selected subset. From [5], the approximate eigenvectors of W are calculated as

$$\bar{U} = \begin{bmatrix} U \\ B^T U \Lambda^{-1} \end{bmatrix} \quad (4)$$

where $U = A U \Lambda^{-1}$ and the ordering of the rows of \bar{U} is such that those corresponding to rows of A are first, followed by those corresponding to rows from B^T . To get an intuition for why this works, imagine that you have performed clustering using the information in A and you would like to add some new data to the problem without starting over. Since $A = A^T$ we can rewrite equation 4 as

$$\bar{U} = \begin{bmatrix} A^T U \Lambda^{-1} \\ B^T U \Lambda^{-1} \end{bmatrix} \quad (5)$$

So this process can be thought of as estimating the rows of \bar{U} by projecting the new data from B into the same space where the old data in A were projected.

There are two more details when using this approximation method. First of all, the normalization of A and B using row sums of W must be performed somehow before calculating \bar{U} . In [5] this is shown to be possible without computing W itself. The second detail is that the estimated eigenvectors of W (the columns of \bar{U}) should be orthonormal. In [5] the authors give two methods for orthogonalizing the columns. However we have found this to be unnecessary because in practice, after the columns of \bar{U} are normalized to length one, we have always found $\bar{U}^T \bar{U}$ to be close to the identity matrix.

3 Results

We employed our method to cluster tractographic paths in a population of 5 brains. For each, the input to tractography was a DTI scan with voxel size $0.86 \times 0.86 \times 5$ mm. The total number of tract paths from all brains was 36,003. We randomly selected 1000 tract paths from the population, and compared all other paths to these in order to generate A , B , and \bar{U} . We then performed spectral

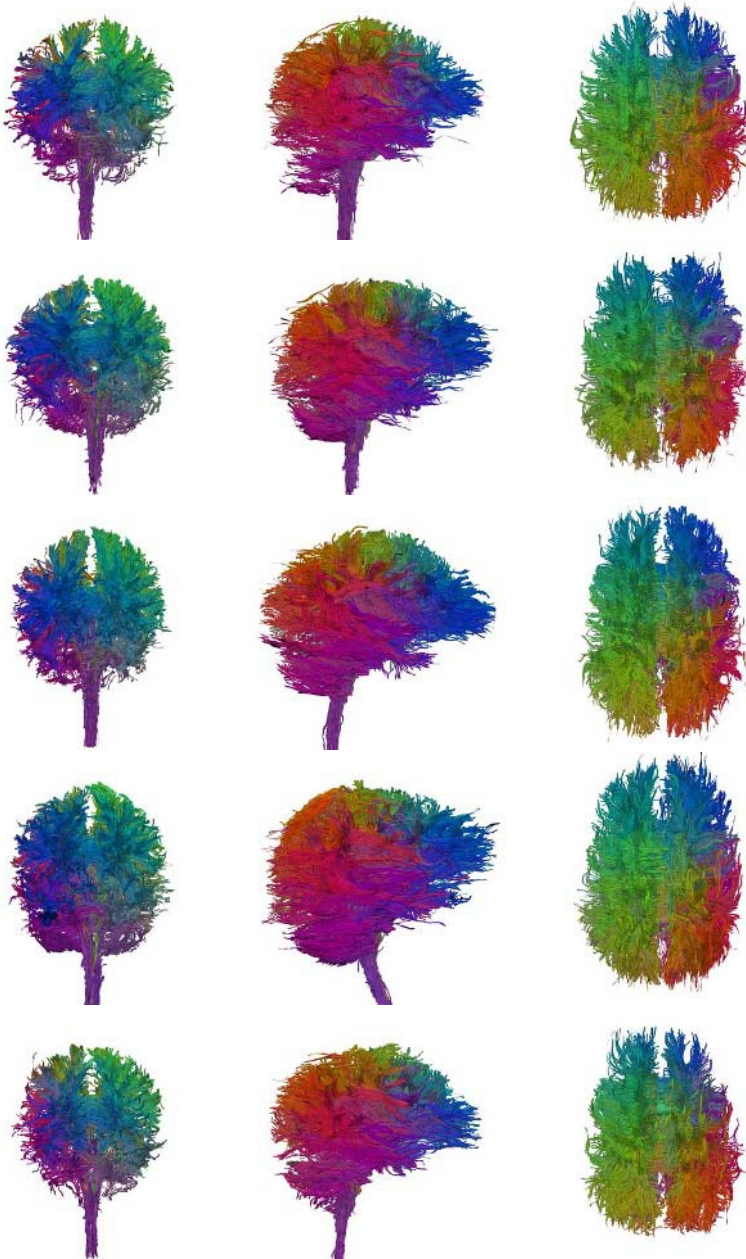


Fig. 2. Anterior, right, and superior views of the five brains in the population. The colors show correspondence via spectral embedding.

clustering to generate 100 clusters, a number that we find in practice is large enough to avoid grouping very dissimilar tracts. Interesting anatomical clusters exist at many size scales so choosing the number of clusters is not easy.

The output clustered/embedded brains are shown in Figure 2. The correspondence of the colors directly shows the correspondence in embedding space.

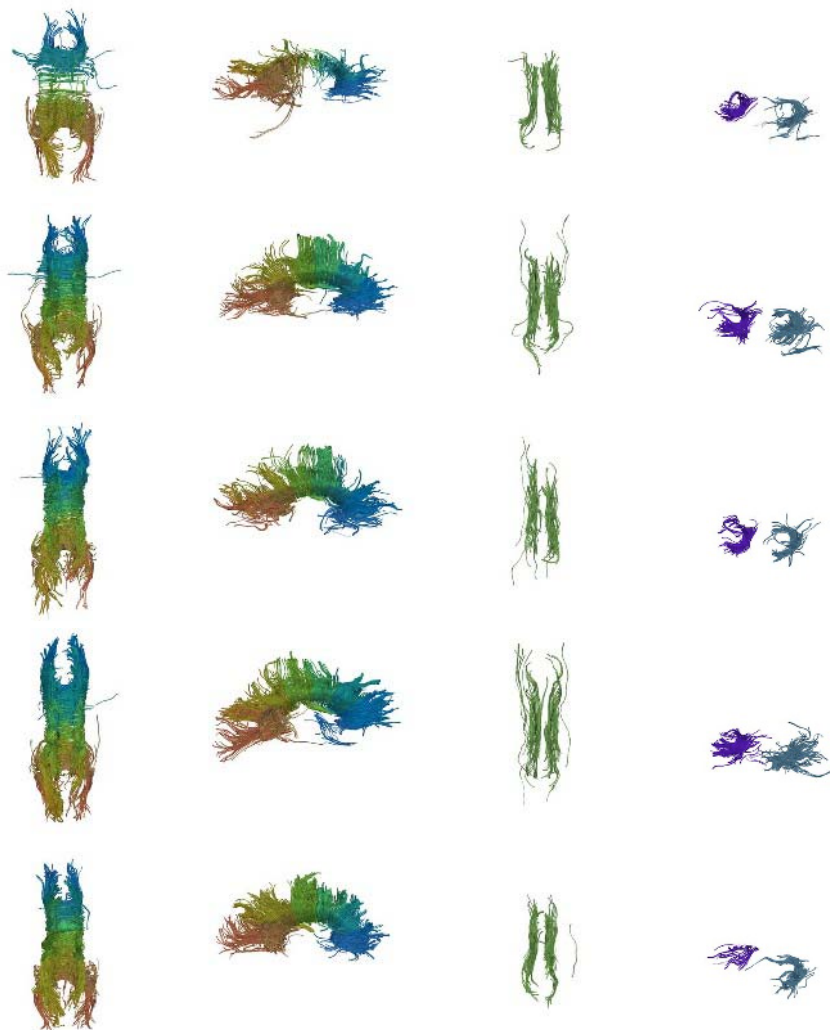


Fig. 3. Anatomical correspondences: selected clusters, displayed in all 5 brains. The two leftmost images show the corpus callosum viewed superiorly and from the right. Of the 100 clusters found, 10 were manually chosen as belonging to the corpus callosum. The third images from the left show a single cluster containing the cingulum bundles, viewed superiorly. Finally, the rightmost images show the two clusters that contain the left and right uncinate fasciculi, viewed anteriorly and from the right.

To generate this visualization, the paths in each cluster were colored based on the centroid of that cluster in the embedding space. Mapping the centroid vectors to colors was performed as described in [1], where the first three components of the embedding vector are used to define red, green, and blue.

In Figure 3 we show more detailed anatomical correspondences. Using the same 100 clusters, one of the brains in the population was visualized in order to identify clusters of interest. The cluster indices aren't in themselves significant, but clusters 3, 8, 12, 14, 19, 33, 54, 66, 77, and 81 were found to correspond to the corpus callosum, cluster 31 happened to correspond to both cingulum bundles, and clusters 29 and 42 represented the uncinate fasciculi. After selecting these clusters in one brain, the same clusters were then displayed in all brains to demonstrate correspondence.

4 Discussion

The first issue with the presented method is that it depends on paths from tractography, a process prone to errors due to noise and partial voluming, with limited validity in regions of low tensor anisotropy. The alternative approach of working directly in voxel space by defining a suitable voxel-to-voxel similarity measure is possible, however capturing long range connectivity may be more difficult. Also in principle our method could be applied to paths produced by another algorithm or paths through data with another diffusion model.

Another issue of interest is that because the scan protocol was chosen to acquire the images in a consistent way, in this work we explored performing clustering without rigid alignment. This can be seen especially in the fourth brain from the top in Figure 2, where the brain stem is angled very differently from the rest of the brains. This may be the reason for the less-structured uncinate fasciculus clusters for that brain in Figure 3. The robustness of the method is shown by the overall success of the correspondence, however we believe correspondence in the population will improve with rigid and potentially nonrigid alignment and perhaps normalization for brain size.

A question that might be posed about our choice of method is whether we could instead use a sparse W and sparse eigensolvers. Our W matrix is not particularly sparse, unlike the case of a W matrix that comes from similarities between pixels, because neighborhood relationships between paths are not limited by a rectangular grid. In addition, in [5] it is shown that the Nystrom approximation has comparable performance and much better running time when compared to a sparse solver.

It also makes sense to ask how this method will scale to a large population of brains. Our current implementation would likely scale up to a small clinical study, but for population sizes in the 100's a more memory-careful implementation would be necessary. One difficulty is holding all embedding vectors in memory to perform k-means clustering. But if a subset of the embedding vectors is chosen for clustering, and the subset is representative enough of the population, possibly the cluster centroids would be close to those obtained by doing the full clustering.

Another potential extension is the incorporation of additional information into the creation of the affinity matrix. For example if there already exist segmentations of the brains in question, the affinity should increase for two paths that pass through similar segmented anatomy. Another important point about the affinity matrix is that the choice of Hausdorff distance is somewhat arbitrary (though reasonable) and we hope to be able to compare various distance measures now that we have a working framework for clustering studies.

5 Conclusion

We have presented a novel method for obtaining anatomical clusters and correspondences across brains using paths from tractography through DTI data. To our knowledge, this is the first method for obtaining dense correspondences (path-to-path correspondences), as well as cluster-to-cluster correspondences across a population of brains. We believe there are many interesting future applications for this method, including studies of anatomy in populations and DTI visualization using automatic segmentation and labeling of tract clusters.

Acknowledgments

We would like to thank the HST Neuroimaging Training Grant, NIH P41-RR13218 (NAC), NIH 1-R01-NS051826-01, and NIH NCCR Morphometry Biomedical Informatics Research Network (BIRN) grant U24 RR021382. This work is part of the National Alliance for Medical Image Computing (NAMIC), funded by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant U54 EB005149.

References

1. A. Brun, H. Knutsson, H.-J. Park, M. E. Shenton, and C.-F. Westin. Clustering fiber traces using normalized cuts. In *MICCAI*, pages 368–375, 2004.
2. A. Brun, H.-J. Park, H. Knutsson, and C.-F. Westin. Coloring of DT-MRI fiber traces using Laplacian eigenmaps. In *EUROCAST*, pages 564–572, 2003.
3. I. Corouge, S. Gouttard, and G. Gerig. Towards a shape model of white matter fiber bundles using diffusion tensor MRI. In *ISBI*, pages 344–347, 2004.
4. Z. Ding, J. C. Gore, and A. W. Anderson. Classification and quantification of neuronal fiber pathways using diffusion tensor MRI. *MRM*, 49:716–721, 2003.
5. C. Fowlkes, S. Belongie, F. Chung, and J. Malik. Spectral grouping using the Nystrom method. *PAMI*, 26(2):214–225, 2004.
6. G. Gerig, S. Gouttard, and I. Corouge. Analysis of brain white matter via fiber tract modeling. In *EMBS*, page 426, 2004.
7. L. Jonasson, P. Hagmann, J.-P. Thiran, and V. J. Wedeen. Fiber tracts of high angular resolution diffusion MRI are easily segmented with spectral clustering. In *ISMRM*, 2005.
8. A. Ng, M. Jordan, and Y. Weiss. On spectral clustering: Analysis and an algorithm. In *NIPS*, volume 14, 2001.
9. J. Shi and J. Malik. Normalized cuts and image segmentation. *PAMI*, 22(8):888–905, 2000.
10. S. Zhang and D. H. Laidlaw. DTI fiber clustering and cross-subject cluster analysis. In *ISMRM*, 2005.

76-Space Analysis of Grey Matter Diffusivity: Methods and Applications

Tianming Liu^{1,2}, Geoffrey Young^{1,2}, Ling Huang^{1,2}, Nan-Kuei Chen²,
and Stephen TC Wong^{1,2}

¹ Center for Bioinformatics, Harvard Center for Neurodegeneration and Repair,
Harvard Medical School

² Department of Radiology, Brigham and Women's Hospital

Abstract. Diffusion Weighted Imaging (DWI) and Diffusion Tensor Imaging (DTI) are widely used in the study and diagnosis of neurological diseases involving the White Matter (WM). However, many neurological and neurodegenerative diseases (e.g., Alzheimer's disease and Creutzfeldt-Jakob disease) are generally considered to involve the Grey Matter (GM). Investigation of GM diffusivity of normal aging and pathological brains has both scientific significance and clinical applications. Most of previous research reports on quantification of GM diffusivity were based on the manually labeled Region of Interests (ROI) analysis of specific neuroanatomic regions. The well-known drawbacks of ROI analysis include inter-rater variations, irreproducible results, tediousness, and requirement of a priori definition of interested regions. In this paper, we present a new framework of automated 76-space analysis of GM diffusivity using DWI/DTI. The framework will be evaluated using clinical data, and applied for study of normal brain, Creutzfeldt-Jakob disease and Schizophrenia.

1 Introduction

DWI/DTI has been widely used to study the WM abnormality associated with disease progression [1], as DTI yields quantitative measures reflecting the integrity of WM fiber tracts, by taking advantage of the intrinsic directionality of water diffusion. However, many neurological diseases, including Alzheimer's disease, are considered to involve the GM [2, 5]. The diffusivity of GM is widely believed to be able to reflect the neuropathological changes in the neurological and neurodegenerative diseases [2, 5], and can be quantified by the apparent diffusion coefficient (ADC), which provides information on the degree of restriction of water molecules.

To date, applications of DWI/DTI in quantification of GM diffusivity have been done via manual ROI analysis. However, ROI analysis has several notable drawbacks, including inter-rater variations, irreproducible results, tediousness, and requirement of a priori definition of interested regions. In this paper, we present a computational framework of automated 76-space analysis of GM diffusivity. The framework has been evaluated against clinical data, and has been applied to study the GM diffusivity of normal brains, as well as the GM diffusivity abnormalities occurring in Creutzfeldt-Jakob disease and schizophrenia.

2 Methods

2.1 Overview

The computational framework of 76-space analysis of GM diffusivity is composed of seven steps, as summarized in Figure 1. The first two steps automatically segment the SPGR brain image into CSF/GM/WM tissues [3], and the GM map is further parcellated into 76 fine-detailed neuroanatomic regions (called 76-space) using the high-dimensional hybrid registration method [3]. The third step performs pre-processing in the DWI/DTI space, e.g., eddy current correction, ADC/FA (Fractional Anisotropy) image generation, reslicing, and co-registration of the B0/ADC/FA images with SPGR image. However, due to potential problems such as EPI geometric distortion, partial volume effect, image reslicing errors, and inaccuracy of co-registration algorithm, the warped B0/ADC/FA images would not be in exact anatomic correspondences with SPGR image. The inaccurate alignment of ADC/FA images with SPGR image prevents us from applying the GM parcellation in the SPGR space directly to the ADC/FA images, as the GM in SPGR space may correspond to heterogeneous tissues e.g., CSF and WM, in the DWI/DTI space. To overcome this problem, Step 4 and 5 segments the brain into CSF/GM/WM tissues in the DWI/DTI space by utilizing the tissue contrasts existing in the ADC image and FA image. Afterwards, Step 6 combines the tissue segmentation results from both SPGR space and DWI/DTI space, and takes the union of the GM maps in both spaces. This AND operation of GM maps results in a GM-Union map, which is classified as GM by both SPGR image segmentation and DWI/DTI image segmentation. Finally, we perform the automated 76-space analysis of GM diffusivity, represented by Step 7 in Figure 1.

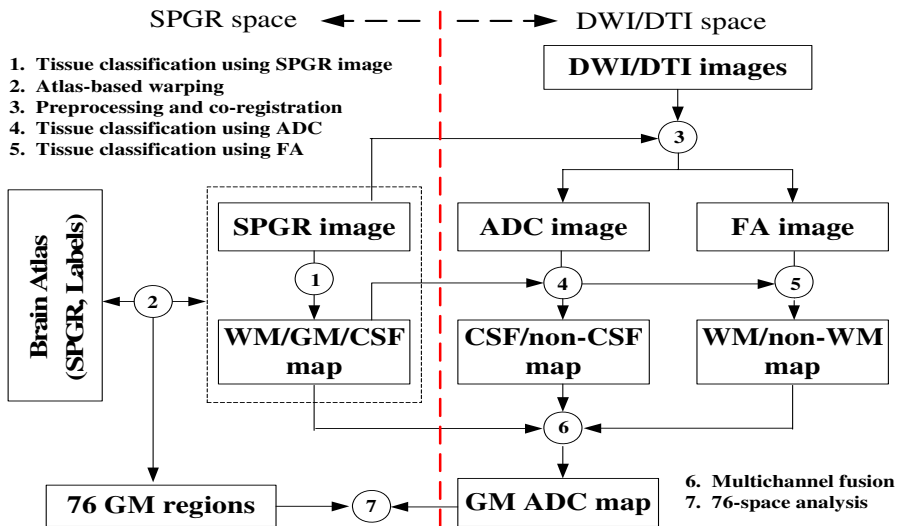


Fig. 1. Illustration of the computational framework of 76-space analysis

2.2 GM Parcellation in SPGR Space

We employ a high-dimensional atlas-based warping method [3] to automatically segment the subject brain SPGR image into 76 neuroanatomic structures. Meanwhile, we apply tissue segmentation on the subject SPGR image, and use the resulted GM map to mask the automatically labeled SPGR image, generating the labeled GM map. Finally, we remove other non-GM tissues by looking at the GM neuroanatomy table, and obtained the 76 GM structures for the following step of 76-space analysis.

2.3 Tissue Classification in DWI/DTI Space

2.3.1 Pre-processing

For the DWI/DTI data, firstly, we perform the eddy current correction by using the Oxford FSL FDT tools. We calculate the diffusion tensor, and generate the ADC and fractional anisotropy (FA) images using our in-house built tools. Then, we co-register the B0 (no diffusion gradient) image, ADC image and FA image with the SPGR structural image using the linear multimodality registration method of Oxford FSL Flirt. To obtain better co-registration, we further use a non-linear registration method, UCLA-AIR, to produce a more accurate co-alignment between ADC and FA images and SPGR structural image.

2.3.2 Motivation

Although the Oxford FSL Flirt and UCLA-AIR algorithms can produce reasonably good co-registration results, a couple of resources still render the possibility that GM regions in SPGR space may include heterogeneous tissues (e.g., CSF and WM) in the DWI/DTI space, and measurements of GM diffusivity may fail to reveal real changes occurring in the GM tissue, if we directly apply the GM segmentation in the SPGR space into the DWI/DTI space. Given that the GM is a lamina with thickness of about 3.0 mm, and the ADC values of the CSF are more than twice as high as the GM values, small errors in co-registration may lead to significant deviation of the measured ADC value in the GM region.

There are multiple reasons causing the problem of heterogeneous tissues. First, there could be geometrical distortions in EPI-generated diffusion tensor images, due to the rapid sampling of the gradient echo train. Second, there is the partial volume effect in MRI. Thirdly, we usually need to reslice the DWI/ADC/FA images, and this reslicing and interpolation process would inevitably generate certain inaccuracies. Fourthly, the accuracy of co-registration methods could be limited. Our solution is to remove the heterogeneous tissues in the GM ADC map by utilizing tissue contrasts existing in the ADC and FA images. The ADC values in the CSF are more than twice as high as the GM and WM values, because water diffusion in CSF is much less restricted than that in the GM and WM tissues. Therefore, it is straight-forward to use the ADC image to segment CSF from non-CSF tissues. Meanwhile, the FA image can be used to separate WM from non-WM tissues, since highly directional white matter structures have much larger fractional anisotropy values. Importantly, the ADC image and FA image are intrinsically in the same DWI/DTI space, and we can combine the results of CSF/non-CSF and WM/non-WM segmentation results into a complete CSF/GM/WM

segmentation map in the DWI/DTI space, which is independent of the SPGR space. Then, after taking the union of the GM map in SPGR space and the one in DWI/DTI space, we can substantially remove heterogeneous tissues on the GM ADC map.

2.3.3 Tissue Classification

2.3.3.1 HMRF-EM Algorithm. Other than segmenting brain tissues into three classes as in structural MR images, we classify brain tissues into two classes: CSF and non-CSF in ADC images, and WM and non-WM in FA images. We apply a Hidden Markov Random Field (HMRF) model and the Expectation-Maximization (EM) algorithm for the two-class tissue segmentation, which is akin to that in [4].

2.3.3.2 Initialization. As both EM model fitting and MRF ICM labeling converge locally [4], the selection of the initial parameter set is thus very important. In the literature, automated selection of initial centroids methods using k-means classification is widely used for the initial estimation [4]. However, applying k-means method to ADC images for automated selection of initial centroids might be problematic. Figure 2 shows the distribution of the ADC values in the whole brain, in which there are no clear boundaries between different tissues. The k-means method might fail to select the desired initial centroids, rendering difficulty for the MRF ICM labeling and EM model fitting to converge desirably. This may result in the HMRF-EM algorithm being trapped in a local minimum. Recall that we already have the tissue segmentation using SPGR image. We assume that the volume percentage of each tissue (CSF, GM, and WM) segmented in SPGR space should be the same, or close to, that obtained in DWI/DTI space. Hence, the volume percentage of each tissue class obtained in SPGR space can be used as *a priori* knowledge to guide the automated selection of initial centroids for HMRF-EM segmentation of the ADC image, as illustrated in Figure 2. In Figure 2, we know that the WM and GM occupy 82.4 % of the whole brain volume. This percentage is used to set the ADC threshold, below which are the GM and WM, and their volume sum is the same as 82.4 %, as shown in Figure 2. In a similar way, the volume percentage of CSF and GM obtained in SPGR segmentation can be used to set the FA threshold.

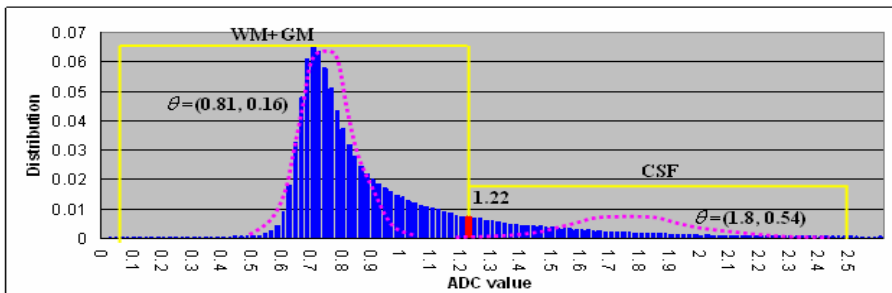


Fig. 2. Initialization for HMRF-EM segmentation of ADC image. ADC scale is 10^{-3} mm^2/sec . The threshold is set such that the total volume percentage below it is exactly the same as 0.824. The estimated parameters and fitted Gaussian models are overlaid.

2.4 Multi-channel Fusion

The SPGR channel has complete segmentation of CSF, GM and WM (Figure 3b), while the ADC channel has segmentation of CSF and non-CSF (Figure 3f), and the FA channel has segmentation of WM and non-WM (Figure 3j). The aim of multi-channel fusion is to take the union of the SPGR GM map (Figure 3c), the ADC non-CSF map (Figure 3g), and the FA non-WM map (Figure 3k), by performing an AND-like operation, and generating a GM-Union map, as shown in Figure 3o. The AND-like operation ensures that the GM-Union map is the consensus of all three channels. To demonstrate that multi-channel fusion greatly removes heterogeneous tissues, Figure 3d shows the overlap of SPGR GM map (Figure 3c) and ADC CSF map (Figure 3h). Clearly, there are large overlapping areas between CSF tissue in DWI/DTI space and GM tissue in SPGR space. Likewise, Figure 3p shows the overlap of the SPGR GM map (Figure 3c) and the FA WM map (Figure 3l). There are also large overlapping areas of WM tissue in the DWI/DTI space and GM tissue in the SPGR space. By using

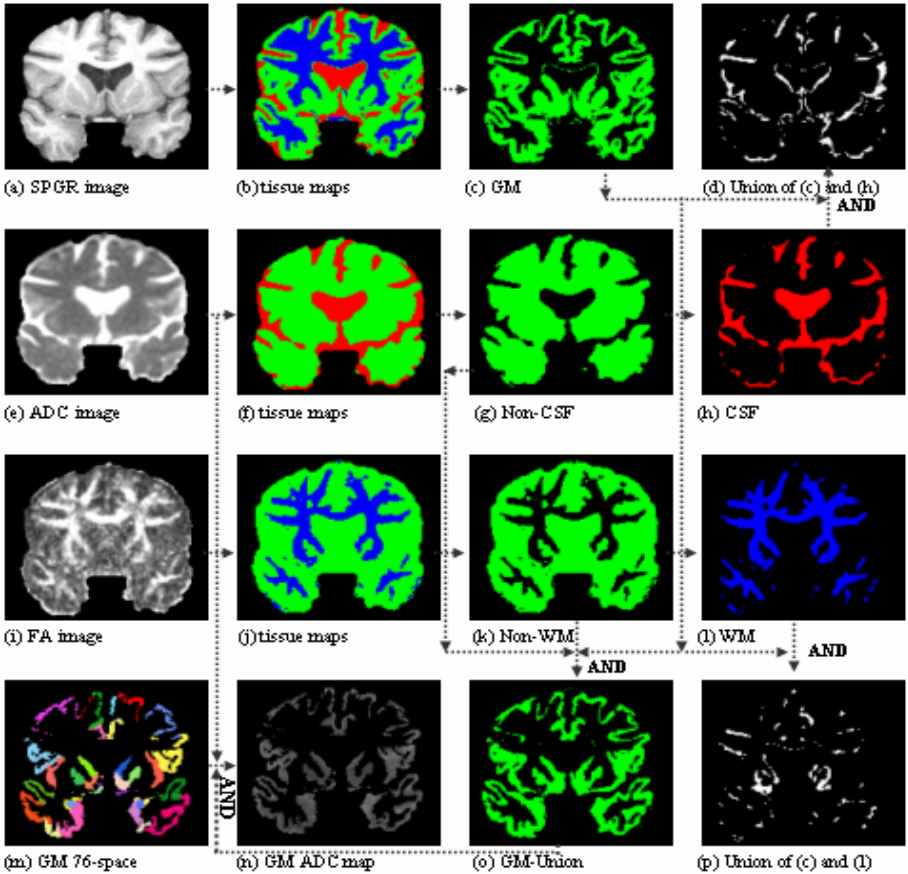


Fig. 3. Multi-channel data fusion. The “AND” means taking the union of two maps

the method in Subsection 2.2, we automatically segment the GM into 76 spaces, as shown in Figure 3m. Then, we combine the GM-Union map (Figure 3o), the labeled GM map (Figure 3m), and the original ADC map (Figure 3e) together, and we have the GM ADC map (Figure 3n) for 76-space analysis.

3 Evaluation and Validation

3.1 Tissue Classification

In the absence of digital DTI phantoms, we evaluate tissue classification on ADC and FA images by measuring volume agreement between the segmentation result in DWI/DTI space and that in SPGR space. The average volume agreements for WM, GM and CSF over fifteen cases are 0.88, 0.94 and 0.92, respectively. These high agreements indicate that the tissue segmentation in DWI/DTI space has good consensus with that in SPGR space. Notably, the volume agreement for GM is the highest, which is mostly desired for GM diffusivity analysis.

3.2 Heterogeneous Tissue Removal

A side-effect of the proposed multi-channel fusion is that certain GM regions will be shrunk after the heterogeneous tissue removal. The average volume-shrinking percentage caused by heterogeneous CSF and WM removals over fifteen cases are 15% and 17%, respectively. This result further confirms that the problem of heterogeneous tissues in co-registration of DWI/DTI and SPGR images is very common. Not surprisingly, the heterogeneous CSF tissues could significantly increase the average ADC values. Averaging over the 76 GM structures, the heterogeneous CSF caused an increase of 9% of the ADC value. For particular GM structures, e.g., the right thalamus in one case, the increased ADC value caused by heterogeneous CSF reaches 29%. After both the heterogeneous CSF and WM removal, the average ADC value slightly increases by 2%. This is expected, as the ADC value of WM is lower than that of GM, and removal of WM would decrease the average ADC in GM.

3.3 Manual Labeling

Without digital DTI phantoms, we evaluate our automatic measurements of GM diffusivity by the manual labeling. We randomly selected two normal subjects in our dataset, and two neuroanatomy experts manually painted ten GM ROIs for the two normal brains. The manual tracing was performed on the ADC image directly, and avoidance of including heterogeneous tissues in GM was especially taken care of. In the first case, the averaged differences between computerized results and two manual labeling results over the ten GM ROIs are 1% and 6% respectively. As for case 2, the average differences are 6% and 8%, respectively. Overall, the averaged difference between computerized result and manual labeling is about 5%, which is reasonably low considering the inter-rater difference of 5%.

4 Applications

4.1 GM Diffusivity Study of Normal Brains

We applied the 76-space analysis method to fifteen normal brains. The ages of the fifteen normal controls are between 29-51. All of them are male. Figure 4 shows the color-coded GM ADC distribution, where the ADC values of GM structures are mapped onto GM/WM surface of an atlas. The visualization shows that the parietal lobe has higher ADC values, while the temporal lobe has lower ADC values. Also we observed that deep GM structures have much lower ADC values. Figure 4 exhibits that there is no visible difference between the ADC on the right and left hemispheres.

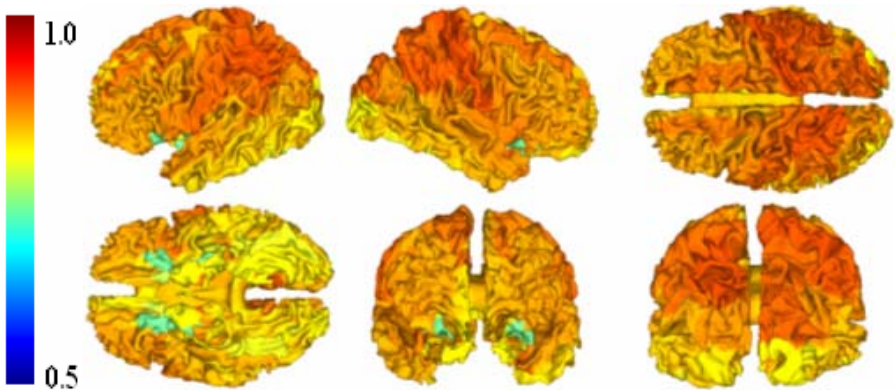


Fig. 4. Color-coded ADC distribution of GM. ADC scale is 10^{-3} mm²/sec

4.2 GM Diffusivity Study of Creutzfeldt-Jakob Disease

We automatically measured the ADC values of the 76 GM structures of 4 CJD patients. We found eight GM structures have significant differences (p -value < 0.05) between CJD and normal brains. Our results show that basal ganglia are frequently involved in these CJD cases, which is in agreement with other research reports [5]. Specifically, the average ADC values of putamen, thalamus, and globus palladus of CJD patients are much lower than those of normal brains. The average ADC values of CJD patients' right and left putamen dropped 29% and 26% respectively, compared to those of normal brains. The ADC droppings are confirmed by expert manual tracing.

4.3 GM Diffusivity Study of Schizophrenia

We applied the 76-space analysis to study the GM diffusivity of six schizophrenia patients. Their ages are in the range of 28-40. Their genders are all male. Our results show that there are no significant difference (p -value > 0.05) between ADC values of GM structures of schizophrenia patients and those of normal brains, except that the right and left temporal pole of schizophrenia patients have significant lower ADC values than those of normal brains (p -value: 0.0186 and 0.012). These ADC droppings

in temporal pole of schizophrenia patients are confirmed by expert manual tracing. In the literature, it is widely reported that temporal pole might be involved in schizophrenia [6]. However, our results need to be verified by future studies involving larger datasets, and its biological meanings are to be further investigated.

5 Conclusion

We proposed a new framework of 76-space analysis of GM diffusivity, in which structural information in SPGR image and diffusivity information in DWI/DTI images are integrated through two enabling technologies: high-dimensional hybrid registration and multi-channel fusion. The framework has been applied to study normal brains, CJD and schizophrenia, and produced meaningful results.

Acknowledgements

We would like to thank Ms. Yi-ru Lin for manual labeling of selected datasets, Dr. Susumu Mori for sharing DTI datasets, Dr. Kabani Noor for sharing brain atlas, and Dr. Martha Shenton (PIs of NIMH R01 MH 50740 and NIH K05 MH 01110) and Dr. Robert McCarley (PIs of NIMH R01 MH 52807 and NIMH R01 MH 40799) for sharing the schizophrenia dataset. This research is funded by a research grant to Dr. Stephen TC Wong by Harvard Center for Neurodegeneration and Repair, Harvard Medical School.

References

1. M. A. Horsfield, D. K. Jones. Applications of diffusion-weighted and diffusion tensor MRI to white matter diseases - a review. *NMR Biomed*, 5(7-8):570-7. 2002.
2. P. C. Sundgren, et al, Diffusion tensor imaging of the brain: review of clinical applications, *Neuroradiology*, 46 (5), 2004.
3. Tianming Liu, D. Shen, C. Davatzikos, Deformable registration of cortical structures via hybrid volumetric and surface warping. *NeuroImage*, 22 (4) 2004.
4. Y. Zhang et al, Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm, *IEEE TMI*, 20(1), 2001.
5. Geoffrey S. Young, et al, Diffusion Weighted Imaging (DWI) and FLAIR in Creutzfeldt-Jakob disease (CJD): High Sensitivity and Specificity for Diagnosis, in press, *American Journal of Neuroradiology*.
6. K Kasai, M. E. Shenton, et al. Differences and similarities in insular and temporal pole MRI gray matter volume abnormalities in first-episode schizophrenia and affective psychosis, *Arch Gen Psychiatry*, 60(11):1069-77, 2003.

Fast Orientation Mapping from HARDI

Evren Özarslan¹, Timothy M. Shepherd², Baba C. Vemuri¹,
Stephen J. Blackband², and Thomas H. Mareci³

¹ Department of Computer and Information Science and Engineering,
University of Florida, P.O. Box 116120, Gainesville, FL 32611, USA
{evren, vemuri}@cise.ufl.edu

² Department of Neuroscience, University of Florida,
P.O. Box 100244, Gainesville, FL 32610, USA

³ Department of Biochemistry and Molecular Biology, University of Florida,
P.O. Box 116120, Gainesville, FL 32611, USA

Abstract. This paper introduces a new, accurate and fast method for fiber orientation mapping using high angular resolution diffusion imaging (HARDI) data. The approach utilizes the Fourier relationship between the water displacement probabilities and diffusion attenuated magnetic resonance (MR) signal expressed in spherical coordinates. The Laplace series coefficients of the water displacement probabilities are evaluated at a fixed distance away from the origin. The computations take under one minute for most three-dimensional datasets. We present orientation maps computed from excised rat optic chiasm, brain and spinal cord images. The developed method will improve the reliability of tractography schemes and make it possible to correctly identify the neural connections between functionally connected regions of the nervous system.

1 Introduction

The diffusional attenuation of MR signal has been exploited to characterize diffusional anisotropy in fibrous tissues such as white-matter of the central nervous system (CNS) [1]. When the duration of the applied diffusion sensitizing gradients (δ) is much smaller than the time between the two pulses (Δ), the fundamental relationship between the MR signal attenuation $E(\mathbf{q})$ and average displacement probabilities $P(\mathbf{R})$ is given by a Fourier integral [2]:

$$P(\mathbf{R}) = \int E(\mathbf{q}) \exp(-2\pi i \mathbf{q} \cdot \mathbf{R}) d\mathbf{q}, \quad (1)$$

where \mathbf{R} is the displacement vector and \mathbf{q} is the reciprocal space vector defined by $\mathbf{q} = \gamma \delta \mathbf{G} / 2\pi$, where γ is the gyromagnetic ratio and \mathbf{G} is the gradient vector.

In fibrous tissues the orientations specified by large displacement probabilities are expected to coincide with the fiber orientations. One could in principle estimate the values of $P(\mathbf{R})$ by using Eq. 1 and fast Fourier transform (FFT), however, this would require data points all across the space spanned by the diffusion gradients (or \mathbf{q} vectors) necessitating very high gradient strengths and long acquisition times that are difficult to achieve in clinical settings.

More than a decade ago, Basser et al. introduced diffusion tensor imaging (DTI) [3] that employed a symmetric, positive-definite, second order tensor to characterize the signal attenuation. This model enabled simple measurement and estimation of diffusional anisotropy, and predicted a fiber orientation specified by the principal eigenvector of the diffusion tensor. Despite its modest requirements, DTI has been very successful in most areas of white-matter in the CNS, and enabled the mapping of anatomical connections [4].

DTI assumes a displacement probability characterized by an oriented Gaussian probability distribution function (PDF) whose covariance matrix is proportional to the diffusion tensor. Such a PDF has only one orientational mode and as such can not resolve more than one fiber orientation inside a voxel. This shortcoming of DTI has prompted interest in the development of more sophisticated models. Tuch et al. introduced the HARDI method [5] which suggested that apparent diffusivities could be evaluated along many orientations independently. The result is an angular distribution of diffusivities, $D(\mathbf{u})$, henceforth referred to as the diffusivity profile. It has been shown that the diffusivity profile has a complicated structure in voxels with orientational heterogeneity, and it is possible to represent the diffusivity profile using a Laplace series (LS) expansion [6, 7]. However, the peaks of the diffusivity profile do not necessarily yield the orientations of the distinct fiber populations. In [8] it was shown that the spherical harmonic transform (SHT) approach could be thought of as a generalization of DTI since the coefficients of the LS (obtained from the SHT of the diffusivity profile) are related to the components of higher order Cartesian tensors. Later Özarslan et al. proposed to use this generalization to generate signal values on the three-dimensional q -space, and evaluated a FFT to approximate the displacement probabilities [9]. Jansons and Alexander proposed a method to compute a displacement probability map [10] also from HARDI datasets by enforcing the condition that the probabilities are nonzero only on a spherical shell. Although the results are encouraging, both of the schemes are computationally expensive.

In this work, we introduce a new method, called the diffusion orientation transform (DOT), that describes how the diffusivity profiles can be transformed into probability profiles. Our approach is described in more detail in Ref. [11]. The constructed function is a LS expansion of the displacement PDF at a fixed radius. Our method is robust, fast and does not employ any model fitting. The requirements of our method makes it suitable for the clinical environment.

2 Theory

The Fourier transform as given in (Eq. 1) can be expressed in spherical coordinates. This is a consequence of the pointwise convergent expansion of the plane wave in spherical coordinates [12] given by

$$e^{\pm 2\pi i \mathbf{q} \cdot \mathbf{R}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l (\pm i)^l j_l(2\pi q R) Y_{lm}(\mathbf{u})^* Y_{lm}(\mathbf{r}), \quad (2)$$

Table 1. $A_l(\mathbf{u})$ and $B_l(\mathbf{u})$ functions up to $l = 6$. In this table, β stands for $\beta(\mathbf{u})$

l	$A_l(\mathbf{u})$	$B_l(\mathbf{u})$
0	1	0
2	$-(1 + 6\beta^{-2})$	3
4	$1 + 20\beta^{-2} + 210\beta^{-4}$	$\frac{15}{2}(1 - 14\beta^{-2})$
6	$-(1 + 42\beta^{-2} + \frac{1575}{2}\beta^{-4} + 10395\beta^{-6})$	$\frac{105}{8}(1 - 36\beta^{-2} + 396\beta^{-4})$

where $\mathbf{q} = q\mathbf{u}$ and $\mathbf{R} = r\mathbf{r}$, with $q = |\mathbf{q}|$ and $r = |\mathbf{R}|$. Here $j_l(x)$ denotes the l -th order spherical Bessel function whereas $Y_{lm}(\mathbf{r})$ are the spherical harmonics. Inserting this expression into Eq. 1, we get

$$P(R_0\mathbf{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l (-i)^l Y_{lm}(\mathbf{r}) \int d\mathbf{u} Y_{lm}(\mathbf{u})^* I_l(\mathbf{u}), \tag{3}$$

where

$$I_l(\mathbf{u}) = 4\pi \int_0^{\infty} dq q^2 j_l(2\pi q R_0) \exp(-4\pi^2 q^2 t D(\mathbf{u})). \tag{4}$$

Here t is the diffusion time, r was set to a particular radius R_0 and it is assumed that signal attenuation along each radial line in q -space is dictated by the Stejskal-Tanner relationship [13] where the rate of this attenuation is characterized by the diffusivity profile $D(\mathbf{u})$. Note that the function $P(R_0\mathbf{r})$ is the probability of finding the particle, initially at the origin, at the point $R_0\mathbf{r}$, i.e., we will be interested in the probability values on a sphere of radius R_0 . The integral in Eq. 4 can be evaluated and it is given by

$$I_l(\mathbf{u}) = \frac{R_0^l \Gamma(\frac{l+3}{2})}{2^{l+3} \pi^{3/2} (D(\mathbf{u})t)^{(l+3)/2} \Gamma(l+3/2)} {}_1F_1\left(\frac{l+3}{2}; l + \frac{3}{2}; -\frac{R_0^2}{4D(\mathbf{u})t}\right), \tag{5}$$

where ${}_1F_1$ is the confluent hypergeometric function. Using the recurrence relations of the confluent hypergeometric functions [14] iteratively, these functions can be written in the following form:

$$I_l(\mathbf{u}) = A_l(\mathbf{u}) \frac{\exp(-\beta(\mathbf{u})^2/4)}{(4\pi D(\mathbf{u})t)^{3/2}} + B_l(\mathbf{u}) \frac{\text{erf}(\beta(\mathbf{u})/2)}{4\pi R_0^3}, \tag{6}$$

where

$$\beta(\mathbf{u}) = \frac{R_0}{\sqrt{D(\mathbf{u})t}}. \tag{7}$$

$A_l(\mathbf{u})$ and $B_l(\mathbf{u})$ values for functions up to $l = 6$ are given in Table 1. Note that only the even order terms will be included which is a consequence of the antipodal symmetry of the diffusivity profiles as well as the displacement PDFs.

In Figure 1 we plot the square root of I_l as a function of R_0 calculated with double precision using Eq. 6 with $D = 1.5 \times 10^{-3} \text{mm}^2/\text{s}$ and $t = 25 \text{ms}$. Very

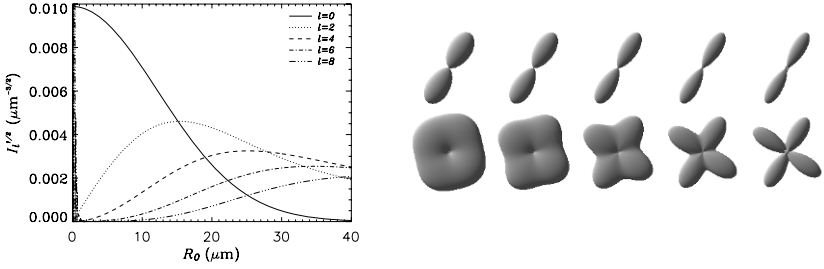


Fig. 1. a. Dependence of the square root of the radial integral, $\sqrt{I_l}$, on R_0 . The curve is drawn for l values ranging from 0 to 8. **b.** Probability maps estimated on a sphere of radius 8 to $16\mu\text{m}$ in equal steps of $2\mu\text{m}$ (from left to right) for voxels with a single orientation (top row) and two distinct orientations (bottom row).

large values taken by the higher order terms near the origin are due to round-off errors. However, this is not a big concern because we will mostly be interested in the values of this function in the $10 - 20\mu\text{m}$ range. Note that the contribution from higher order terms is rapidly collapsing for R_0 values in this range.

In order to estimate the probability on the surface of a sphere of radius R_0 , we go back to Eqs. 3 and 4, and expand $I_l(\mathbf{u})$ in Laplace series, i.e.,

$$I_l(\mathbf{u}) = \sum_{l'=0}^{\infty} \sum_{m'=-l'}^{l'} \alpha_{ll'm'} Y_{l'm'}(\mathbf{u}), \quad (8)$$

where

$$\alpha_{ll'm'} = \int Y_{l'm'}(\mathbf{u})^* I_l(\mathbf{u}) d\mathbf{u}. \quad (9)$$

Comparing the integration over \mathbf{u} in Eq. 3 with the expression in Eq. 9, it is easy to see that

$$P(R_0\mathbf{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l (-i)^l \alpha_{llm} Y_{lm}(\mathbf{r}), \quad (10)$$

which is just a LS expansion of $P(R_0\mathbf{r})$. Note that coefficients of this LS for some l value come from the l -th order LS coefficients of $I_l(\mathbf{u})$.

2.1 Implementation Aspects

In summary, the estimation of the probability of finding the particle at the point $R_0\mathbf{r}$ away from the origin involves the following steps:

1. Compute the diffusivity profile $D(\mathbf{u})$ using the Stejskal-Tanner relation.
2. Compute $I_l(\mathbf{u})$ using Eq. 5 or 6 with Table 1.
3. Compute α_{llm} , the l -th order SHT of $I_l(\mathbf{u})$.
4. Evaluate Eq. 10.

Implementation of the items 1, 2 and 4 above are trivial. Our data acquisition scheme involves sampling the sphere on the vertices of a tessellated icosahedron. This way, 46 or 81 points are sampled on the unit hemisphere from second or third order tessellations respectively. Following [15], we compute the SHT by discretizing the integrals on the sphere with integration weights calculated from the areas of the polygons specified by the dual tessellation. The computation of the α_{lm} coefficients take 25 to 60s for the entire dataset depending on the matrix size and number of sample points using a modest AMD Athlon XP 1800 processor.

3 Simulations

We have applied the scheme described above to the simulations of single fiber and crossing fiber systems. The simulations employed the exact form of the MR signal from particles diffusing inside cylindrical boundaries [16]. Fig. 1b shows that increasing R_0 gives rise to the sharpening of the displacement PDFs. This could be predicted from Fig. 1a which clearly indicates that for small R_0 the largest contribution comes from I_0 which upon the SHT forms the isotropic part of the constructed probabilities. Therefore, when R_0 is small compared to the characteristic length $\sqrt{6Dt}$ associated with the diffusion process and the radii of the cylinders, distribution of probability on the surface becomes more uniform.

We have also computed the probability surfaces for a simulated image of fiber crossing shown in Figure 2. The surfaces are consistent with the underlying known fibrous structure. The circular and linear fiber bundles were chosen so that a distribution of crossing angles is achieved across the region with orientational heterogeneity. We notice that distinct fiber orientations are better resolved when the different fiber bundles make larger angles with each other. Fig. 2a shows the probability profiles when there is no noise added to the signal values. Similar to the simulations in [10], we added Gaussian noise of increasing standard deviation to the real and complex parts of the signal. When the signal intensity in the image with no diffusion weighting is taken to be centered around 1, and Gaussian noise

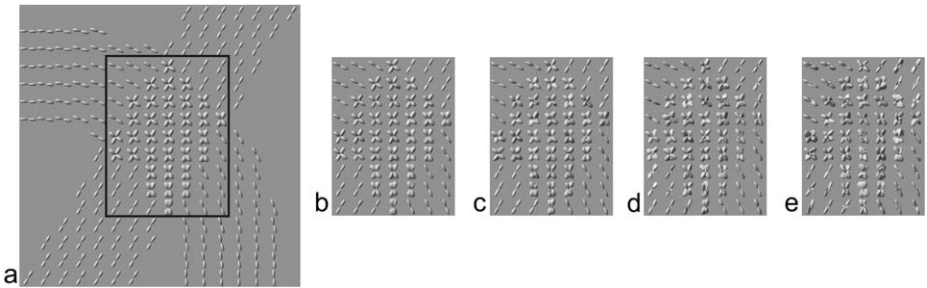


Fig. 2. **a.**The probability maps from a simulated image of two crossing fiber bundles computed using the DOT method. **b-e.**Surfaces in the framed area of panel a recomputed under increasing levels of noise.

of standard deviations 0.02 through 0.08 is added (in equal steps), the probability profiles shown in Fig. 2b-e are obtained. Note that in our standard HARDI protocol, we obtain SNR values in excess of 30 in diffusion weighted scans and about 100 in non-diffusion weighted images. Therefore, in real experiments, one can expect to achieve results that will be of similar or better quality with the image presented in Fig. 2b.

4 Results

In order to test the performance of the DOT method on real tissue, we computed the orientation probabilities on data from three anatomical regions (all from excised rat tissue): optic chiasm, brain and spinal cord. The images were acquired at $17.6T$ (brain) and $14.1T$ (spinal cord and optic chiasm) using Bruker Avance imaging systems. Diffusion weighted images were acquired along 81 (brain) and 46 (spinal cord and optic chiasm) directions with a b -value of $1500\text{mm}^2/\text{s}$ (brain and spinal cord) and $1250\text{mm}^2/\text{s}$ (optic chiasm) along with a single image acquired at $b \approx 0$. Resolution of the images were $33.6 \times 33.6 \times 200\mu\text{m}^3$, $150 \times 150 \times 300\mu\text{m}^3$, $60 \times 60 \times 300\mu\text{m}^3$ for optic chiasm, brain and spinal cord data respectively. The optic chiasm images were signal averaged to $67.2 \times 67.2 \times 200\mu\text{m}^3$ resolution prior to probability calculations.

For visualization purposes, we have overlaid the orientation surfaces on generalized anisotropy (GA) maps [17] computed from the displacement probabilities. The directionality of the probability profiles on the image plane is readily available from the surfaces. In order to visualize the peakedness through the image plane, we color coded the surface so that as the orientation of the surface varies

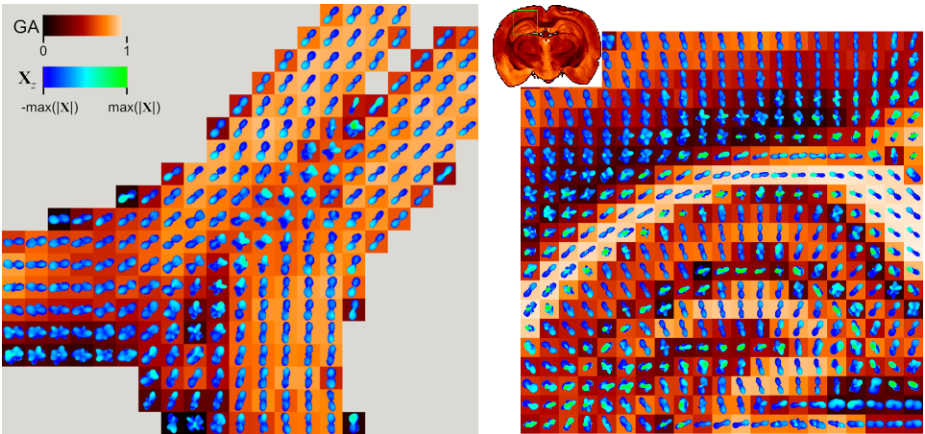


Fig. 3. **a.** Probability maps computed from a rat optic chiasm dataset. Every other pixel of the image is included for clarity. **b.** Probability maps computed from a diffusion-weighted dataset acquired from an excised rat brain.

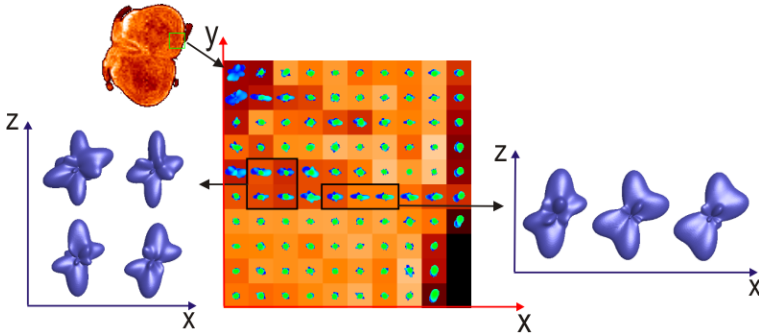


Fig. 4. Probability maps computed from a diffusion-weighted dataset acquired from an excised rat spinal cord

from $-z$ -axis to $+z$ -axis, its color changes from blue to green. In all calculations R_0 was set to $16\mu m$.

Optic chiasm is a distinct white-matter structure with well-known crossing fiber architecture. Therefore, it provides a perfect sample for a validation of the DOT method. Probability maps reconstructed as shown in Figure 3a appear to detect decussating fibers from the temporal visual fields in the center of the optic chiasm. Figure 3b shows the displacement probabilities computed from the excised rat brain data set. At the top left is a diffusion-weighted image that shows the selected ROI. The probability maps computed on this ROI demonstrate several interesting cytoarchitectural features of the rat brain. For example, the border between cortex and subcortical white matter contains fibers penetrating from the corpus callosum that cross the radial diffusivities of the cortex.

Finally, we show the probability maps computed from the excised spinal cord data set in Figure 4. Orientation maps from selected pixels of this image are enlarged at the left and rightmost sections of the figure. These surfaces were rotated by -90° about the x -axis so that the up-and-down direction in the individual surfaces shown in blue correspond to in-and-out direction in the images in red. The magnified surfaces clearly indicate the secondary fiber orientations that represent subtle differences in the connectivity of white matter in those regions.

5 Conclusion

The DOT method introduced in this work provides a direct estimation of displacement probability surfaces within each voxel from HARDI data. The technique is robust and fast. It provides a parametric description of the probability surfaces. As a result high resolution probability surfaces can be reconstructed easily from the calculated Laplace series coefficients. The potential applications of our approach include more accurate estimates of anisotropy and fiber orientations that will improve the existing fiber tractography schemes.

Acknowledgments. This research was supported by the National Institutes of Health Grants R01-NS42075, R01-NS36992 and P41-RR16105, and the National High Magnetic Field Laboratory (NHMFL), Tallahassee.

References

1. Chenevert, T.L., Brunberg, J.A., Pipe, J.G.: Anisotropic diffusion in human white matter: demonstration with MR techniques in vivo. *Radiology* **177** (1990) 328–329
2. Callaghan, P.T.: *Principles of Nuclear Magnetic Resonance Microscopy*. Clarendon Press, Oxford (1991)
3. Basser, P.J., Mattiello, J., LeBihan, D.: MR diffusion tensor spectroscopy and imaging. *Biophys J* **66** (1994) 259–267
4. Conturo, T.E., Lori, N.F., Cull, T.S., Akbudak, E., Snyder, A.Z., Shimony, J.S., McKinstry, R.C., Burton, H., Raichle, M.E.: Tracking neuronal fiber pathways in the living human brain. *Proc Natl Acad Sci* **96** (1999) 10422–10427
5. Tuch, D.S., Weisskoff, R.M., Belliveau, J.W., Wedeen, V.J.: High angular resolution diffusion imaging of the human brain. In: *Proc. of the 7th Annual Meeting of ISMRM, Philadelphia*. (1999) 321
6. Frank, L.R.: Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magn Reson Med* **47** (2002) 1083–1099
7. Alexander, D.C., Barker, G.J., Arridge, S.R.: Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. *Magn Reson Med* **48** (2002) 331–340
8. Özarslan, E., Mareci, T.H.: Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *Magn Reson Med* **50** (2003) 955–965
9. Özarslan, E., Vemuri, B.C., Mareci, T.H.: Multiple fiber orientations resolved by generalized diffusion tensor imaging. In: *Proc. of the 12th Scientific Meeting of ISMRM*. (2004) 89
10. Jansons, K.M., Alexander, D.C.: Persistent angular structure: new insights from diffusion magnetic resonance imaging data. *Inverse Problems* **19** (2003) 1031–1046
11. Özarslan, E., Shepherd, T.M., Vemuri, B.C., Blackband, S.J., Mareci, T.H.: Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). Technical Report TR05-004, Department of CISE, University of Florida, Gainesville, FL (2005)
12. Schwabl, F.: *Quantum Mechanics*. Springer-Verlag, Berlin (1989)
13. Stejskal, E.O., Tanner, J.E.: Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *J Chem Phys* **42** (1965) 288–292
14. Abramowitz, M., Stegun, I.A.: *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York (1977)
15. Ritchie, D.W., Kemp, G.J.L.: Fast computation, rotation, and comparison of low resolution spherical harmonic molecular surfaces. *J Comput Chem* **20** (1999) 383–395
16. Söderman, O., Jönsson, B.: Restricted diffusion in cylindrical geometry. *J Magn Reson A* (1995) 94–97
17. Özarslan, E., Vemuri, B.C., Mareci, T.H.: Generalized scalar measures for diffusion MRI using trace, variance and entropy. *Magn Reson Med* **53** (2005) 866–876

An Automated Approach to Connectivity-Based Partitioning of Brain Structures

P.A. Cook¹, H. Zhang², B.B. Avants², P. Yushkevich², D.C. Alexander¹,
J.C. Gee², O. Ciccarelli³, and A.J. Thompson³

¹ Centre for Medical Image Computing, Department of Computer Science,
University College London, UK

² Departments of Computer & Information Science and Radiology,
University of Pennsylvania, Philadelphia PA 19104, USA

³ Department of Headache, Brain Injury and Neuroinflammation,
Institute of Neurology, University College London, UK

Abstract. We present an automated approach to the problem of connectivity-based partitioning of brain structures using diffusion imaging. White-matter fibres connect different areas of the brain, allowing them to interact with each other. Diffusion-tensor MRI measures the orientation of white-matter fibres *in vivo*, allowing us to perform connectivity-based partitioning non-invasively. Our new approach leverages atlas-based segmentation to automate anatomical labeling of the cortex. White-matter connectivities are inferred using a probabilistic tractography algorithm that models crossing pathways explicitly. The method is demonstrated with the partitioning of the corpus callosum of eight healthy subjects.

1 Introduction

Diffusion-weighted MRI (DW-MRI) offers insight into the structure of white-matter fibres in the brain. From DW-MRI we can infer the local fibre orientation [1], and “tractography” algorithms use this information to reconstruct entire fibre pathways [2]. The possibility to parcellate brain structures based on their anatomical connectivity has important implications since it may help to identify functionally distinct subregions of the brain [3, 4]. Correspondence between anatomical connectivity and functional activation has been shown in the thalamus [5].

In this paper, we present an automated approach to brain region parcellation based on connectivity to cortical grey-matter inferred from fibre tractography. Compared to earlier efforts [3, 4, 6], our work’s original contributions can be outlined as follows. Firstly, we leverage advances in atlas-based image segmentation to produce cortex parcellation of subjects automatically. Secondly, we resolve crossing fibres and use a novel model of uncertainty when two fibre bundles are present within a voxel. We begin by reviewing the basic concepts of diffusion imaging, tractography and connectivity-based parcellation. We then present methods and results of our approach, followed by a discussion of future works.

1.1 Diffusion-Tensor MRI

In diffusion-tensor MRI (DT-MRI) [7], we fit a Gaussian model of the displacement density

$$p(\mathbf{r}, t) = G(\mathbf{r}; \mathbf{D}, t) = ((4t)^3 \det(\mathbf{D}))^{-\frac{1}{2}} \exp(-\mathbf{r}^T \mathbf{D}^{-1} \mathbf{r} (4t)^{-1}) \quad (1)$$

of diffusing water molecules after time t , where \mathbf{r} is the displacement and \mathbf{D} is a second-order symmetric tensor, with eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ and corresponding eigenvectors \mathbf{e}_1 , \mathbf{e}_2 and \mathbf{e}_3 . A useful statistic from \mathbf{D} is the fractional anisotropy [8], which quantifies the anisotropy in the diffusion with a value between 0 (isotropic diffusion) and 1 (displacement is nonzero only along \mathbf{e}_1).

In voxels containing a single bundle of axonal fibres with a common orientation, $\lambda_1 \gg \lambda_2 \simeq \lambda_3$, and the principal direction \mathbf{e}_1 provides an estimate of the direction of the fibre bundle. However, when the tissue structure in the voxel is more complex, such as at fibre crossings, the Gaussian model is a poor approximation of p , and \mathbf{e}_1 is not a reliable indicator of the fibre orientation. A variety of techniques exist to resolve the orientations of crossing fibres [9].

1.2 Tractography

Diffusion-tensor MRI provides estimates of the local fibre orientation in each voxel, which typically occupies a volume on the order of 10^{-9}m^3 . Tractography uses the local fibre-orientation measurements to reconstruct entire axonal paths.

Simple tractography calculates streamlines, which follow the local fibre-orientation estimate from voxel to voxel. Streamline trajectories are sensitive to noise, and authors often reduce false-positive connections by imposing limits on the streamline curvature, which can be as low as 45 degrees, and by specifying a minimum anisotropy for tracking to continue, since the fibre orientation is not well defined at low anisotropy [10].

Monte-Carlo based probabilistic methods, such as those proposed by Behrens et al [6] and Parker et al [11], define a probability density function (PDF) on the true fibre orientation in each voxel from the DW-MRI data. Streamlines are tracked repeatedly from a single seed point. Each time, fibre orientations are drawn from the PDF in each voxel. Over a large number of iterations, the process yields a connection probability from the seed point to any other voxel v , which is the fraction of streamlines that pass through v . The PDF is less concentrated when there is high uncertainty in the fibre orientation, for example in voxels with low anisotropy. This means that the set of streamlines passing through regions of high uncertainty disperse rapidly, which results in lower connection probability. Since probabilistic algorithms give a measurement of the confidence in each potential connection, they do not require restrictive thresholds.

1.3 Connectivity-Based Parcellation of the Brain

Tractography, combined with anatomic labeling of cortex, makes possible the non-invasive study of anatomical connectivity between distinct cortical structures. Connectivity-based parcellation is a method for segmenting a region of

interest in the brain based on anatomical connectivity to other brain regions. Given a segmentation of the brain into distinct labelled regions, the method uses tractography to determine which labelled region is most likely to be connected to each voxel in the region of interest.

The method can be performed using either deterministic or probabilistic algorithms. However, the anisotropy threshold in deterministic algorithms limits their ability to identify pathways near to grey-matter areas, such as cortex, and the tight curvature restriction penalizes genuine pathways demonstrating high curvature. Although such algorithms have been applied to parcellate the corpus callosum [3, 4], their known limitations require extra care in interpreting the results they produce.

Behrens et al [6] use a probabilistic approach to partition the thalamus, a deep grey-matter structure. Behrens uses a Bayesian method to define a single-fibre PDF. The cortex is segmented into several functional zones, and probabilistic streamlines are tracked from seed points in the thalamus. The cortical zone with the highest connection probability to the seed point is the zone that contains the most probabilistic streamlines. The thalamic segmentation is consistent with previous postmortem histological studies of connectivity.

2 Method

We will refer to the brain region of interest as the seed region and each voxel in the seed region as a seed point. We define the “connectivity map” for each seed point as the collection of probabilistic streamlines emanating from the seed.

The algorithm to partition the seed region is as follows:

1. Label the grey-matter regions of the subject in the space of the diffusion data.
2. Generate the connectivity map for each seed point from the diffusion data.
3. For each seed point, determine the labelled grey-matter region with the highest connection probability.

Steps 1 and 2 of the algorithm are described in detail below. In step 3, we use the same approach as Behrens et al, which we described in section 1.3.

2.1 Cortical Region Labeling

We require cortical regions to be defined in the same space as the connectivity maps. However the diffusion-weighted images lack the resolution for accurate delineation of cortex. Thus a high resolution T1-weighted image of the subject is acquired, on which the cortical regions of interest are defined. These regions are then transferred onto the diffusion data after co-registration of the T1-weighted to a non-diffusion-weighted image acquired as part of the diffusion MRI sequence, hereafter referred to as the $[b=0]$ image.

One feature of our method is the automation of the cortical region labeling step by warping an atlas into alignment with the T1-weighted image. Figure 1 depicts the brain atlas and its warped version after registration to the T1-weighted

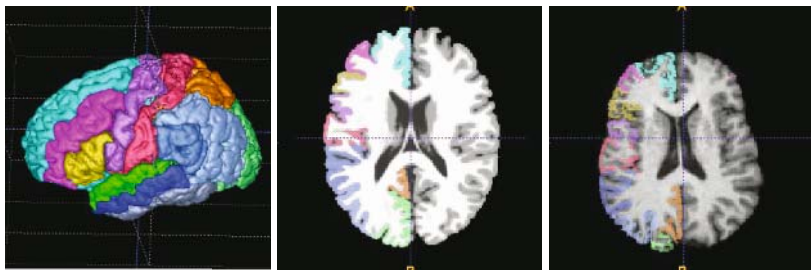


Fig. 1. Atlas-based brain image segmentation. (Left) A surface rendering of the labeled atlas used in this work. (Middle) The grey-matter labels for one hemisphere are shown superimposed on the underlying structural image of the atlas. (Right) The atlas is registered to the T1-weighted image of one subject, and the warped grey-matter labels for one hemisphere are shown superimposed on the subject’s structural image. The partition results of the subject’s corpus callosum is shown in figure 2, which also contains the color-coded list of the cortical regions delineated in the atlas, further details of which can be found in [12].

image of one subject. To improve the quality of the co-registration between the T1-weighted and $[b=0]$ images, we also acquire a high-resolution EPI image, which is used as an intermediate representation of the brain configuration in the diffusion data space. Specifically, the T1-weighted image is non-rigidly aligned to the EPI image, which in turn is non-rigidly registered to the $[b=0]$ image. The corresponding transformations are computed to enable atlas labels defined on the T1-weighted image to be transferred to the $[b=0]$ image. The non-rigid registration algorithm used optimizes an intensity cross-correlation metric under the constraints of a diffeomorphic transformation model in multiresolution fashion [13].

2.2 Generation of Probabilistic Streamlines

Before we generate probabilistic streamlines, we must complete three stages of pre-processing of the DW-MRI data: first classifying voxels as containing zero, one, or multiple fibre bundles, then determining fibre orientations in each voxel, and finally calibrating fibre-orientation PDFs. The first two steps must be completed for each subject, while the last is performed once.

Fibre-Orientation Estimation. In a similar way to Parker and Alexander [14], we use the spherical-harmonic classification and fibre-orientation estimation algorithm of Alexander et al [15] to determine which voxels contain zero, one, or multiple fibre populations. When the algorithm detects multiple fibres, we use a mixture of two zero-mean Gaussian distributions to model p , otherwise we use the single Gaussian as in DT-MRI (Eq. 1). We use Levenberg-Marquardt optimization to fit the parameters of the models to the data. We reduce the complexity of the two-fibre model by assuming two Gaussian compartments in

the voxel, mixed in equal proportion, and that the diffusion tensors are both cylindrically symmetric, so $\lambda_1 \geq \lambda_2 = \lambda_3$. The principal eigenvector of each tensor provides a separate fibre-orientation estimate.

Calibration of the Fibre-Orientation PDF. In single-tensor voxels, we use an existing PDF proposed by Cook et al [16], based on the Watson distribution [17]:

$$f(\mathbf{x}) = W(\pm\mathbf{x}; \boldsymbol{\mu}, \kappa) = M\left(\frac{1}{2}, \frac{3}{2}, \kappa\right)^{-1} \exp(\kappa(\boldsymbol{\mu}^T \mathbf{x})^2), \quad (2)$$

where M denotes the confluent hypergeometric function of the first kind and \mathbf{x} is a unit axis. The parameter κ determines the concentration of samples about the mean axis $\boldsymbol{\mu}$. We construct a lookup table that predicts κ from the tensor shape parameters $\frac{\lambda_1}{\lambda_3}$ and $\frac{\lambda_2}{\lambda_3}$. These indices are sufficient to fully specify a tensor, because the trace $\lambda_1 + \lambda_2 + \lambda_3$ is approximately constant in brain tissue. For each entry in the lookup table we synthesize measurements predicted by $p(\mathbf{r}) = G(\mathbf{r}, D, t)$, add complex Gaussian noise and fit the diffusion tensor to the noisy measurements. We repeat this process to obtain a large sample of noisy principal directions, \mathbf{x}_i , $1 \leq i \leq 10,000$, from which we estimate κ numerically [17].

In voxels containing two fibre populations, we use a Watson PDF for each fibre bundle. We calibrate the PDF by adding noise to the synthetic signal from $p(\mathbf{r}) = 0.5(G(\mathbf{r}, D_1, t) + G(\mathbf{r}, D_2, t))$, but we compute fewer samples than for the single-fibre PDF because generating two-fibre samples is more computationally expensive. The resulting samples \mathbf{x}_i , $1 \leq i \leq 4,000$ contain both fibre-orientation estimates from each trial. We estimate the concentration of both distributions simultaneously by maximising the log-likelihood l of the axes:

$$l(W_1, W_2) = \sum_{i=1}^{4000} \log\left(\frac{1}{2} [W_1(\mathbf{x}_i; \boldsymbol{\mu}_1, \kappa_1) + W_2(\mathbf{x}_i; \boldsymbol{\mu}_2, \kappa_2)]\right). \quad (3)$$

Previously, Parker and Alexander [14] constructed a two-fibre PDF by sorting the samples into two groups and fitting a Gaussian model to each. The objective function in Eq. 3 requires no assignment of axes to a particular fibre, and hence avoids any potential bias from incorrect assignment. We optimize $l(W_1, W_2)$ with repeated runs of the Levenberg-Marquardt algorithm. We construct a lookup table that gives κ_1 and κ_2 as a function of the fractional anisotropy of each tensor, and the angle at which the two principal directions cross.

Tractography. We use no anisotropy threshold in the tractography, and we apply a minimal curvature threshold to prevent streamlines from looping back on themselves and inflating the connection probability. We track by following the principal direction in each voxel, without interpolation. In voxels containing two fibres, we make a probabilistic choice of whether to sample from W_1 or W_2 based on the previous direction of the streamline, \mathbf{x}_p . The probability of sampling W_1 is $p_1 = W_1(\mathbf{x}_p; \boldsymbol{\mu}_1, \kappa_1) [W_1(\mathbf{x}_p; \boldsymbol{\mu}_1, \kappa_1) + W_2(\mathbf{x}_p; \boldsymbol{\mu}_2, \kappa_2)]^{-1}$, and the probability of choosing W_2 is $(1 - p_1)$. We track fibres from the seed points in both directions,

and we stop tracking if the streamline either curves by more than 80° over the length of one voxel, reaches the brain surface, or intersects itself.

3 Results and Discussions

We demonstrate the method by partitioning the corpus callosum (CC), the structure through which the two cerebral hemispheres communicate, using the brain images from 8 healthy volunteers. The diffusion imaging sequence uses 60 diffusion weighted measurements in each voxel. The first 54 measurements are at a fixed $|\mathbf{q}|$ giving $b = 1050\text{mm}^{-2}$, with the gradient directions spread evenly on the hemisphere to minimise dependence of the tensor upon the orientation of the tissue relative to the gradient directions. Six of the 54 diffusion-weighted measurements are repeated with $b = 260\text{mm}^{-2}$, and six measurements are made with $b = 0$. We estimate the signal to noise ratio in each $[b = 0]$ image to be approximately 14 in white-matter.

Because the fibre bundles originating from the CC extend to both the left and right cortex, we construct two connectivity maps for each seed point. Each map represents the connections to one side of the brain and is used to derive the partitioning for the corresponding hemisphere. An example of the partitioning is shown in figure 2.

A useful statistic is the relative sizes of the partitions which are computed for all subjects and shown in fig. 2. Evidently there are four primary partitions in the CC. Although this is consistent with the earlier study [3], in our study the cortical connectivity is better localized because we have chosen a significantly finer cortical labeling. The voxels within these partitions are most likely to connect to the superior frontal gyrus, the superior parietal gyrus, the occipital gyrus and the parahippocampal gyrus, respectively. On average, over all the controls, the sizes of the largest partitions found, computed as a percentage of

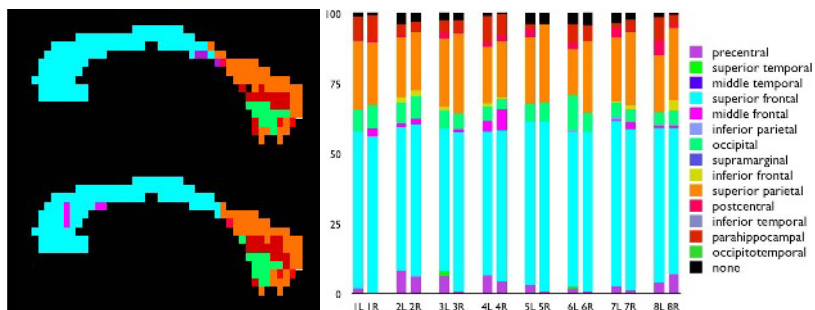


Fig. 2. The partitioning of the corpus callosum. On the left is the partitioning of the subject used in figure 1 for the left (top) and right (bottom) hemispheres. In the centre are the relative sizes of each partition as a percentage of the total area of the corpus callosum, for the left and right connectivity maps from all 8 subjects. On the right is the color-coded list of the cortical regions delineated in the atlas shown in fig. 1.

the size of the corpus callosum, are 53%, 22%, 7% and 5% for the left hemisphere, and are 55%, 26%, 5% and 5% for the right hemisphere. Despite the finer cortical labelling scheme, it appears that the majority of the connections through the corpus callosum are restricted to several large cortical regions. Additionally, several regions, notably the middle and inferior temporal gyrus, the occipitotemporal gyrus and the supramarginal gyrus, are not found in the partitions. It is possible that the connections to these regions are mediated through other regions, either by ipsilateral or U-fiber connections. Equally likely is that other connections exist across the corpus callosum but are more difficult to track reliably using current methods. Certain connections may be relatively difficult to track for several reasons, for example the fibre bundles may be smaller, or the bundles may extend through complex fibre architecture such as fibre crossings.

Also worth noting is the difference in the partitioning results from the left and the right hemispheres. postmortem studies [18] show that there are widespread heterotopic commissure connections across the corpus callosum. Partitioning studies such as this have the potential to offer new evidence of the existence of the heterotopic connections and, furthermore, possibly locate where such connections traverse the corpus callosum.

4 Conclusion

We have presented a method for parcellating brain structures based on connectivity maps derived from diffusion MRI data. The method is demonstrated with the partitioning of the corpus callosum. We construct connectivity maps using a novel probabilistic tractography algorithm that resolves crossing fibres. One feature of our method is the automated cortical region labeling based on our atlas-based and registration-driven segmentation technique. Future work includes quantitatively assessing the quality of the automated labeling compared to a manual delineation of the cortex.

Acknowledgement

PAC is funded by the MIAS IRC, EPSRC GR/N14248/01. OC and AT thank the MS Society of Great Britain and Northern Ireland for their generous support of the NMR Unit, Institute of Neurology. OC is a Wellcome Trust Advanced Fellow. JG thanks the NIH for their generous support through grants NS044189, DA015886 and NS045839. The authors thank Claudia Wheeler-Kingshott (Institute of Neurology, UCL), for the acquisition of the MR images.

References

1. Pierpaoli, C., Jezzard, P., Basser, P.J., Barnett, A., Chiro, G.D.: Diffusion tensor imaging of the human brain. *Radiology* **201** (1996) 637–648
2. Mori, S., van Zijl, P.C.M.: Fiber tracking: principles and strategies – a technical review. *NMR in Biomed.* **15** (2002) 468–480

3. Huang, H., Zhang, J., Jiang, H., Wakana, S., Poetscher, L., Miller, M.I., van Zijl, P.C.M., Hillis, A.E., Wytike, R., Mori, S.: DTI tractography based parcellation of white matter: Application to the mid-sagittal morphology of corpus callosum. *NeuroImage* (2004) doi:10.1016/j.neuroimage.2005.01.019
4. Gee, J.C., Zhang, H., Dubb, A., Avants, B.A., Yushkevich, P.A., Duda, J.T.: Anatomy-based visualizations of diffusion tensor images of brain white matter. In Welckert, J., Hagen, H., eds.: *Visualization and Image Processing of Tensor Fields*. Berlin:Springer (in press)
5. Johansen-Berg, H., Behrens, T.E., Sillery, E., Ciccarelli, O., Thompson, A.J., Smith, S.M., Matthews, P.M.: Functional-anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus. *Medical Image Analysis* **15** (2005) 31:39
6. Behrens, T.E., Johansen-Berg, H., Woolrich, M.W., Smith, S.M., Wheeler-Kingshott, C.A., Boulby, P.A., Barker, G.J., Sillery, E.L., Sheehan, K., Ciccarelli, O., Thompson, A.J., Brady, J.M., Matthews, P.M.: Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nat. Neurosci.* **6** (2003) 750–757
7. Basser, P.J., Mattiello, J., Bihan, D.L.: MR diffusion tensor spectroscopy and imaging. *Biophys J* **66** (1994) 259–267
8. Basser, P.J., Pierpaoli, C.: Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *JMR* **B111** (1996) 209–219
9. Alexander, D.C.: An introduction to computational diffusion MRI: the diffusion tensor and beyond. In Weichert, J., Hagen, H., eds.: *Visualization and Image Processing of Tensor Fields*. Springer (2005)
10. Stieltjes, B., Kaufmann, W.E., van Zijl, P.C.M., Fredericksen, K., Pearlson, G.D., Solaiyappan, M., Mori, S.: Diffusion tensor imaging and axonal tracking in the human brainstem. *NeuroImage* **14** (2001) 723–735
11. Parker, G.J.M., Wheeler-Kingshott, C.A.M., Haroon, H.A.: A framework for a streamline-based probabilistic index of connectivity (PICO) using a structural interpretation of MRI diffusion measurements. *JMRI* **18** (2003) 242–254
12. Dubb, A., Yushkevich, P.A., Xie, Z., Gur, R.C., Gur, R.E., Gee, J.C.: Regional structural characterization of the brain of schizophrenia patients. In: *Proc. MIC-CAI*. (2004)
13. Avants, B.A., Schoenemann, P.T., Gee, J.C.: Lagrangian frame diffeomorphic image registration: Morphometric comparison of human and chimpanzee cortex. *MIA* (in press)
14. Parker, G.J.M., Alexander, D.C.: Probabilistic monte-carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. In: *Proc. IPMI*. (2003) 684–695
15. Alexander, D.C., Barker, G.J., Arridge, S.R.: Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data. *MRM* **48** (2002) 331–340
16. Cook, P.A., Alexander, D.C., Parker, G.J.M.: Modelling noise-induced fibre-orientation error in diffusion-tensor MRI. In: *Proc. ISBI*. (2004) 332–336
17. Mardia, K.V., Jupp, P.E.: *Directional Statistics*. Wiley (2000)
18. Virgilio, G.D., Clarke, S.: Direct interhemispheric visual input to human speech areas. *HMP* **5** (1997) 347–354

Deformable Registration of Diffusion Tensor MR Images with Explicit Orientation Optimization

Hui Zhang¹, Paul A. Yushkevich², and James C. Gee²

¹ Department of Computer & Information Science, University of Pennsylvania, USA

² Department of Radiology, University of Pennsylvania, USA

Abstract. In this paper we present a novel deformable registration algorithm for diffusion tensor (DT) MR images that enables explicit analytic optimization of tensor reorientation. The optimization seeks a piecewise affine transformation that divides the image domain into uniform regions and transforms each of them affinely. The objective function captures both the image similarity and the smoothness of the transformation across region boundaries. The image similarity enables explicit orientation optimization by incorporating tensor reorientation, which is necessary for warping DT images. The objective function is formulated in a way that allows explicit implementation of analytic derivatives to drive fast and accurate optimization using the conjugate gradient method. The optimal transformation is hierarchically refined in a subdivision framework. A comparison with affine registration for inter-subject normalization of 8 subjects shows that our algorithm improves the alignment of manually segmented white matter structures (corpus callosum and cortico-spinal tracts).

1 Introduction

Diffusion tensor magnetic resonance imaging (DT-MRI)[1] is a water diffusion imaging technique that has been used to provide unique insight into the white matter (WM) organization in human brains [2, 3]. Diffusion describes the random movement of molecules. The rate of diffusion at a point, as a function of spatial direction, is referred to as a *diffusion profile*. This function can be sampled by acquiring a series of MR images sensitized to diffusion in a set of selected directions. In DT-MRI, diffusion profiles are assumed to be Gaussian and are defined by 2nd-order symmetric tensors. A DT image is produced by fitting this model to sampled diffusion profiles in each voxel of the MR series. The way in which DT describes local water diffusion can be best understood through its eigenvectors and eigenvalues. The eigenvectors of a DT coincide with the three principal directions of diffusion while the corresponding eigenvalues measure the rates of diffusion in those directions. In fibrous tissue, such as WM, although water is free to diffuse along the axis of the axon fiber bundles, diffusion is hindered in perpendicular directions by myelin that bound the axons. Consequently, the principal eigenvectors of the DTs measured in these regions tend to be parallel to the axis of the axon bundles. Compared to conventional MR images, this

unique ability of DT images to reveal the orientation of axon bundles makes the modality an ideal choice for understanding and analyzing WM structures.

The ability to detect WM differences between pairs and groups of subjects is essential in studying the pathology of various neurological disorders that are associated with WM abnormalities and may prove useful in early detection. To enable such group comparison analysis using DT images, a registration method is required [4, 5]. Compared to registering scalar images, the registration of DT images is particularly challenging not only due to the multi-dimensionality of the data, but also because one must ensure that the DT orientations remain consistent with the anatomy after image transformations [6]. Current DT registration techniques either circumvent tensor reorientation by registering scalar images derived from DT images, thus discarding the orientation component of the data, or perform tensor reorientation iteratively rather than analytically, thus precluding direct applications of gradient methods (see [7] for a survey).

In this paper we describe a novel algorithm for deformable registration of DT images that incorporates explicit optimization of tensor reorientation in an analytic manner. The optimization seeks a piecewise affine transformation that divides the image domain into uniform regions, each of which is transformed affinely. The objective function optimizes both the image similarity and the smoothness of the transformation across contiguous regions. The piecewise nature of the candidate transformations allows us to express the image similarity as the sum of the region similarities. Individual region similarity is computed using our previously published affine registration algorithm [8] which incorporates tensor reorientation in its objective function analytically, allowing explicit orientation optimization. The smoothness criterion imposed on the transformation is adapted from classic optical flow estimation [9], and it takes an analytic form as well. The resulting algorithm is based on analytic derivatives and uses the conjugate gradient method for optimization. By applying our algorithm to inter-subject registration, we demonstrate the algorithm is computationally efficient and improves the quality of image alignment.

In section 2, we will first briefly review the properties of DT images and the affine registration algorithm, then present our new piecewise affine formulation, and finally discuss the details of registration evaluation. The results of applying our algorithm to eight different subjects and their evaluations are shown and discussed in section 3.

2 Method

A diffusion tensor \mathbf{D} is a symmetric and positive-definite (SPD) 2nd-order tensor, which is related to a Gaussian diffusion profile $d_{\mathbf{D}}(\hat{k})$ by the quadratic form

$$d_{\mathbf{D}}(\hat{k}) = \hat{k}^T \mathbf{D} \hat{k}, \quad (1)$$

where \hat{k} is a unit vector defined over the unit sphere \mathbb{S}^2 . By measuring values of $d_{\mathbf{D}}(\hat{k})$ for different \hat{k} , \mathbf{D} can be determined through (1) by least square fitting.

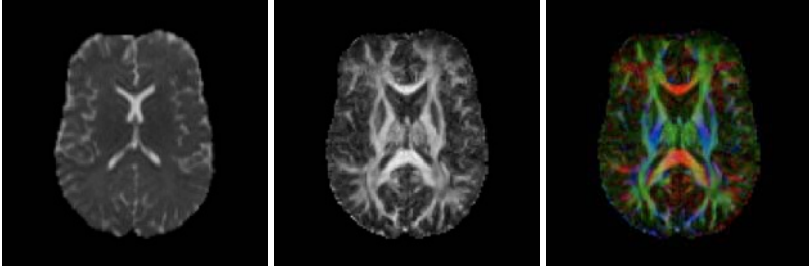


Fig. 1. The axial slice 24 of the DT image chosen as the template in this study. From left to right are the trace, the FA and the color-coded principal eigenvector maps. The latter is scaled by the FA map and the directions encoded by each color channel are mediolateral for red, anteroposterior for green and superoinferior for blue.

Because \mathbf{D} has six independent components, measurements from a minimum of 6 independent directions have to be acquired.

Two useful transformation-invariant scalar indices for \mathbf{D} are its trace and fractional anisotropy (FA). The trace is proportional to the mean diffusivity. Because the cerebrospinal fluid (CSF) has the most freely diffusing water, it appears the brightest in trace maps. The FA values vary from 0 for isotropic diffusion, such as in CSF and grey matter (GM), to 1 for anisotropic diffusion, such as in WM. Thus the WM is highlighted in FA maps. To illustrate WM fiber bundle orientation, a popular method is to use the principal eigenvector map [10] scaled by FA in which the orientation is encoded in terms of color. The stark difference between FA and principal eigenvector maps is the best illustration of the rich orientation information contained in DT images. An example of these maps is shown in fig. 1.

When measuring the similarity of two DTs, various metrics are available. Given no clear consensus on what the best metric may be, we have chosen to compare three different metrics that are defined using the following inner products:

$$\langle \mathbf{D}_1, \mathbf{D}_2 \rangle = \begin{cases} Tr(\mathbf{D}_1)Tr(\mathbf{D}_2) & \text{(for trace distance)} \\ Tr(\mathbf{D}_1\mathbf{D}_2) & \text{(for Euclidean distance) ,} \\ 2Tr(\mathbf{D}_1\mathbf{D}_2) + Tr(\mathbf{D}_1)Tr(\mathbf{D}_2) & \text{(for } L^2 \text{ distance)} \end{cases}$$

with the actual distance given by $\|\mathbf{D}_1 - \mathbf{D}_2\| = \sqrt{\langle \mathbf{D}_1 - \mathbf{D}_2, \mathbf{D}_1 - \mathbf{D}_2 \rangle}$. The trace distance (TD) between two tensors is the absolute value of the difference of their traces and is proportional to the difference of their mean rates of diffusion. The Euclidean distance (ED) is the Frobenius norm of the difference of two tensors. The ED reflects the relative orientation of the tensors and it is shown to perform better than the TD for DT registration [7]. The L^2 distance (LD) is defined in the functional space of diffusion profiles which is, generically, an infinite dimensional Hilbert space [8]. When applied to DTs, the LD is a weighted sum of the ED and the TD and our prior work suggests that it affords more robust registration over small regions [8] than each of those metrics does on its own.

2.1 Affine Registration Algorithm

The unique feature of our affine registration algorithm [8] is that the tensor reorientation is incorporated into the analytic objective function for explicit orientation optimization. This is accomplished by the combination of a special way of parameterizing the affine transform and a particular tensor reorientation strategy, which are described below.

An affine transformation F is parameterized as

$$F(x) = Mx + T = (QS)x + T, \quad (2)$$

where M , the Jacobian matrix of F , is parameterized based on its polar decomposition in terms of Q , an orthogonal matrix with determinant 1 representing pure rotation, and S , a SPD matrix representing pure deformation. The matrix Q can be represented using the 3 Euler angles, the matrix S has 6 independent components and the translation T has 3 components. We use these 12 variables to parametrize the affine transformation F , and denote them by vector \mathbf{p} .

For tensor reorientation, we use the finite strain (FS) reorientation strategy [6]. We choose FS over the more accurate method, the preservation of principal directions (PPD), because the PPD method is not analytic but algorithmic, computationally expensive, and the difference in accuracy between the two methods is minor [7]. When an orthogonal transformation Q is applied to a tensor \mathbf{D} , the corresponding reorientation is given by $Q\mathbf{D}Q^T$ [6]. For a non-orthogonal transformation M , the FS strategy finds the best orthogonal approximation, Q_M , to M and uses it for reorientation. Our method leverages the fact that the pure rotation component of the polar decomposition of M is precisely the best orthogonal approximation of M . Thus, in the framework of our method, FS reorientation can be formulated analytically, rather than using eigen-decomposition, as done in other methods.

The objective function of registration is then

$$O(\mathbf{p}) = \int_{\mathbb{R}^3} \|\mathbf{D}_s((QS)x + T) - Q\mathbf{D}_t(x)Q^T\|^2 dx, \quad (3)$$

where \mathbf{D}_t and \mathbf{D}_s are the template (fixed) and subject (moving) DT images respectively. The derivatives of $O(\mathbf{p})$ can be computed analytically; for example, the derivative with respect to t_i , the i -th component of the translation T is

$$\frac{\partial O}{\partial t_i} = \int_{\mathbb{R}^3} 2 \left\langle \frac{\partial \mathbf{D}_s}{\partial x_i}, \mathbf{D}_s((QS)x + T) - Q\mathbf{D}_t(x)Q^T \right\rangle dx.$$

2.2 Piecewise Affine Algorithm

The piecewise algorithm we propose involves using our affine algorithm for region-wise matching, enforcing the overall smoothness of the warp via smoothness constraints on interfaces of regions.

We subdivide the template \mathbf{D}_t into equal size regions denoted Ω_i . In general, each region, Ω_i , has 6 neighboring regions and thus 6 different interfaces. For

each region Ω_i in the template, the goal of the piecewise algorithm is to find an affine transformation F_i that gives the best match with the subject, under certain smoothness constraints that are described below.

We will refer to the collection of F_i over all possible regions as a piecewise affine transformation, denoted as \mathbb{F} . Because the transformation within each region is affine, the smoothness within a region is guaranteed. The smoothness of the piecewise affine transformation thus needs to be imposed only on region interfaces. Following the standard approach in optical flow estimation [9], we minimize the transformation discontinuities across interfaces, which is formulated for neighboring regions Ω_i and Ω_j as

$$\int_{\Omega_i \cap \Omega_j} \|F_i(x) - F_j(x)\| dx. \quad (4)$$

Similar to (3), analytic derivatives can be derived for (4).

If the number of regions in each dimension is n , the parameter space of this optimization problem has a dimension of $12n^3$. We subdivide the template hierarchically with n being 4, 8 and 16. At the finest subdivision level, the dimension of the parameter space is 49152. The ability to compute derivatives of (4) analytically allows us to take advantage of the conjugate gradient method, which is generally more efficient than optimization techniques that do not use derivatives, such as the Powell's direction set.

By construction, discontinuities across interfaces in the piecewise affine transformation can be minimized but not eliminated. Therefore, after the piecewise affine approximation to the underlying transformation is estimated at the finest level, it is interpolated using the standard approach [11] to generate a smooth warp field which is then used to deform the subject into the space of the template with the PPD reorientation strategy discussed in section 2.1.

2.3 Registration Evaluation

Here we outline two voxel-based measures that are specific to DT images to quantify the quality of image alignment. They are used to evaluate our inter-subject registration results.

To evaluate the overall quality of matching, we compute the average overlap of eigenvalue-eigenvector pairs (AOE) [12] of all the WM voxels ($FA \geq 0.3$) in the template and the corresponding voxels in the image to compare. The AOE measures, on average, the extent to which two tensors at each voxel are aligned; it is defined as

$$\frac{1}{N} \sum_{i=1}^N \frac{\sum_{j=1}^3 \lambda_j^i \lambda_j'^i (\epsilon_j^i \cdot \epsilon_j'^i)^2}{\sum_{j=1}^3 \lambda_j^i \lambda_j'^i}, \quad (5)$$

where λ_j^i , ϵ_j^i and $\lambda_j'^i$, $\epsilon_j'^i$ are the j -th eigenvalue-eigenvector pair at the i -th voxel location in the pair of images, and N is the total number of voxel locations to compare.

To assess the quality of local matching, we compute the average angular separation of the principal eigenvector (AAS) [6] in a region-specific manner as

in [13]. The AAS measures specifically how well the principal eigenvectors are aligned; it is defined as

$$\frac{\sum_{i=1}^N \sqrt{\mu_i \mu'_i} \arccos |\epsilon_1^i \cdot \epsilon'_1{}^i|}{\sum_{i=1}^N \sqrt{\mu_i \mu'_i}}, \quad (6)$$

where μ_i, ϵ_1^i and $\mu'_i, \epsilon'_1{}^i$ are the FA and the principal eigenvector at the i -th voxel location in both images. The two regions we have chosen are the corpus callosum (CC) and the cortico-spinal tracts (CST), which are manually segmented from the template.

3 Results

Here we report the results of inter-subject registration of eight subjects to the template shown in fig. 1. The DT images are of the size $128 \times 128 \times 48$, with the voxel spacings $1.72 \times 1.72 \times 3.0 \text{ mm}^3$. Each subject is registered to the template first using our affine registration algorithm and followed by the new piecewise affine algorithm. The transformation estimated from the affine registration is used to initialize the following piecewise affine registration.

For visual inspection, the color-coded principal eigenvector maps of one subject after affine registration and after piecewise affine registration are shown together with the template in fig. 2. Notice that the genu, splenium and internal capsules are better aligned after piecewise registration. Moreover, the alignment of the CC after piecewise registration is significantly better than after affine registration.

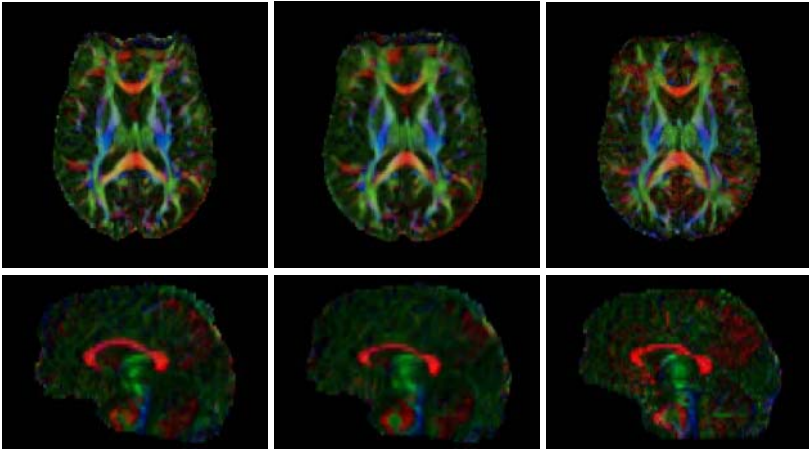


Fig. 2. The color-coded principal eigenvector maps of the axial slice 24 (top row) and the sagittal slice 64 (bottom row) of one of the subjects in this study together with the template. From left to right are the subject after affine registration and after piecewise affine registration, and the template.

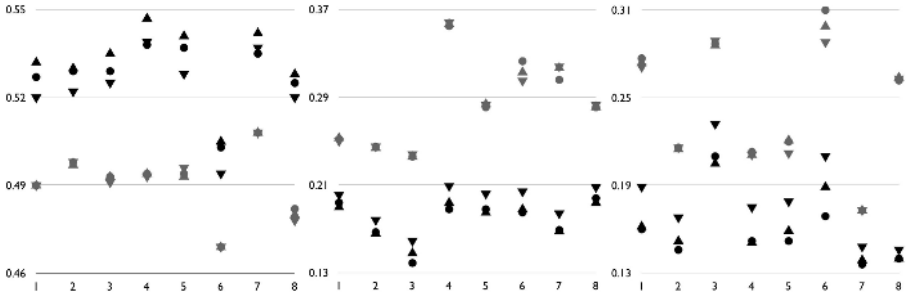


Fig. 3. Quantitative evaluation of registration results: In each graph, the x-axis represents the list of 8 subjects and the y-axis is some evaluated quantity. From left to right, the quantity evaluated are the AOE of all the WM voxels, the AAS of the voxels in the manually segmented CC, and the AAS of the voxels in the manually segmented CST. For each subject, there are six data points in each graph that correspond to six different registration methods, the affine (grey symbols) and piecewise affine (black symbols) registrations with three different DT metrics: TD (inverted triangles), ED (circles) and LD (triangles), and larger AOE or smaller AAS correspond to better alignment.

Quantitative evaluations are done as outlined in section 2.3. The results are summarized in fig. 3. It is evident that, consistent with the observation from visual inspection, the new piecewise algorithm outperforms its affine counterpart consistently for all the metrics tested. The two tensor metrics perform similarly well and they do slightly better than the scalar metric TD.

Finally, the algorithm is computationally efficient. Running on a 3.0GHz Pentium 4 Xeon processor, the computation time of registering each subject is less than 10 minutes.

4 Conclusion

In conclusion, we have presented a piecewise affine algorithm that demonstrates explicit orientation optimization required for optimal matching of DT imagery can be accommodated in deformable registration. Moreover, our novel formulation enables fast and accurate optimization using analytic derivatives. Results from inter-subject registration demonstrate the algorithm improves image alignment in a region-specific manner over affine registration. Future work includes more quantitative assessment of the algorithm using larger datasets and analyzing the effect of smoothing of the piecewise affine transformations.

Acknowledgement

This work was supported by the USPHS via NIH grants NS044189, DA015886 and NS045839.

References

1. Basser, P.J., Mattiello, J., Bihan, D.L.: Estimation of the effective self-diffusion tensor from the NMR spin echo. *JMR* **103** (1994) 247–254
2. Jones, D.K., Simmons, A., Williams, S.C.R., Horsfield, M.A.: Non-invasive assessment of axonal fibre connectivity in the human brain via diffusion tensor MRI. *MRM* **42** (1999) 37–41
3. Wakana, S., Jiang, H., Nagae-Poetscher, L.M., van Zijl, P.C., Mori, S.: Fiber tract-based atlas of human white matter anatomy. *Radiology* **230** (2004) 77–87
4. Jones, D.K., Griffin, L.D., Alexander, C.C., Catani, M., Horsfield, M.A., Howard, R., Williams, S.C.R.: Spatial normalization and averaging of diffusion tensor MRI data sets. *NeuroImage* **17** (2002) 592–617
5. Park, H.J., Kubicki, M., Shenton, M.E., Guimond, A., McCarley, R.W., Maier, S.E., Kikinis, R., Jolesz, F.A., Westin, C.F.: Spatial normalization of diffusion tensor MRI using multiple channels. *Med Image Anal* **6** (2002) 143–161
6. Alexander, D.C., Pierpaoli, C., Basser, P.J., Gee, J.C.: Spatial transformations of diffusion tensor magnetic resonance images. *TMI* **20** (2001) 1131–1139
7. Gee, J.C., Alexander, D.C.: Diffusion-tensor image registration. In Welckert, J., Hagen, H., eds: *Visualization and Image Processing of Tensor Fields*. Berlin:Springer (2005)
8. Zhang, H., Yushkevich, P., Gee, J.C.: Registration of diffusion tensor images. In: *Proc. CVPR.* (2004)
9. Hellier, P., Barillot, C., Memin, E., Perez, P.: Hierarchical estimation of a dense deformation field for 3-d robust registration. *TMI* **20** (2001) 388–402
10. Pajevic, S., Pierpaoli, C.: Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data: application to white matter fiber tract mapping in the human brain. *MRM* **42** (1999) 526–540
11. Little, J.A., Hill, D.L.G., Hawkes, D.J.: Deformations incorporating rigid structures. *Computer Vision and Image Understanding* **66** (1997) 223–232
12. Basser, P.J., Pajevic, S.: Statistical artifacts in diffusion tensor MRI (DT-MRI) caused by background noise. *MRM* **44** (2000) 41–50
13. Curran, K.M., Alexander, D.C.: Orientation coherence optimisation in tensor image registration. In: *Proc. MIUA.* (2004)

A Hamilton-Jacobi-Bellman Approach to High Angular Resolution Diffusion Tractography

Eric Pichon¹, Carl-Fredrik Westin², and Allen R. Tannenbaum¹

¹ Georgia Institute of Technology, Atlanta GA 30332, USA
`{eric, tannenba}@ece.gatech.edu`

² Harvard Medical School, Boston, MA 02115, USA
`westin@bwh.harvard.edu`

Abstract. This paper describes a new framework for white matter tractography in high angular resolution diffusion data. A direction-dependent local cost is defined based on the diffusion data for every direction on the unit sphere. Minimum cost curves are determined by solving the Hamilton-Jacobi-Bellman using an efficient algorithm. Classical costs based on the diffusion tensor field can be seen as a special case. While the minimum cost (or equivalently the travel time of a particle moving along the curve) and the anisotropic front propagation frameworks are related, front speed is related to particle speed through a Legendre transformation which can severely impact anisotropy information for front propagation techniques. Implementation details and results on high angular diffusion data show that this method can successfully take advantage of the increased angular resolution in high b-value diffusion weighted data despite lower signal to noise ratio.

1 Introduction

The development of Diffusion Tensor MRI has raised hopes in the neuroscience community for in vivo methods to track fiber paths in white matter. Diffusion Tensor Magnetic Resonance Imaging (DT-MRI) measures the self-diffusion of water in biological tissue [1]. The utility of this method stems from the fact that tissue structure locally affects the Brownian motion of water molecules and will be reflected in the DT-MRI diffusion measurements. In classical theory, diffusion follows a Gaussian process which can be described locally by a second order tensor.

A simple and effective method for tracking nerve fibers using DT-MRI is to follow the direction of maximum diffusion at each voxel [2, 3, 4, 5]. Although this method is widely spread and used in various ways the fiber trajectory is based solely on local information which makes it very sensitive to noise. Moreover the major direction of diffusion can become ill-defined for example at fiber crossings.

It has been proposed to shift from the Lagrangian, particle¹ based streamline approach described above to a Eulerian front propagation approach that can use

¹ In this work, “particle” refers to the position of a fictitious evolving point

full tensor information and is more robust to noise [6, 7]. This can be set in a Riemannian framework [8, 9].

The Gaussian assumption of diffusion tensor imaging does not hold for example if several fibers with different directions co-exist within the same voxel. Extensions to the tensor model have been proposed [10]. High angular resolution diffusion modalities such as Q-Ball imaging [11] acquire diffusion information in potentially hundreds of directions thus measuring direction information in a non-parametric way. Hagmann *et al.* [12] extend the streamline technique to multi-valued vector fields and Campbell [13] proposes a front evolution approach based on high angular resolution data.

In this work tractography is set in a continuous minimum cost framework. This is different from [12, 13] who do not propose variational (cost minimizing) techniques. Local costs are defined for every direction on the unit sphere based on high angular resolution diffusion imagery. Equivalently this can be considered a minimum arrival time framework in which the speed of fictitious particles would be the inverse of the cost. We show that while the minimum arrival time and anisotropic front propagation frameworks are deeply related, front speed is related to particle speed through a Legendre transformation [14] which can severely impact anisotropy information for front propagation techniques (Section 2). Implementation details for our technique are discussed (Section 3) along with considerations on the definition of the cost and the use of prior information (Section 4). Finally, the technique is illustrated on high angular resolution diffusion imagery (Section 5). Due to space constraints, full details will be published in [15].

2 Theory

2.1 Direction-Dependent Local Costs

Minimum cost approaches have been used extensively in image segmentation. In active contour models, an initial curve $\mathbf{I}(t = 0)$ is continually deformed using calculus of variations in a way that optimally minimizes a global cost $C(\mathbf{I}(t))$. This global cost is obtained by integrating along the curve a local cost, $\Psi : \mathbb{R}^n \rightarrow \mathbb{R}^+$ based on image information. The Live-Wire algorithm [16] determines all the optimal curves between a seed region and any pixel of the image using dynamic programming.

We propose to extend these two techniques to directional data by considering direction-dependent local costs of the form² $\Psi : \mathbb{R}^n \times \mathbb{S}^{n-1} \rightarrow \mathbb{R}^+$. The global cost of a given curve \mathbf{I} , is then defined to be

$$C(\mathbf{I}) \triangleq \int_{\mathbf{I}} \Psi(\mathbf{I}, \hat{\mathbf{T}}_s) ds, \quad (1)$$

which depends locally not only on the position of the curve but also on the direction of its unit tangent $\hat{\mathbf{T}}_s$.

² A position \mathbf{p} in n -dimensional space \mathbb{R}^n is written in bold, and a hat denotes a direction $\hat{\mathbf{d}}$ of the unit sphere \mathbb{S}^{n-1} .

2.2 Calculus of Variations

The first variation of the functional (1) can be computed which leads to the optimal minimizing deformation

$$\frac{\partial \Gamma}{\partial t} = -P_{\hat{\Gamma}_s^\perp} (\nabla_{\Gamma} \Psi - \frac{\partial}{\partial s} \nabla_{\hat{\Gamma}_s} \Psi) + \Psi \Gamma_{ss}, \tag{2}$$

where $P_{\hat{\Gamma}_s^\perp}$ is the projection onto the plane normal to $\hat{\Gamma}_s$ and any user-defined direction-dependent cost $\Psi(\Gamma, \hat{\Gamma}_s)$ can be used. This result extends minimizing flows for functional of the form $C(\Gamma) = \int_{\Gamma} \mathbf{v} \cdot \hat{\Gamma}_s \, ds$, proposed independently for tractography [13, 17].

2.3 Dynamic Programming

This technique is based on defining a value function $C^*(\mathbf{p})$ that represents, for any point \mathbf{p} in the domain, the minimum cost over all curves between \mathbf{p} and the region S . From any \mathbf{p} , the optimal curve back to S is obtained by descent on C^* . It can be shown that the value function C^* satisfies the Hamilton-Jacobi-Bellman equation

$$\max_{\hat{\mathbf{d}} \in \mathbb{S}^{n-1}} \{ \nabla C^*(\mathbf{p}) \cdot \hat{\mathbf{d}} - \Psi(\mathbf{p}, \hat{\mathbf{d}}) \} = 0. \tag{3}$$

The cost Ψ can be interpreted as the inverse of the speed of a particle traveling along the curve. $C^*(\mathbf{p})$ is then the minimum arrival time at \mathbf{p} . Alternatively, an infinite number of particle departing from S at the same time would propagate as a front geometrically evolving in the direction of its normal with a speed F . The corresponding equation is $\|\nabla C^*(\mathbf{p})\| F(\mathbf{p}, \nabla C^*(\mathbf{p}) / \|\nabla C^*(\mathbf{p})\|) = 1$. The two speeds F and $1/\Psi$ are not identical because particles are not restricted to moving along the normal of the evolving front. Front speed and cost (or its inverse, particle speed) are related through the Legendre transformations [14]

$$F(\mathbf{p}, \hat{\mathbf{n}}) = 1 / \min_{\hat{\mathbf{d}} \cdot \hat{\mathbf{n}} > 0} \left\{ \frac{\Psi(\mathbf{p}, \hat{\mathbf{n}})}{\hat{\mathbf{d}} \cdot \hat{\mathbf{n}}} \right\} \text{ and } \Psi(\mathbf{p}, \hat{\mathbf{d}}) = \max_{\hat{\mathbf{d}} \cdot \hat{\mathbf{n}} > 0} \{ (\hat{\mathbf{d}} \cdot \hat{\mathbf{n}}) / F(\mathbf{p}, \hat{\mathbf{n}}) \}. \tag{4}$$

Anisotropic front propagation techniques are not set in a Hamilton-Jacobi-Bellman framework. Consequently it is the front speed F which is defined from the diffusion data. Curves will then be determined that are optimal for C , which can only be interpreted as a cost (or an arrival time) in terms of Ψ . The *max* operator present in the definition of Ψ from F will tend to filter out the highest values of the front speed F that correspond to the preferred spatial directions (and therefore potential fibers) while preserving the slowest directions. Simulations and further analysis show that this distortion affects particularly speed functions with very localized direction information such as those encountered in high angular resolution diffusion imagery and can even result in loss of information. This problem can be avoided by setting directly the problem in a minimum cost framework (as proposed) or taking into account the Legendre transformation (4) when defining F .

Given the boundary condition $C^* = 0$ on some seed region S the Hamilton-Jacobi-Bellman equation (3) can be solved for the value function C^* as well as the characteristic direction $\hat{\mathbf{d}}^*$. Then, from any point \mathbf{p} , an optimal curve $\Gamma^*(\mathbf{p})$ can be determined back to S by following the characteristics $\hat{\mathbf{d}}^*$. By construction, $C(\Gamma^*(\mathbf{p}))$ equals the optimal cost $C^*(\mathbf{p})$. For these C -optimal curves, the value of another global cost, K , corresponding to a different local cost Φ can be computed by solving the transport equation

$$\nabla K(\mathbf{p}) \cdot \hat{\mathbf{d}}^*(\mathbf{p}) = \Phi(\mathbf{p}, \hat{\mathbf{d}}^*(\mathbf{p})) \quad (5)$$

with boundary condition $K = 0$ on the seed region S . In particular the length $L(\Gamma^*(\mathbf{p})) = \int_{\Gamma^*(\mathbf{p})} 1 \, ds$ of these optimal curves corresponds to $\Phi = 1$. The cost per unit length K/L can be used to define a validity index and rank curves that are optimal for one criterion using another criterion as in [6, 7].

3 Implementation

The HBJ equation (3) can be solved using a fast sweeping numerical scheme proposed by Kao *et al.* [18], similar to the one used in [7]. The algorithm

Algorithm 1. Sweeping algorithm to solve the Hamilton-Jacobi-Bellman equation (3), see [18]

Require: seed region S , direction-dependent local cost Ψ

- 1: Initialize $C^*(\cdot) \leftarrow +\infty$, except at starting points $s \in S$ where $C^*(s) \leftarrow 0$
 - 2: **repeat**
 - 3: **sweep** through all voxels \mathbf{p} , in all possible grid directions
 - 4: $\hat{\mathbf{d}}' \leftarrow \arg \min_{\mathbf{d} \in \mathbb{S}^{n-1}} f_{C^*, \Psi}(\mathbf{p}, \hat{\mathbf{d}})$
 - 5: **if** $f_{C^*, \Psi}(\mathbf{p}, \hat{\mathbf{d}}') < C^*(\mathbf{p})$ **then** $C^*(\mathbf{p}) \leftarrow f_{C^*, \Psi}(\mathbf{p}, \hat{\mathbf{d}}')$ and $\hat{\mathbf{d}}^*(\mathbf{p}) \leftarrow \hat{\mathbf{d}}'$
 - 6: **end if**
 - 6: **end sweep**
 - 7: **until** convergence of C^*
 - 8: **return** value function C^* , characteristics $\hat{\mathbf{d}}^*$
-

sweeps through all points \mathbf{p} in search of the least expensive direction. The cumulated cost to reach \mathbf{p} from direction $\hat{\mathbf{d}}$ is $f_{C^*, \Psi}(\mathbf{p}, \hat{\mathbf{d}}) \triangleq (\sum_{k=1}^n \alpha_k C^*(\mathbf{p} + \delta_k) + \Psi(\mathbf{p}, \hat{\mathbf{d}})) / (\sum_{k=1}^n \alpha_k)$, where the n neighbors³ $\mathbf{p} + \delta_1, \dots, \mathbf{p} + \delta_n$ of \mathbf{p} in direction $\hat{\mathbf{d}}$ are interpolated using the components of the vector $\alpha \triangleq [\delta_1 \mid \dots \mid \delta_n]^{-1} \hat{\mathbf{d}}$. For example, if $\hat{\mathbf{d}} = \delta_k / \|\delta_k\|$ (i.e., $\hat{\mathbf{d}}$ points directly at one of the neighboring voxels) then $f = C^*(\mathbf{p} + \delta_k) + \Psi(\mathbf{p}, \hat{\mathbf{d}}) \|\delta_k\|$. While this is reminiscent of Dijkstra's algorithm, the search for the optimal direction is not restricted to

³ In 3D, this is $n = 3$ neighbors among 26.

discrete grid directions. In our implementation the minimization is performed over 100 directions sampled uniformly on the sphere and the coefficients $\alpha(\hat{\mathbf{d}})$ are pre-computed.

Once the vector field $\hat{\mathbf{d}}^*$ is known, a slightly modified⁴ version of Algorithm 1 is used to solve transport equations.

4 Application to High Angular Diffusion MRI Tractography

4.1 Constructing the Direction-Dependent Cost

Most front propagation techniques for diffusion tensor tractography use some ad hoc function f of the quadratic form $\hat{\mathbf{d}}^t D \hat{\mathbf{d}}$. If the Gaussian assumption holds, the diffusion weighted images follow $S(\mathbf{p}, \hat{\mathbf{d}}) \simeq S(\mathbf{p}, \mathbf{0}) \exp(-b \hat{\mathbf{d}}^t D(\mathbf{p}) \hat{\mathbf{d}})$. Tensor based techniques can formally be extended to high angular resolution diffusion datasets by setting

$$\Psi(\mathbf{p}, \hat{\mathbf{d}}) \triangleq f\left(-\frac{1}{b} \log\left(\frac{S(\hat{\mathbf{d}})}{S(\mathbf{0})}\right)\right) \quad (6)$$

Notice that Q-Ball datasets [11] and direction-dependent local costs Ψ are both defined on the same space $\mathbb{R}^3 \times \mathbb{S}^2$.

However, due to the low signal to noise ratio of these datasets, it is desirable to consider more than one value at a time. The anisotropic cost can be defined by some decreasing function f' of the Funk-Radon transform⁵ of the attenuation $S(\mathbf{p}, \cdot)/S(\mathbf{p}, \mathbf{0})$.

$$\Psi'(\mathbf{p}, \hat{\mathbf{d}}) \triangleq f'\left(\int_{\hat{\mathbf{v}} \perp \hat{\mathbf{d}}} \frac{S(\hat{\mathbf{v}})}{S(\mathbf{0})} d\hat{\mathbf{v}}\right) \quad (7)$$

The cost $\Psi'(\mathbf{p}, \hat{\mathbf{d}})$ will therefore be small if and only if there is limited diffusion loss over the corresponding equator, i.e., if diffusion does not occur normal to $\hat{\mathbf{d}}$.

4.2 Using Prior Knowledge

Mumford [19] showed that variational techniques, such as the one proposed here, can be set in an elegant and principled Bayesian framework by considering the cost $\Psi = \Psi_{\text{data}} + \Psi_{\text{prior}}$. The extreme simplicity of this construction constitutes another advantage over non-variational front propagation approaches. Here, Ψ_{data} would be as described above and Ψ_{prior} could be obtained from an atlas of neural tracts.

The problems of generating such an atlas and registering it to the dataset at hand are well beyond the scope of this paper. Note however that masking off (with infinite cost values) the non white matter regions of the brain is a trivial and widely employed use of prior knowledge.

⁴ No minimization needs to be performed.

⁵ Interestingly, the FRT is also central to the Q-Ball technique [11].

5 Results

Here we show results obtained by applying the methodology described in the above sections to diffusion weighted data sets acquired using a single-shot diffusion-weighted EPI sequence, with 31 different gradient directions with b-values of 500, 1000, and 1500 s/mm^2 , on a 1.5 Tesla GE Echospeed system. The data was acquired with different b-values to enable comparisons of the results. Traditional eigenvector based tractography is normally carried out in data with b-values in the range of 700-1000 s/mm^2 . Higher b-values give data with higher angular contrast but at the expense of more noise.

Cost per unit length, which can be interpreted as a validity index for the putative tracts was determined for all b-values Fig. 1. All curves are optimal given their starting point. The cost per unit length is a measure of how good

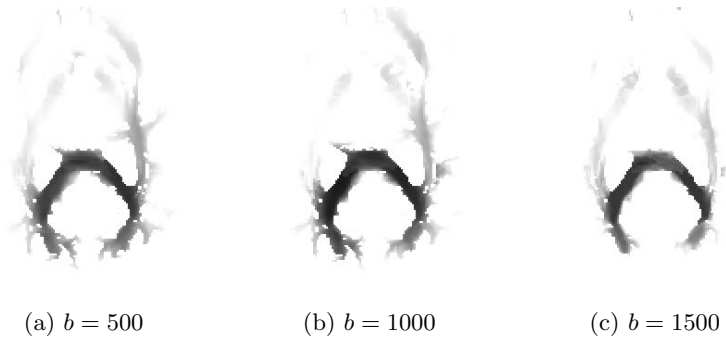


Fig. 1. Cost per unit length of end points of optimal curves for different b-values is a validity index (see text). Best results are achieved for the highest b-value.

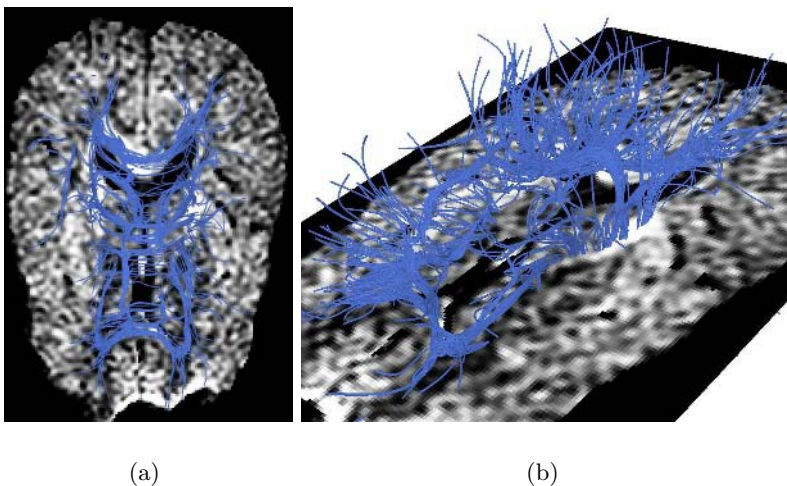


Fig. 2. Fiber tracking from high angular resolution dataset ($b=1500 s/mm^2$)

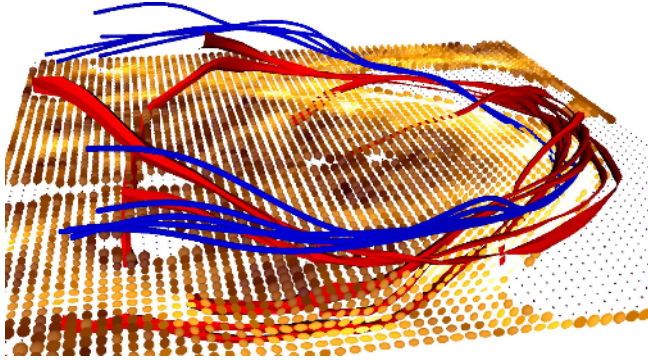


Fig. 3. Proposed technique on high angular resolution data (blue) compared with streamline technique on tensor field (red) ($b=1500 \text{ s/mm}^2$)

they are compared to each other. The best contrast (corresponding to the most coherent set of “super-optimal” tracts for a given seed point posterior of the corpus callosum) was obtained at the highest b -value available. This could indicate that the algorithm was able to take advantage of the higher angular contrast in spite of the lower SNR.

Tract results for several user defined seed points are presented on Fig. 2.

Finally the proposed technique was compared to a streamline technique for the needs of which the tensor field was computed (Fig. 3). While validation is a very challenging task due to the unavailability of ground truth, it can be noted that both algorithm give similar results even though their inputs are different. The tracts of the proposed technique tend to be more coherent as any noise in the data might set the streamline off course whereas the proposed technique is more global.

6 Conclusion

A novel technique has been proposed for fiber tractography from high angular resolution diffusion imagery. In difference to [12,13] this is based on a principled minimum cost (or arrival time) approach. By setting front propagation techniques in a variational light the proposed technique gives some insights on the interpretation of front speed versus particle speed. Preliminary results show that the technique performs better for high b -values when directional resolution is higher. Further research topics include detailed validation as well as the use of prior information.

Acknowledgments

This work is part of the National Alliance for Medical Image Computing (NAMIC), funded by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant U54 EB005149. It is also supported by NIH grant NAC P41 RR-13218.

References

1. Basser, P., Mattiello, J., LeBihan, D.: MR diffusion tensor spectroscopy and imaging. *Biophys. J.* **66** (1994) 259–267
2. Mori, S., Crain, B., Chacko, V., van Zijl, P.: Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann Neurol.* **45** (1999) 265–269
3. Conturo, T., Lori, N., Cull, T., Akbudak, E., Snyder, A., Shimony, J., McKinstry, R., Burton, H., Raichle, M.: Tracking neuronal fiber pathways in the living human brain. In: *Proc. Natl. Acad. Sci. USA.* (1999) 10422–10427
4. Westin, C.F., Maier, S.E., Khidhir, B., Everett, P., Jolesz, F.A., Kikinis, R.: Image processing for diffusion tensor magnetic resonance imaging. In: *MICCAI.* (1999) 441–452
5. Basser, P., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A.: In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine* **44** (2000) 625–632
6. Parker, G., Wheeler-Kingshott, C., Barker, G.: Estimating distributed anatomical connectivity using fast marching methods and diffusion tensor imaging. *IEEE Transactions on Medical Imaging* **21** (2002) 505–512
7. Jackowski, M., Kao, C.Y., Qiu, M., Constable, R.T., Staib, L.: Estimation of anatomical connectivity by anisotropic front propagation and diffusion tensor imaging. In: *MICCAI.* (2004) 663–671
8. O’Donnell, L., Haker, S., Westin, C.F.: New approaches to estimation of white matter connectivity in diffusion tensor MRI: Elliptic PDEs and geodesics in a tensor-warped space. In: *MICCAI.* (2002)
9. Lenglet, C., Deriche, R., Faugeras, O.: Inferring white matter geometry from diffusion tensor MRI: Application to connectivity mapping. In: *ECCV.* (2004)
10. Liu, C., Bammer, R., Acar, B., Moseley, M.: Characterizing non-Gaussian diffusion by using generalized diffusion tensors. *Magnetic Resonance in Medicine* **51** (2004) 924–937
11. Tuch, D.: Q-ball imaging. *Magn Reson Med.* **52** (2004) 1358–1372
12. Hagmann, P., Reese, T.G., Tseng, W.Y.I., Meuli, R., Thiran, J.P., Wedeen, V.J.: Diffusion spectrum imaging tractography in complex cerebral white matter: an investigation of the centrum semiovale. In: *ISMRM.* (2004)
13. Campbell, J.S.: Diffusion Imaging of White Matter Fibre Tracts. PhD thesis, McGill University (2004)
14. Strang, G.: Introduction to Applied Mathematics. Wellesley-Cambridge Press (1986)
15. Pichon, E. PhD thesis, Department of Electrical and Computer Engineering, Georgia Institute of Technology (2005)
16. Mortensen, E., Morse, B., Barrett, W., Udupa, J.: Adaptive boundary detection using live-wire two-dimensional dynamic programming. In: *IEEE Proceedings of Computers in Cardiology.* (1992) 635–638
17. Pichon, E., Sapiro, G., Tannenbaum, A.: Segmentation of Diffusion Tensor Imagery. Number 286 in *LNCIS.* In: *Directions in Mathematical Systems Theory and Optimization.* (2003) 239–247
18. Kao, C., Osher, S., Tsai, Y.: Fast sweeping methods for static Hamilton-Jacobi equations. Technical Report 03-75, UCLA CAM (2003)
19. Mumford, D.: The Bayesian Rationale for Energy Functionals. In: *Geometry-driven Diffusion in Computer Vision.* Kluwer Academic Publisher (1994) 141–153

Automated Atlas-Based Clustering of White Matter Fiber Tracts from DTMRI*

Mahnaz Maddah^{1,2}, Andrea U.J. Mewes², Steven Haker^{2,3},
W. Eric L. Grimson¹, and Simon K. Warfield^{1,2,3,4}

¹ Computer Science and Artificial Intelligence Laboratory,
Massachusetts Institute of Technology, Cambridge, MA 02139, USA

² Computational Radiology Laboratory,

Brigham and Women's Hospital, Boston, MA 02115, USA

³ Surgical Planning Lab., Brigham and Women's Hospital, Boston, MA 02115, USA

⁴ Children's Hospital and Brigham and Women's Hospital,
Harvard Medical School, Boston, MA 02115, USA

{mmaddah, welg}@mit.edu, {mewes, haker, warfield}@bwh.harvard.edu

Abstract. A new framework is presented for clustering fiber tracts into anatomically known bundles. This work is motivated by medical applications in which variation analysis of known bundles of fiber tracts in the human brain is desired. To include the anatomical knowledge in the clustering, we invoke an atlas of fiber tracts, labeled by the number of bundles of interest. In this work, we construct such an atlas and use it to cluster all fiber tracts in the white matter. To build the atlas, we start with a set of labeled ROIs specified by an expert and extract the fiber tracts initiating from each ROI. Affine registration is used to project the extracted fiber tracts of each subject to the atlas, whereas their *B*-spline representation is used to efficiently compare them to the fiber tracts in the atlas and assign cluster labels. Expert visual inspection of the result confirms that the proposed method is very promising and efficient in clustering of the known bundles of fiber tracts.

1 Introduction

Diffusion tensor MR imaging (DT-MRI) non-invasively measures the diffusivity of water molecules within the tissue [1]. It thus provides some knowledge about the direction and density of the fiber tracts in the brain, as the water diffusion is restricted in the direction normal to the fibers. While anisotropy measures are being used to assess the density of the fiber tracts and or the degree of myelination in different regions of the brain, tractography methods have been

* The authors would like to thank Dr. Polina Golland and Dr. Kilian Pohl for their helpful discussions and suggestions. This investigation was supported in part by NSF ITR 0426558, a research grant from the Whitaker Foundation, a research grant from CIMIT, grant RG 3478A2/2 from the NMSS, and by NIH grants R21 MH67054, R01 LM007861, R01 CA109246, P41 RR13218, 1U54 EB005149, and P01 CA67165.

developed to track the fibers (e.g. [2]), extract the fiber bundles, and estimate the interconnectivity of different regions of the brain [3].

To date, significant work has been devoted to providing tools for processing the images [4], tractography schemes [2, 4], visualization [5], and recently quantitative analysis of the extracted tracts [6]. However, limited work has been done to address the inter-subject similarity/variability of the fiber tracts in the human brain, which is our ultimate goal. Toward this goal and to quantitatively analyze the fiber tracts in a population, the first step is to identify the fiber tract clusters in each case. This process would be very inefficient if an expert were needed to specify all regions of interest in every case under study. An alternative approach is to extract all of the tracts in the whole brain and then automatically cluster the desired bundles.

Clustering of fiber tracts has attracted significant attention recently. The tracts have been successfully clustered into bundles using spectral clustering [7, 8], normalized cuts [1], agglomerative hierarchical clustering [9], fuzzy *c*-means [10], *k*-nearest neighbors [11]. However, in the absence of any expert knowledge or atlas to assign the actual anatomical structures to the clusters, there is no guarantee that a particular clustering algorithm yields the desired bundles. The algorithm might easily over- or under-cluster the fiber tracts. So, a clustering scheme is desired that is able to automatically cluster large numbers of tracts to the bundles of interests, and takes into account available anatomical knowledge. Such an algorithm should be also flexible enough to cluster either to the whole anatomical structure or into sub-structures. Accuracy and time-efficiency are also required as in any other application.

In this paper, we present a new method which satisfies the above requirements. We propose to use a labeled atlas of the fiber tracts of the whole brain to perform clustering on the subjects under study. The essence of the proposed method is as follows: First, all tracts in the white-matter tissue are extracted and then projected to the atlas. The applied transformation matrix is obtained by registering the baseline image of the subject (the images obtained without applying any gradient in the magnetic field) to that of the atlas. The final step is to assign a cluster label to each projected tract. This is done by taking the cluster label of the most similar tract in the atlas. The details of each step will be described in the next sections. The preliminary results show the efficiency of the proposed method. The next section describes how we develop the atlas, while the clustering method is explained in Section 3 along with the results obtained at each step. Conclusions and future work are discussed in Section 4.

2 Constructing the Fiber Tract Atlas

2.1 Specifying the Regions of Interest

High-resolution DT-MRI data were achieved from healthy volunteers on a 3T General Electrics Signa scanner (Milwaukee, WI) with repetition time/echo time of 2500/70 ms, b-factors 50 of 5 and 750 s/mm², and six gradient directions. Raw data were converted to derive tensor information, and were loaded into 3D slicer

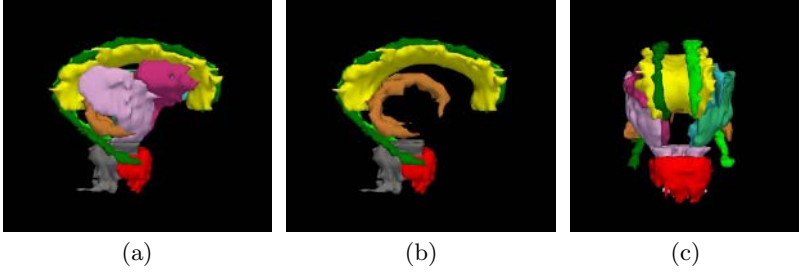


Fig. 1. Sagittal (a),(b) and anterior (c) view of the ROIs for some of the fiber tract bundles specified by an expert. The internal capsule are removed in (b) to show the inner structures.

(www.slicer.org). The region of interests (ROIs) were outlined by an expert based on information from the map of the fractional anisotropy (FA), and a map, which color-codes the direction of the main eigenvector, both generated in the slicer. The expert specified major association, projection and commissural tracts as the ROIs which are used as starting points for tractography (See Fig. 1).

2.2 Tractography

The tractography approach we implemented is categorized as a stochastic approach. Various stochastic tractography methods have been proposed to alleviate the shortcomings of deterministic approaches in crossing regions by adding some randomness to the deterministic tractography. Instead of defining an ad-hoc stochastic component, we used the assumption of a Gaussian distribution for finding a particle at a distance \mathbf{r} from its initial position when diffusing in an anisotropic medium, as described by Basser [12]. If we represent the diffusion tensor at each voxel, \mathbf{r} , by $\mathbf{D}(\mathbf{r})$, with eigenvalues and corresponding eigenvectors denoted by λ_i and \mathbf{E}_i for $i = 1, 2, 3$, we get a Gaussian displacement distribution [12]:

$$\lim_{\mathbf{r} \rightarrow \infty} \Pr(\mathbf{r}, \Delta) = \frac{1}{\sqrt{|\mathbf{D}|(4\pi\Delta)^3}} \exp\left(-\frac{\mathbf{r}^T \mathbf{D}^{-1} \mathbf{r}}{4\Delta}\right) \quad (1)$$

where Δ is the diffusion time. Diagonalization of the tensor \mathbf{D} gives:

$$\lim_{\mathbf{r} \rightarrow \infty} \Pr(\mathbf{r}, \Delta) = \frac{1}{\sqrt{\lambda_1 \lambda_2 \lambda_3 (4\pi\Delta)^3}} \exp\left(-\frac{r_1^2}{4\Delta\lambda_1} - \frac{r_2^2}{4\Delta\lambda_2} - \frac{r_3^2}{4\Delta\lambda_3}\right) \quad (2)$$

where the r_i 's are the components of \mathbf{r} in the base described by the \mathbf{E}_i 's. This means that the stochastic term is composed of three Gaussian distribution functions along each \mathbf{E}_i , whose variance is equal to the corresponding λ_i .

Using the proposed tractography method, we extract all the tracts starting from all voxels of the ROIs obtained in section 2.1. Fig. 2 shows the extracted fiber tracts, colored based on the cluster to which they belong. This data, i.e. the labeled fiber tracts, is used as the reference model or atlas throughout the work.

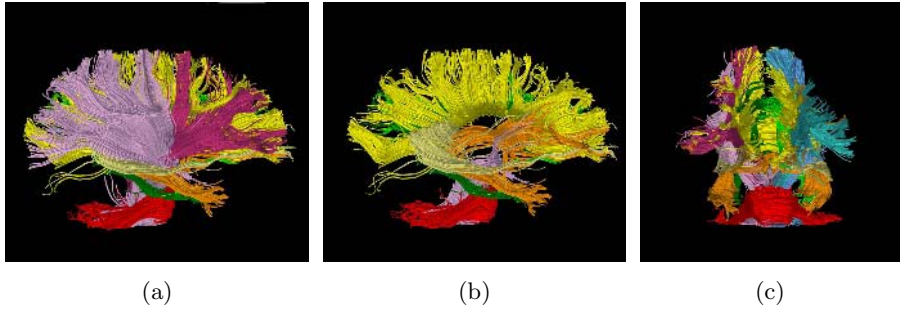


Fig. 2. The constructed atlas of fiber tracts shown for the obtained bundles in the sagittal (a),(b), and anterior view. Internal capsula's tracts have been removed in (b) to show the inner bundles.

3 Clustering Framework

Having constructed the atlas, the whole brain fiber tracts are extracted from DT-MRI of each subject and then will be clustered using the atlas as will be described in this section. The method was applied to 10 DT-MRI data sets acquired from healthy volunteers with voxel size of $1.054 \times 1.054 \times 2$ mm.

3.1 Whole Brain Tractography

The white matter of the whole brain is segmented using a k -NN algorithm [13], and used as the starting point for the tractography. The tracts are then extracted through the stochastic approach described in Section 2.2. From each starting point, two tracts are extracted in opposite directions and connected together to form the complete fiber tract. The tractography algorithm stops for each tract when an FA value of less than 0.2 is reached.

3.2 Projection

The second step is to project the extracted tracts to the reference subject (atlas). To obtain the transformation matrix, we register the baseline images of the subject to that of the atlas. Fig. 3 shows the results of affine registration using a block matching algorithm [14] for one of the subjects under study.

3.3 Clustering

The main and final step is to assign each projected tract to the proper cluster, considering the fact that the tracts of the atlas are already labeled based on the ROI to which they belong (Sec. 2.2). There are several possibilities: The simplest idea is to use the nearest neighborhood technique such that for each projected tract, the nearest tract in the atlas is identified and its cluster label is assigned to the projected tract. The nearest tract can be defined as the tract

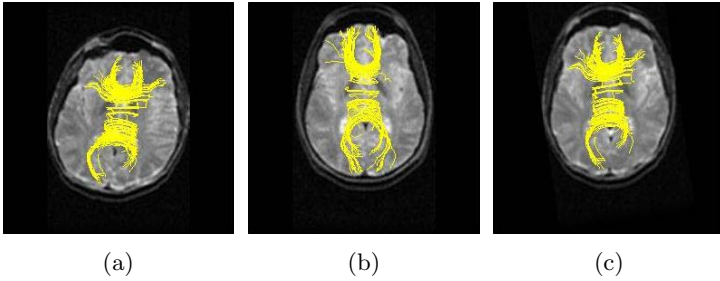


Fig. 3. (a) A set of fiber tracts extracted from one of the subjects under study, overlaid on an axial slice of its baseline image. (b) Same as (a) for the reference subject. (c) The projected tracts overlaid on the registered baseline image of the subject. The pairwise comparison is performed between tracts in (b) and (c) to assign cluster labels to the tracts in the subject.

which minimizes either the closest point distance, the mean of the closest distance, or the Hausdorff distance between the two tracts [6]. Such methods are usually computationally expensive. An alternative approach is to use a curvature and torsion of the 3D curves (tracts), which has the advantage of being invariant to translation and rotation [15]. Such approaches in fact compare a shape model of the curves and neglect the spatial distance between the two curves. Thus, this is not be the best choice for our purpose where spatial location of the tracts is also important. Also, computing the curvature and torsion requires 3rd order differentiation of the data, which is prone to noise and quantization errors. Quintic splines are sometimes used to smooth out the data and to perform the differentiation through convolution with spline kernels instead of simple differentiation [15].

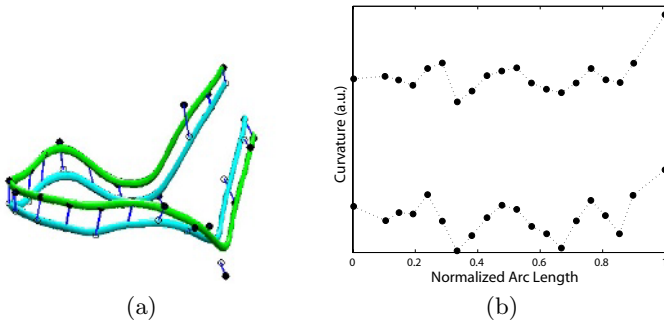


Fig. 4. (a) Two neighboring fiber tracts and their corresponding 3D spline representation. Knot points are marked with circles and connected with a short line to guide the eye. (b) The corresponding curvature along each of the two tracts calculated based on their spline representation. The curvature values are shifted with an arbitrary offset for ease of comparison.

In our clustering approach, we use the B -spline representation [16] of the fiber tracts for pairwise comparison of the fiber tracts extracted from the subject to those from the atlas. The spline representation contains information of both spatial location and shape of the tracts. We used quintic splines to preserve the information needed to calculate the curvature and torsion of the 3D curves. A 20-knot representation proved sufficient. Figure 4(a) shows the spline representation of two neighboring tracts from the corpus-callosum, while Figure 4(b) shows the corresponding curvatures computed along the tracts.

Once the spline representation is computed, each tract from the subject is compared to the tracts in the atlas. In doing so, a hierarchical scheme was used so that the projected tracts are first compared with the major bundles in the atlas and then with the smaller ones. For the atlas shown in Fig. 2, this leads to a significant drop in computation time once the corpus-callosum and capsula-interna bundles are clustered. The maximum number of bundles clustered is equal to the number of labeled bundles in the atlas. Also, note that since our atlas does not include all of the structures, some tracts are left unclustered.

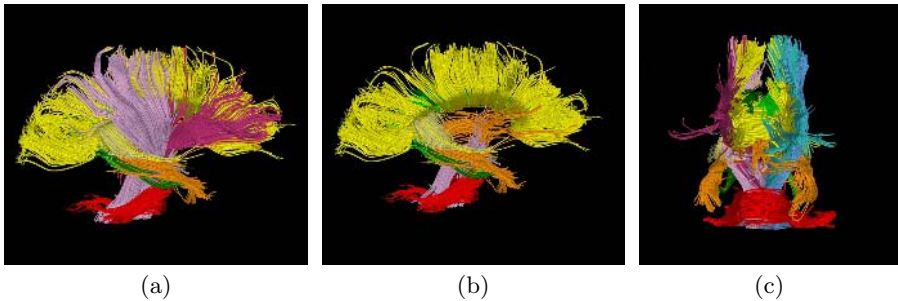


Fig. 5. Sagittal (a),(b) and anterior (c) view of the clustered fiber tracts in of one of the subjects under study



Fig. 6. Sagittal view of the clustered corpus callosum in four different subjects

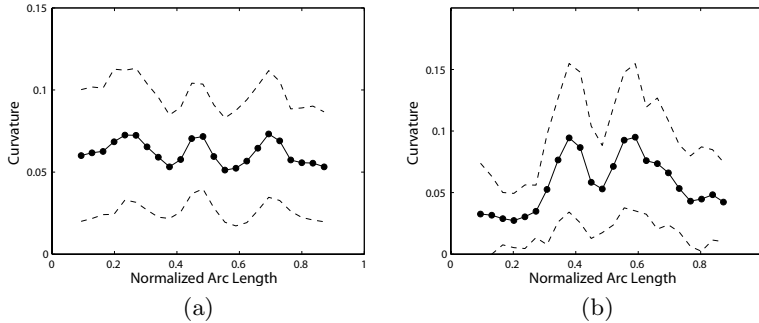


Fig. 7. The average and variation of the curvature along the tracts for the (a) corpus callosum and (b) pons. Dotted lines represent the average \pm the standard deviation.

These tracts can be further clustered with a non-atlas based algorithm to reveal those structures not present in the atlas, or the atlas can be extended to include more structures. Fig. 5 shows different views of the fiber tracts from a subject clustered to the bundles based on the atlas of Fig. 2. Finally, Fig. 6 shows the sagittal view of the clustered corpus-callosum bundle in four different subjects.

In our framework it is also possible to include shape information, such as curvature and torsion when clustering the bundles. For example, for a structure such as the corpus callosum that can be described by a single shape model [17], the tracts with such shape characteristics can be clustered based on their shape representation. This can be performed even without the need to project the subject tracts to the atlas and by comparing only to a single shape representation obtained as the average shape of the tracts for that particular bundle in the atlas. Fig. 7 shows the curvature along the tract for the corpus callosum and the pons.

4 Conclusions and Future Work

In this paper, we presented for the first time an atlas-based clustering framework for fiber tracts of the whole human brain. Preliminary results prove the efficiency of the proposed method to cluster the fiber tracts into anatomically known bundles. The correspondence of clusters in different subjects is defined by default thus unlike most clustering methods proposed for this application, no post-processing is required. The proposed framework has also the flexibility to use different similarity measures, such as spatial distance or shape similarity or a combination of them, for different structures.

Since the atlas has a significant impact on the clustering results, modification and improvement of the atlas, in terms of accuracy of ROIs, and the inclusion of more cases when constructing the atlas are of great importance. Quantitative analysis and study of the variability of particular fiber tract bundles in a population is currently underway.

References

1. Brun, A., Knutsson, H., Park, H.J., Shenton, M.E., Westin, C.F.: Clustering fiber traces using normalized cuts. In: MICCAI. (2004) 368–375
2. Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A.: In vivo fiber tractography using DT-MRI data. *Magn. Reson. Med.* **44** (2000) 625–632
3. Behrens, T.E.J., et al.: Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nature Neuroscience* (2003) 750–757
4. Westin, C.F., Maier, S.E., Mamata, H., Nabavi, A., Jolesz, F.A., Kikinis, R.: Processing and visualization of diffusion tensor MRI. *Med. Image Anal.* **6** (2002) 93–108
5. Fillard, P., Gerig, G.: Analysis tool for diffusion tensor MRI. In: MICCAI. (2003) 967–968
6. Gerig, G., Gouttard, S., Corouge, I.: Analysis of brain white matter via fiber tract modeling. In: Proc. IEEE Int. Conf. EMBS. (2004)
7. Jonasson, L., Hagmann, P., Thiran, J.P., Wedeen, V.J.: Fiber tracts of high angular resolution diffusion MRI are easily segmented with spectral clustering. In: Proc. ISMRM. (2005) 1310
8. O'Donnell, L., Westin, C.F.: White matter tract clustering and correspondence in populations. In: MICCAI. (2005) to be published.
9. Zhang, S., Laidlaw, D.H.: DTI fiber clustering in the whole brain. In: IEEE Visualization. (2004) 28p
10. Shimony, J.S., Snyder, A.Z., Lori, N., Conturo, T.E.: Automated fuzzy clustering of neur-sonal pathways in diffusion tensor tracking. In: Proc. ISMRM. (2002) 10
11. Ding, Z., Gore, J.C., Anderson, A.W.: Classification and quantification of neuronal fiber pathways using diffusion tensor MRI. *Mag. Reson. Med.* **49** (2003) 716–721
12. Basser, P.J.: Relationships between diffusion tensor and q-space MRI. *Magn. Reson. Med.* (2002) 392–397
13. Warfield, S.K.: Fast k -NN classification for multichannel image data. *Pattern Recognition Lett.* **17** (1996) 713–721
14. Ourselin, S., Roche, A., Prima, S., Ayache, N.: Block matching: A general framework to improve robustness of rigid registration of medical images. In: MICCAI. (2000) 557–566
15. Kehtarnavaz, N., deFigueiredo, R.J.P.: A 3-D contour segmentation scheme based on curvature and torsion. *IEEE Trans. Pattern Analysis and Machine Intell.* **10** (1988) 707–713
16. Cohen, F.S., Huang, Z., Yang, Z.: Invariant matching and identification of curves using B -splines curve representation. *IEEE Trans. Image Proc.* (1995) 1–10
17. Corouge, I., Gouttard, S., Gerig, G.: Towards a shape model of white matter fiber bundles using diffusion tensor MRI. In: Int. Symp. Biomed. Imag. (2004) 344–347

MR Diffusion-Based Inference of a Fiber Bundle Model from a Population of Subjects

V. El Kouby^{1,3}, Y. Cointepas^{1,3}, C. Poupon^{1,3}, D. Rivière^{1,3}, N. Golestani^{1,2},
J.-B. Poline⁴, D. Le Bihan^{1,3}, and J.-F. Mangin^{1,3}

¹ Service Hospitalier Frédéric Joliot, CEA, 91401 Orsay, France
elkouby@shfj.cea.fr, <http://brainvisa.info>

² INSERM U562 Cognitive Neuro Imaging, Orsay, France

³ Institut Fédératif de Recherche 49 (Imagerie Neurofonctionnelle), Paris

Abstract. This paper proposes a method to infer a high level model of the white matter organization from a population of subjects using MR diffusion imaging. This method takes as input for each subject a set of trajectories stemming from any tracking algorithm. Then the inference results from two nested clustering stages. The first clustering converts each individual set of trajectories into a set of bundles supposed to represent large white matter pathways. The second clustering matches these bundles across subjects in order to provide a list of candidates for the bundle model. The method is applied on a population of eleven subjects and leads to the inference of 17 such candidates.

1 Introduction

In the recent past, the development of diffusion tensor imaging (DTI) has led to a new field of opportunities for the study of brain white matter [11]. A possible approach for such studies lies in the usual spatial normalization framework. For this purpose, the development of methods for warping tensor fields using full tensor information is a promising undertaking, because it may improve the normalization computed from standard anatomical images [10, 20]. Spatial normalization can be used for instance for voxel-based statistical analysis. Combining normalization with fiber tracking techniques can also lead to statistical parametric maps reflecting variability of bundle location and thickness. Other morphometric approaches can be developed from dedicated parameterizations of the bundles [8, 6].

One of the difficulties for bundle morphometry stems from the poor anatomical knowledge about white matter organization of the human brain. Most of the anatomical techniques used to trace tracts can not be used in the human brain. Therefore, the current knowledge built from Klinger method or observation of Wallerian degeneration after lesion is relatively sparse [7, 13]. There is a need for computational methods aiming at inferring a model of the main bundles making up the human brain white matter from MR diffusion imaging. This paper explores this direction of research.

A lot of methods have been proposed to compute fiber related pathways from diffusion-weighted MR images. The simplest approaches follow trajectories of maximum diffusion [14, 5, 1]. More sophisticated methods try to overcome partial volume problems induced by fiber crossing using either high angular resolution acquisitions [18, 15],

regularization [19, 16], front propagation [12, 9] or Monte Carlo sampling [3, 2]. Most of these techniques convert the raw MR diffusion data into a huge set of trajectories supposed to correspond to putative fascicles. These fascicles can be used to create connectivity matrices related either to functional or anatomical regions of interest (ROIs) [5]. These fascicles can also be organized into larger bundles using clustering algorithms [4], which is addressed in this paper. This last approach is interesting because the potential influence of erroneous definitions of the ROIs is discarded.

The proposed method aims at the inference of a model of the bundles that can be detected in most of the subjects. Some of these bundles should be related to the current anatomical knowledge, but the ultimate goal is to enrich the current understanding of white matter organization. The method is made up of two nested clustering algorithms. The first clustering is performed subject by subject in order to reduce the complexity of the full set of fascicles to a small set of large bundles (cf. Fig. 1). The second clustering is performed across subjects in order to match similar bundles that may correspond to a generic anatomical pathway (cf. Fig. 3 and 4). This second stage relies on a simple affine spatial normalization computed from the diffusion-free T2-weighted image included in the dataset. The work described in this paper is exploratory. Therefore, we do not address the current weaknesses of the tracking algorithms that may provide spurious fascicles. Moreover, we use a very simple clustering algorithm in order to get a first insight into the problem before designing more sophisticated dedicated methods.

2 Method

2.1 Clustering the Fascicles

The first clustering is performed subject by subject according to the following scheme:

Global fascicle tracking. The computation of the initial set of fascicles is performed with brainVISA (<http://brainvisa.info>), an open software implementing Mori's algorithm [14]. Simply speaking, the fascicles are trajectories of highest diffusion reconstructed step by step from the diffusion tensor field. A mask of the brain is computed from the T2-weighted image and one trajectory is obtained for each voxel, leading to a set of about 500000 fascicles. In this paper, we have chosen to focus on long bundles which led us to discard the fascicles whose length is less than 5cm. The final set includes about 20000 fascicles (see Fig.1.A).

ROI map. The clustering is not performed directly on the fascicles but on a set of Regions of Interest (ROIs) defined from Talairach proportional system. The Talairach grid is split into cubic ROIs that are transformed to the subject referential using affine spatial normalization of the T2-weighted image performed via SPM2 (<http://www.fil.ion.ucl.ac.uk/spm>) (see Fig.1.B). For the study described in this paper, the spatial resolution of the grid of ROIs is 5mm. After masking this grid with a mask of the subject white matter (see Fig.1.D), the final set is made up of about 5000 ROIs.

Connectivity matrix. An anatomical connectivity matrix A is computed for the set of ROIs mentioned above. For each pair of ROIs (i, j) , A_{ij} is the number of fascicles

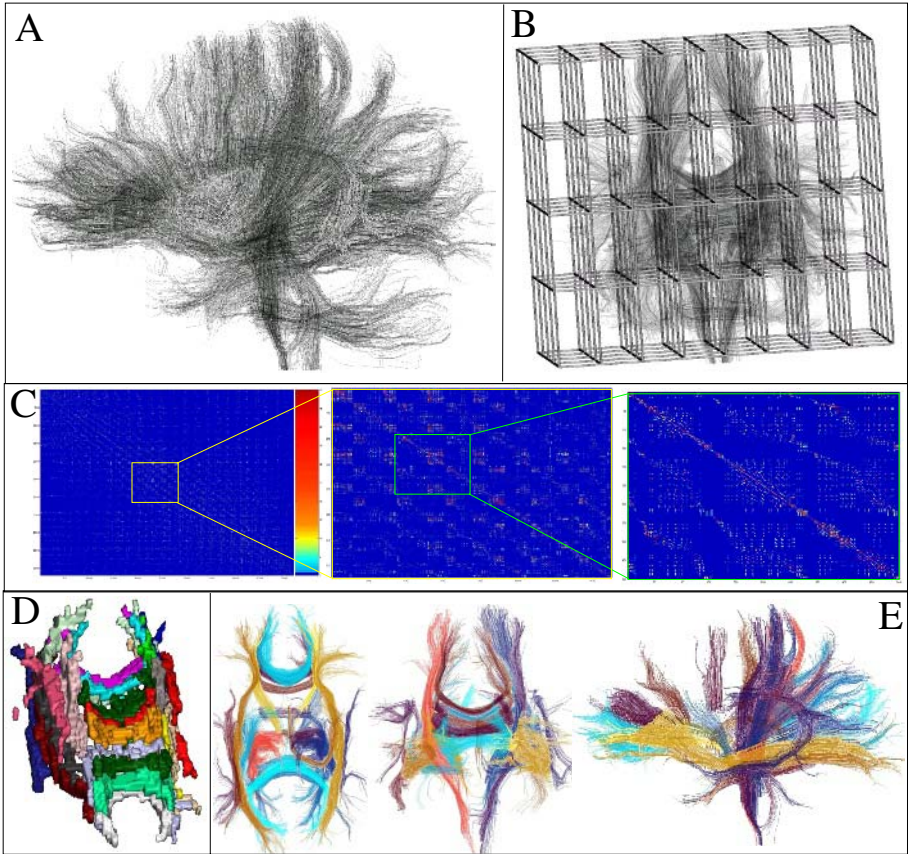


Fig. 1. Clustering the fascicles of one subject: **A:** The set of long fascicles to be clustered into bundles. **B:** The same set of fascicles embedded into the Talairach grid used to define cubic ROIs of 5mm resolution. **C:** The connectivity matrix of the ROIs with 2 levels of zoom (about 5000 ROIs). Each point in the matrix is the number of fascicles linking the two related ROIs. **D:** Result of the clustering for the ROIs. Each color denotes a cluster. Each ROI shape is a cube in Talairach space masked by the subject white matter. **E:** Result of the clustering at the level of the fascicles. The garbage clusters have been filtered out.

crossing i and j (see Fig.1.C). In order to focus the clustering on the strongest connectivity information, this matrix is binarized using a threshold on the number of fascicles. For the experiment described further, this threshold has been set to the mean of non null coefficients plus one standard deviation, namely about 40 fascicles depending on the subject.

Clustering. In the matrix A , each row stands for the connectivity of one ROI with the other ROIs. The ROIs belonging to the same bundle should have similar connectivity patterns. Therefore, we perform a clustering among the ROI set using a distance between ROI corresponding to the Euclidean distance between the rows of A . The

clustering is performed using the standard k-means with random initialization of the centers (see Fig.1.D). K-means algorithm assigns each ROI to a cluster, which lead us to assume that the largest cluster in a kind of garbage collector that is discarded before further processing.

Fascicle bundles. The final step of the clustering consists in moving back to the fascicle world. Each cluster of ROIs is used to extract a subset of fascicles from the global set. For robustness purpose, a fascicle only needs to be partly included in the ROI cluster to be selected. In practice, a fascicle is selected when at least 60% of its sampling is included (see Fig.1.E). A last processing aims at splitting the resulting set of fascicles into connected components, which is required to deal with some unpredictable behaviour of the K-means. The clustering, indeed, gathers sometimes two bundles symmetric across interhemispheric plane, which calls for the use of better algorithms in the future. This postprocessing can not be performed directly at the level of the ROI clusters because the ROIs including fiber crossing can be assigned to only one cluster, leading sometimes to disconnect a cluster of ROIs corresponding to a reasonable bundle. The analysis of the connectedness of a fascicle set is performed via embedding of the fascicles into a 3D volume. It should be noted that this postprocessing implies that the final number of fascicle bundles depends on the subject.

2.2 Clustering the Bundles Across Subjects

The goal of the second stage of clustering is the inference of a model of the bundles that should be found in any human brain. The underlying idea is that a simple clustering performed to match similar bundles could provide a list of candidates for such a model. These candidates would not have to be found in any brain during this inference stage, but should be included in the model of the white matter used to design further pattern recognition systems. What we have in mind is a model of the bundles used to label the fascicles of any new brain in a way similar to what has been developed to deal with cortical sulci [17].

Similarity measure for bundles. In order to match bundles across subjects, a shape-based similarity measure is required. Relying on iconic representations of the bundles allows the definition of discriminant measures stemming from the spatial normalization

$$C = \begin{bmatrix} 0_{00} & M^{12} & \dots & \dots & M^{1N} \\ M^{21} & & \ddots & & \vdots \\ \vdots & \ddots & 0_{ii} & \ddots & \vdots \\ \vdots & & \ddots & & M^{N-1N} \\ M^{N1} & \dots & \dots & M^{NN-1} & 0_{NN} \end{bmatrix}$$

Fig. 2. The block-based structure of the matrix of similarity measures between bundles. Each M^{ij} block corresponds to the one to one correlation coefficients computed between the bundles of the subject i and the bundles of the subjects j . The O_{ii} blocks would be zero-based without smoothing of the binary representations.

towards Talairach space. For this purpose, each fascicle bundle is converted first into a binary representation lying into Talairach space. Then, to reduce the influence of the non perfect affine spatial normalization, these binary representations are convoluted with a Gaussian kernel with 3mm standard deviation. Finally, for each pair of bundles (i, j) , the similarity C_{ij} is computed as the correlation coefficient between their respective smoothed representations. With a subject-based ordering of the bundles, the similarity matrix C has a block structure (see Fig. 2 and 3).

Clustering. The second stage of clustering is also performed with a K-means algorithm. Each bundle is represented by a row of the similarity matrix C . It should be noted that

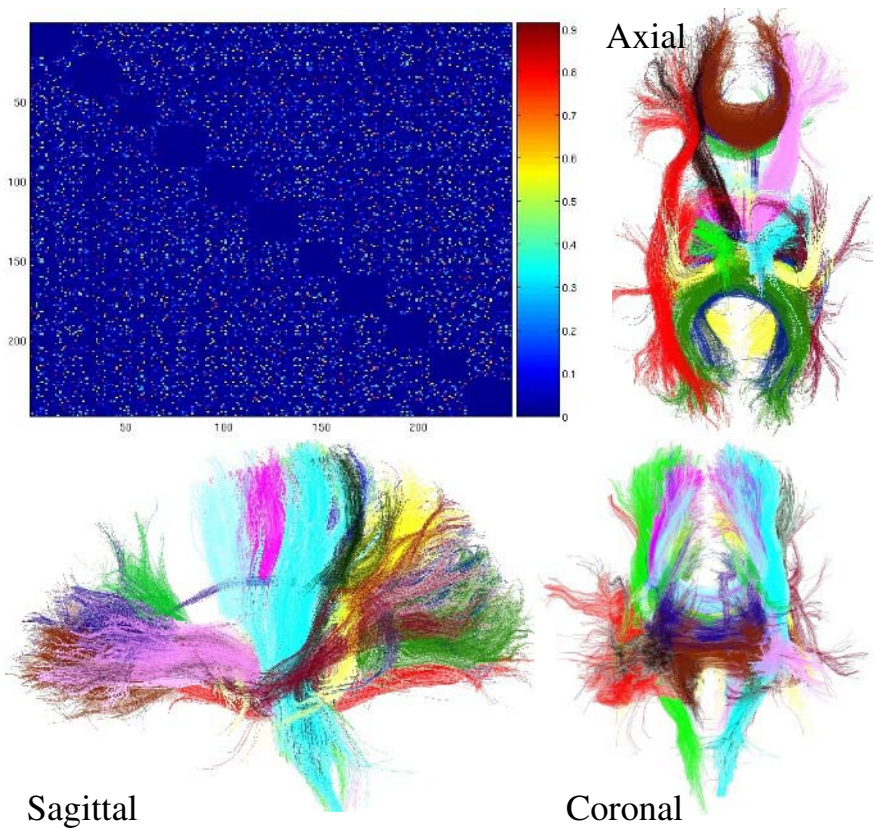


Fig. 3. Clustering the bundles across subjects: The matrix of correlation coefficient computed for all pairs of bundles and three orthogonal views of the result of the clustering. All subjects bundles have been gathered into Talairach space. Each color denotes a cluster of bundles that may correspond to a generic anatomical entity. Garbage clusters have been filtered out. Among the clusters that are candidates to stand for some known anatomical pathways, we noticed the following: (black) ALIC: Anterior Limb of Internal Capsule; (pink) ATR: Anterior Thalamic Radiation; (red) ILF: Inferior Fronto-Occipital; (red) IFO: Inferior Longitudinal Fasciculus; (braun and dark blue) GCC: Genu of Corpus Callosum; (dark green and violet blue) SCC: Splenium of corpus callosum; (light yellow) FX: Fornix.

the K-means can merge several bundles of the same subject into the same cluster. It should be noted also that a cluster does not have to include a bundle of each subject. It should be noted finally that the K-means has to assign each bundle to a cluster. These remarks highlight the fact that a postprocessing is required to clean up the clustering before building the bundle models. The idea is first to get rid of the clusters that are not reasonable representation of a bundle, second to get rid of the bundles that look like outliers in the cluster they belong to.

Let N be the number of subjects. For the results presented in the following, the clusters gathering less than $N/2$ bundles were discarded because of a lack of reproducibility across subjects. The clusters made up of more than $2 * N$ bundles were discarded as garbage collector clusters. Finally, inside each clusters, the statistics of the distance to the cluster center were computed. The outlier bundles were defined as those beyond one mean plus one standard deviation. The remaining clusters are considered as candidates to represent one of the building block of the human brain white matter organization.

3 Results

3.1 Clustering the Fascicles

Eleven subjects were processed. Their diffusion data have been acquired on a 1.5T with 41 different directions for the diffusion-weighting gradient. The resulting 3D volumes have a 128x128x60 resolution with 1.875x1.875x2 voxel sizes. BrainVISA software was triggered with 42x42x20 5mm cubic ROIs. The K-means was applied with 20 classes for each subject. After the connectivity-based postprocessing, we obtained between 20 and 25 bundles for each subject. An example of result can be found in figure 1.

3.2 Clustering the Bundles Across a Population

The second clustering was applied on the eleven subjects with a 25 classes K-means. After cleaning up the result, we obtained 17 reasonable clusters. The result can be visualized in figures 3 and 4. Some of these clusters seem to fit some well known bundles described in anatomical atlases, while the various bundles resulting from the split of the corpus callosum may have less anatomical meaning.

4 Discussion

The inference of a model of the white matter organization will become mandatory to fully exploit the information embedded into MR diffusion-weighted images. While a straightforward approach for this purpose lies in the iconic atlas and spatial normalization principles, this paper advocates an alternative which aims at inferring a higher level model of the bundle organization. The work proposed in this paper is still very preliminary. However, it provides an overview of the various difficulties that have to be addressed to tackle the inference of such a model. A lot of the ad hoc choices presented above could be questioned and more sophisticated methods should be derived in the future.

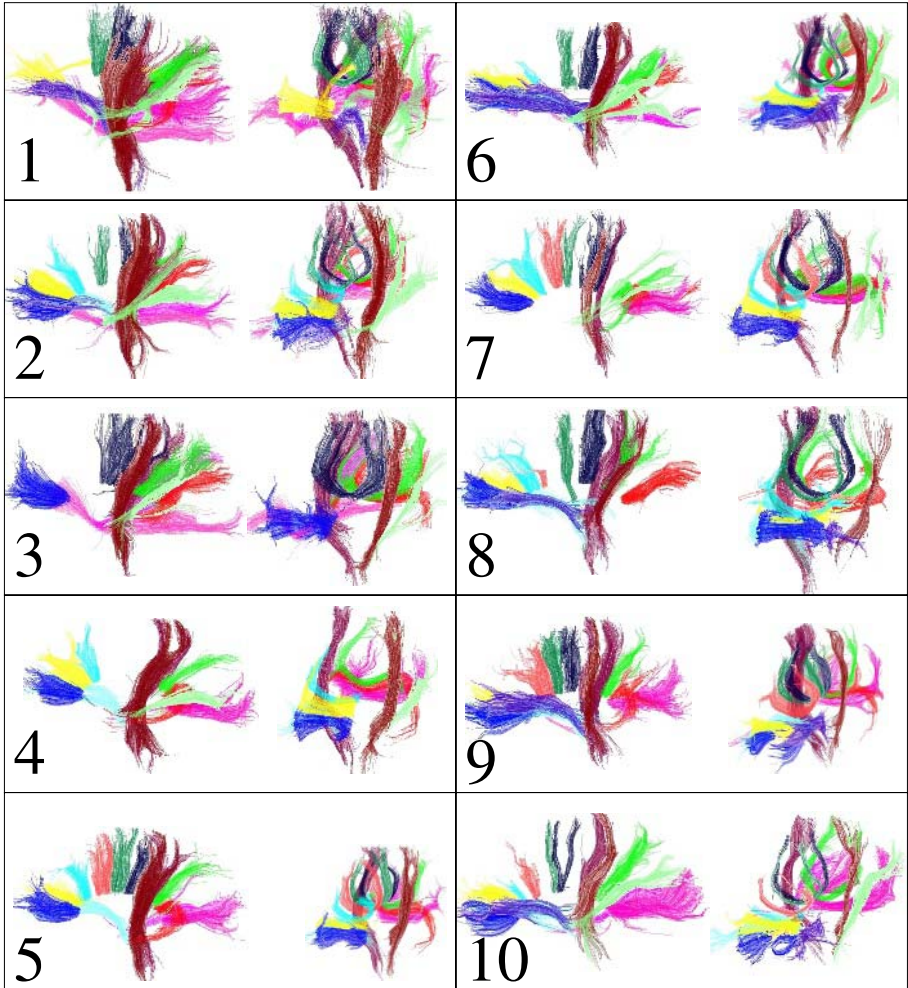


Fig. 4. Clustering the bundles across subjects: The result of the clustering for ten of the eleven subjects. For each subject two views are presented. Each color denotes a cluster that gathers bundle across most of the subjects.

The key idea proposed in our preliminary work is the two stage inference strategy: building first a bundle-based representation before trying to match bundles across subjects. This idea allows the matching to deal with reasonable data sizes. The first clustering, however, can not be questioned during the matching stage which is problematic. Therefore, future approaches may lead to build multiscale representations of the individual data. For instance, each individual bundle may be split further into smaller bundles using one more non supervised clustering. Some other multiscale representations could stem from selecting other ranges of fascicle length in order to deal with short pathways like U-fiber bundles.

As mentioned above, the model inference does not have to find each anatomical entity in every brain. This could be the job of a dedicated pattern recognition system developed to process large databases of diffusion-weighted data. Such a system would face a simpler situation where brains are processed one by one. Hence, the matching with the bundle model could be performed at the fascicle level. Each fascicle would have to be labelled with the name of one of the entities of the model. Such a system would open the door to massive morphometric analysis of the white matter bundles, which would lead to numerous applications in neurosciences.

References

1. P. J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi. In vivo fiber tractography using DT-MRI data. *MRM*, 44(4):625–32, 2000.
2. T.E.J Behrens, M.W. Woolrich, M. Jenkinson, H. Johansen-Berg, R.G. Nunes, S. Clare, P.M. Matthews, J.M. Brady, and S.M. Smith. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine*, 50:1077–1088, 2003.
3. M. Björnemo, A. Brun, R. Kikinis, and C.-F. Westin. Regularized stochastic white matter tractography using diffusion tensor MRI. In *MICCAI'02, LNCS 2488*, pages 435–442, 2002.
4. A. Brun, H. Knutsson, H.-J. Park, M. E. Shenton, and C.-F. Westin. Clustering fiber traces using normalized cuts. In *MICCAI 2004, LNCS 3216*, pages 368–375, 2004.
5. T. E. Conturo, N. F. Lori, T. S. Cull, E. Akbudak, A. Z. Snyder, J. S. Shimony, R. C. McKinstry, H. Burton, and M. E. Raichle. Tracking neuronal fiber pathways in the living human brain. *Proc. Natl. Acad. Sci. USA*, 96:10422–10427, August 1999.
6. I. Corouge, S. Gouttard, and G. Gerig. A statistical shape model of individual fiber tracts extracted from diffusion tensor MRI. In *MICCAI 2004, LNCS 3217*, pages 671–679, 2004.
7. J. Dejerine. *Anatomie des centres nerveux*. Paris: Rueff, 1901.
8. P. Fillard, J. Gilmore, W. Lin, and G. Gerig. Quantitative analysis of white matter fiber properties along geodesic paths. In *MICCAI 03, LNCS 2879*, pages 16–23, 2003.
9. M. Jackowski, C. Y. Kao, R. T. Constable, and L. H. Staib. Estimation of anatomical connectivity by anisotropic front propagation and diffusion tensor imaging. In *MICCAI04, LNCS 3217*, pages 663–670, 2004.
10. D. K. Jones, L. D. Griffin, D. C. Alexander, M. Catani, M. A. Horsfield, R. Howard, and S. C. Williams. Spatial normalization and averaging of diffusion tensor mri data sets. *Neuroimage*, 17(2):592–617, 2002.
11. D Le Bihan, J.-F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko N, and H. Chabriat. Diffusion tensor imaging: Concepts and applications. *Journal of Magnetic Resonance Imaging*, 13(4):536–546, 2001.
12. C. Lenglet, R. Deriche, and O. Faugeras. Inferring white matter geometry from diffusion tensor mri: Application to connectivity mapping. In *8th ECCV*, 2004.
13. E. Ludwig and L. Klinger. *Atlas cerebri humani*. Basel: Karger, 1956.
14. S. Mori, B.J. Crain, V. P. Chacko, and P. C. M. van Zijl. Three dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann. Neurol.*, 45:265–269, 1999.
15. G. J. M Parker and D. C. Alexander. Probabilistic Monte Carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. In *IPMI*, pages 684–95, 2003.

16. C. Poupon, C.A. Clark, V. Frouin, J. Régis, I. Bloch, D. Le Bihan, and J-F. Mangin. Regularization of diffusion-based direction maps for the tracking of brain white matter fascicles. *NeuroImage*, 12(2):184–195, 2000.
17. D. Rivière, J.-F. Mangin, D. Papadopoulos, J.-M. Martinez, V. Frouin, and J. Régis. Automatic recognition of cortical sulci of the human brain using a congregation of neural networks. *Medical Image Analysis*, 6(2):77–92, 2002.
18. D. S. Tuch. *Diffusion MRI of complex tissue structure*. PhD thesis, MIT, 2002.
19. D. Weinstein, G. Kindlmann, and E. Lundberg. Tensorlines: Advection-diffusion based propagation through diffusion tensor fields. In *IEEE Visualization*, pages 249–254, 1999.
20. D. Xu, S. Mori, D. Shen, P. C. van Zijl, and C. Davatzikos. Spatial normalization of diffusion tensor fields. *Magn Reson Med*, 50(1):175–182, 2003.

Knowledge-Based Classification of Neuronal Fibers in Entire Brain

Yan Xia, And U. Turken, Susan L. Whitfield-Gabrieli, and John D. Gabrieli

Psychology Department, Stanford University
yanxia@stanford.edu, andt@stanford.edu, suewhit@stanford.edu

Abstract. This work presents a framework driven by parcellation of brain gray matter in standard normalized space to classify the neuronal fibers obtained from diffusion tensor imaging (DTI) in entire human brain. Classification of fiber bundles into groups is an important step for the interpretation of DTI data in terms of functional correlates of white matter structures. Connections between anatomically delineated brain regions that are considered to form functional units, such as a short-term memory network, are identified by first clustering fibers based on their terminations in anatomically defined zones of gray matter according to Talairach Atlas, and then refining these groups based on geometric similarity criteria. Fiber groups identified this way can then be interpreted in terms of their functional properties using knowledge of functional neuroanatomy of individual brain regions specified in standard anatomical space, as provided by functional neuroimaging and brain lesion studies.

1 Introduction

Delineation of functional brain networks is an important issue for understanding brain function. Most of what is known about functional organization of the brain is based on contributions of gray matter regions distributed across the brain. Relating this knowledge to structural organization of white matter has become possible with recent DTI methods, [1]. Tractography allows characterization of connectivity patterns across cortical and subcortical regions by specifying extent and orientations of the neural fibers in the white matter. With this information, normal brain function, its development, pathologies, and the effects of normal aging can be better understood if the principles of white matter organization can be described [2]; the nature of the interactions between functionally related brain regions but geometrically separated can be investigated. Conversely, functional interpretation of white matter fiber bundles can be made in reference to the gray matter regions that they connect.

An important goal in this area of research is to achieve systematic identification and characterization of axonal fiber tracts. Current attempts for manually and automatic grouping and labeling tend to focus on similarities between geometric properties of fibers. Ding *et al.* [3] combined a corresponding segment ratio with the mean distance over the segments to delineate the similarity between two streamlines generated from a region of interest (ROI), and grouped the geometrically similar streamlines with nearby seeding points into the bundle. Corouge *et al.* [4, 5] defined three distance measures based on the point sets on the pair of streamlines, and classified a set of streamlines

within specified ROI(s) into meaningful bundles. However, variations across subjects in white matter pathway organization can make extracting and labeling the fiber paths by hand time-consuming and error-prone. The task becomes more challenging when classifying the paths over the entire brain instead of a ROI since no human input about the path information will be implied. Zhang *et al.* [6] explored an unsupervised classifying algorithm on a dense set of paths over the entire brain, and used a serial of strategies in setting seeding points, path constraints, culling distance and setting distance metrics to minimize the misclassifications. Brun *et al* [7] compared the fiber path pairwise to create a weighted undirected graph that was partitioned into coherent sets using the normalized cut criterion.

Our approach starts by identification of commonalities in terminations of white fibers in gray matter. For this, we use a standard brain atlas which provides a parcellation scheme for brain gray matter [Talairach and Tournoux, 1986], which is used by most functional neuroimaging studies for assigning functional roles to individual brain regions in a common standard space. The rationale is that fibers that connect neighboring gray matter regions that are considered to form functional units are most likely to be part of the same white matter tract. At this stage of analysis, alternate pathways between the same regions may appear to form a single tract, and fibers whose gray matter terminations are nearby but labeled differently in the atlas might be assigned to different clusters. After this initial classification, subsequent steps eliminate unlikely fibers, determine subgroups of fibers based on pairwise comparisons of geometric properties and blend groups of fibers with similar trajectories. This approach also considerably reduces computational load, since not all possible comparisons between fibers need to be computed. Additional anatomical constraints related to global organization of brain structures, as well as functional neuroimaging results, can easily be incorporated into this procedure for fiber classification.

Another reason why it is desirable to use the spatial organization of gray matter as a reference point for identifying white matter pathways is that this might make easier to combine tractography data across subjects, and to make comparisons across group. Once pathways within an individual brain have been identified, there is also the problem of deciding which pathways across individuals correspond to the same anatomical tract so that quantitative comparisons can be made. Variability of spatial extents of pathways across individual brains is currently not well known. Correspondences in the relative positions of gray matter structures have been better characterized, and addressed by spatial normalization methods. More concretely, a tract connecting two cortical regions, such as dorsal frontal and posterior parietal cortices, might show considerable variability in its trajectory, so that spatial overlap of its extent across individuals might be poor. Choosing tracts based on the regions that they connect rather than the regions that they cover in space might be a better basis for comparing the properties of tracts across individuals.

2 Materials and Methods

2.1 Tracking Streamlines

The extraction of fiber paths for the entire brain is performed with the tractography tool described in [8]. The tracking algorithm used is based on the STT algorithm. In brief, tensor components are calculated from a voxel size of $1.85 \times 1.85 \times 1.85 \text{ mm}^3$

diffusion weighted images acquired along 13 gradient directions. Streamlines following the principal diffusion directions are produced from seed points along a regular 1.85 mm grid, constrained only to voxels with fractional anisotropy (FA) value exceeding 0.15, and following a local continuity constraints. The integration stops at points where FA value falls below 0.15, which indicate low diffusion directionality.

Figure 1 shows whole-brain streamline sets extracted by this method from a test dataset, and transformed to a common stereotactic space using the Montreal Neurological Institute (MNI) T2 template with SPM2 software. There are more than 46,700 paths which are longer than 20mm. Since the tracking results are not constrained by geometric criteria, a significant portion these 40K paths can be due to effect of noise and include multiple outliers.

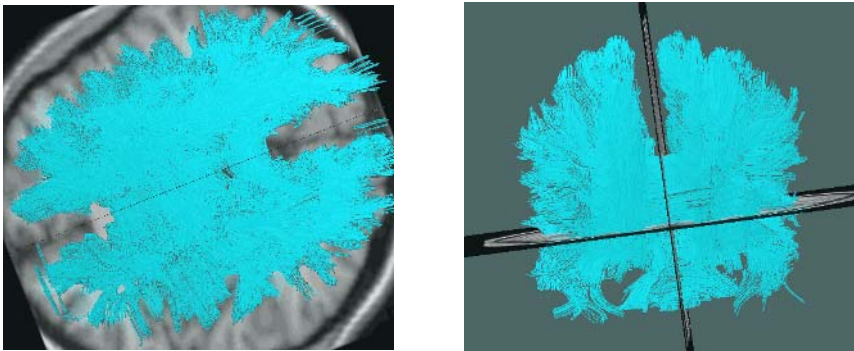


Fig. 1. Whole-brain streamline sets extracted by the method from a test dataset

2.2 Atlas-Based Labeling and Classification of the Fiber Bundles Based on Initial Labels

As a first step, Talairach brain atlas is used to provide labels for each streamline or path and for each fiber bundle separately on the basis of coordinates of two end-points. The labeling of each streamline is done as following:

- If both of two end-points of the streamline fall within regions labeled as gray matter in the atlas space, the streamline is labeled with the names of the regions that it connects (eg., “BA 7_Thalamus”).
- If only one end-point of the streamline falls into labeled gray matter, the second termination is not identified in the atlas as gray matter, the streamline is marked as the name of the gray matter and unidentified (eg., “BA7_unidentified”).
- If both end-points of the streamline are within regions not identified in the atlas as gray matter, it is marked as “unidentified_unidentified”.

Streamlines that can actually be interpreted as corresponding to axonal fibers should be connecting functional brain regions, and organized in groups (fiber bundles) that correspond to white matter tracts. After the first label assignment, the streamlines are classified into the groups of fiber bundles based on correspondence of atlas labels for their terminations. Figure 2 shows the fiber bundles that connect left and right

hemisphere and classified into more than 150 bundles in the step. Since the robustness of fiber tracking remains limited at junctions and in noisy regions, the extracted fiber set contains outlier paths, paths that are part of other anatomical tracts, and spurious paths linking between anatomically unrelated tracts. These paths should be removed or classified into the correct bundles during the initial classifying stage:

- All streamlines linking between identified anatomic or functional gray matter structures are grouped into the different fiber bundles according to their names. For instance, all streamlines labeled as “BA7_Thalamus” are incorporated into a single group.
- The streamlines named as ‘unidentified _unidentified’ are assigned to one of the set of the fiber bundles defined for the whole brain based on the similarity measures defined in the next section. Figure 3 shows that red color fiber bundle is the source bundle that is within left hemisphere regions not identified in atlas as gray matter, and incorporated into the white color fiber bundle based on the similarity.
- The streamlines labeled as “gray matter _unidentified” are incorporated into one of the set fiber bundles which share the identified gray matter label for this streamline as a termination based on the similarity.
- Streamlines that are shorter than 30 mm or that cannot be assigned to any fiber group based on the similarity measure are removed.

2.3 Classifying Fiber Bundles

The way the fibers bundles are clustered in 3D space, the degree to which fibers in the same bundle run parallel, and their distinct physical and geometric properties should match the properties that brain white matter tracts have in general to be biologically plausible. Low spatial resolution of DT images, limited robustness of fiber tracking methods at junctions and in noisy regions, and interindividual variability of white matter anatomy introduce errors into the fiber tracking procedure. Therefore, the fiber bundles generated by the steps above may contain biologically implausible outlier paths and paths forming spurious bridges between fiber bundles that correspond to actual pathways. The next classification stage is introduced to address these problems. The goals of classifying are removing the outliers, partitioning the fibers in the bundle into natural clusters, and agglomerating the clusters into correct bundles according to the position and shape similarity of fibers. Two measures of similarity between pairs of single fibers are defined below.

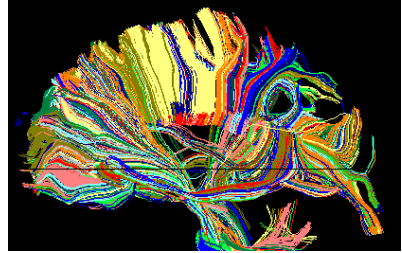


Fig. 2. Fibers connecting the left and right hemisphere are classified into the bundles viewed with different colors

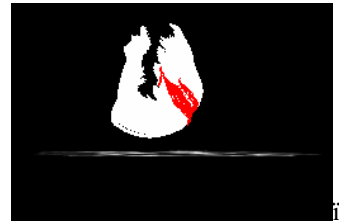


Fig. 3. Red-color fiber bundle is incorporated into the white-color fiber bundle

Similarity Measure. Two fibers are considered similar when they have comparable length, similar shape, and are separated by a short distance. Two pairwise distances between fibers F_i and F_j are used for the similarity measure:

1. Mean distance d_M :

$$d_M(F_i, F_j) = \min(d_m(F_i, F_j), d_m(F_j, F_i))$$

with $d_m(F_i, F_j) = \text{mean}_{p_l \in F_i} \min_{p_k \in F_j} \|p_k - p_l\|^c$ $\|\dots\|$ being point-by-point Euclidean distance \triangleright

2. Hausdorff distance d_H :

$$d_H(F_i, F_j) = \max(d_h(F_i, F_j), d_h(F_j, F_i))$$

with $d_h(F_i, F_j) = \max_{p_l \in F_i} \min_{p_k \in F_j} \|p_k - p_l\| \triangleright$

The mean distance d_M is a good indicator of shape similarity and spatial closeness of a pair of fibers. When two fibers that have comparable length are similar in shape and close in location, d_M is small; when the distance between two fibers is large, or their shapes are different, d_M is large. On the contrary, the Hausdorff distance d_H being a worst-case distance, it is a useful metric to reject outliers and prevent classifying paths with high dissimilarity.

Classifying Fiber Bundles. Classifying the fibers in the bundle involves finding subsets of fibers that correspond to natural anatomical groups by identifying fibers that are clustered closely together in 3D space, running parallel to each other, and have distinct physical and geometric properties. If a bundle formed in Section 2.2 contains more than one such group, it is split into natural sub-clusters. After splitting of bundles, similarities between neighboring sub-clusters are assessed in order to identify those that are likely to belong together in the same tract, so that they are regrouped into a new bundle, and to reject sub-clusters containing small numbers of fibers and showing low similarity with other groups as outliers that are likely to be due to artifacts.

Splitting approach begins with the calculation of pairwise d_M of the fibers in the same bundle. It is not a time-demanding task since the number of fibers in a bundle is not high generally. Supposed calculating the pairwise d_M between the fiber F_i and the fiber F_j in the bundle C , $j \neq i$. If $d_M(F_i, F_j) < t$ (a threshold to be chosen), F_i and F_j are in the same sub-cluster, and vice versa. After calculating a table of pairwise d_M , the approach propagates from neighboring fiber to neighboring fiber. The threshold t depends on the position and shape of the fiber bundle. The larger t value is, the lower the number of sub-classes will be, and vice versa.

Finding the neighbor candidates of the fiber bundle, and limiting the similarity measurement between sub-cluster fibers and the candidate-bundle fibers are very

important to searching approach. The pairwise d_M comparison of all fiber paths is a time-consuming task because the number of fiber paths in entire brain is huge. So the fiber bundle properties of anatomic position, direction, shape, length, end-points, the mass center, and connectivity are introduced for the finding of neighbor bundle candidates in order to limit the number of similarity measurements.

The following steps are involved to reject a sub-cluster or group it into the neighbor fiber bundle or keep it as a natural fiber bundle:

- 1) For a fiber F_i in the sub-cluster, the fiber candidates are found in the candidate bundles. These candidates and fiber F_i have similar in shape and direction, and the mass centers of them are the neighbors.
- 2) Similarity (mean distance d_M) between F_i and each of the fiber candidates are calculated, and a similarity threshold value is set.
- 3) One fiber in candidates is located, whose similarity is greatest (d_M is smallest).
- 4) If the similarity in step 3 is accepted, the sub-cluster is merged into the bundle with the found fiber in step 3 and the processing is stopped. If the similarity is unaccepted, another fiber F_j in the sub-cluster will be chosen and steps 1-4 are repeated.
- 5) When all fibers F in the sub-cluster are checked and the sub-cluster cannot be merged into other bundle, if the average fiber length of the sub-cluster is less than 30 mm and the number of fibers is less than 3, the sub-cluster is rejected forcibly. Otherwise, the sub-cluster is kept and waited for checking manually by an expert.

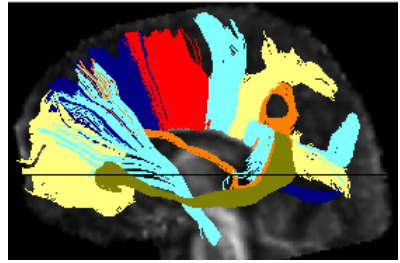


Fig. 4. The processed results of the fiber bundles that connect left and right hemisphere

Repeating the above steps for all sub-clusters and fiber bundles, the algorithm partitions the set of fibers into different fiber bundles. Figure 4 shows the processed results of the fiber bundles that connect left and right hemisphere, there are 14 fiber bundles left.

3 Results

The neuronal fiber classification and quantification DTI method was implemented in Matlab, and tested on the fiber paths from real DTI datasets. Figure 1 shows the streamlines with the length greater than 20 mm before classifying and the path set without the constraint. There are about 46,700 paths; some are short paths stopping within the ambiguous white matter regions and being the part of other paths, which could potentially create artificial links between fiber bundles. Figure 5A shows the

bundles formed after the first step, based on common gray matter projections determined by atlas labels. The final results, derived by redistribution and pruning of fibers based on similarity measurements are shown in Fig 5B.

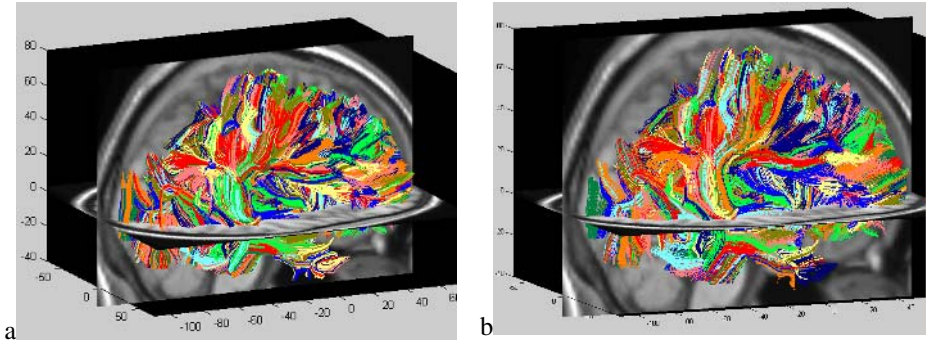


Fig. 5. (a) shows the bundles formed after the first step, there are more than 3000 bundles. (b) The final results, derived by redistribution and pruning of fibers based on similarity measurements, there are 264 bundles left.

4 Discussion and Conclusion

The focus of this work is on classification fiber paths derived from DTI data into bundles associated with distinct anatomical and functional structures. The fiber classification procedure is centered around knowledge of anatomical and functional subdivisions of gray matter obtained from the Talairach atlas. Fibers are first grouped based on common projections to parts of gray matter designated as distinct regions in the atlas. Subsequent stages refine the initial distribution of fiber groups, so that the final set of bundles contain fibers that share similar physical and geometrical properties, and connect gray anatomically and functionally related regions of gray matter. This approach reduces the computational load of whole brain fiber analysis, and automatically produces groups of fibers that can be interpreted in an intuitive manner.

The algorithm extracts whole-brain streamlines from naive dataset, transforms the streamlines to a common stereotactic space using the MNI T2 template with SPM2 software, and names the fiber and fiber bundle with the region names of two end-points of the fiber and fiber bundle. The transform accuracy and interindividual variability do influence the name of one fiber and may put the fiber in an incorrect bundle at initial stage, the algorithm re-classifies the fibers into natural anatomical bundles by identifying fibers that are clustered closely together in 3D space, running parallel to each other, and have distinct physical and geometric properties. If the difference between the individual brain and Atlas is large, some fiber bundles connecting small cortical areas may have an incorrect label names.

Low spatial resolution of DT images and limited robustness of fiber tracking methods at junctions and in noisy regions may issue tracts that are locally consistent but incorrectly connected. The algorithm cannot handle the case currently, and the

fiber bundles have to remove manually. Our future work will focus to validate the fiber tracts and to explore the relationship between the physical and geometric properties of the fiber tracts and the brain diseases and brain development.

References

1. Mori, S., Crain, B. J., Chacko, V.P., and van Zijl, P. C. M. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann. Neurol.* Vol 45, pp.265-269, 1999.
2. O'Sullivan, M., Jones, D.K., Summers, P.E., Morris, R.G., Williams, S.C.R., and Markus, A.S. 2001. Evidence for cortical disconnection as a mechanism of age-related cognitive decline. *Neurology* 57:632-638.
3. Zhaohua Ding, John C. Gore, and Adam W. Anderson. Classification and quantification of neuronal fiber pathways using diffusion tensor mri. *Magnetic Resonance in Medicine*, 49:716–721, 2003.
4. Isabelle Corouge, Sylvain Gouttard, and Guido Gerig. Towards a shape model of white matter fiber bundles using diffusion tensor MRI. In *International Symposium on Biomedical Imaging*, pages 344–347, 2004.
5. Guido Gerig, Sylvain Gouttard, and Isabelle Corouge. Analysis of brain white matter via fiber tract modeling. *EMBS* 2004.
6. Song Zhang, and David H. Laidlaw. DTI fiber clustering in the whole brain.
7. Anders Brun, Hans Knutsson, Hse-Jeong Park, Martha E. Shenton, and Carl-Fredrik Westin. Clustering fiber traces using normalized cuts. *Miccai* 2004, pp. 368-375.
8. David Akers, Anthony Sherbondy, Rachel Mackenzie, Robert Dougherty, and Brian Wandell. Exploration of the brain's white matter pathways with dynamic queries. In *Proceedings IEEE Visualization*, pp.77-384, 2004.
9. Dongrong Xu, Susumu Mori, Meiyappan Solaiyappan, Peter C.M. van Zijl, and Christos Davatzikos. A framework for callosal fiber distribution analysis. *NeuroImage* vol. 71, pp. 31-143, 2002.

A Physical Model for DT-MRI Based Connectivity Map Computation^{*}

Erdem Yörük¹, Burak Acar¹, and Roland Bammer²

¹ Department of Electrical&Electronics Engineering, Boğaziçi University, Turkey
erdem.yoruk, acarbu@boun.edu.tr

² LUCAS MRS/I Center, Department of Radiology, Stanford University, CA, USA

Abstract. In this study we address the problem of extracting a robust connectivity metric for brain white matter. We defined the connectivity problem as an energy minimization task, by associating the DT-field to a physical system composed of nodes and springs, with their constants defined as a function of local structure. Using a variational approach we formulated a fast and stable map evolution, which utilizes an anisotropic kernel smoothing scheme equivalent to a diffusion PDE. The proposed method provides connectivity maps that correlate with normal anatomy on real patient data.

1 Introduction

Magnetic Resonance Diffusion Tensor Imaging (DT-MRI) is the only known technique that provides information about the geometry of fibrous structures, such as the fibers in the human brain, in vivo. DT-MRI is based on measuring the diffusion of water molecules. The diffusion process is approximated by a second order tensor. It was shown that at fiber positions, the principal eigenvector of this tensor is correlated with the underlying fibrous structures' orientation [3]. So, DT-MRI has been accepted as a promising tool to get information about the connections in human brain. Despite its promise, DTI has its drawbacks: Low order approximation to the diffusion process and insufficient spatial resolution.

There are two basic approaches in utilizing the information DTI data provides: Fiber tractography and connectivity mapping. The former attempts to answer the following question: Can we reconstruct the fiber that passes through a given seed point? Among different approaches, the Runge-Kutta method has been widely accepted [1]. However, fiber tracking is prone to cumulative errors and can not overcome the partial volume effect [2, 3]. The second group of approaches attempts to utilize the nature of the DTI data, i.e. the diffusion process, by estimating a connectivity map. They consider each and every possible connection with weights set by the DTI data. Koch et al. used a Monte-Carlo simulation of the random walk model [4], yet their approach uses successive jumps whose curvature is restricted to be less than 90 degrees and requires multiple runs for

^{*} This work was in part supported by EU 6th Framework SIMILAR NoE (www.similar.cc), TÜBİTAK KARIYER-DRESS Project (104E035) and Boğaziçi University Research Funds (SK/DTIcad).

generating the final connectivity map. Hagmann et al. extended this method in a supervised fashion by incorporating the white matter fiber tract curvature [5], again treating the problem as a random process. Chung et al. applied adaptive kernel smoothing of the transition probabilities of the underlying diffusion process, which is governed by the DTI tensor field [6]. Their approach relatively being faster and more deterministic than the first two, is shown to be equivalent to the Monte Carlo method. It remains to be a numerical diffusion simulation, and needs to be masked with the FA map in order to be a heuristic approximation of connectivity. Batchelor et al. solved the anisotropic diffusion equation set by the DTI data with the initial conditions being a set of seed points set to an initial nonzero (unity) concentration of virtual diffusing particles [7]. Lenglet et al., on the other hand, recasted the connectivity problem to Riemannian differential geometry framework where they defined their local metric tensor using the DTI data and solved for geodesics [8].

In this work, we propose a physical setup of unit-mass nodes (voxels) and springs. The constants of spring linking adjacent nodes are set according to the local connectedness of the node pair. The local connectedness is measured using the so called *mutual diffusion coefficient*. It approximates the ellipsoidal volume overlap of a given tensor pair in \mathbb{R}^3 . The system is perturbed by setting the seed node to constant potential and computing the stationary potential map, which is the connectivity map, evoked. The method is evaluated on real patient data.

2 Method

In general, connectivity can be interpreted as a measure which is proportional to the qualitative similarity and spatial proximity of the units contained in data to be analyzed. In the case of DT images, where we are trying to reveal a functional connectivity, qualitative assessment of tensors becomes crucial in order to construct an appropriate model. We have to consider some certain features embedded in a DT image as indicators of connectivity. In general, the connectivity measure between two tensors can be visualized in terms of the volume overlap of their ellipsoids centered at the corresponding voxel coordinates. This interpretation will take the relative features of tensors into consideration such as their respective locations, sizes, orientations and shapes. A similar but simpler metric, that we proposed and used in this study is the so-called distance scaled mutual diffusion coefficient K . Given two tensors \mathbf{D}_1 and \mathbf{D}_2 , located at \mathbf{r}_1 and \mathbf{r}_2 , respectively, we define their connectivity K_{12} as:

$$K_{12} = \frac{[(\mathbf{v}^T \mathbf{D}_1 \mathbf{v})(\mathbf{v}^T \mathbf{D}_2 \mathbf{v})]^\gamma}{\delta^2} \quad (1)$$

where $\mathbf{v} = (\mathbf{r}_1 - \mathbf{r}_2)/\delta$ and $\delta = \|\mathbf{r}_1 - \mathbf{r}_2\|_2$. Thus, K reflects the mutual influence of tensor pairs by giving the distance scaled product of their diffusion coefficients evaluated in the unit direction of their Euclidean link and raised to the power γ , which will be used as a tuning parameter. For the time being we take γ to be 1. We can construct a physical system, based on K_{ij} 's which reflect the connectivity pattern within DTI data.

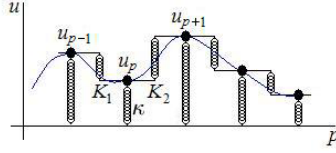


Fig. 1. Illustration of proposed spring system for a 1D curve (solid line), dots represent the nodes, spring constants are given for the specific node u_p

Let $\Omega = [0, a] \times [0, b] \times [0, c] \subseteq \mathbb{R}^3$ be our image domain, and let $\mathbf{D} : \Omega \rightarrow \mathbb{R}^3 \times \mathbb{R}^3$ be the given tensor field. We propose that the sought connectivity map $u : \Omega \rightarrow \mathbb{R}$ with respect to a given seed point $(x_0, y_0, z_0) \in \Omega$ is the stationary pattern of a physical spring system defined as follows: Each voxel of the domain corresponds to a node and potential at seed node $u(x_0, y_0, z_0)$ is set to 1, which is kept constant in time. For each pair of adjacent nodes, a spring is associated with a stiffness, that is set to the mutual diffusion coefficient (Eqn. 1) evaluated for the adjacency of that particular pair. Thus, if we consider 6-neighborhood (N_6) on a regular 3D grid with δ -spacing in all directions, this model will correspond to 6 "neighbor springs" acting on each node and a node at $\mathbf{r}_i = (x, y, z)^T$ would share the following spring constant with respect to its immediate neighbor $\mathbf{r}_j = (x - \delta, y, z)^T$:

$$K_1(x, y, z) = \frac{(\mathbf{e}_1^T \mathbf{D}(x, y, z) \mathbf{e}_1)(\mathbf{e}_1^T \mathbf{D}(x - \delta, y, z) \mathbf{e}_1)}{\delta^2} \quad (2)$$

where $\mathbf{e}_1 = (1, 0, 0)^T / \|\mathbf{r}_i - \mathbf{r}_j\|$ is the unit vector of the orthonormal basis. Remaining constants of neighbor springs K_2, \dots, K_6 are defined similarly with $\mathbf{e}_2 = (0, 1, 0)^T$ and $\mathbf{e}_3 = (0, 0, 1)^T$. We set spring elongations as the local directional differences of u (for instance as $u(x, y, z) - u(x - \delta, y, z)$ for K_1), such that they force u to be equal to the corresponding immediate neighbor. Finally, an extra spring, the so-called "ground spring" with a constant stiffness κ , is also attached to all nodes, with the elongation being u and forcing u to be 0. This is necessary to justify our model for revealing connectivity, which will be explained next. The resulting spring system can be visualized in Figure 1 for a 1D curve. With this model, total potential energy stored will be given as:

$$V_{springs} = \frac{1}{2} \kappa u^2 + \sum_{n=1}^6 V_n \quad (3)$$

where $V_1 = \frac{1}{2} K_1 [u(x, y, z) - u(x - \delta, y, z)]^2$ and V_2, \dots, V_6 are defined similarly. In order to associate this potential energy to the point (x, y, z) rather than to the springs, the latter six energy terms due to the neighbor connections should further be halved, as if they are equally shared with the corresponding neighbor. We let $\delta \rightarrow 0$ assuming that we have infinite resolution, such that V can be approximated to a continuous function, enabling us to interpret the system better. Inserting K_n 's common denominator δ^2 and taking the limit, we obtain:

$$V(u, u_x, u_y, u_z) = \frac{1}{2}(\kappa u^2 + d_{11}^2 u_x^2 + d_{22}^2 u_y^2 + d_{33}^2 u_z^2) \tag{4}$$

where d_{jj} ($j = 1, 2, 3$) are the diagonal elements of \mathbf{D} as can easily be derived from Equation 2. We can put the current problem into a variational one, where the connectivity map u with respect to the seed at (x_0, y_0, z_0) can be found by minimizing the following energy functional:

$$J(u) = \int_{\Omega} V(u, u_x, u_y, u_z) d\Omega \tag{5}$$

assuming Neumann boundary conditions on $\partial\Omega$, and an extra seed condition $u(x_0, y_0, z_0) = 1$. Corresponding Euler-Lagrange equation is:

$$\kappa u - \frac{d}{dx}(d_{11}^2 u_x) - \frac{d}{dy}(d_{22}^2 u_y) - \frac{d}{dz}(d_{33}^2 u_z) = 0 \tag{6}$$

It is easy to show that fixing $u(x_0, y_0, z_0)$ to 1 for an interior point of Ω , does not change the Euler-Lagrange equation since $J(u)$ can be rewritten as a sum of several integrals evaluated at Cartesian subregions, each having (x_0, y_0, z_0) at their boundaries and sharing the same integrand. The physical interpretation of our model is as follows: A connected spring system, which is only lifted at (x_0, y_0, z_0) to a constant level of 1 and kept there, achieves its stationary pattern by minimizing its total potential energy. Nodes other than the seed try to come to zero due to the ground springs, but they will be lifted up as well, in proportion to their connectivity to the stationary seed node. Hence, the stationary pattern will be equivalent to the sought connectivity map with respect to the seed node.

Temporally evolving a given initial u_0 at a rate determined by the negative of the left term of Equation 6, by a steepest descent scheme, we obtain the following PDE:

$$u_t = -\kappa u + \frac{d}{dx}(d_{11}^2 u_x) + \frac{d}{dy}(d_{22}^2 u_y) + \frac{d}{dz}(d_{33}^2 u_z) = -\kappa u + \nabla \bullet (\tilde{\mathbf{D}} \nabla u) \tag{7}$$

where $\tilde{\mathbf{D}}$ is a diagonal matrix with squared diagonal entries of \mathbf{D} . The same equation could be obtained with another $\tilde{\mathbf{D}}$, using different neighborhood (and/or spring) definitions. In any case, resulting PDE can be considered as a modified diffusion process, with a seed condition $u(x_0, y_0, z_0) = 1$ and an extra term $-\kappa u$, keeping u attached to the ground. Thus, as $t \rightarrow \infty$, u does not get totally flat. If we interpret the model in 2D, a structure sensitive tent-like pattern will be obtained with its maximum at the seed.

Numerical implementation of such a map evolution scheme can easily be accomplished by anisotropic diffusion filters with the specified modifications and an initial map:

$$u_0(x, y, z) = \begin{cases} 1 & \text{if } (x, y, z) = (x_0, y_0, z_0) \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Keeping the model in its discrete form from the beginning, we can get the discrete version of Equation 5 as:

$$J(u) = \sum_p \{ \kappa u_p^2 + \frac{1}{2} \sum_{n=1}^N K_{pn} (u_p - u_{pn})^2 \} \quad (9)$$

where p is the voxel index and N is the number of neighbors. u_{pn} and K_{pn} stand for the n^{th} neighbor of the p^{th} node and its associated spring constant, respectively. $\frac{1}{2}$ scales the potential energy of neighbor springs, since they are counted twice in the outer summation. Now, writing the discrete Euler-Lagrange equation of this functional, we obtain:

$$\kappa u_p + \sum_n^N K_{pn} (u_p - u_{pn}) = 0 \quad (10)$$

This equation says that along an extremal, the total spring force applied to each node is zero. This could also be observed by rewriting Equation 6 in a discrete form where space derivatives are replaced with central differences.

Seeking for the connectivity as a PDE evolution with diffusion filters, will bring the common trade-off between stability and rate of convergence. We propose, another map evolution technique, which mimics the diffusion process in the light of the balanced force condition imposed by discrete Euler-Lagrange equation. The idea is, not to calculate the update u_t , but iteratively solve for a new u_p , which makes u_t vanish at p . "Balance in the neighborhood" condition dictated by Equation 10, locally and explicitly gives:

$$u_p = \frac{\sum_n^N K_{pn} u_{pn}}{\kappa + \sum_n^N K_{pn}} \quad (11)$$

This scheme with $\kappa = 0$ would be the same as assigning a weighted neighborhood average to each u_p , but with a spatially varying kernel coming from the local image structure. In fact, this adaptive directional smoothing constitutes the essence of the diffusion filtering, hence it can be interpreted to be equivalent to the diffusion process [6]. Equation 11 has a similar computational cost to Equation 7, but a higher convergence rate. Moreover, stability is naturally provided by the bounds specified by local neighborhood of u_{pn} 's.

3 Results

In order to validate the computed connectivity maps, we used a simple tracking module, that follows coherence directions in converged u . As in [9], we first compute the structure tensor of u , which is obtained by taking the tensor product of its gradient and componentwise smoothing the resulting matrix:

$$\mathbf{S}(u) = G_\sigma * (\nabla u \otimes \nabla u) = [\omega_1 \ \omega_2 \ \omega_3] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} \omega_1^T \\ \omega_2^T \\ \omega_3^T \end{bmatrix} \quad (12)$$

where G_σ is a Gaussian kernel with variance σ^2 . $\mathbf{S}(u)$ is a symmetric positive semidefinite matrix, which is invariant under sign changes of the gradient. Its magnitude ordered eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ give the amount of u 's fluctuation, along the corresponding eigenvectors. Thus, as ω_1 will be parallel to the smoothed gradient, the smallest eigenvector ω_3 , which is also called the coherent direction, will correspond to the orientation with least decrease in the connectivity. Starting from the seed, we can easily follow ω_3 vectors for tracking the candidate fiber.

We used a real patient data in our experiments. The scans were single-shot EPI scans with diffusion encoding along 12 non-collinear directions plus one reference without diffusion-weighting. The FOV was 25-26cm, TE was minimum with partial k-space acquisition. TR was $\sim 10s$ and b-value was $\sim 850s/mm^2$.

Experiments are comparatively conducted based on 26 neighborhood for different seed points. Note that, with N_{26} , spring constants for corners of a $3 \times 3 \times 3$

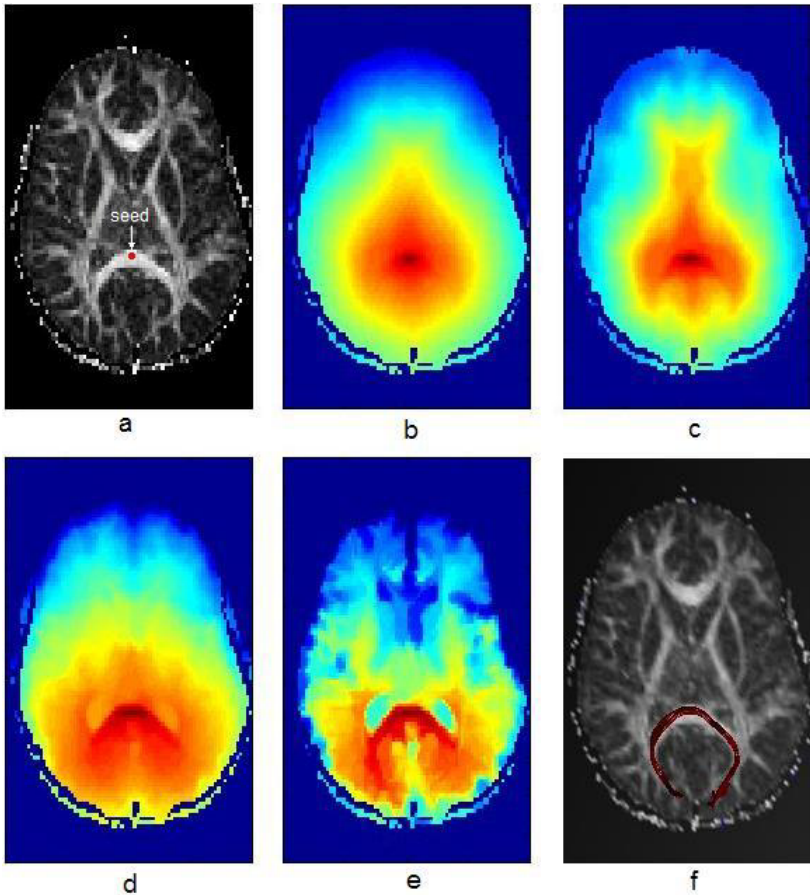


Fig. 2. a) Axial slice of FA map with the seed indicated, b) Converged connectivity map ($\gamma = 1$), c-e) Converged maps for $\gamma = 2, 4, 10$, f) computed tracts for $\gamma = 10$

neighborhood cube, appear with weights $\frac{1}{2}$ and $\frac{1}{3}$ due to the denominator δ^2 . Stiffness for the ground springs is empirically taken to be $\frac{1}{100}$ of the mean neighbor stiffness of the whole domain and converged maps are visualized in the log-scale. Convergence of the iterations was assumed when the mean absolute change in u dropped below 10^{-4} with $u \in [0, 1]$. Qualitative evaluation of the results were done based on the computed connectivity maps and the tracts computed from u maps.

Figure 2 shows axial cross-sections of the computed 3D connectivity maps with different γ values, ranging from 2 to 10 and the computed tracts. Note the correlation between the connectivity map and the brain anatomy as well as the assigned connectivity values with respect to the seed point. The anatomy is enhanced as the γ increases. The tracts in Figure 2f were computed based on the coherence directions of u as defined in [9]. They also correlate with the underlying fiber structure.

The overall computation time is independent of the number of seed points. It took less than 100 seconds with Matlab R14 using a PC with P4 2.4 GHz processor and 2GB RAM to process a $128 \times 128 \times 38$ volume.

4 Discussion

In this study we proposed a novel method for computing the connectivity in white matter. We associated the DT-image to a physical spring system, and interpreted the underlying connectivity to be its stationary pattern. We followed a variational approach, which aims to minimize the total spring potential energy and showed that the PDE evolution scheme coming from the Euler-Lagrange equation corresponds to a modified diffusion process. Reformulating the energy minimization problem in the discrete domain enabled us to replace the PDE with an alternative directional averaging method, which guarantees stability and exhibits a higher convergence rate.

The proposed method is computationally inexpensive and provides several advantages. The model is valid for different definitions of neighborhood and spring constants. It provides a tuning parameter γ that allows one to continuously trade-off between using the complete DTI data with using the principal diffusion direction only. At one end, our method corresponds to the tractography approach (large γ) and at the other hand it corresponds to complete connectivity model ($\gamma = 1$).

The proposed method allows to set multiple seed points, even regions. Thus, one can easily incorporate a-priori information about the anatomy into the connectivity map computation. Potentially, one can search for the most probable path between two given points by simply setting them as the seed points, computing the associated connectivity map and running a tractography algorithm on this map.

As mentioned above ground springs are necessary for preventing total flattening of u , such that we obtain meaningful maps that monotonically decrease as we get away from the seed. The stiffness κ of the ground springs does not

make much difference as long as it is nonzero. κ should be sufficiently small, such that the farthest point of the brain can also attain a nonzero connectivity value and should be large enough so that the converged map does not become nearly flat. We empirically set it to be $\frac{1}{100}$ of the mean of neighbor spring constants.

Further research is required to evaluate the method.

References

1. Tench, C.R., Morgan, P.S., Wilson, M., Blumhardt, L.D.: White Matter Mapping Using Diffusion Tensor MRI. *Magn. Res. Med.* 47 (2002) 967-972
2. Alexander, A.L., Hasan, H.M., Lazar, M., Tsuruda, J.S., Parker, D.L.: Analysis of Partial Volume Effects in Diffusion-Tensor MRI. *Magn. Res. Med.* 45 (2001) 770-780.
3. Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A.: In vivo tractography using DT-MRI data. *Magn. Res. Med.* 44 (2000) 625-632
4. Koch, M.A., Norris, D.G., and Hund-Georgiadis, M.: An Investigation of Functional and Anatomical Connectivity using Magnetic Resonance Imaging. *NeuroImage* 16 (2002) 241-250.
5. Hagmann, P., Thiran, J.-P., Vandergheynst, P., Clarke, S., Meuli, R.: Statistical Fiber Tracking on DT-MRI Data as a Potential Tool for Morphological Brain Studies. *ISMRM Workshop on Diffusion MRI: Biophysical Issues* (2000).
6. Chung, M.K., Lazar, M., Alexander, A.L., Lu, Y., Davidson, R.: Probabilistic Connectivity Measure in Diffusion Tensor Imaging via Anisotropic Kernel Smoothing. University of Wisconsin, Dept. of Statistics (2003) Technical report no: 1081
7. Batchelor, P.G., Hill, D.L.G., Atkinson, D., Calamante, F., Atkinson, D.: Study of Connectivity in the Brain using the Full Diffusion Tensor from MRI. *IPMI 2001, Lecture Notes in Computer Science* 2082 (2001) 121-133.
8. Lenglet, C., Deriche, R., Faugeras, O.: Diffusion Tensor Magnetic Resonance Imaging: Brain Connectivity Mapping. *INRIA* (2003) Technical report no: 4983
9. Weickert, J.: Coherence-Enhancing Diffusion Filtering. *International Journal of Computer Vision* 31(2/3) (1999) 111-127

A Novel 3D Partitioned Active Shape Model for Segmentation of Brain MR Images

Zheen Zhao^{1,2}, Stephen R. Aylward¹, and Eam Khwang Teoh²

¹ Computer Aided Display and Diagnosis Laboratory,
University of North Carolina at Chapel Hill, Chapel Hill, NC
kurtzhao@unc.edu, aylward@unc.edu <http://caddlab.rad.unc.edu>

² School of Electrical & Electronics Engineering,
Nanyang Technological University, Singapore
ekteoh@ntu.edu.sg

Abstract. A 3D Partitioned Active Shape Model (PASM) is proposed in this paper to address the problems of the 3D Active Shape Models (ASM). When training sets are small. It is usually the case in 3D segmentation, 3D ASMs tend to be restrictive. This is because the allowable region spanned by relatively few eigenvectors cannot capture the full range of shape variability. The 3D PASM overcomes this limitation by using a partitioned representation of the ASM. Given a Point Distribution Model (PDM), the mean mesh is partitioned into a group of small tiles. In order to constrain deformation of tiles, the statistical priors of tiles are estimated by applying Principal Component Analysis to each tile. To avoid the inconsistency of shapes between tiles, training samples are projected as curves in one hyperspace instead of point clouds in several hyperspaces. The deformed points are then fitted into the allowable region of the model by using a curve alignment scheme. The experiments on 3D human brain MRIs show that when the numbers of the training samples are limited, the 3D PASM significantly improve the segmentation results as compared to 3D ASMs and 3D Hierarchical ASMs.

1 Introduction

Object segmentation is a prerequisite of many medical image analysis applications. The 2D Active Shape Model (ASM) [1] provide a promise to improve the robustness of the model to local minimum by restricting the deformation within the allowable region.

However, in 3D segmentation, 3D ASM often constrains itself from catching details during deformations. This is because the number of eigenvectors/eigenmodes cannot exceed the number of training samples, in the case of the dimension of the model being large, which is usually the case in 3D segmentation, as the dimension of the model is typically two or three orders of magnitude higher than the number of training samples. It is difficult to estimate a high-dimensional probability distribution from a small training set. Therefore, the allowable region spanned by the relatively few eigenvectors limits the deformation of ASM to catch details. A solution is by using large training sets. However, it requires

manually segmenting 3D images slice by slice, which is very laborious. Therefore it is inconvenient to build a large training set.

There have been various attempts to address this limitation. ASM combined with elastic models has been proposed in [1], but the detailed deformations are regulated by elastic forces, which do not reflect true shape variability. In [2], Davatzikos *et al.* proposed 2D Hierarchical Active Shape Model (HASM), a hierarchical scheme based on 2D curve partition and 2D curve wavelet decomposition to keep details. Their experiments have shown promising results in this “short of samples ” scenario.

However, in [2], the spatially partitioned representations of objects, both curve segmentation and wavelet decomposition, introduce shape inconsistency. Because in [2], the allowable regions/plausible areas of bands are independent to each other, illegal shapes could be tolerated during model fitting. As a result, this model could be vulnerable to noises and low contrasts. The authors use a hierarchical representation and a deformation scheme to subdue this problem. However the approaches only partially solve this problem.

Another issue of HASM in [2] is that the distribution is assumed as Gaussian distribution. In 3D applications, this assumption is debatable because the number of training samples is usually so limited that the distribution is sparse. For non-Gaussian distributions, Cootes uses a mixture of Gaussians approximation to a kernel density estimate of the distribution in [1]. But for partitioned representations, sole non-linear estimation is not adequate, because the estimations are still independent between tiles/bands.

In this work, a 3D Partitioned Active Shape Model, which uses curve alignment to fit models during deformations, is proposed. In 3D PASM, each training sample and deformed model is represented as a curve instead of a point as ASM and HASM. After the model has deformed to find the best match in the image, the deformed model is projected back. The curve representing the deformed model is aligned with the closest training sample by undergoing an affine transformation. Furthermore, to investigate the strength and weakness of 3D PASM, the 2D HASM in [2] is extended to 3D HASM in this work based on the mesh partitioning algorithm proposed in this paper.

2 A 3D Partitioned Active Shape Model

A 3D Partitioned Active Shape Model is described in this section. It is assumed the vertices of manual segmentations have been corresponded to construct a Point Distribution Model (PDM). It means that N training samples are available as sets of K corresponding landmarks in the 3D space.

2.1 Construction of Allowable Regions

In conventional ASMs, the training samples are aligned by using Procrustes alignment. The vectors X_n , $n = 1, \dots, N$ are then formed by concatenating the coordinates of K landmark points of N samples. Therefore, the dimension of X

is $3K \times N$. To reduce the dimension of X , Principal Component Analysis (PCA) is applied to X . The eigenvectors e_0, \dots, e_{N-1} , which are corresponding to the nonzero eigenvalues of the covariance matrix of X , are calculated. In the 3D case, it is typical that $N \ll 3K$. Therefore it is likely that S , spanned by e_0, \dots, e_{N-1} , cannot include the full range of shape variation. As a result, 3D ASM, which projects the tentative shape by using S , tends to reconstruct a shape without fine details.

3D PASM solves this problem by a partitioned representation of 3D ASMs. Because a 3D ASM is represented as a mean mesh and its eigen variations of individual tiles, the partitioned representation of the ASM is the partitioned mean mesh and eigen variations of tiles. Mesh partition means that a mesh is partitioned into a group of surface patches, which are called tiles in this study.

If only the statistical prior of a small tile, which comprises of K' vertices, of the whole mesh is required, N samples will be adequate to include the variation of this tile, as long as the K' is small enough. Thus fine details of this tile could be captured in the allowable region.

Therefore, the mesh is partitioned into tiles so that the tiles cover all faces of the mesh and each tile consists of roughly K' vertices. PCA is then independently applied to each tile to form a statistical prior for each tile.

2.2 Mesh Partitioning

In this section, an algorithm is described to partition a mesh M into a group of tiles. A result of applying this algorithm is shown in Fig. 1, where the tiles are shown in red and the boundaries are shown in blue. This algorithm segments mesh M into tiles τ_1, \dots, τ_s which cover all faces of M , given a set of sites positioned at the centroids of the *site faces* $S = f_1, \dots, f_s$. A face of M is randomly selected as the initial site. Once the tile associated with the site stops growing, another face of M from the faces not covered by grown tiles is selected as the next site. The procedure is repeated until all faces of M are covered by tiles. In other words, the algorithm stops when all faces of M are included in tiles.

A tile τ_i is a collection of faces for which the closest site face is f_i . The measure of distance between faces is an approximation of the geodesic distance over the mesh. It is defined by constructing a dual directed graph of the mesh M , *i.e.* the nodes of the graph correspond to faces of M . The edges of the graph

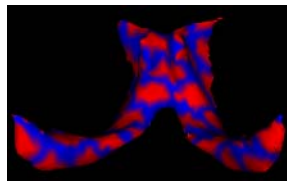


Fig. 1. A partitioned mesh. Tiles are shown in red; boundaries are shown in blue. The blue boundaries are enlarged for visualization.

connect nodes of adjacent faces. The cost of edges in this directed graph is set to the distance between centroids of the corresponding faces. This distance is defined as the length of the path in this directed graph.

Constructing a tile is a single-source shortest path problem in the graph, which is solved by a variant of Dijkstra's algorithm. The algorithm grows a tile until the size of the tile reaches the maximum. The size of a tile is defined as the number of vertices included in the faces of the tile. The maximum is set by $K'_{max} = \sqrt{K} \times 1.2$, where the K is the number of vertices of mesh M . The K'_{max} is set so that the number of tiles is not much bigger than the sizes of individual tiles. More partitioned objects are available at www.unc.edu/~kurtzhao.

2.3 Model Fitting Scheme

In ASM, after applying PCA, each training sample is represented as one point in a hyperspace. The training set constructs a cloud in this space. The allowable region can be estimated from the distribution of this cloud of points/samples.

In PASM, PCA is independently applied to the coordinates of vertices of each tile. Each training sample is then represented as points in H hyperspaces, one point in each hyperspace. In order to represent samples in one hyperspace, the indices of tiles are introduced as another dimension, the H hyperspaces are then combined to one hyperspace. As a result, each training sample is represented as a curve in the new hyperspace instead of points in many hyperspaces. The curves are called sample curves. During deformation, the model points move to find the best match in a test image. The deformed model is then projected back to this hyperspace as a curve called the model curve. The closest sample curve, in terms of Euclidean distance, is chosen as the target. The model curve is then aligned with the target by undergoing an affine transformation. The affine invariant alignment can be estimated by using the Least Square method in [3]. The affine transformation is chosen because it strike a balance between keeping the shape of the model plausible and allowing the model approaching the target.

The model fitting scheme of the 3D PASM is shown in the left figure in Fig. 2, where solid lines indicate models, the broken lines stand for training samples. Assume that the model/mesh is partitioned into 3 tiles and the dimension of the 3 hyperspaces is 1D after individual PCA applications. The indices of the tiles are introduced as another dimension. The two training samples are then represented by two curves. To fit the deformed model points, the curve representing the deformed model is aligned with the curve representing training sample 1 by using an affine transformation, which is the closest sample to the model curve. As a result, the shape of the fitted model cannot be far from the closest sample. The inconsistency between tiles is avoided, because the relationship between the shapes of tiles from the samples are enforced. On the contrary, the model fitting of the traditional HASM [2] and the 3D HASM shown in the right figure in Fig. 2 treats training samples as individual points in individual hyperspaces of the tiles. The deformed model points are fitted to the model by a truncation. It could tolerate the inconsistency of shapes because the shape of the fitted model could be different from any training sample.

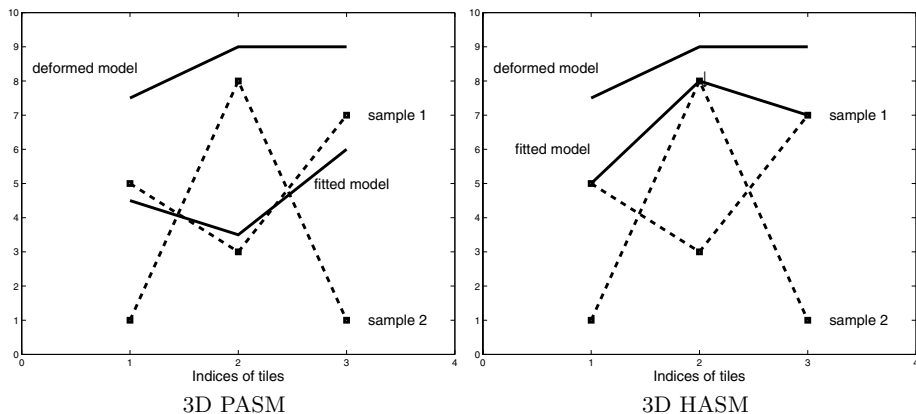


Fig. 2. The model fitting schemes of the 3D PASM and the 3D HASM

There are two forms of 2D HASM. One uses curve segmenting and the other one use wavelet decomposition. The 3D extension of the 2D HASM uses the mesh partitioning instead of wavelet decomposition because the wavelet based decomposition of a mesh is difficult to implement and unreliable.

3 Experimental Results and Discussion

In this section, the experimental results and quantitative analysis of the performance of 3D ASM, 3D PASM and 3D HASM are presented. The 2D HASM in [2] is extended to 3D HASM based on mesh partitioning in this study to compare with 3D PASM. The 3D HASM uses overlapping tiles and statistics in coarser scales to avoid the inconsistency of shape between tiles.

3.1 Image Data

The training sets and images used in this work are from the Internet Brain Segmentation Repository (IBSR) [4], which is an open online database. There are two sets of human brain MR images and their manual segmentations used in this work. The first set is IBSR v1.0. They are 20 normal T1-weighted MR brain 3D images and their manual segmentations. The second set is IBSR v2.0. They are 18 normal T1-weighted MR brain 3D images and their manual segmentations. The objects to segment are Lateral Ventricles extracted from IBSR v1.0, Lateral Ventricles extracted from IBSR v2.0, Left Thalamus Propers extracted from IBSR v2.0 and Left Hippocampuses extracted from IBSR v2.0. For simplicity, they are called LV1, LV2, LTP and LH in this study, respectively.

These three objects and two sets of images are chosen because of their diversity. Images of IBSR v1.0 are noisy. Images of IBSR v2.0 have higher quality and are free from noise. The contrast of the region around lateral ventricles is

sharp. On the contrary, the contrast of the regions around thalamus proper and hippocampus is relatively vague.

3.2 A Comparative Study of Segmentation Using Various Models

Standard 3D ASMs, 3D HASMs and 3D PASM are applied to segment LV1, LV2, LTP and LH from test images. For comparison purpose, the initialization and parameters of the three models are exactly the same. The “leave-one-out” method is used in this experiment. The “segmentation error” is used to measure the accuracy of the segmentations. The error is defined as the distances between the segmentation results and the corresponding manual segmentations. This metric is defined in [5]. In short, it is the shortest distance between two normalized meshes. If a segmentation error is bigger than 1/3 of the initialization error, this segmentation is considered as a “failure” in this study.

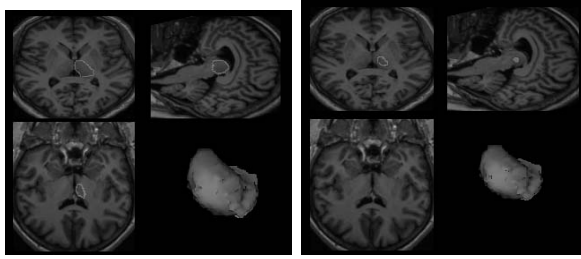
Table 1. The average segmentation errors (in voxels)

	LV1	LTP	LH	LV2
3D ASM	2.6364	4.6983	9.1490	1.5470
3D HASM	1.6282	2.6014	4.4104	1.5609
3D PASM	1.2469	1.1717	4.2943	1.5104

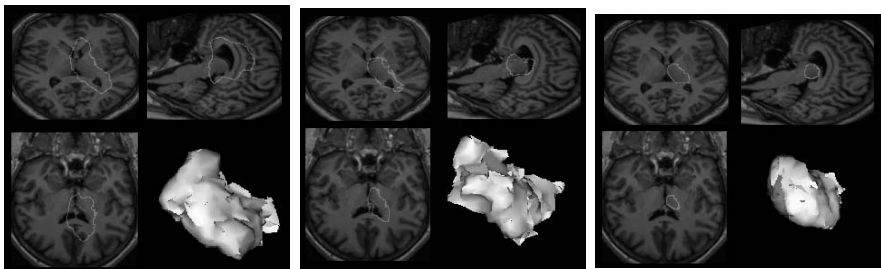
One example of the manual segmentation, initialization and the segmentation results of models are illustrated in Fig. 3. The segmentation by models are overlaid with cross-sections (in gray contours) and the 3D surface of manual segmentations, where the dark surface is the segmentation by the models and the bright one is the manual segmentation. The quantitative results are shown in Fig. 4. Their average is listed in Table 1. More results are available at www.unc.edu/~kurtzhao.

In several cases, the standard 3D ASMs failed to find the desired boundaries, because 19 or 17 samples are not sufficient for ASMs to accurately estimate the distribution of samples. During iterations, the rigidity of ASMs tends to keep model points from moving to approximate appropriate shape details. The accumulation of errors in iterations could lead to segmentation failures. Furthermore, model points have difficulty to find their best match in noisy images. As a result, during iterations, when a few points find their best matches, the shape of the model cannot change accordingly to guide other points, because 3D ASMs are too restrictive.

3D HASMs have partitioned representation, so they do not have this “rigidity” problem. But they still failed to accurately segment in a few cases because they tolerate illegal deformations, *i.e.* their allowable regions are not plausible. On the contrary, the 3D PASM improved the segmentation accuracy significantly because of the partitioned representation and the model fitting scheme using curve alignment. The improvements by 3D PASM on LV2 are not as significant as other objects because of the images of LV2 are very clear and the



(a) Manual Segmentation (b) Initialization: 5.5668



(c) ASM: 10.2965 (d) 3D HASM: 5.4502 (e) 3D PASM:1.1381

Fig. 3. The comparative study of the accuracy of segmentations for LTP by using different models

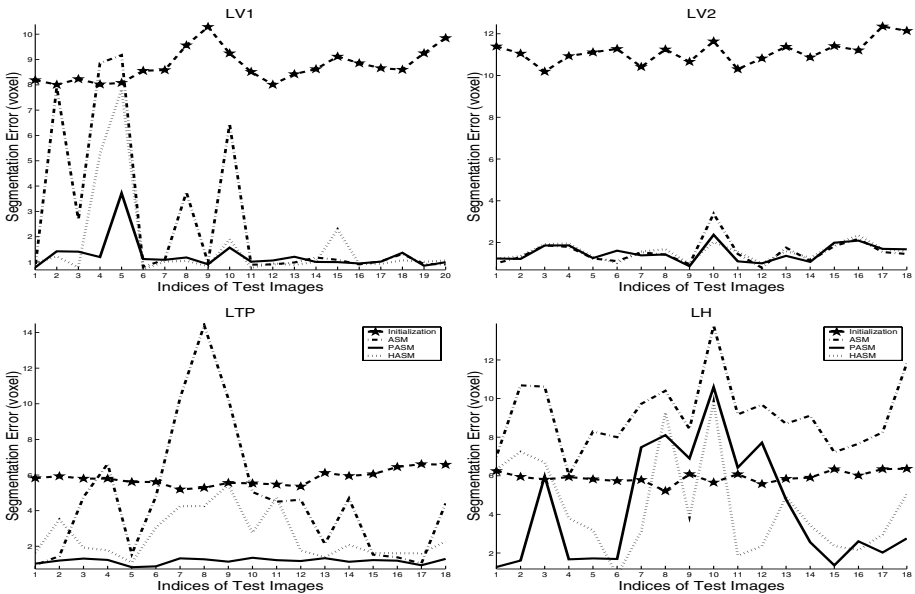


Fig. 4. The quantitative results of segmentations using different models

lateral ventricles have strong contrasts. However, if the images are noisy or the contrasts are low, 3D PASM brings significant improvements over 3D ASM and 3D HASM.

4 Conclusion and Discussion

A 3D Partitioned Active Shape Model is proposed in this work to solve the “short of training samples” problem. The experimental results show that 3D PASM brought significant improvement on the accuracy of segmentation over ASM and HASM. The PASM differs from HASM in how they avoid the inconsistency/discontinuities of shape caused by the mesh partitioning. The HASM uses overlapping tiles, as well statistics in coarser scales. The PASM uses a more systematic approach: considering the statistics of all tiles simultaneously. As the results show, the PASM is more robust to noisy images and low contrasts. Currently, the tiles are constructed so that they have similar numbers of faces. However, it is worth to investigate to take into consideration the shape variation of tiles. There are two extreme conditions: 1) a mesh consisting of N faces is partitioned into N tiles; 2) a mesh is partitioned into one tile only. In these two extreme, the PASM is not expected to improve the performance over ASM. However, K' , the number of tiles is set to avoid these two extremes.

References

1. Cootes, T.F., Taylor, C.J. In: Statistical Models of Appearance for Computer Vision. Technical Report, University of Manchester (2001)
2. Davatzikos, C., Tao, X., Shen, D.: Hierarchical active shape models, using the wavelet transform. *IEEE Transactions on Medical Imaging* **22(3)** (March 2003) 414–423
3. Ip, H.H.S., Shen, D.: An affine-invariant active contour model (ai-snake) for model-based segmentation. *Image and Vision Computing* **16 No.2** (1998) 125–146
4. available at <http://neuro-www.mgh.harvard.edu/cma/ibsr>: (The internet brain segmentation repository (ibsr))
5. Zhao, Z., Teoh, E.K.: A novel framework for automated 3D PDM construction using deformable models. In: Proc. Medical Imaging 2005: Image Processing, Proceedings of SPIE 2005 (SPIE MI 05) (in press). (2005)

Cross Entropy: A New Solver for Markov Random Field Modeling and Applications to Medical Image Segmentation

Jue Wu and Albert C.S. Chung

Bioengineering Program and Computer Science Department,
Lo Kwee-Seong Medical Image Analysis Laboratory,
The Hong Kong University of Science and Technology, Hong Kong

Abstract. This paper introduces a novel solver, namely cross entropy (CE), into the MRF theory for medical image segmentation. The solver, which is based on the theory of rare event simulation, is general and stochastic. Unlike some popular optimization methods such as belief propagation and graph cuts, CE makes no assumption on the form of objective functions and thus can be applied to any type of MRF models. Furthermore, it achieves higher performance of finding more global optima because of its stochastic property. In addition, it is more efficient than other stochastic methods like simulated annealing. We tested the new solver in 4 series of segmentation experiments on synthetic and clinical, vascular and cerebral images. The experiments show that CE can give more accurate segmentation results.

1 Introduction

As the core concept of Bayesian image analysis, Markov random field (MRF) theory has aroused great interest in the field of medical image processing. Researchers have applied MRF to various applications such as image enhancement [1], object detection [2], data modeling [3], tissue classification (segmentation) [4], etc. The reasons why MRF modeling has so many successful applications in medical image analysis are that it can easily incorporate spatial interaction and convert a problem in image processing or computer vision into a mathematical optimization problem by means of maximum *a posteriori* (MAP). The former property allows us to consider contextual constraints in images and the latter makes many delicate problems computationally tractable.

The most important part of the MRF modeling is the problem formulation in which we specify the posterior probability to be maximized or the energy function to be minimized. However, it remains a major open problem in MRF theory to optimize the objective function. Due to the large number of pixels in usual images, the configuration space of MRF in image analysis is huge. This makes the brute force methods to search for the optima infeasible in practice. Actually, it is proved that obtaining the global optimum of an arbitrary objective function is NP-hard [5]. Therefore, it has been an active research topic to design a "good" solver for MRF models over the past two decades. The goodness of a

solver lies in whether it can efficiently find a local optimum which is as optimal as possible (e.g., the lower energy, the better).

One of the earliest efforts to optimize MRF objective functions was made by Kirkpatrick *et al.* [6] who proposed the solver, simulated annealing (SA). SA can guarantee to converge to a global minimum as long as the temperature is decreasing slowly enough which makes SA too slow for practical use especially for clinical data. Another pioneering work was done by Besag [7], where the iterated conditional modes (ICM) was presented. This is a fast solver at the cost that it finds local optima in a neighborhood where only one site label is allowed to change. After those two methods, quite a few solvers were introduced [8], such as mean field approximation (MFa), relaxation labeling (RL), graduated non-convexity (GNC), etc. Recently two efficient and fairly accurate solvers, belief propagation (BP) [9] and graph cuts (GC) [5], were proposed. These two solvers are now often used for MRF models because they give good accuracy in an efficient way, which means they can find "global" optima within a rather large neighborhood while maintaining acceptable time complexity. However, since the perfect solver is not existing unless $P = NP$, there is still space to get more accurate results. Moreover, BP and GC are not applicable to all types of objective functions. They obtain their accuracy at the cost of function form restrictions. For example, standard BP is only proper for pairwise MRFs and generalized BP is either not for all functions [9]. So we cannot solve by BP such MRF models as FRAME [10] or multi-level logistic (MLL) with more than two sites in a clique [8]. The same situation occurs for graph cuts because GC will work only when the energy function is regular [11]. These may considerably limit the usage of the two popular solvers.

In this paper, we proposed a new simple stochastic solver for MRF modeling, called cross entropy (CE). This idea is originated from the field of operations research to simulate rare events [12]. This paper combines the idea of CE with MRF theory for the first time and applies the whole model to medical image segmentation. The CE solver is a general and stochastic optimization method that can be applied to any kind of MRF formulation. Unlike BP and GC, CE makes no assumption on the form of the objective function so it is able to solve more complicated MRF models. The efficient CE solver is completely insensitive to initialization and more importantly, as a stochastic method, CE tends to find more global optimum than deterministic solvers like BP. This statement is supported by the experiments on four sets of synthetic and clinical data, which shows CE has higher segmentation accuracy than BP. It is believed that the accuracy of BP is comparable to that of GC [13]. Although CE is stochastic, it is efficient but expends more time than deterministic solvers like BP and GC.

2 The Cross Entropy Method

The cross entropy (CE), also known as Kullback-Leibler cross entropy, has been used in combinatorial and multi-extremal optimization, and rare event simulation. Owing to its simplicity and accuracy, it has quite a few successful applica-

tions in operations research and machine learning [14]. In this section, we will present the CE method as a novel simple accurate solver for MRF modeling.

Consider the following general energy minimization problem of MRFs. Let \mathcal{F} be the configuration space of MRF, F , and f is one configuration of F . The energy minimization of MRFs is formulated by

$$f^* = \arg \min_{f \in \mathcal{F}} E(f), \tag{1}$$

where $E(\cdot)$ is the energy function to be minimized and f^* is the wanted F configuration. CE method associates an estimation problem with the optimization problem (Eq. 1). We first define an indicator function $I_{\{\text{event}\}}$, which is equal to 1 when the event is true otherwise 0. Then, suppose $p(\cdot; v)$ is a family of discrete probability density functions (pdf) on \mathcal{F} and v is its parameter. Let us estimate the following probability

$$\mathcal{P}_v(E(F) \leq e) = \sum_x I_{\{E(f) \leq e\}} p(f; v), \tag{2}$$

where \mathcal{P}_v is the probability measure and F is a vector of configurations that has pdf $p(\cdot; v)$. If $e = \min_{f \in \mathcal{F}} E(f)$ and $p(\cdot; v)$ is a uniform density on \mathcal{F} , Eq. 1 and Eq. 2 are connected. Note that $\mathcal{P}_v(E(F) \leq e)$ is typically $1/|\mathcal{F}|$, which is very small. This is similar to the situation of rare event simulation. Thus, we can borrow the idea of CE from rare event simulation to construct a multi-level optimization approach for MRF energy, where we generate a sequence of levels e_1, e_2, \dots, e_T and parameter vectors v_1, v_2, \dots, v_T such that e_T is close to the optimal e^* and v_T is the density that assigns high probability mass to the configuration which corresponds to a low energy.

Suppose m is the size of the label space of the MRF model and there are n sites altogether. The CE solver for MRF labeling can be described as follows.

CE Algorithm for MRF energy minimization

1. Set level $t = 1$ and the initial parameter vector $\mathbf{v}_0 = \{v_{0,1}, \dots, v_{0,n}\}$. Each $v_{t,i} = \{v_{t,i}^1, \dots, v_{t,i}^m\}$ is a vector with m elements for site i .
2. Generate a collection of samples F_1, \dots, F_N ($F = \{f_1, \dots, f_n\}$ is one MRF configuration) from the density $p(\cdot; v)$ and compute the energy $E_i(F_i)$ for every $i \in \{1, \dots, N\}$.
3. Sort all the $E_i(F_i)$ in a non-increasing order to $\{E_1, \dots, E_N\}$. Then pick $e_t = E_{\lceil(1-\rho)N\rceil}$.
4. Use the samples F_1, \dots, F_N to update \mathbf{v}_t by

$$v_{t,i}^j = \frac{\sum_{k=1}^N I_{\{E_k(F_k) \leq e_t\}} I_{\{F_{ki}=j\}}}{\sum_{k=1}^N I_{\{E_k(F_k) \leq e_t\}}}, \tag{3}$$

for $i = 1, \dots, n$ and $j = 1, \dots, m$.

5. If e_t remains unchanged for several iterations, go to step 6; else, set $t = t+1$ and go to step 2.

6. The final $E_N(F_N)$ of T -th iteration is the estimated minimal MRF energy. The corresponding configuration is embodied by the parameter vector \mathbf{v}_T , where

each element $v_{T,i}^j$ assigns most probability mass to a preferable label among m labels for site i . \square

Function $p(\cdot; v)$ can be any kind of pdf but the simple m -point Bernoulli distribution is usually enough. This means each label j is randomly chosen for site i according to the probability of $v_{T,i}^j$. Two parameters need to be pre-defined, ρ and N . Usually, ρ is a small value between 1% and 10%. When the site number n is large, we tend to choose a large value of ρ . Regarding the sample size N , we set $N = cn$, where c is a constant and often between 1 and 10. Notice that there are other alternative stopping criteria, such as when the parameter v_t converges to a binary (0 or 1) vector.

3 Experimental Results

In this section, we use CE and BP as MRF solvers to solve the same model for medical image segmentation and compare the results. We perform experiments on four sets of data of synthetic and clinical, vascular and cerebral images. Before that, we need to formulate the MRF energy function first.

3.1 MRF Formulation

We adopt the MAP-MRF framework for maximizing the posterior probability $P(X|Y) \propto P(X)P(Y|X)$, where X and Y are the labeling MRF and the observed data, respectively. We use multi-level logistics (MLL) and Gaussian distribution as the MRF prior and likelihood energy, which is one of the most often used models for medical image segmentation [8, 15, 16]. The MRF prior energy is expressed by

$$U(X) = \sum_{i \in S} \sum_{j \in N_i} I_{\{x_i \neq x_j\}}, \quad (4)$$

where N_i is the neighborhood of site i (we use the 4-neighborhood system) and S is the set of all sites. The likelihood energy is defined as

$$U(Y|X) = \sum_{i \in S} \sum_{j=0}^{m-1} \delta(x_i - j) \cdot \frac{(y_i - \mu_j)^2}{2\sigma^2}, \quad (5)$$

where if $x = 0$, $\delta(x) = 1$, else $\delta(x) = 0$, and μ_j represents the mean intensity of region j . In the comparisons below, all the parameters, e.g., μ , σ , are set the same for both solvers.

3.2 Synthetic Data

We first test the model on two sets of synthetic data, one simulates blood vessels following the style in [4] and the other is obtained from BrainWeb, a widely-used simulated brain database [17].

It is a binary segmentation problem to extract blood vessels. We synthesize alternate bars and circles to simulate the vessels and add various levels of Gaussian noise to produce corrupted images with different signal-to-noise ratios (SNR) (Fig. 1). The results of minimal energy values found by two solvers and segmentation errors are listed in Tab. 1. Regarding brain images, we segment the BrainWeb T1-weighted data into four classes {white matter, gray matter, cerebral fluid, others} and perform several experiments with different levels of noise. Results are shown in Tab. 2. In both experiments, CE algorithm is repeated for 5 times to give means and standard deviations because it is stochastic. We do not repeat BP because it is deterministic. It can be found that the CE solver can reach lower energy than BP and the segmentation accuracy is also higher than BP for both synthetic images thanks to the ability of CE to find more global minima.

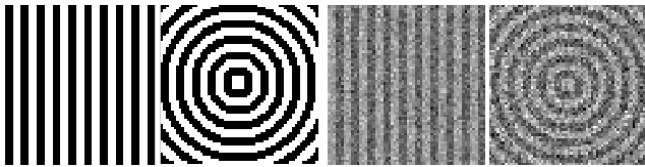


Fig. 1. Synthetic images for binary segmentation. The left two images are truth patterns and the right two images are corrupted images.

Table 1. Results of the MRF model for binary segmentation on synthetic images

image	SNR	BP		CE	
		minimum	error	minimum	error
bar (width=3)	2	2215505	18.65%	1656942±7707	9.01%±1.53%
	3	1848604	7.54%	1577403±17069	3.52%±0.69%
	4	1598662	2.25%	1523299±17271	1.44%±0.96%
	5	1526652	0.59%	1501826±2126	0.22%±0.28%
bar (width=6)	2	2101877	19.29%	1319552±4612	1.20%±0.37%
	3	1659417	8.30%	1242550±1923	0.59%±0.20%
	4	1408863	3.13%	1232195±2309	0.03%±0.06%
	5	1250346	0.68%	1216728±2223	0.07%±0.12%
circle (width=3)	2	2241872	17.58%	1721610±6633	11.45%±1.04%
	3	2000580	8.59%	1691143±3865	6.72%±0.54%
	4	1817165	4.49%	1672465±2259	3.37%±0.40%
	5	1774078	2.34%	1675028±926	1.91%±0.20%
circle (width=6)	2	2178246	18.65%	1348983±3880	4.45%±1.11%
	3	1715610	8.30%	1295023±1162	1.98%±0.25%
	4	1516673	4.39%	1278273±3337	1.86%±0.08%
	5	1380865	1.86%	1269815±634	1.37%±0.08%

Table 2. Results of the MRF model for multi-class segmentation on BrainWeb data

noise level	BP		CE	
	minimum	error	minimum	error
3%	348740	4.25%	211878.3 ± 2737.2	$2.77\% \pm 0.08\%$
5%	488391	6.00%	257021.0 ± 1505.7	$4.14\% \pm 0.06\%$
7%	487083	6.58%	229624.5 ± 1184.7	$3.57\% \pm 0.02\%$
9%	691409	11.65%	525039.5 ± 768.6	$9.08\% \pm 0.01\%$

3.3 Clinical Data: Vascular and Cerebral Datasets

We then apply the model to clinical vascular and cerebral datasets, and compare the performance of the two solvers.

The first dataset used is phase contrast magnetic resonance angiographic (PCMRA) images (size of 308×355) obtained from the University Hospital of Zurich, Switzerland. We test on 8 sets of data and the overall mean energy obtained by the CE solver is 174338.81 ± 765.84 , which is smaller than the energy 177475.56 ± 769.93 obtained by BP. Fig. 2 shows one example of the experiments. Although we do not have the ground truth to compare the two solvers quantitatively, we still can find by visual inspection that the CE acquires a little better segmentation than BP which is consistent to the smaller energy values. BP may be trapped in those local minima caused by the noise which can be seen in the middle subfigure in Fig. 2 and CE can alleviate the problem although not completely solve it.

The second dataset used is acquired from the Internet Brain Segmentation Repository (IBSR) [18] which we can get some segmented data as ground truth. There are three labels {white matter, gray matter, others} and the image is 141×149 large. We again test the model on 8 datasets and calculate the energy and accuracy obtained by the two solvers, which is shown in Tab. 3. We repeat the CE algorithm for 5 times and calculate the means and standard deviations considering its stochastic property.

Results on both synthetic and clinical data show that the CE solver is able to find lower energy (which means more global) than BP. The improvement is usually from 5% to 45% and accuracy can be increased by 1 to 7 percentage points. It is worth pointing out that lower energy does not always give better segmentation results unless the formulation is good enough, but this is not the



Fig. 2. The left is a region of interest of an original PCMRA image (contrast enhanced). The middle and right are the segmented images solved by BP and CE, respectively.

Table 3. Results of the MRF model for multi-class segmentation on IBSR brain data

dataset	BP		CE	
	minimum	error	minimum	error
1	19361	3.32%	15103 \pm 39	2.29% \pm 0.08%
2	28049	4.85%	24665 \pm 24	3.70% \pm 0.14%
3	34641	7.72%	30657 \pm 28	6.87% \pm 0.24%
4	31483	4.98%	25458 \pm 12	2.94% \pm 0.02%
5	22104	3.08%	18008 \pm 15	1.80% \pm 0.08%
6	35652	8.27%	31274 \pm 59	5.50% \pm 0.16%
7	37561	8.07%	31646 \pm 39	6.23% \pm 0.24%
8	29782	4.72%	25317 \pm 24	2.86% \pm 0.16%

goal of solvers, which should find energy as low as possible. The running time of CE algorithm for one of the clinical vascular images is around 300 minutes (BP needs about 15 minutes) and for one of the clinical brain images around 60 minutes (BP needs about 9 minutes) on a computer with 1.3GHz CPU and 500 MB memory. Decreasing the parameter ρ will reduce the required number of iterations and thus the computation time of CE at the cost of likely increasing the energy value found in the end.

4 Discussion and Conclusion

In general, the CE algorithm is very simple and easy to implement. It is an iterative procedure and in each iteration, a sequence of samples is generated according to a certain probability distribution. The method chooses one threshold of objective function value and just focuses on those samples whose performance (e.g., lower energy) is better than this threshold. Then CE updates the distribution parameters according to these good samples. This completes one iteration. We can see that CE is stochastic and no deterministic decision is made, which gives CE the ability to find more global optima than deterministic methods. Compared with other stochastic solvers like simulated annealing (SA), CE is obviously more efficient because it concentrates on a few high-performance samples among a large collection of random samples and quickly converges to states which have good performance. Moreover, the CE algorithm only evaluates the energy values and requires no specific energy form. Thus it can be applied to any type of objective functions. This makes CE a more general solver than BP and GC. Another advantage of CE lies in its insensitivity to initialization because CE's initialization is unchanged for any inputs, i.e., the parameter vector \mathbf{v}_0 is always set to uniform distribution and every element is $1/m$.

The study on the computational complexity of CE algorithm is still an open problem partly because CE can apply to all kinds of different applications and the necessary total number of iterations is different. For some applications like max-

cut and partition problems [14], the theoretical complexity of CE is $\mathcal{O}(n^3 \ln n)$ and its empirical complexity is $\mathcal{O}(n \ln n)$. The CE solver is not as efficient as BP or GC. But, since the architecture of CE algorithm is inherently parallel and its steps are all simple, it has large potential to speed up. Regarding space complexity, CE can occupy a lot of space if we store all the samples. However, the problem can be solved by keeping a small portion of them since CE just uses the high-performance part.

In this paper, we have introduced a new MRF solver, namely cross entropy (CE), applied it to medical image segmentation and shown some advisable properties of it compared to an existing popular MRF solver, belief propagation (BP). CE algorithm is a general solver that can be applied to any type of MRF models and it is stochastic and iterative, which endows it with the capability to find more global optima than BP. This makes CE useful in medical image segmentation or any other tasks which can be formulated as energy minimization problems. However, CE requires more computation time and space than deterministic optimization methods. Since it has parallel architecture and is very simple in itself, CE has potential to be accelerated by parallel processing or algorithm optimization. This will be the major work to be done in the future.

References

1. N.Villain, Y.Goussard, J.Idier, M.Allain: Three-dimensional edge-preserving image enhancement for computed tomography. *IEEE TMI* **22** (2003) 1275–1287
2. Y.Chen, A.Amini: A map framework for tag line detection in spamm data using markov random fields on the b-spline solid. *IEEE TMI* **21** (2002) 1110–1122
3. M.Svensen, F.Kruggel, D.Cramon: Probabilistic modeling of single-trial fmri data. *IEEE TMI* **19** (2000) 25–35
4. A.C.S.Chung, J.A.Noble, P.Summers: Vascular segmentation of pcmra based on statistical mixture modeling and local phase coherence. *IEEE TMI* **23** (2004) 1490–1507
5. Y.Boykov, O.Veksler, R.Zabin: Fast approximate energy minimization via graph cuts. *IEEE on PAMI* **23** (2001) 1222–1239
6. S.Kirkpatrick, D.Gellatt, M.Vecchi: Optimization by simulated annealing. *Science* **220** (1983) 671–680
7. J.Besag: On the statistical analysis of dirty pictures. *J. of the Roy. Stat. Soc. Series B* **48** (1986) 259–302
8. S.Z.Li, ed.: *Markov Random Field Modeling in Image Analysis*. Springer-Verlag Tokyo (2001)
9. J.S.Yedidia, W.T.Freeman, Y.Weiss: Understanding belief propagation and its generalizations. Technical Report, Mitsubishi Electric Research **TR-2001-22** (2001)
10. S.Zhu, Y.Wu, D.Mumford: Frame: filters, random fields, and minimax entropy towards a unified theory for texture modeling. In: *IEEE CVPR*. Volume 6. (1996) 686 – 693
11. V.Kolmogorov, R.Zabih: What energy functions can be minimized via graph cuts? *IEEE Trans. PAMI* **26** (2004) 147–159
12. R.Rubinstein: Optimization of computer simulation models with rare events. *European Journal of Operations Research* **99** (1997) 89–112

13. M.Tappen, W.Freeman: Comparison of graph cuts with belief propagation for stereo, using identical mrf parameters. In: IEEE ICCV. (2003) 900–906
14. R.Rubinstein, D.Kroese, eds.: The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning. Springer-Verlag, New York (2004)
15. K.Held, *et al*: Markov random field segmentation of brain mr images. IEEE TMI **16** (1997) 878–886
16. S.Ruan, *et al*: Brain tissue classification of magnetic resonance images using partial volume modeling. IEEE TMI **19** (2000) 1179–1187
17. BrainWeb: (<http://www.bic.mni.mcgill.ca/brainweb/>)
18. IBSR: (<http://www.cma.mgh.harvard.edu/ibsr/>)

Semi-automated Basal Ganglia Segmentation Using Large Deformation Diffeomorphic Metric Mapping

Ali Khan¹, Elizabeth Aylward², Patrick Barta³,
Michael Miller⁴, and M. Faisal Beg¹

¹ School of Engineering Science, Simon Fraser University,
8888 University Drive, Burnaby BC, V5A 1S6, Canada

² Department of Radiology and Psychiatry,
University of Washington, Box 357115, Seattle, WA 98195, USA

³ Department of Psychiatry and Behavioral Sciences,
Division of Psychiatric Neuro-Imaging,
The Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA

⁴ Center for Imaging Science,
The Johns Hopkins University, 301 Clark Hall, Baltimore, MD 21218, USA

Abstract. This paper investigates the techniques required to produce accurate and reliable segmentations via grayscale image matching. Finding a large deformation, dense, non-rigid transformation from a template image to a target image allows us to map a template segmentation to the target image space, and therefore compute the target image segmentation and labeling. We outline a semi-automated procedure involving landmark and image intensity-based matching via the large deformation diffeomorphic mapping metric (LDDMM) algorithm. Our method is applied specifically to the segmentation of the caudate nucleus in pre- and post-symptomatic Huntington's Disease (HD) patients. Our accuracy is compared against gold-standard manual segmentations and various automated segmentation tools through the use of several error metrics.

1 Introduction

Shape and volume change of the basal ganglia structures, which include the caudate nucleus and putamen, have been the focus of investigation in clinical studies of Huntington's Disease [1] and other neuro-degenerative disorders. In order to study shape and volume changes, segmentation of the basal ganglia structures in an MRI image is the first processing step. Manual segmentation by a trained rater is the current anatomic gold standard, but this technique requires a sizable amount of time from an anatomic expert. Manual segmentations by different anatomic experts also suffer from inter-rater reliability issues. Reliable and accurate semi-automated segmentation of the caudate nucleus is limited by the following factors: limited resolution of MRI scans, inhomogeneous intensities throughout the caudate, and ill-defined boundaries. Successful segmentation algorithms built on automatic to semi-automatic algorithms making use of some form of expert prior knowledge are likely to overcome the mentioned difficulties.

The LDDMM tool [2] provides non-rigid registration between two grayscale anatomical images, which we use to develop our system for automated segmentation. Given a template image, target image, and template segmentation, the non-rigid mapping of the template image to the target image can be used to propagate the template segmentation, generating a target segmentation. In this way, we are able to use anatomic expert priors, namely a manually generated template segmentation, to produce target segmentations. In this paper, we present the validation of this method for accuracy and reliability.

2 Method

2.1 Data

MRI scans used for this study were all acquired on a GE 1.5T scanner. Three types of scans were analyzed: (1) 1.5 mm SPGR (spoiled gradient recalled echo in steady state) coronal series, TR = 35; TE = 5, NEX = 1, flip angle = 45° , voxel size = 0.9735 mm in x and y direction, from subjects with known caudate atrophy (patients with Huntington's disease); (2) 1.5 mm SPGR axial series, TR = 18 TE = 3, NEX = 2, field of view = 24 cm, flip angle = 20° , voxel size = 0.9735 mm in x and y direction, from subjects with possible caudate atrophy (pre-symptomatic carriers of the HD gene); and (3) 1.2 mm SPGR sagittal series, TR = 11.1, TE = 2.2, NEX = 1, field of view = 24 cm, flip angle = 25° , voxel size = 0.9735 mm in x and y direction, from children with no known caudate atrophy.

Scans were of varying levels of quality for the segmentation program (with the NEX=2 scans have the most defined borders between caudate and white matter) and represented groups will have varying levels of caudate volume. Scans were stripped of all identifying information.

Computation of automated segmentations via image matching involves the use of a template and a target image. Our Series 1 data consisted of patients scanned on two separate occasions, therefore one scan was chosen as a template image, and the other scan was chosen as a target image. Series 2 and Series 3 data did not include multiple scans from the participants, restricting us from using the same patient as template and target. For the Series 2 data, one patient image was chosen as a template to compute segmentations for the other Series 2 patient images. Series 3 segmentations were computed similarly. In total, 17 automated segmentations were produced, five Series 1, seven Series 2, and five Series 3.

2.2 Registration and Pre-processing

Global alignment of the basal ganglia structures is required before intensity-based image matching can take place; rigid landmark-based registration was used to accomplish this task. We chose to place landmarks on the surfaces of the lateral ventricles because of their unique properties: adjacency to the caudate nucleus during all stages of atrophy, well-defined boundaries, and homogeneous intensities within these boundaries. Furthermore, placement of landmarks on a

3D surface as opposed to a 3D volume limits the degrees of freedom in placement hence resulting in lower variability. Aside from registration, landmarks placed on the ventricular surface are subsequently used in image matching to calculate an initial template to target mapping which is later refined by image intensity-based alignment. The ventricle segmentation used for surface generation is produced by global thresholding with user guidance and takes a trained user approximately five minutes to complete.

Landmarks are placed interactively by the user on the inferior surface of the ventricles, which is adjacent to the superior surface of the caudate nucleus. Landmarks are placed on the ventricular surface of the template first, then the target; the user attempting to place the landmarks in the same relative location on the surface. After placing 6-7 landmarks on each side of the ventricles, the target landmarks on the left target ventricle are rigidly deformed to the template landmarks on the left template ventricle, generating the target image deformed in accordance with the left set of landmarks. The landmarks on the right ventricles are transformed in the same manner producing another target image deformed in accordance with the right set of landmarks. A six-parameter translation/rotation least squares landmark matching algorithm is used to rigidly rotate and translate the landmarks and the corresponding images.

After the left and right rigid transformations, the left template and target caudates, and the right template and target caudates are globally aligned. We then define a region of interest (ROI) to include the caudate nuclei. We isolate our ROI by inspecting the boundaries of the caudates in the template image and cropping a bounding box surrounding these areas. The template and target images are then preprocessed with edge-preserved smoothing [3], histogram matching, and intensity stretching for noise reduction, intensity matching and contrast enhancement respectively.

2.3 Image Matching

Image matching is achieved through the solution of the large deformations diffeomorphic metric mapping (LDDMM); a detailed derivation and discussion of this can be seen in [2]. Template and target images, I_0 and I_1 , represented by functions $I : \Omega \rightarrow \mathbb{R}$, are mapped via the transformation $\varphi : \Omega \rightarrow \Omega$, where $\Omega \subseteq \mathbb{R}^3$. The diffeomorphic transformation generated is smooth and has a smooth inverse, hence, smoothness of anatomical features is preserved and coordinates are transformed consistently.

Globally registered and image intensity equalized images are input into the image matching (LDDMM) program. Our computations were carried out on a 64-bit symmetric multi-processor, with computation times not exceeding one hour. Composition of the target segmentation is achieved by composing the template binary segmentation with the appropriate LDDMM mapping. Rigidly deforming the target segmentation back to the patient space and combining the left and right sides produces the final automated segmentation.

A graphical summary of our method for obtaining a mapping between a template and target image is shown in Figure 1.

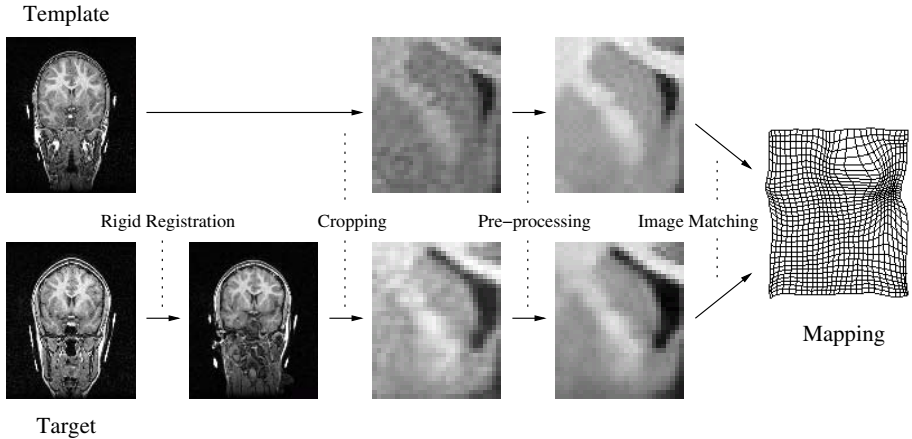


Fig. 1. Procedural flow outlining the generation of a template to target diffeomorphic mapping for one side (left) of the brain. The target is deformed rigidly to the template, both images are cropped and pre-processed, and LDDMM image matching is used to compute the final mapping. Images shown are representative slices taken from a Series 3 patient.

3 Results

To facilitate comparison of our continuous automated segmentations with the binary manual segmentations, the manual segmentations are slightly smoothed. The L1-distance between segmentations is reported as the L1 error, and caudate volumes are also computed.

We have also developed a boundary distance metric which aims to provide correspondence between two surfaces by computing the LDDMM mapping from an automated segmentation to a manual segmentation. Applying this mapping to the template surface displaces each point on the template to its corresponding point on the target surface. Since the LDDMM matching is in the inexact framework, a template surface point may not map exactly to a target surface point. However since the deformed template surface closely approximates the target surface, choosing the nearest point on the target surface to a point on the template surface is likely to provide very good correspondence. Using this correspondence we can compute the boundary distances for each point on the template surface to a point on the target surface. The average for the surface is calculated and is reported as the mean boundary distance.

Another estimate of surface distance, namely the maximum surface distance, is given by the Hausdorff distance. Calculation of this metric is carried out as described in [4].

Segmentation accuracy results were compiled by computing the metrics described above for each automated segmentation. We report the average and standard deviation of these metrics for each dataset series. These results are shown in Table 1. L1 error and volume error results for individual Series 1 pa-

Table 1. Caudate nucleus LDDMM segmentation results

Error Metric	Series 1	Series 2	Series 3
Volume Error (%)	5.10 ± 3.60	8.08 ± 3.72	7.72 ± 6.18
L1 Error (%)	27.94 ± 6.51	36.21 ± 5.51	30.17 ± 6.99
95% Sym. Hausdorff (mm)	2.10 ± 0.98	2.27 ± 0.75	2.19 ± 0.70
Mean Boundary Dist. (mm)	1.12 ± 0.57	1.17 ± 0.47	0.94 ± 0.20

This table shows the computed results for our automated segmentations, reporting the mean and standard deviation ($\mu \pm \sigma$) for each dataset. High L1 error is mainly due to discrepancies on the boundaries of the automated segmentations. Small boundary distance metrics (Hausdorff, mean boundary distance) indicate that the overall shape is accurate

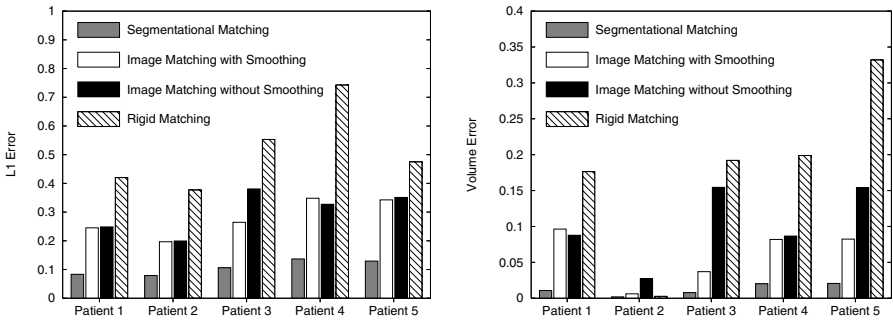


Fig. 2. L1 errors (left) and volume errors (right) for Series 1 segmentations using different transformations. Segmentational matching refers to using binary segmentations as template and target images, image matching with smoothing is our regular system, image matching without smoothing skips the noise-reduction step, and rigid matching does not use any image matching transformation.

tients are shown in Figure 2. We compare results we get choosing from different mappings strategies: rigid rotation/translation, intensity-based image matching with and without edge-preserving smoothing, and pure segmentational matching. Representative slices for two different Series 1 patients displaying automated and manual caudate segmentations are shown in Figure 3.

Our caudate nucleus segmentation results are in the same range as other comparable methods reported in literature [5], with less requirements on parameter tuning and user input.

Proven reliability of our segmentation system is essential for its use in clinical studies. Our reliability study consisted of performing the segmentation procedure 8 times each for two different Series 2 patients, generating a new caudate nucleus segmentation each time. Ideally, for an inter-rater reliability calculation, 8 different raters would rate each patient, however, we were limited to using a single rater, computing segmentations on different days. For each image matching computation (total of 16), the ventricles were segmented, landmarked, and the

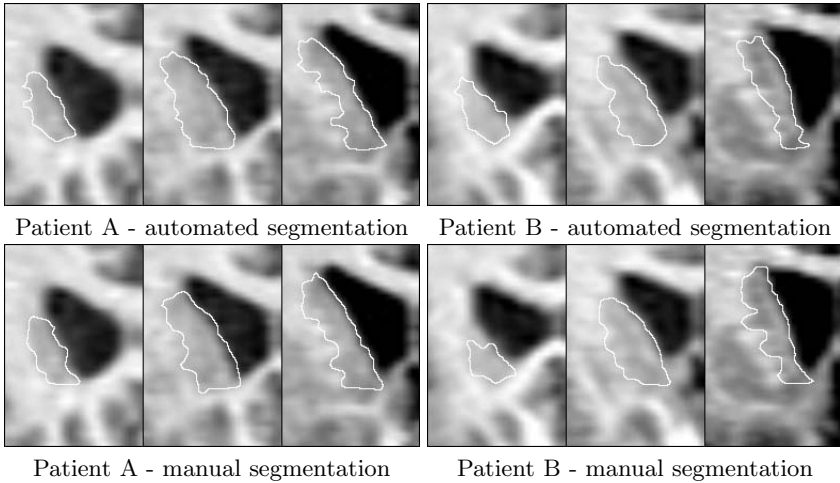


Fig. 3. Three coronal slices of the left caudate nucleus of two Series 1 patients shown with automated binary threshold segmentations (Top) and manual segmentations (Bottom) outlined in white. (Left) Patient A: L1 error = 26.45%, mean boundary distance = 0.6223 mm. (Right) Patient B: L1 error = 34.82%, mean boundary distance = 1.3409 mm.

average caudate nucleus intensity was chosen; as usual, the same template image was used for all segmentations. Segmentations for the left and right caudates were derived for each of these cases, and their volumes were used to calculate the intra-class correlation coefficient. Although technically an intra-rater reliability coefficient, because each segmentation was done on a different day, it is our best approximation to inter-rater reliability. This is because the day to day variability for a single rater segmenting the ventricles, landmarking and choosing the average caudate intensity should be similar to the variability between different raters, especially when the system as a whole is considered.

The simple replication reliability study uses a one-way random effects model to calculate the intra-class correlation coefficients (ICC) and confidence intervals for the caudate nucleus volumes as dictated by [6]. The calculated ICC for Series 2 caudates is 0.937 with a 95% confidence interval of (0.759, 0.994). According to [7], $ICC > 0.74$ corresponds to excellent reliability, which suggests our segmentation system is sufficiently reliable.

4 Discussion

One difficulty in producing automated segmentations of the caudate nucleus is determining the location of the most inferior axial slice. This is defined to be the most inferior slice when the caudate and putamen are visibly separated by the internal capsule with the brain aligned in AC-PC orientation. Generating an automated segmentation which follows the same strict guidelines is difficult at

best, and determining the inferior axial cut-off is a task better suited for the user. Using a template segmentation which includes the caudate nucleus, putamen, and the nucleus accumbens would produce a target segmentation that can be manually bisected to produce caudate and putamen segmentations cut-off at the correct axial plane.

User involvement in our semi-automated segmentation scheme is minimal and does not require expertise in brain anatomy. The user is necessary for four tasks: 1. automated thresholding-based segmentation of the ventricles, 2. landmarking of the ventricular surfaces, 3. definition of the bounding box to delineate the image region for matching, and 4. identification of average caudate intensities for intensity equalization. As discussed earlier, our procedure for lateral ventricle segmentation is fast and easy since high accuracy is unnecessary as only a small area of the surface is used for landmarking. Placing landmarks is also an easy task once landmark positions have been standardized. Definition of a bounding box is only necessary for template images, as image registration should ensure that the target images will fit in the template bounding box by always padding the template bounding box by a tolerance of about 10 voxels or such. Identifying the average intensity for the target structures is subject to some variability, but as our reliability study shows, these variations do not amount to notable changes in the segmentations.

The automated and the manual segmentations are found to differ on the exterior boundary of the caudate. Due to the elongated, narrow shape of the caudate nucleus, the ratio of the number of voxels on the surface of the caudate to the total number of caudate voxels is very high, on average being $71 \pm 9\%$. Therefore, the partial volume effects are likely to be a heavy influence in calculation of the L1 distance which measures the difference in voxel labeling between the automated and the manual segmentation, especially since there are a large number of these on the exterior of the caudate that are not exactly matched due to regularization constraints placed in computing dense diffeomorphic transformations. Examination of Figure 3 reveals that accurate delineation of the caudate nuclei is still present in segmentations possessing relatively high L1-distance errors. Further improvement in accuracy is likely through the use of a template image representing the average of the study population. We are currently working on geodesic shooting methods in LDDMM to produce an average image.

Concluding, the results we have shown demonstrate that our semi-automated image matching system reliably segments the caudate nucleus with accuracy comparable to current automated methods. Segmentation of the basal ganglia via image matching is the initial step in shape analysis as the diffeomorphic mapping defines the necessary correspondence between images. Application of our procedure to other structures should also prove successful, as our findings indicate that the general image matching segmentation concept is robust.

Acknowledgments

We acknowledge funding support from NSERC 31-611387 for this work.

References

1. Aylward, E., Rosenblatt, A., Field, K., Yallapragada, V., Kieburtz, K., McDermott, M., Raymond, L., Almqvist, E., Hayden, M., Ross, C.: Caudate volume as an outcome measure in clinical trials for huntington's disease: a pilot study. *Brain Research Bulletin* **62** (2003) 137–141
2. Beg, M.F., Miller, M.I., Trouvé, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *IJCV* **61** (2005) 139–157
3. Whitaker, R., Xue, X.: Variable-conductance, level-set curvature for image denoising. (2001)
4. Gerig, G., Jomier, M., Chakos, M.: Valmet: A new validation tool for assessing and improving 3D object segmentation. (2001)
5. Pitiot, A., Delingette, H., Thompson, P., Ayache, N.: Expert knowledge guided segmentation system for brain MRI. *NeuroImage* **23** (2004) S85–S96 Special Issue: Mathematics in Brain Imaging.
6. Fleiss, J.L.: *The Design and Analysis of Clinical Experiments*. Wiley-Interscience (1999)
7. Fleiss, J.L.: *Statistical Methods for Rates and Proportions*. Wiley (1981)

Particle Filters, a Quasi-Monte-Carlo-Solution for Segmentation of Coronaries

Charles Florin¹, Nikos Paragios², and Jim Williams¹

¹ Imaging & Visualization Department,
Siemens Corporate Research, Princeton, NJ, USA
² CERTIS - Ecole Nationale des Ponts et Chaussees,
Champs-sur-Marne, France

Abstract. In this paper we propose a Particle Filter-based approach for the segmentation of coronary arteries. To this end, successive planes of the vessel are modeled as unknown states of a sequential process. Such states consist of the orientation, position, shape model and appearance (in statistical terms) of the vessel that are recovered in an incremental fashion, using a sequential Bayesian filter (Particle Filter). In order to account for bifurcations and branchings, we consider a Monte Carlo sampling rule that propagates in parallel multiple hypotheses. Promising results on the segmentation of coronary arteries demonstrate the potential of the proposed approach.

1 Introduction

Cardio-vascular diseases are the leading cause of deaths in the USA (39%) and therefore there is a demand for constant improvement of diagnostic tools to detect and measure anomalies in the cardiac muscle. Coronary arteries are thin vessels that feed the heart muscle in blood. Therefore, their segmentation provides a valuable diagnostic tool for clinicians interested in detecting calcifications, and stenosis. Because of the low contrast conditions, and the coronaries vicinity to the blood pool, segmentation is a difficult task. Computer Tomography (CT) and Magnetic Resonance (MR) imaging of the heart have become standard tools to medical diagnosis resulting to a substantial number of patients being imaged.

On one hand, vessel segmentation techniques consist of model-free and model-based methods. Vessel enhancement approaches [6] and differential geometry-driven methods [11] do not segment vessels per se, but allow a better visualization. Region growing [21], flux maximization [2], morphological operators [5] and skeleton-based techniques [19] are more advanced vessel segmentation techniques. On the other hand, model-based techniques use prior knowledge and features to match a model with the input image and extract the vessels. Prior knowledge refer either to the whole structure, or to the local vessel model. Tracking approaches recover the vessel centerline - given a starting condition - through processing information on the vessel cross section [10]. Vessel template matching [17], generalized cylindrical models [15] as well as parametric/geometric deformable models [18] are alternatives to vessel tracking and seek to minimize an objective function computed along the model.

Level sets [16] is an established method to address such minimization [12]. One can refer to the fast marching algorithm and its variant for vessel segmentation using the minimal path principle [1]. To discourage leaking, a local shape term that constrains the diameter of the vessel was proposed in [14].

One can claim that existing approaches suffer from certain limitations. Local operators, region growing techniques, morphological filters as well as geometric contours might be very sensitive to local minima and fail to take into account prior knowledge on the form of the vessel. Parallel to that, cylindrical models, parametric active contours and template matching techniques may not be well suited to account for the non-linearity of the vessel structure, and require particular handling of branchings and bifurcations. Tracking methods can often fail in the presence of missing and corrupted data, or sudden changes. Level sets are very computational time-consuming and the Fast Marching algorithm loses all the local implicit function properties.

In this paper, we propose a particle-based approach to vessel segmentation where we re-formulate the problem of recovering successive planes of the vessel in a probabilistic fashion with numerous possible states. One can consider the problem of vessel segmentation as a tracking problem of tubular structures in 3D volumes. Thus, given a starting position, the objective is to consider a feature vector that, upon its successful propagation, provides a complete segmentation of the coronaries. In the proposed technique, unlike standard techniques where the most probable hypothesis is maintained, a discrete number of states (possible solutions) remain active and are associated with a probability density function. The final paradigm consists of a fast multiple hypothesis adaptive propagation technique where the vessel structure and its appearance are successfully recovered. Such a framework allows to naturally address the non-linearities of the geometry and the appearance of coronaries and is compared in a favorable fashion with the existing approaches.

The remainder of this paper is organized as follows. In section 2, we motivate vessel segmentation and introduce the concept of the proposed approach and Particle Filters, while vessel segmentation is presented in section 3. Implementation and validation are part of section 4, while discussion is part of the last section.

2 Preliminaries and Particle Filters

To explain our method at a concept level, let us assume that a segment of the vessel has been detected: a 2D shape on a 3D plane. Similar to region growing and front propagation techniques, our method aims to segment the vessel in adjacent planes. To this end, one can consider the hypotheses ω of the vessel being at a certain location (\mathbf{x}), having certain orientation (Θ), and referring to certain shape - an elliptic model is a common choice (ϵ) - with certain appearance characteristics (\mathbf{p}_{vessel}).

$$\omega = \left(\underbrace{\mathbf{x} = (x_1, x_2, x_3)}_{\text{position}}, \underbrace{\Theta = (\theta_1, \theta_2, \theta_3)}_{\text{orientation}}, \underbrace{\epsilon = (\alpha, \beta, \phi)}_{\text{shape}}, \underbrace{\mathbf{p}_{vessel}}_{\text{appearance}} \right)$$

Then, segmentation consists in finding the optimal parameters of ω given the observed 3D volume. Let us consider a probabilistic interpretation of the problem with $\pi(\omega)$

being the posterior distribution that measures the fitness of the vector ω with the observation. Under the assumption that such a law is present, segmentation consists in finding at each step the set of parameters ω that maximizes $\pi(\omega)$. However, since such a model is unknown, one can assume an autoregressive mechanism that, given prior knowledge, predicts the actual position of the vessel and a sequential estimate of its corresponding states. To this end, we define:

- a state/feature vector ω ,
- an iterative process to predict the next state and update the density function, that can be done using a Bayes sequential estimator and is based on the computation of the present state ω_t pdf of a system, based on observations from time 1 to time t $z_{1:t}$: $\pi(\omega_t|z_{1:t})$. Assuming that one has access to the prior pdf $\pi(\omega_{t-1}|z_{1:t-1})$, the posterior pdf $\pi(\omega_t|z_{1:t})$ is computed according to the Bayes rule:

$$\pi(\omega_t|z_{1:t}) = \frac{\pi(z_t|\omega_t)\pi(\omega_t|z_{1:t-1})}{\pi(z_t|z_{1:t-1})}$$

The recursive computation of the prior and the posterior pdf leads to the exact computation of the posterior density.

- a distance between prediction and actual observation, based on the observation.

Kalman filter is the most popular variant of this model, a well known linear approach able to track vessels with limited variation in appearance and geometry. Cardiac vessel trees are highly irregular. Random bifurcations, branches of variable width, non-linear visual properties because of the presence of calcifications, stents, stenosis and diseased vessel lumen are some examples demonstrating the non-linearity of the vessel tree as shown in [Fig. (1)].

Consequently simple parametric statistical models will fail to account for the statistical and geometric properties of the vessel leading to the consideration of more complex distributions. To this end, instead of one single prediction, a collection of hypotheses can be generated at each step and being evaluated using the distance between prediction and actual observation. Nevertheless, in practical cases, it is impossible to compute exactly the posterior pdf $\pi(\omega_t|z_{1:t})$, which is to be approximated. An elegant approach to implement such a technique refers to the use of particle filters where each given hypothesis is a state in the feature space (or particle), and the collection of hypothesis is a sampling of the feature space.

Particle Filters [3, 9] are sequential Monte-Carlo techniques that are used to estimate the Bayesian posterior probability density functions (pdf) [7, 20]. In terms of a mathematical formulation, such a method approximates the posterior pdf by M random measures $\{\omega_t^m, m = 1..M\}$ associated to M weights $\{\lambda_t^m, m = 1..M\}$, such that

$$\pi(\omega_t|z_{1:t}) \approx \sum_{m=1}^M \lambda_t^m \delta(\omega_t - \omega_t^m), \tag{1}$$

where each weight λ_t^m reflects the importance of the sample ω_t^m in the pdf. The samples ω_t^m are drawn using the principle of *Importance Density* [8], of pdf $q(\omega_t|x_{1:t}^m, z_t)$, and it is shown that their weights λ_t^m are updated according to

$$\lambda_t^m \propto \lambda_{t-1}^m \frac{\pi(z_t|\omega_t^m)\pi(\omega_t^m|\omega_{t-1}^m)}{q(\omega_t^m|\omega_{t-1}^m, z_t)}. \tag{2}$$

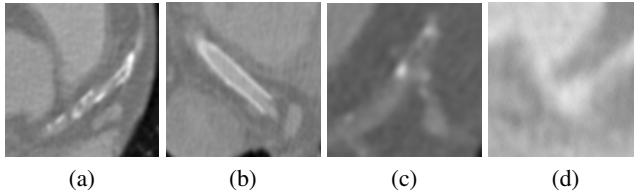


Fig. 1. (a) calcification, (b) stent (high intensity prosthesis), (c) branching with obtuse angles, (d) stenosis (sudden reduction of vessel cross section diameter)

Once a set of samples has been drawn, $\pi(\omega_t^m | \omega_{t-1}^m, z_t)$ can be computed out of the observation z_t for each sample, and the estimation of the posteriori pdf can be sequentially updated.

3 Vessel Segmentation

We now consider the application of such a non linear model to vessel segmentation and tracking. Without loss of generality, one can assume that the root of a coronary is known, either provided by the user or through some prior automatic procedure. Simple segmentation of that area can provide an initial guess on the statistical properties of the vessel appearance. It is reasonable to assume irregularity in the appearance \mathbf{p}_{vessel} of the vessel because of the presence of calcifications, stents, stenosis and diseased vessel lumen [FIG. (1)]. Therefore simple parametric statistical models on the appearance space will fail to account for the statistical properties of the vessel and more complex distributions are to be considered. We consider a Gaussian mixture model that consists of two components to represent the evolving distribution of the vessel, the contrast enhanced blood (P_B, μ_B, σ_B) and the high density components, such as calcifications or stent, (P_C, μ_C, σ_C) subject to the constraint $[P_C + P_B = 1]$ leading to the following state vector:

$$\omega = (\mathbf{x}, \Theta, \epsilon, (P_B, \mu_B, \sigma_B), (P_C, \mu_C, \sigma_C)) \quad (3)$$

The vessel state vector consists of the 3D location of the vessel \mathbf{x} , the tangent vector Θ , its shape model at a given cross-section (the model used here is an ellipse (α (major axis radius), β (minor axis radius), ϕ (orientation))) and the appearance \mathbf{p}_{vessel} , mixture of two gaussians.

Once such a recursive paradigm is built, the next and last issue to be addressed is the definition of a measure between a prediction and the actual observation. To this end, we are using mostly the image terms, and in particular the intensities that do correspond to the vessel in the current cross-section. The observed distribution of this set is approximated using a Gaussian mixture model according to the expectation-maximization principle.

Let us now consider a random state vector ω , that refers to a certain segmentation hypothesis that is to be evaluated ($p(\omega | \mathcal{D})$) where \mathcal{D} is the observed 3D volume. Such a hypothesis should refer to a region that has consistent visual properties with the ones expected (\mathbf{p}_{vessel}). While the separation of the vessels from the cardiac muscle is a rather tedious task (blood is present in both organs), one can claim that their separation

from the liquid of the vascular structure is possible and can be used to validate the goodness of a hypothesis.

For the vessel lumen pixels distribution \mathbf{p}_{vessel} , the probability is measured as the distance between the hypothesized distribution and the distribution actually observed. The distance we use is the symmetrized Kullback-Leibler distance D_{ap} between the model $p(\omega) = \mathbf{p}_{vessel}$ and the observation $q(\omega)$:

$$D_{ap} = \int p(\omega) \log \left(\frac{p(\omega)}{q(\omega)} \right) + q(\omega) \log \left(\frac{q(\omega)}{p(\omega)} \right) d\omega,$$

which have important values when the distance between these two distributions is significant. Therefore, one can consider the following measure $\left[p(\omega|D_{ap}) = e^{-\frac{|D_{ap}|}{\sigma_{ap}}} \right]$ where σ_{ap} is a normalization factor. Toward discriminating the vessel from the vascular liquid one can consider a ribbon measure

$$D_{rb} = \begin{cases} -\infty & , \quad \mu_{int} \leq \mu_{ext} \\ \frac{\mu_{int} - \mu_{ext}}{\mu_{int} + \mu_{ext}} & , \quad otherwise \end{cases}$$

where μ_{in} is the mean within the ellipse and μ_{ext} is the mean within a ring centered at the ellipse center with greater radius (the ring area is equal to the inner circle area). Such a measure aims at maximizing the distance between the mean values of the interior and the exterior region [13], based on the fact that the coronary arteries are brighter than the background, and can also be used to measure the fitness of the segmentation: $\left[p(\omega|D_{rb}) = e^{-\frac{|D_{rb}|}{\sigma_{rb}}} \right]$. We assume that the two conditions are independent and therefore one can multiply the two measures to determine the goodness of the hypothesis under consideration.

Given a starting point and a number of particles, one now performs random perturbations to each particle in the feature space. Once a perturbation has been applied, the corresponding hypothesis is evaluated using the visual matching and the ribbon measure introduced earlier. At each step of the process, segmentation refers to a weighted linear combination of the state vectors (the particles) [EQ. (1)].

4 Implementation

After certain iterations, such a process will remove most of the particles and only the ones that express the data will present significant weights. Consequently the model will lose its ability to track significant changes on the pdf. At the same time, in the presence of bifurcations, new hypotheses are to be introduced in order to capture the entire vessel tree. Therefore, a resampling procedure has to be executed on a regular basis. Such a process will preserve as many samples as possible with respectful weights. One can find in the literature several resampling techniques. We chose the most prominent one, Sampling Importance Resampling, for its simplicity to implement, and because it allows more hypothesis with low probability to survive, compared to more selective techniques such as Stratified Resampling [4].

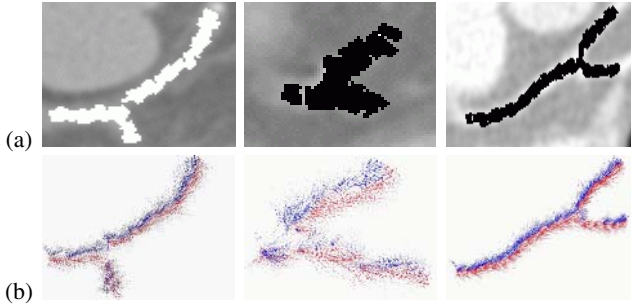


Fig. 2. (a) branching points between LCX and LAD for three patients with the particles' mean state overlaid, (b) the particles, clustered using K-means, follow up the two branches

The Sampling Importance Resampling (SIR) algorithm [7] consists in choosing the prior density $\pi(\omega_t|\omega_{t-1})$ as importance density $q(\omega_t|\omega_{1:t}^m, z_t)$. This leads to the following condition, from [EQ. (2)]: $\lambda_t^m \propto \lambda_{t-1}^m \pi(z_t|\omega_t^m)$.

The samples are updated by setting $\omega_t^m \propto \pi(\omega_t|\omega_{t-1}^m)$, and perturbed according to a random noise vector. The SIR algorithm is the most widely used resampling method because of its simplicity from the implementation point of view. Nevertheless, the SIR uses mostly the prior knowledge $\pi(\omega_t|\omega_{t-1})$, and does not take into account the most recent observations z_t . Such a strategy could lead to an overestimation of outliers. On the other hand, because SIR resampling is performed at each step, fewer samples are required, and thus the computational cost may be reduced with respect to other resampling algorithms.

Particular attention is also to be paid during the resampling process to address branching and bifurcations. When a branching occurs, the particles split up in the two daughter branches, and then track them separately (see [Fig. (2)]). Although Particle Filters track the two branches, experiments have shown a branching detection heuristics improves the results. To this end, a simple K-means approach on the joint space (position+orientation) of the particles is considered. When the two clusters are well separated, the number of particles is doubled and equally dispatched in the two branches.

Regarding the initial configuration, the use of approximately 1,000 particles gave sufficient results for our experiments. We perform a systematic resampling according to the Sampling Importance Resampling every time the effective sampling size $N_{eff} = \sum_i 1/\lambda_i^2$ (where λ_i is the weight of the i^{th} particle) falls below half the number of particles. The preference for SIR, compared to Stratified Resampling [4], is motivated by the robustness of the segmentation.

5 Discussion

In this paper, we have proposed a particle-filter based approach to vascular segmentation. Experiments were conducted on several healthy and diseased patients CTA data sets, segmenting both the *Left Main Coronary Artery* and the *Right Coronary Artery*.

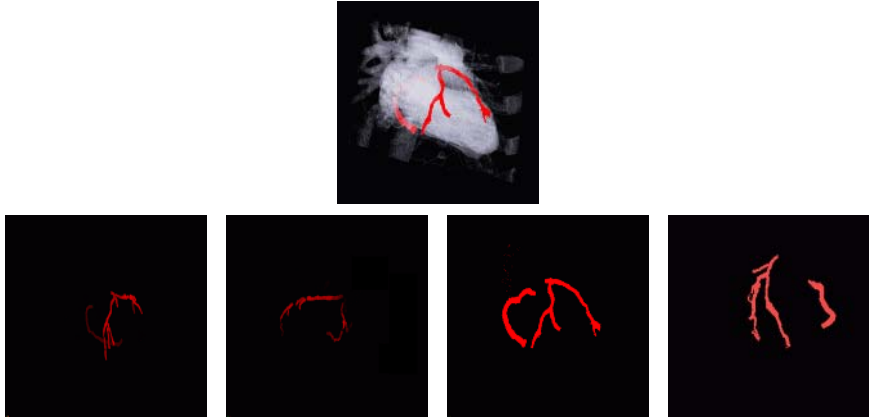


Fig. 3. Segmentation of the Left anterior descending coronary artery and Right coronary artery in CTA (in red) for four patients, and full view with heart

Validation is a challenging but required step for any coronary segmentation method. The algorithm has been evaluated on 34 patients, and has successfully recovered all the main arteries (RCA, LAD, LCX) for each patient as shown in the following table, while

vessel name	RCA	Acute Marginal	LAD	First Septal	LCX	Obtuse Marginal
% of cases segmented	100%	85.3%	100%	94%	100%	94%

a small portion of visual results are also presented in [Fig. (3)]. The table indicates the number of branches (in percentage) successfully segmented.

These results were achieved with a one-click initialization. From the first point provided by the user, the initial direction is determined as the direction of minimal gradient variation. All patients presented some kind of artery pathologies in one, at least, of their coronary vessels. This means the Particle Filter successfully segmented both healthy and unhealthy coronaries. The method successfully detects all the main branchings, while in some cases smaller branchings at the lowest parts of the vessel tree, have been missed. Nevertheless, one can argue that their clinical use is of lower importance.

In this paper, we have shown that Particle Filters can be used for vascular segmentation. In the context of vascular segmentation, Particle Filters sequentially estimate the pdf of segmentations in a particular feature space. The case of coronary arteries was considered to validate such an approach, where the ability to handle discontinuities on the structural (branching) as well as appearance space (calcifications, pathological cases, etc.) was demonstrated. The main advantage of such methods is the non-linearity assumption on the evolution of samples. The use of an image term and a statistical model makes the probability measure both robust to pathologies, and yet, drives the segmentation toward the most probable solution given the statistical prior.

Future work consists in learning the variation law that rules the feature space toward better tests for hypotheses validation, as well as the one that controls process noise, to better guide the resampling stage toward an intelligent reduction of the required number of particles.

References

1. B. Avants and J. Williams. An adaptive minimal path generation technique for vessel tracking in cta/ce-mra volume images. In *MICCAI*, pages 707–716, 2000.
2. M. Descoteaux, L. Collins, and K. Siddiqi. Geometric Flows for Segmenting Vasculature in MRI: Theory and Validation. In *Medical Imaging Computing and Computer-Assisted Intervention*, pages 500–507, 2004.
3. A. Doucet, J. de Freitas, and N. Gordon. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, 2001.
4. P. Fearnhead and P. Clifford. Online inference for well-log data. *Journal of the Royal Statistical Society*, 65:887–899, 2003.
5. M. Figueiredo and J. Leita. A nonsmoothing approach to the estimation of vessel contours in angiograms. *IEEE Transactions on Medical Imaging*, 14:162–172, 1995.
6. A. Frangi, W. Niessen, P. Nederkoorn, O. Elgersma, and M. Viergever. Three-dimensional model-based stenosis quantification of the carotid arteries from contrast-enhanced MR angiography. In *Mathematical Methods in Biomedical Image Analysis*, pages 110–118, 2000.
7. N. Gordon. Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation. *IEE Proceedings*, 140:107–113, 1993.
8. N. Gordon. On Sequential Monte Carlo Sampling Methods for Bayesian Filtering. *Statistics and Computing*, 10:197–208, 2000.
9. N. Gordon. A Tutorial on Particle Filters for On-line Non-linear/Non-Gaussian Bayesian Tracking. *IEEE Transactions on Signal Processing*, 50:174–188, 2002.
10. M. Hart and L. Holley. A method of Automated Coronary Artery Tracking in Unsubtracted Angiograms. *IEEE Computers in Cardiology*, pages 93–96, 1993.
11. K. Krissian, G. Malandain, N. Ayache, R. Vaillant, and Y. Troussset. Model based detection of tubular structures in 3d images. *Computer Vision and Image Understanding*, 80:130–171, 2000.
12. R. Malladi and J. Sethian. A Real-Time Algorithm for Medical Shape Recovery. In *International Conference on Computer Vision*, pages 304–310, 1998.
13. D. Mumford and J. Shah. Boundary detection by minimizing functionals. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 22–26, 1985.
14. D. Nain, A. Yezzi, and G. Turk. Vessel Segmentation Using a Shape Driven Flow. In *Medical Imaging Computing and Computer-Assisted Intervention*, 2004.
15. T. O’Donnell, T. Boult, X. Fang, and A. Gupta. The Extruded Generalized Cylinder: A Deformable Model for Object Recovery. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 174–181, 1994.
16. S. Osher and N. Paragios. *Geometric Level Set Methods in Imaging, Vision and Graphics*. Springer Verlag, 2003.
17. R. Petrocchi, K. Manbeck, and J. Elion. Three Dimensional Structure Recognition in Digital Angiograms using Gauss-Markov Models. *IEEE Computers in Radiology*, pages 101–104, 1993.
18. D. Rueckert, P. Burger, S. Forbat, R. Mohiadin, and G. Yang. Automatic Tracking of the Aorta in Cardiovascular MR images using Deformable Models. *IEEE Transactions on Medical Imaging*, 16:581–590, 1997.
19. E. Sorantin, C. Halmai, B. Erbohelyi, K. Palagyi, K. Nyul, K. Olle, B. Geiger, F. Lindbichler, G. Friedrich, and K. Kiesler. Spiral-CT-based assesment of Tracheal Stenoses using 3D Skeletonization. *IEEE Transactions on Medical Imaging*, 21:263–273, 2002.
20. W. West. Modelling with mixtures. In J. Bernardo, J. Berger, A. Dawid, and A. Smith, editors, *Bayesian Statistics*. Clarendon Press, 1993.
21. P. Yim, P. Choyke, and R. Summers. Grayscale Skeletonization of Small Vessels in Magnetic Resonance Angiography. *IEEE Transactions on Medical Imaging*, 19:568–576, 2000.

Hybrid Segmentation Framework for Tissue Images Containing Gene Expression Data

Musodiq Bello¹, Tao Ju², Joe Warren², James Carson³, Wah Chiu³,
Christina Thaller³, Gregor Eichele⁴, and Ioannis A. Kakadiaris¹

¹ Computational Biomedicine Lab, Dept. of Computer Science,
University of Houston, Houston TX, USA

² Dept. of Computer Science, Rice University, Houston TX, USA

³ Verna and Marrs McLean Dept. of Biochemistry,
Baylor College of Medicine, Houston TX, USA

⁴ Max Planck Institute of Experimental Endocrinology, Hannover, Germany

Abstract. Associating specific gene activity with functional locations in the brain results in a greater understanding of the role of the gene. To perform such an association for the over 20,000 genes in the mammalian genome, reliable automated methods that characterize the distribution of gene expression in relation to a standard anatomical model are required. In this work, we propose a new automatic method that results in the segmentation of gene expression images into distinct anatomical regions in which the expression can be quantified and compared with other images. Our method utilizes shape models from training images, texture differentiation at region boundaries, and features of anatomical landmarks, to deform a subdivision mesh-based atlas to fit gene expression images. The subdivision mesh provides a common coordinate system for internal brain data through which gene expression patterns can be compared across images. The automated large-scale annotation will help scientists interpret gene expression patterns at cellular resolution more efficiently.

1 Introduction

With mammalian genomes of over 20,000 genes [1] now sequenced, the next challenge facing the biomedical community is to determine the function of these genes. Knowledge of gene function is important for a better understanding of diseases and the development of potential new therapies. The mouse is a well-established model system for exploring gene function and disease mechanisms. Consequently, determining where genes are active in different mouse tissues can lead to a greater understanding of how gene products affect human diseases. Non-radioactive *in situ* hybridization (ISH) is a histological method that can be applied to reveal cellular-resolution gene expression in tissue sections [2]. This is an appropriate resolution for addressing the questions about the role of genes in cell identity, differentiation, and signaling. Robotic ISH enables the systematic acquisition of gene expression patterns in serially sectioned tissues [3]. By organizing a large collection of gene expression patterns into a digital atlas, ISH data can be used to make great advances in functional genomics as DNA sequence databases have done.

A major step toward efficient characterization of gene expression patterns is the automatic segmentation of gene expression images into distinct anatomical regions and

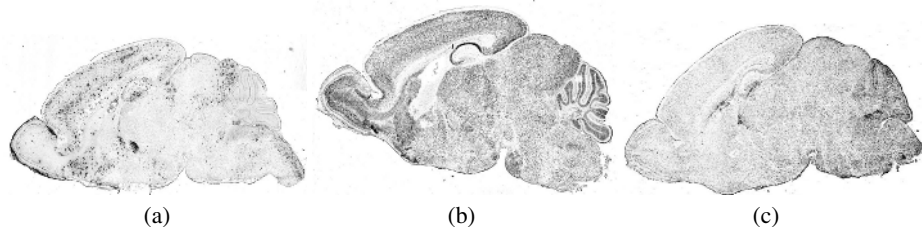


Fig. 1. Variation in the shape of the brain and expression pattern of (a) *Npy*, (b) *Cbfat2t1h*, and (c) *Neurog2* genes in mouse brain images

subregions. This is a challenging task, mainly because each gene is expressed differently from region to region leading to a substantial variation in the appearance of each image. There is also a natural variation in the shape of anatomical structures, compounded by the non-linear distortion introduced during sectioning of the brain. Moreover, there are many regions where no edges or intensity variation can be visually observed. Figure 1 depicts typical gene expression images.

To compare gene expression patterns across images, Ju *et al.* [4] constructed a deformable atlas based on subdivision surfaces which provide a common coordinate system when fitted to sagittal sections of the mouse brain. The 2D brain atlas is represented as a quadrilateral subdivision mesh, as shown in Fig. 2(a). Subdivision is a fractal-like process of generating a smooth geometry from a coarse shape [5]. Starting from an initial mesh, subdivision generates a sequence of refined meshes with increasing smoothness. In our application, the mesh is partitioned by a network of crease edges into sub-meshes, each modeling a particular anatomical region of the brain. The atlas was fitted to images using affine transformation to account for rotation and translation, and local deformation based on iterated least-squares to account for shape variation. However, the accuracy of the local fitting, and interior coordinate system resulting from the segmentation, is limited by its reliance on tissue boundary detection only. Thus, manual deformation of the internal regions of the atlas must still be performed.

In our previous work [6], we have extended that approach by identifying selected anatomical landmarks in expression images and using them to guide the fitting of internal regions of the mesh. Our method improved the general fitting of the internal regions,

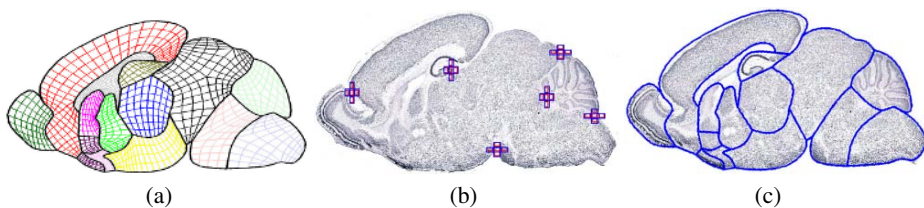


Fig. 2. (a) The atlas at subdivision level 2. (b) Feature-extracting templates overlaid on selected anatomical landmarks. (c) A typical gene expression image manually segmented into anatomical regions.

ensuring that specific landmarks were placed in appropriate regions. However, the region boundaries did not always match those drawn by neuroanatomists. In this paper, we propose a new hybrid segmentation framework that combines texture variation at region boundaries with textural features of specific landmarks to deform the subdivision atlas. In the rest of the paper, we explain the hybrid segmentation framework in detail in Section 2 and present results from using our algorithms in Section 3. Section 4 summarizes our work.

2 Hybrid Segmentation Framework

Our hybrid model is a triplet $\{S, B, L\}$ where S represents the shape of the subdivision mesh, B represents the appearance of the quads on the boundaries of anatomical regions, and L models the texture features of selected anatomical landmarks. The shape and boundary quad appearance are obtained for multiple mesh subdivision levels. Our framework consists of training and deployment stages. Training is performed on several mouse brain gene expression images which were previously fitted with a standard subdivision atlas by neuroanatomists (Fig. 2(c)). Deployment involves fitting the atlas to new gene expression images in order to segment them into anatomical regions.

2.1 Training

Shape: The shape term, S , defines the geometry and topology of the subdivision atlas (Fig. 2(a)) that will be fitted to each image. The geometry is a collection of the coordinates of the vertices of the mesh at a given subdivision level while the topology denotes the relationships between the vertices to form anatomical regions. The geometry is modeled as $\mathbf{x}^k = [x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n]^T$ for a mesh at subdivision level k , where $[x_i, y_i]$ are the Euclidean coordinates of vertex i . For all N meshes in the training set, the mean shape is obtained as: $\bar{\mathbf{x}}^k = \frac{1}{N} \sum_{i=1}^N \mathbf{x}^k$. A training instance that is close to the mean was selected as a standard mesh. The shape is obtained for different subdivision levels of the mesh in a multi-resolution approach.

Boundary quad features: The second element, B , of our hybrid model captures information about the features at the anatomical region boundaries. It can be observed that the cell density pattern in the cerebellum is different from that of its neighbors in most images. A texture variation can similarly be observed along the boundaries of the cortex, septum, and thalamus. This slight variation in the texture patterns of anatomical regions is utilized to model the boundary quads at subdivision level k as $B^k = [B_1^k B_2^k \dots B_s^k]$ for s selected segments in the mesh boundary. A boundary segment is a collection of adjacent crease edges and has quads from no more than two anatomical regions attached to it (Fig. 3). By separating the regional boundaries into segments, optimal features for each segment can be chosen, since no set of features will be equally optimal for all region boundaries. For each segment, $B_j = \{Q^j, F^j, p^j\}$ where Q^j is the set of all quads attached to segment j and distinguished by the side of the segment they belong to, F^j is the set of optimal features, and p^j is the set of classifier parameters to distinguish

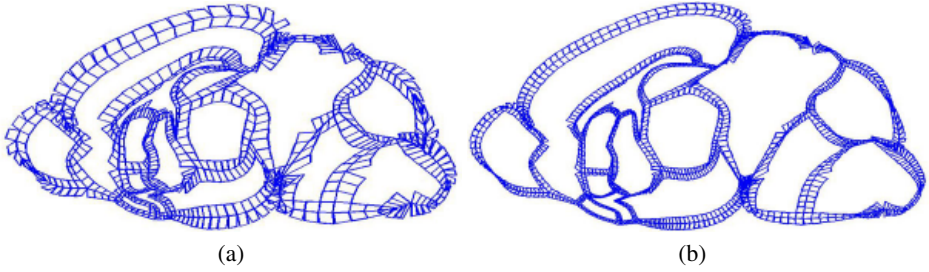


Fig. 3. Boundary quads at subdivision level (a) 2, and (b) 3

between quads on either side of the boundary segment. For purposes of this research, we have selected the Support Vector Machine (SVM) [7] classifier due to its ability to obtain a good decision boundary even on sparse data sets. The optimal feature set, F^j , and classifier parameters, p^j , for each segment are obtained as follows:

Step 1: Extract features from quads. Each image was filtered with Laws' texture filters [8] to obtain a bank of 25 texture energy maps. The 1st, 2nd, and 3rd moments of the distribution of intensity values of all pixels in each quad on either side of the boundary were used as feature for the quad in each filtered image [4].

Step 2: Feature normalization. The feature values vary widely, necessitating normalization. A few feature normalization techniques were considered, including linear scaling to unit range, linear scaling to unit variance, transformation to uniform random variable, and rank normalization [9]. We obtained the best performance with rank normalization in which the feature values f_1, f_2, \dots, f_m are first ordered to obtain their ranks (1 . . . m), where m is the number of samples in the boundary segment. Each feature value is then replaced by $\tilde{f}_i = \frac{\text{Rank}(f_i)-1}{m-1}$.

Step 3: Optimal feature selection. The relevance of each feature f is computed using the Information Gain (IG) metric [10] after discretization using Fayyad and Irani's minimum description length algorithm [11]. The features were then sorted according to the relevance indices assigned by IG and each feature is included one-at-a-time in a feature set, F^j . The average error, E_c , of classifying with the feature set F^j is then obtained in a 10-fold cross-validation and the smallest set of features, F^j , with a stable and sufficiently low value of E_c is selected.

Step 4: Model parameter computation. For each segment j , a SVM classifier was trained to distinguish between quads on either side of the boundary segment based on the optimal features. Best performance was obtained by using the Radial Basis Function (RBF) kernel [12] with SVM. The optimal values for the kernel parameter and error penalty parameter are obtained by cross validation and used to compute SVM model parameters, p^j , for each segment as part of the hybrid model.

Anatomical landmark features: In addition to the region boundary quads, a few anatomical landmarks were modeled with respect to their texture features. Each landmark is associated with a vertex and they are used to guide the general position and

orientation of the mesh during fitting. Each landmark i is modeled as $L_i = \{v_i, F^i, p^i\}$, where v_i is the coordinates of the vertex that it is attached to at subdivision level 3, the highest level used in the model. The set of optimal features, F^i , that can be used to distinguish a landmark from its surrounding area and the set of classifier parameters, p^i , are computed as described above. To extract the features for the landmarks, a rectangular template was overlaid on the landmark in each of the texture maps and summary statistics for sub-windows in the template were used as features. Similar features were extracted from a 4-neighborhood (Fig. 2(b)) of the landmark to serve as non-landmark examples as described in our previous work [6].

2.2 Deployment

Given a new image, the model is fitted to the image by minimizing a quadratic energy function $E(\mathbf{x}^k)$ of the form: $E(\mathbf{x}^k) = E_f(\mathbf{x}^k) + E_d(\mathbf{x}^k)$ using a linear solver such as conjugate gradient. The energy term $E_f(\mathbf{x}^k)$ measures the fit of the mesh at subdivision level k to the image and $E_d(\mathbf{x}^k)$ measures the energy used in deforming the mesh. The fitting term $E_f(\mathbf{x}^k)$ is: $E_f(\mathbf{x}^k) = \alpha E_U(\mathbf{x}^k) + \beta E_B(\mathbf{x}^k) + \gamma E_L(\mathbf{x}^k)$, where $E_U(\mathbf{x}^k)$ is the fitting error of the outer boundary of the mesh to the outer boundary of the image, $E_B(\mathbf{x}^k)$ measures the fitting error of the regional boundaries resulting from the classification of the boundary quads, and $E_L(\mathbf{x}^k)$ measures the error of fit of the anatomical landmarks. The formulation of $E_U(\mathbf{x}^k)$ and the deformation energy term $E_d(\mathbf{x}^k)$ are the same as in [4]. The method for obtaining the other terms $E_L(\mathbf{x}^k)$ and $E_B(\mathbf{x}^k)$ is described below.

Step 1 - Shape Initialization: First, a global alignment of the reference shape to the image is performed. The image was segmented from the background using a flood-filling approach after a simple intensity threshold. Principal Component Analysis was applied to obtain the principal axes of the segmented image. The principal axes of the reference mesh are also obtained and an affine transformation of the mesh is performed to align the two pairs of axes [4].

Step 2 - Using the Landmarks for Fitting: Second, the fitting error, $E_L(\mathbf{x}^k)$, of the mesh to the landmarks is computed as $\sum_i (l_i - v_i)^2$, where v_i is the vertex of the mesh associated with landmark i detected at location l_i . Specifically, for each landmark, features are extracted for the pixels around the expected location and classified using the model parameters obtained at the training stage. For efficiency, classification is performed on every third pixel initially before conducting a pixel-by-pixel search around the area with the highest SVM decision values. This is possible because the decision values were found to monotonically increase towards the expected ground truth in all the images tested [6].

Step 3 - Using Boundary Quads for Fitting: Finally, for each segment j on the boundary at subdivision level 1, optimal features F^j are extracted for the quads Q^j in the segment and the model parameters p^j are used to classify each quad. There are four possibilities when two quads on opposite sides of a crease edge are classified with respect to the model. When the classification of both quads is in agreement with the model (Fig.

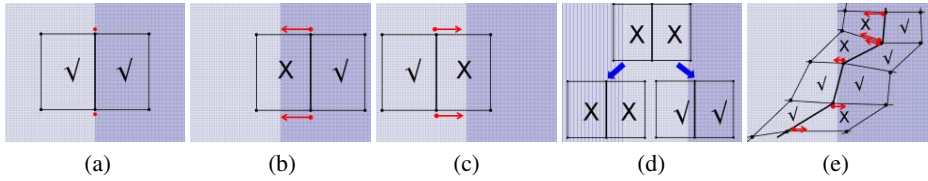


Fig. 4. Classifying opposite quads on a boundary crease edge (the background depicts two regions): (a) both are classified in accordance with the model, no displacement at the vertices. (b,c) Displacement of the boundary edge towards the misclassified quad. (d) Rare case of opposite quads being simultaneously incorrectly classified with respect to the model: position the quads on either side and select best match. (e) Various scenarios of quad classification and the resulting displacement at the vertices.

4(a)), the force exerted by the corresponding vertices is zero. When either of the quads is classified contrary to the model (Fig. 4(b,c)), the vertices exert a force pulling the boundary edge in the direction of the misclassified quad. The decision value returned by SVM gives an estimate of the confidence in the classification. In the event that both quads are incorrectly classified with respect to the model (this is very rare since the mesh boundaries are already quite close to the image boundaries after global fitting), the two quads are temporarily positioned on both sides of the segment and the position that results in correct classification is retained (Fig. 4(d)). If both positions still result in misclassified quads, the vertices are left unchanged. The various possibilities are illustrated in Fig. 4(e).

This process is performed iteratively until a specified ratio of the quads (e.g., 95%) are correctly classified, in which case the process is repeated at a finer subdivision level, up to level 3. With increasing subdivision level, the size of the quads decreases. This reduces the displacement of the vertices, resulting in a smooth fit.

3 Results and Discussion

Our experimental data are 2D images of sagittally sectioned (level 9) postnatal day 7 (P7) C57BL/6 mouse brains on which *in situ* hybridization has been performed to reveal the expression of a single gene. For computational efficiency, the images were scaled down by 25% from their original size of approximately 2000x3500 pixels. We trained our framework on 36 images manually fitted with subdivision meshes by neuroanatomists, and tested on 64 images. The weights for the terms $E_U(\mathbf{x}^k)$, $E_B(\mathbf{x}^k)$, and $E_L(\mathbf{x}^k)$ in the energy minimization equations were selected experimentally.

To quantify the quality of fit using our hybrid segmentation approach, we compared individual anatomical regions as delineated by our framework with those manually delineated by neuroanatomists. The number of overlapping pixels in both meshes was further normalized by the total number of pixels in the manually fitted mesh. In Fig. 5(a-e), the accuracy of fit of five major anatomical regions is compared for all 64 images. In regions where the distinction along the boundaries is pronounced such as the cortex, cerebellum and their adjoining regions, we expect to have higher

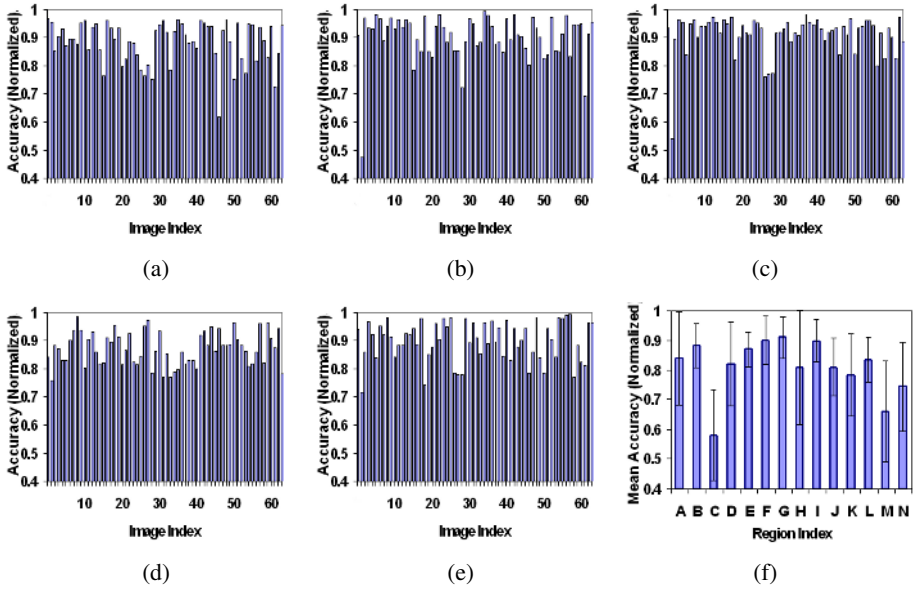


Fig. 5. Comparison of the accuracy of the fit in the (a) cortex, (b) medulla, (c) midbrain, (d) hypothalamus, and (e) pons in 64 expression images. (f) Mean accuracy of the fit in the 64 images across 14 regions using our hybrid segmentation framework. The error bars indicate the standard deviation. The letters indicate the following regions: A-cerebellum, B-cortex, C-basal forebrain, D-hippocampus, E-hypothalamus, F-medulla, G-midbrain, H-olfactory bulb, I-pons, J-septum, K-striatum, L-thalamus, M,N-ventral striatum.

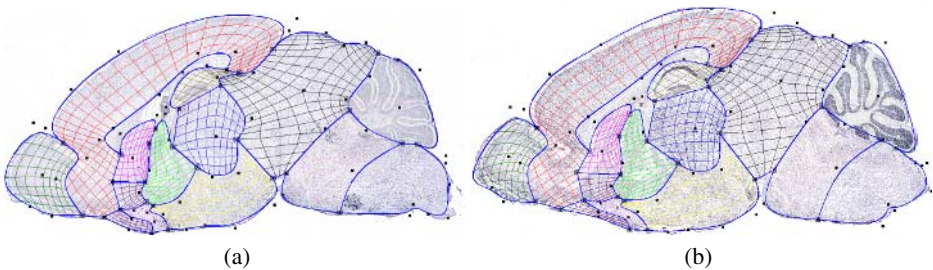


Fig. 6. The result of fitting the standard mesh on (a) *ChAT*, and (b) *BMALI* gene expression images

accuracy in delineating the boundaries. Similarly, we expect lower accuracy in regions where this distinction is minimal, such as in the forebrain and the ventral striatum. This is confirmed by our results summarized for all 14 regions in Fig. 5(f). Note that the two sub regions of the ventral striatum are treated as two separate regions for purposes of comparison. Examples of fitting using our approach are illustrated in Fig. 6.

4 Conclusion

Due to its many advantages, subdivision surface modeling is becoming increasingly popular for geometric modeling and is also starting to appear in medical applications. For example, in gene expression images, subdivision surface modeling facilitates the comparison of expression patterns not only in regions, but also in subregions of the brain. The challenge is to fit a subdivision-based atlas to expression images accurately and automatically. Our approach combines the detection of selected anatomical landmarks with feature differentiation at regional boundaries using trained classifiers. Our results have been very encouraging.

Acknowledgements. This work was supported in part by a training fellowship from the W.M. Keck Foundation to the Gulf Coast Consortia through the Keck Center for Computational and Structural Biology.

References

1. Waterston, R., and the Mouse Genome Sequencing Consortium: Initial sequencing and comparative analysis of the mouse genome. *Nature* **420** (2002) 520–562
2. Albrecht, U., Lu, H.C., Revelli, J.P., Xu, X.C., Lotan, R., Eichele, G.: Studying Gene Expression on Tissue Sections Using In Situ Hybridization. In: *Human Genome Methods*. CRC Press, Boca Raton (1997) 93–119
3. Carson, J., Thaller, C., Eichele, G.: A transcriptome atlas of the mouse brain at cellular resolution. *Curr Opin Neurobiol* **12** (2002) 562–565
4. Ju, T., Warren, J., Eichele, G., Thaller, C., Chiu, W., Carson, J.: A geometric database for gene expression data. In: *Eurographics Symposium on Geometry Processing*, Aachen, Germany (2003) 166–176
5. Warren, J., Weimer, H.: *Subdivision Methods for Geometric Design: A Constructive Approach*. Morgan Kaufmann Publishers, San Francisco, CA (2002)
6. Kakadiaris, I.A., Bello, M., Arunachalam, S., Kang, W., Ju, T., Warren, J., Carson, J., Chiu, W., Thaller, C., Eichele, G.: Landmark-driven, atlas-based segmentation of mouse brain tissue images containing gene expression data. In: *Proc. 7th International Conference on Medical Image Computing and Computer-Assisted Intervention*, Rennes, France (2004) 192–199
7. Vapnik, V.N.: *The Nature of Statistical Learning Theory*. Springer-Verlag (2000)
8. Laws, K.: *Textured Image Segmentation*. PhD thesis, USC (1980)
9. Aksoy, S., Haralick, R.: Feature normalization and likelihood-based similarity measures for image retrieval. *Pattern Recognition Letters* **22** (2001) 563–582
10. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. *Journal of Machine Learning Research* **3** (2003) 1157–1182
11. Fayyad, U.M., Irani, K.B.: On the handling of continuous-valued attributes in decision tree generation. *Machine Learning* **8** (1992) 87–102
12. Schölkopf, B., Smola, A.: *Learning with Kernels*. MIT Press, Cambridge, MA (2002)

Fully Automatic Kidneys Detection in 2D CT Images: A Statistical Approach

Wala Touhami, Djamel Boukerroui, and Jean-Pierre Cocquerez

HEUDIASYC, UMR CNRS #6599, Université de Technologie de Compiègne,
BP 20529 - 60205 Compiègne Cedex, France
{ouala.touhami, djamal.boukerroui, jean-pierre.cocquerez}@hds.utc.fr

Abstract. In this paper, we focus on automatic kidneys detection in 2D abdominal computed tomography (CT) images. Identifying abdominal organs is one of the essential steps for visualization and for providing assistance in teaching, clinical training and diagnosis. It is also a key step in medical image retrieval application. However, due to gray levels similarities of adjacent organs, contrast media effect and relatively high variation of organ's positions and shapes, automatically identifying abdominal organs has always been a challenging task. In this paper, we present an original method, in a statistical framework, for fully automatic kidneys detection. It makes use of spatial and gray-levels prior models built using a set of training images. The method is tested on over 400 clinically acquired images and very promising results are obtained.

1 Introduction

Recently, a clinically operational Content-based image retrieval (CBIR) system, based on a semi-supervised learning approach and dedicated to help osteo-articular diagnosis on Magnetic Resonance (MR) images, is developed by our research team [1]. Needless to emphasize that the retrieval success is highly dependent on the images description stage. The purpose of our actual work is to develop an extension of the system to be applied to computed tomographic (CT) images of the abdomen, specifically to kidneys cysts. Unlike, the osteo-articular MR images, abdominal CT images present a higher variability. Indeed, different tissues (soft and hard) with different sizes and shapes across individuals and across slices may appear. This increases the complexity and the hardness of the image description step in a sense that a description index constructed using the whole image may be inefficient in the retrieval step. Salient features and relevant descriptors, in our application, are those which encode information about the kidney and the cysts. All other features, salient or not, could be considered as outliers. Therefore, the detection of the regions of interest (kidneys in our case) is a prerequisite step for the success of the retrieval process.

Automatically identifying organs from abdominal CT images series is challenging. Typically, the identification task is related to the segmentation problem, and few works treated this problem. For instance, Lee et al [2] solve the identification problem using fuzzy rules established based on knowledge of anatomy.

Overlap information from consecutive slices is also used to guide the recognition process. Kobashi and Shapiro [3] integrate the recognition process with the segmentation one. The authors use anatomy knowledge to dynamically constrain a thresholding based segmentation. Here too shape constraints and overlap information from consecutive slices are used. These methods operate on 3D data and, to our knowledge, on healthy organs. The problem we address in this paper is a *fully automatic detection of kidneys with cysts in a 2D CT slices*. The presence of cysts makes the task more difficult because of the considerable increase of shape variability and intensity inhomogeneity. This discards the use of active shape models (ASM) or even appearance models (AAM) to guide the identification process. The construction of a probabilistic atlas is a potential alternative (see eg. [4]). Unfortunately, we are considering the problem in 2 dimensions¹ and the slice level may differ significantly making the approach unsuitable as different organs may appear/disappear. Besides this type of methods requires elastic registration, which adds a computational complexity to the problem.

In this paper, an original solution to the problem is proposed in a statistical framework. The outline of this paper is as follows. First, we start by the description of the mathematical modeling of the detection method and by specifying the different terms of the optimization problem to be solved. Section 3 describes the preprocessing steps and the models building. Section 4 presents results on clinically acquired data, and the summary and conclusion appear in Section 5.

2 Modeling

Our approach for the automatic kidneys detection makes use of spatial and gray-levels prior models built using a set of training images. Then, the detection problem is solved as a minimization of a cost function that takes into account the observed image and the prior models. This is described in details next.

2.1 Prior Information

Let Ω be the image spatial support and \mathcal{L} denotes the set of image gray-level values. We suppose that each image is a realization of a random field $Y = \{Y_s, s \in \Omega\}$ and we define a binary random field $X = \{X_s, s \in \Omega\}$ where $X_s = 1$ means that the pixel s is in the Kidney and $X_s = 0$ otherwise.

- **Spatial prior model:** Given a set of N training images $\{y^i, i = 1 \dots N\}$, we can construct on each image i , a realization x^i of the random field X^i by segmenting manually the kidney regions. Then, we empirically estimate a spatial prior pdf, $p_K(s)$, of the kidney areas using the mean field as follow:

$$p_K(s) = \frac{\sum_{i=1}^N \frac{x_s^i}{n_i}}{N}, \quad (1)$$

¹ The application (CBIR) imposes the 2 dimensional constraint; as well as the need for a fully automatic solution to the problem.

where n_i is the number of pixels of the i^{th} image such as $x_s^i = 1$. Using the above equation, we can build a spatial prior pdf for the left and right kidneys p_{K_l} and p_{K_r} respectively.

- **Gray-level prior:** We define ω a random variable taking values in \mathcal{L} and we suppose that the random field Y is an ergodic process in the same tissue (ie. $X_s = 1$). In the same manner as the spatial model, we can build empirically a prior pdf of the kidney's gray-level values. Given the above assumptions, the prior gray-level pdf, $p_\omega(w)$ is position independent and is given by:

$$p_\omega(w) = \frac{1}{N} \sum_{i=1}^N \left\{ \frac{1}{n_i} \sum_{s/x_s^i=1} \delta(y_s^i - w) \right\} . \quad (2)$$

Here y_s^i is the gray-level value of pixel s in image i and δ is the Delta function.

2.2 Cost Function Definition

The proposed kidney identification method takes the form of an optimization of a cost function in order to detect two square windows R_l and R_r , of side length $2n + 1$, respectively in the left and right kidneys. The general form of the cost function $J(R_l, R_r)$, to be minimized, is defined as follow:

$$J(R_l, R_r) = \sum_{j=\{l,r\}} \{D_{int}(R_j) + D_{spat}(R_j)\} + S_{sym}(R_l, R_r) . \quad (3)$$

Here, $D_{int}(R_j)$ is the data term for region R_j , $D_{spat}(R_j)$ is the spatial prior probability of region R_j and $S_{sym}(R_l, R_r)$ is a similarity term between the left and the right window. These terms will be detailed in the following paragraphs.

- a) **Data term:** It measures how likely the gray-level distribution, $p_{\omega_j}(w)$, of the given region R_j , $j \in \{l, r\}$ is similar to the prior pdf $p_\omega(w)$ given by (eq. 2). This is a classical goodness-of-fit problem and several measures could be used. Two statistical measures are chosen, namely, Kuiper statistic $D_{KP}(p_\omega, p_{\omega_j})$ and Kullback-Leiber divergence $D_{KL}(p_\omega, p_{\omega_j})$.
- b) **Left/Right similarity measure:** Recall that both square windows belong to the two kidneys and therefore are statistically dependant. Hence, we expect a high similarity between the two regions R_l and R_r . Here too, several criterions could be used to measure this similarity. We select two measures, widely used by the computer vision community: the normalized cross-correlation $NCC(R_r, R_l)$ which measures the linear dependance and the mutual information $MI(R_r, R_l)$, which measures the statistical dependance of the two areas. Note that both measures have to be maximized².
- c) **Spatial prior probability:** Let Ω_j be the spatial support of the region R_j and $S_j = \{s, p_{K_j}(s) > 0\}$, $j \in \{r, l\}$. Using the spatial prior model (eq. 1), the probability that $\Omega_j \subset S_j$ is given by: $P(\Omega_j) = \sum_{s \in \Omega_j} p_{K_j}(s)$.

² Because the left kidney would probably be a reflected version of the right, a left-right flipping operation is performed before similarity calculation.

Now that we defined all the different terms, we rewrite eq. (3) as follow:

$$J(R_l, R_r, \Theta) = \sum_{j=\{l,r\}} \{D.(p_{\omega_j}, p_{\omega}) - \lambda_1 \ln(P(\Omega_j))\} - \lambda_2 S.(R_l, R_r) , \quad (4)$$

where $D.(.) = D_{KP}(.)$ or $D_{KL}(.)$, $S.(.) = NCC(.)$ or $MI(.)$ and $\lambda_1 \in \mathbb{R}^+$, $\lambda_2 \in \mathbb{R}^+$ are hyperparameters. The above cost function is minimized with respect to the variable $\Theta = \{s_l, s_r\}$ representing the respective left and right window centroid.

The optimization algorithm employed uses the Nelder-Mead simplex search algorithm for multidimensional unconstrained minimization [5]. It is deterministic method that attempts to minimize a scalar-valued function without any gradient information. The major problem of such local minimizer is the initialization. For our case, we choose to initialize the algorithm with 8 random candidates points with nonzero spatial pdf.

3 Kidneys Detection Algorithm

A block diagram of our new fully automatic kidneys detection approach is shown in Fig. 1. In this section we give algorithmic details about the preprocessing steps, the model building and the whole kidney delineation algorithm.

3.1 Images Database

All our images are routine acquisition in a public hospital. The images are acquired by several radiologists over 1 year time period (# patients: 35 male, 12 female; age range: 19 to 88 years). In some cases a contrast media is administered, and 1 to 3 acquisitions are performed at different time delays. Consequently, there are images with injection of contrast media (+IV) and other without (-IV). The images are acquired using different spatial resolutions (0.5703 to 0.9746 mm) and slice thicknesses. An experienced radiologist selected 502 images (435 test images and 67 training images), representatives of the different cysts, which form our data set.

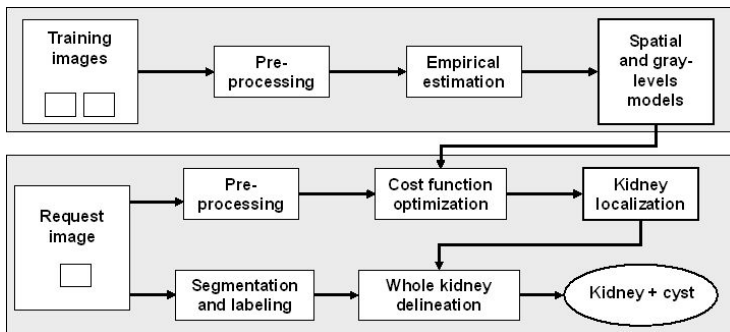


Fig. 1. Block diagram of the proposed automatic kidney detection method

3.2 Preprocessing

In order to build or to make use of the spatial prior model, all the images must be in the same spatial referential. Ideally, all the images have to be elastically registered to a chosen reference image. Unfortunately, 2D elastic registration is not appropriate and will fail in several cases because of the high variability of the images content. Here, the variability is not only due to the soft nature of the tissues but also because of the significative variation, across individuals, of the position of the slice of interest. We found out that an accurate detection of the body area (without the skin and the fatty layer) is sufficient for our application. The detection algorithm and the resizing operation are detailed hereafter:

a) **Region of interest (ROI) delineation:**

A first threshold is used to detect the field of view and creates a binary mask, the exterior mask. Morphological operations (holes closing, opening and erosion) are then used to erase the staff table and reduce the mask in order to erase the skin. A second threshold, higher than the first one is chosen, to eliminate the fatty layer under the skin, followed by morphological operation to remove detected isolated pixels in the fatty layer. Finally, the two binary masks are combined and used to delineate the region of interest (ROI). We have tested this algorithm on the whole date base and found it very efficient with 100% success.

b) **Resizing operation:** Recall that the images in the database have different spatial resolutions. Hence, first a resizing operation is used to make all the ROIs have the same pixel size ($1mm^2$). The result, is a set of ROIs with a high variability of dimensions (x^i, y^i) and more importantly with different aspect ratios. In order to put all the images in the same reference, we have to resize all the ROIs at a given fixed dimension (x, y). This operation will introduce shape distortions (i.e differences in axis resolution). We found that $(\hat{x}, \hat{y}) = (\mathbb{E}[x^i], \mathbb{E}[y^i])$ is close to the optimal minimizer of the expected normalized quadratic error of the resolution differences between the two axis:

$$\mathbb{E} \left[\left(\frac{\frac{x}{x^i} - \frac{y}{y^i}}{\frac{x}{x^i} + \frac{y}{y^i}} \right)^2 \right], \quad \text{where } \left(\frac{x}{x^i}, \frac{y}{y^i} \right) \text{ represent pixel resolutions.}$$

3.3 Models Building

For the construction of the spatial and gray-levels prior models, we use a training set of 67 images (30 (-IV) and 37 (+IV)), randomly selected from the database. The images are manually segmented and preprocessed as detailed above. Then, two spatial models are built separately, for the right and left kidney. Two gray-levels pdfs are also built for the images with and without contrast media.

3.4 Kidneys Detection

Starting from the detected window R_l (or R_r) inside the kidney, we utilize the corresponding segmentation map of the image in order to detect the whole

kidney. The segmentation is performed, in a Bayesian framework, using a maximum a posteriori criteria (MAP) where the data term is modeled using a finite gaussian mixture model and the prior term is modeled by means of a hidden markov random field (HMRF). Specifically, we use the HMRF-Expectation-maximization algorithm, proposed by Zhang et al. [6]. However, the minimization is performed using a coarse to fine strategy based on a combination of a multi-resolution model (for the observed data) and a multiscale model (for the hidden field) [7]. The segmentation is fully automatic.

Our kidney delineation algorithm operates on the connected components of the segmentation map obtained using Suzuki's et al. algorithm [8]. An erosion operation on the fatty cluster - easily detected given its HU (Hounsfield Units) densitometric range - is first applied before the labeling operation in order to disconnect the kidney from other close organs (liver, spleen). This segmentation imperfection happens when the usual fatty layer surrounding the kidneys is too thin because of the closeness of the two organs. Hence, the delineation algorithm takes as input data the set of labels included in the previously detected window R_j , $j \in \{l, r\}$. Then, based on prior knowledge of abdominal anatomy, we define a set of IF THEN rules ($\simeq 7$ rules) operating on a set of region's properties (related to shape, position, densitometric range, relative position, predefined prior models) in order to delineate the whole kidney. Briefly, we search for the bounding box that surely includes the kidney and the cysts, defined as the union of all the regions included in R_j (the regions included in the defined bounding box constitute our candidates set). An adjusting operation utilizing the predefined rules is then performed. The selected regions constitute the kidney+cyst area.

4 Results

The proposed method is tested on the abdominal CT database detailed in section 3.1 (435 images, i.e without the training images). Illustrative detection results are shown in figure 2. The original images are viewed between -135 and 215 HU. The detected windows (R_l, R_r), the rectangle delineating the whole kidney and the rectangle delineating the ROI are highlighted. We can observe images without (first 2 columns) and with (next 2 columns) contrast media. These examples illustrate the robustness of our automatic kidneys detection approach and underline the high variability of image content, in which the kidneys have neither the same appearance nor the same spatial positions. The last column shows typical cases of failure, mainly due to small kidneys (the circles show the correct kidneys location).

The algorithm presented in this paper has three free parameters that needed to be optimized, namely, n for the window size and (λ_1, λ_2) the tradeoffs between the different energy terms of the cost function. Experiments are carried out for different combinations of the statistical measures presented earlier in section 2.2 in order to choose the best parameter setting. To this end, we use the success rate as the selection criterion: we consider a case to be successful if a kidney (left or right) is correctly detected. That means that we calculate the success rate according to the number of kidneys not according to the number

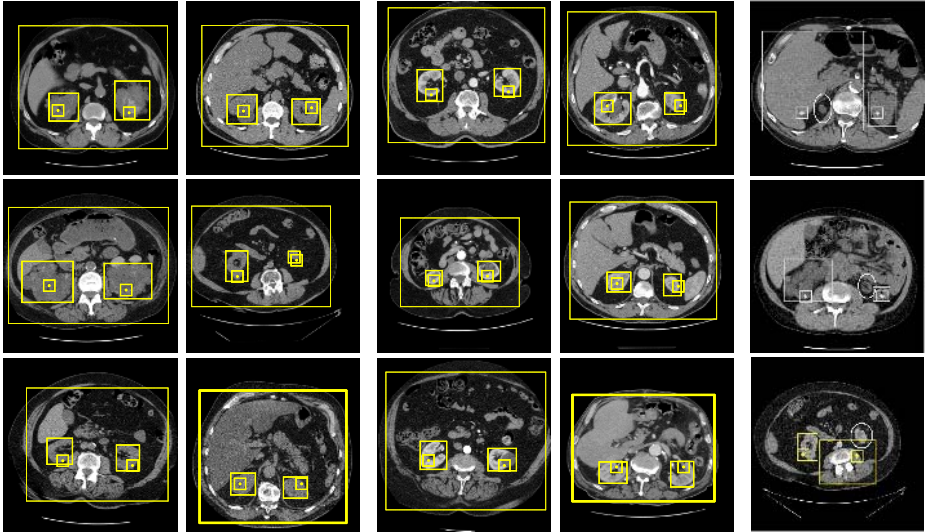


Fig. 2. Results of kidneys detection (see text for description)

Table 1. Best success rate of the window localization: comparison of different combinations of the statistical measures. $n = 16$ is the best window size of all combinations.

	# images		Kuiper test	Kullback test
(-IV)	196	NCC	97.95% $(\lambda_1, \lambda_2) = (1.0, 0.00)$	97.44% $(\lambda_1, \lambda_2) = (2.2, 0.37)$
		MI	97.44% $(\lambda_1, \lambda_2) = (1.3, 0.12)$	97.44% $(\lambda_1, \lambda_2) = (2.2, 0.37)$
(+IV)	239	NCC	95.78% $(\lambda_1, \lambda_2) = (1.6, 0.37)$	94.06% $(\lambda_1, \lambda_2) = (2.2, 0.00)$
		MI	94.91% $(\lambda_1, \lambda_2) = (1.4, 0.37)$	94.72% $(\lambda_1, \lambda_2) = (3.6, 0.25)$

of images. The three free parameters are optimized for each combination of the statistical measures (see table 1). Two conclusions can be drawn from this experiment. First, the success rates for the images (-IV) are significantly better than those of the images (+IV) and this is true for all measures combinations of the statistical measures presented earlier. This result is expected considering the increase of heterogeneity caused by the injected contrast media. Second, there is no significant differences, in terms of success rate, between all the different combinations. However, we can argue that Kullback-Leiber divergence needs a higher weight for the spatial prior (λ_1). Notice that the data term has always a weight of 1 in our formulation. We also present in Table2 preliminary results, however promising, for the delineation of the whole kidney, which of course are expected to be lower than those for the window localization (about -2%; indicating an acceptable success rate for the delineation algorithm of about 95%).

Finally, in order to assess the importance and the efficiency of the preprocessing step, we carried experiments in the same conditions as for the results shown in table1, however, without the preprocessing step. We recorded a drop of performances of about 20% for all combinations.

Table 2. Success rate detection for the Kuiper-NCC combination

	# images	Rate of R_j localization	Rate of kidney detection
Kidneys (-IV)	196	97.49%	95.01%
Kidneys (+IV)	239	95.78%	93.75%

5 Conclusion

In this paper, a *fully automatic* approach for kidneys detection in 2D abdominal CT images is proposed. It involved a two steps approach. A localization step, which make use of a statistical spatial and gray-levels priors, followed by delineation step of the whole organ. An efficient and fast preprocessing step was also proposed in order to make all the images in the same spatial referential. The success rate of the proposed method was satisfactory; especially for images without contrast media (above 97%). Therefore, this approach will be useful to compute relevant descriptors of the regions of interest (kidney + cysts) to be utilized in a CBIR system. Further work is in progress in order to improve the detection on images with contrast media (currently $\simeq 95\%$). We mainly focus on improving the optimization method and on defining a confidence measure for the detection process by automatically identifying false detections. In addition, while in the current work, the main focus was on the localization of the kidney, in future work, we will concentrate on improving the kidney delineation algorithm.

Acknowledgements: We are grateful to Prof. A. Remond and to all the medical staff of the radiology department of CHU of Amiens for the data acquisition. This work was funded by ‘Contrat de Plan Etat/Région Picardie 2000-2006’.

References

1. Najjar, M., Ambrose, C., Cocquerz, J.P.: Image retrieval using mixture models and EM algorithm. In: 13th Scandinavian Conf. SPIE (2003) 1114–1121
2. Lee, C.C., Chung, P.C., Tsai, H.M.: Identifying multiple abdominal organs from CT image series using a multimodule contextual neural network and spatial fuzzy rules. *IEEE Trans. Inf. Technol. Biomed.* **7** (2003) 208–217
3. Kobashi, M., Shapiro, L.G.: Knowledge-based organ identification from CT images. *Pattern Recognit.* **28** (1995) 475–491
4. Park, H., Bland, P.H., Meyer, C.R.: Construction of an abdominal probabilistic atlas and its application in segmentation. *IEEE TMI* **22** (2003) 483–492
5. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, NY (1992)
6. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE TMI* **20** (2001) 45–57
7. Boukerroui, D., Baskurt, A., Noble, J.A., Basset, O.: Segmentation of ultrasound images - multiresolution 2D and 3D algorithm based on global and local statistics. *Pattern Recognit. Lett.* **24** (2003) 779–790
8. Suzuki, K., Horiba, I., Sugie, N.: Linear-time connected-component labeling based on sequential local operations. *Comput. Vis. Image Underst.* **89** (2003) 1–23

Segmentation of Neighboring Organs in Medical Image with Model Competition

Pingkun Yan¹, Weijia Shen¹, Ashraf A. Kassim¹, and Mubarak Shah²

¹ Department of Electrical & Computer Engineering,
National University of Singapore

² School of Computer Science, University of Central Florida, USA

Abstract. This paper presents a novel approach for image segmentation by introducing competition between neighboring shape models. Our method is motivated by the observation that evolving neighboring contours should avoid overlapping with each other and this should be able to aid in multiple neighboring objects segmentation. A novel energy functional is proposed, which incorporates both prior shape information and interactions between deformable models. Accordingly, we also propose an extended maximum *a posteriori* (MAP) shape estimation model to obtain the shape estimate of the organ. The contours evolve under the influence of image information, their own shape priors and neighboring MAP shape estimations using level set methods to recover organ shapes. Promising results and comparisons from experiments on both synthetic data and medical imagery demonstrate the potential of our approach.

1 Introduction

Segmentation of anatomical structures from medical images is often the first step in computer aided diagnosis. Further analysis highly depends on the quality of the segmented structures. In recent years, geometric deformable models, or level set methods [1, 2], have been applied to medical image segmentation. During the past decade, tremendous efforts have been put into different medical image segmentation applications based on level sets [3]. Many new algorithms have been reported to increase the precision and robustness of level set methods.

When segmenting or localizing an anatomical structure, prior knowledge is usually very helpful. The incorporation of more specific prior information into deformable models has received great attention. Several methods of incorporating prior shape information into boundary determination have been developed. Cootes *et al.* [4] made a breakthrough by constructing statistical shape models using corresponding points across a set of training images for image segmentation. Leventon *et al.* [5] incorporate statistical shape influence into the evolution process of geodesic active contours [6] by embedding each shapes in the training data set as the zero level set of level set maps. Qu *et al.* [7] add an enriched speed term incorporating curvature, shape and texture information into the speed function. More recently, Xie *et al.* [8] utilize both texture and shape priors when

defining their energy functional and segmentation is achieved through minimizing the functional. Nevertheless, it should be noticed that while most of level set based methods applied in medical image segmentation deal with only one level set and try to incorporate more prior information to achieve robust results, not much attention has been put into either multiple initializations of level sets or the interactions between multiple distinct level sets. Actually, structures to be segmented are usually surrounded by one or more adjacent anatomical structures. Taking the neighboring structure into consideration may contribute to overcoming the four drawbacks of level sets mentioned in [3], such as leaking through gaps, embedding of the objects.

Zeng *et al.* [9], proposed two coupled surfaces under level set framework to segment cortex from 3D MR images, which is mainly based on the near constant thickness of the cortical mantle. Most recently, Yang *et al.* [10] introduce neighbor-constrained 3D segmentation using a level set based deformable model. However, the neighboring contours may intersect with each other under some conditions because intersection is not penalized in the modeling process.

In medical images, distinct boundaries may not exist between the organs to enable edge detection of organ boundaries. Furthermore, the boundaries may be blurred and ambiguous due to partial volume effects [11]. These problems make the segmentation of neighboring organs a challenging task.

In this paper, we present our method which introduces interactions between neighboring contours when carrying out the segmentation process. A novel approach is proposed for segmenting multiple neighboring objects using both prior shape information and interactions between models. The idea of introducing model competition is inspired by the observation that the radiologists segment organs with poorly defined boundaries by considering the anatomical neighboring structures. In our algorithm, the contours evolve locally based on both image forces and interactions between neighboring shape models. They are also under the influence of their own shape estimates globally to extract the organ contours.

2 Segmentation of Neighboring Organs

2.1 Model Description

Consider an image I that has M objects of interest, which may connect with each other and have similar intensities or texture. Due to partial volume effects, the boundaries may be blurred, which renders many segmentation methods to fail. Our method integrates mutual prior shape information and repulsive interaction between evolving neighboring curves into the segmentation process.

In our approach, curves C_i ($i = 1, \dots, M$) are used to approach the contours of the objects. These curves are allowed to evolve simultaneously to segment the image. These curves evolve partly according to image information as in [12]. Each object has its own shape estimate \hat{C}_i (estimation details are given in Section 2.3), to guide the evolution of the curve C_i in segmentation process. To achieve this, the non-overlapping area of regions enclosed by C_i and \hat{C}_i needs to be minimized. Similarly, the repulsive force between different curves is defined

through minimizing the overlapping area between them. To make the algorithm robust and with lower complexity, the overlapping areas of curve C_i and other objects' shape estimates \hat{C}_j ($j = 1, \dots, M$ and $j \neq i$), instead of other curves C_j , are considered. The overall energy functional E is defined by

$$\begin{aligned}
 E = & \sum_{i=1}^M \left\{ \lambda_{1i} \int_{in(C_i)} |I(x, y) - c_{1i}|^2 dx dy + \lambda_{2i} \int_{out(C_i)} |I(x, y) - c_{2i}|^2 dx dy \right\} \\
 & + \sum_{i=1}^M \mu_i \oint_{Q_i} dq + \sum_{i=1}^M \nu_i A_i + \sum_{i=1}^M \xi_i A(NONoverlap(C_i, \hat{C}_i)) \\
 & + \sum_{i=1}^M \sum_{j=1, j \neq i}^M \omega_{ij} A(overlap(C_i, \hat{C}_j)) \tag{1}
 \end{aligned}$$

where $\lambda_i, \mu_i, \nu_i, \xi_i$, and ω_{ij} are real positive parameters. Parameter ξ_i controls the influence of its own shape estimate. Parameter ω_{ij} controls the repulsive force on the i_{th} object from the j_{th} object, while A denotes area in 2D or volume in 3D.

2.2 Level Set Evolution of the Model

In the level set formulation, contour let C_i denote the contour of the i_{th} object being segmented, which is embedded as the zero level set of a level set map Ψ_i , i.e., $C_i(t) = \{(x, y) | \Psi_i(x, y, t) = 0\}$. The signed distance transform is used to generate the level set function Ψ_i according to contour C_i . Ψ_i is defined to be positive outside C_i and negative inside C_i . Each of the M objects being segmented in the image has its own contour C_i and corresponding Ψ_i .

The energy functional (1) is formulated in level sets using regularized versions of the Heaviside function H and the Dirac function δ , denoted by H_ϵ and δ_ϵ [12], respectively.

$$\begin{aligned}
 E = & \sum_{i=1}^M \left\{ \lambda_{1i} \int_{\Omega} |I(x, y) - c_{1i}|^2 (1 - H_\epsilon(\Psi_i(x, y))) dx dy \right. \\
 & \left. + \lambda_{2i} \int_{\Omega} |I(x, y) - c_{2i}|^2 H_\epsilon(\Psi_i(x, y)) dx dy \right\} \\
 & + \sum_{i=1}^M \mu_i \int_{\Omega} |\nabla \Psi_i(x, y)| \delta_\epsilon(\Psi_i(x, y)) dx dy + \sum_{i=1}^M \nu_i \int_{\Omega} (1 - H_\epsilon(\Psi_i(x, y))) dx dy \\
 & + \sum_{i=1}^M \xi_i \int_{\Omega} \left\{ (1 - H_\epsilon(\Psi_i(x, y))) H_\epsilon(\hat{\Psi}_i(x, y)) \right. \\
 & \left. + H_\epsilon(\Psi_i(x, y)) (1 - H_\epsilon(\hat{\Psi}_i(x, y))) \right\} dx dy \\
 & + \sum_{i=1}^M \sum_{j=1, j \neq i}^M \omega_{ij} \int_{\Omega} (1 - H_\epsilon(\Psi_i(x, y))) (1 - H_\epsilon(\hat{\Psi}_j(x, y))) dx dy \tag{2}
 \end{aligned}$$

where Ω denotes the image domain. Keeping each c_{1i} and c_{2i} fixed and minimizing energy E in (2) with respect to $\Psi_i(x, y)$ ($i = 1, 2, \dots, M$), the associated Euler-Lagrange equation for each unknown level set function $\Psi_i(x, y)$ is deduced. After parameterizing the descent direction by an artificial time $t \geq 0$, the evolution equation in each $\Psi_i(t, x, y)$ is

$$\begin{aligned} \frac{\partial \Psi_i}{\partial t} = \delta_\varepsilon(\Psi_i) & \left\{ \lambda_{1i} |I - c_{1i}|^2 - \lambda_{2i} |I - c_{2i}|^2 + \mu_i \cdot \operatorname{div} \left(\frac{\nabla \Psi_i}{|\nabla \Psi_i|} \right) + \nu_i \right. \\ & \left. + \xi_i \left(2H_\varepsilon(\hat{\Psi}_i) - 1 \right) + \sum_{j=1, j \neq i}^M \omega_{ij} \left(1 - H_\varepsilon(\hat{\Psi}_j) \right) \right\} \end{aligned} \quad (3)$$

2.3 MAP Shape Estimation

In order to incorporate the influence of prior shape model to the evolution process as in (3), the shape and the pose of each model must be correctly estimated. Let $\hat{\Psi}_i$ denote the estimated curve of the i_{th} object. At each step of the curve evolution, it is estimated by

$$\hat{\Psi}_{i\text{MAP}} = \operatorname{argmax}_{\hat{\Psi}_i} p \left(\hat{\Psi}_i | \Psi_i, \hat{T}_i, \nabla I \right) \quad (4)$$

where $\hat{T}_i = \{\hat{\Psi}_k | 1 \leq k \leq M \text{ and } k \neq i\}$. To compute the MAP curve, we expand (4) using Bayes' Rule.

$$p \left(\hat{\Psi}_i | \Psi_i, \hat{T}_i, \nabla I \right) = \frac{p \left(\Psi_i, \hat{T}_i | \hat{\Psi}_i \right) p \left(\nabla I | \Psi_i, \hat{T}_i, \hat{\Psi}_i \right) p \left(\hat{\Psi}_i \right)}{p \left(\Psi_i, \hat{T}_i, \nabla I \right)} \quad (5)$$

The normalization term in the denominator of (5) can be discarded since it does not depend on the estimated shape of the i_{th} object. In addition, by assuming that other estimated shapes are independent of each other, we have

$$\begin{aligned} p \left(\Psi_i, \hat{T}_i | \hat{\Psi}_i \right) &= p \left(\Psi_i | \hat{\Psi}_i \right) \prod_{k=1, k \neq i}^M p \left(\hat{\Psi}_k | \hat{\Psi}_i \right) \\ p \left(\Psi_i | \hat{\Psi}_i \right) &= \exp \left(-A(\text{inside } C_i \text{ and outside } \hat{C}_i) \right) \\ p \left(\hat{\Psi}_k | \hat{\Psi}_i \right) &= \exp \left(-A(\text{inside } C_i \text{ and inside } \hat{C}_k) \right) \end{aligned} \quad (6)$$

The second term in (5) computes the probability of seeing certain image gradients given the current curve and the estimated curves. Since the gradient is a local feature, it is reasonable to assume that this probability does not depend on other estimated curves. Thus, we have

$$p \left(\nabla I | \Psi_i, \hat{T}_i, \hat{\Psi}_i \right) = p \left(\nabla I | \Psi_i, \hat{\Psi}_i \right) \quad (7)$$

3 Experimental Results

In our experiments, we set $\lambda_{1i} = \lambda_{2i} = 0.2$, $\mu_i = 0.05 \cdot 255^2$, and $\nu_i = 0$ so that only parameters ξ_i and ω_{ij} need to be adjusted to balance the influence of self shape prior and model competition.

3.1 Evaluation of Model Competition by Simulations

Our method was first applied on a 2D synthetic image which consists of one black arrow composed of a triangle and a rectangle, with no defined boundaries between them. They are of same intensity and Gaussian noise is added. For each object, a training set of 12 sample contours are used. It is difficult to avoid leakage using only prior shape information. Fig. 1 shows an example of severe leakage where the contours of the triangle and the rectangle leak into other's region. In Fig. 2, model competition is taken into account where the same parameters are adopted except for $\omega_{12} = \omega_{21} = 0.1$. In Fig. 1 and Fig. 2, yellow curves illustrate the initial MAP shape and pose. Green curves illustrate the final MAP shape and pose. Red curves show the initial contours (small red circles) and final contours. In Fig. 2, leakage is avoided by introducing competition between two statistical models.

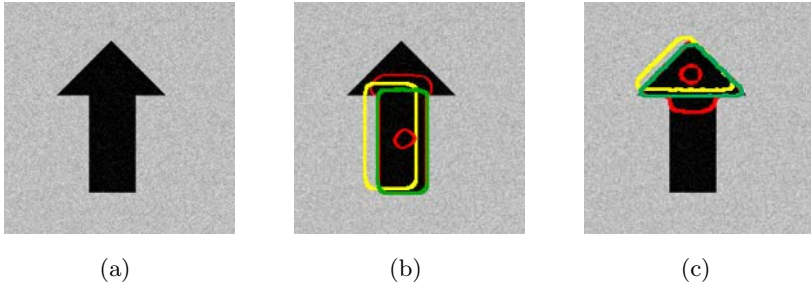


Fig. 1. Results without model competition ($\xi_1 = \xi_2 = 0.2$): (a) the original synthetic image, (b) the segmentation results for rectangle, and (c) the results for triangle

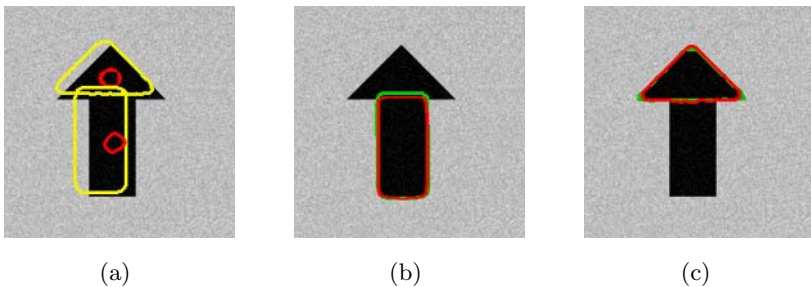


Fig. 2. Results with model competition. Parameters are the same as in Fig. 1 except $\omega_{12} = \omega_{21} = 0.1$: (a) the initial curves and estimated shapes. (b) the segmentation results of rectangle, and (c) the results for triangle.

3.2 Applications to Medical Imagery

Fig. 3 shows the segmentation of the amygdalae and hippocampus in a 2D MR image. In Fig. 3, the first column shows the initial contours; the second column shows intermediate evolving contours. The third column shows the final segmentation results. The top row, which are the results without any prior knowledge, shows that the two evolving contours become undistinguishable due to extremely poorly defined boundaries and very similar intensities. The middle row shows the results of using only prior shape information but without any model competition (refer to [5] for details). It is clear that the final results, which incorporates shape priors are more meaningful. However, there exist overlapping areas marked by red and green contours due to ill defined boundaries. The overlapping is unavoidable unless the neighboring organ imposes constraint on the evolution of the contour. The third row presents the results of using both shape priors and model competition. Other parameters are the same as for the middle row except

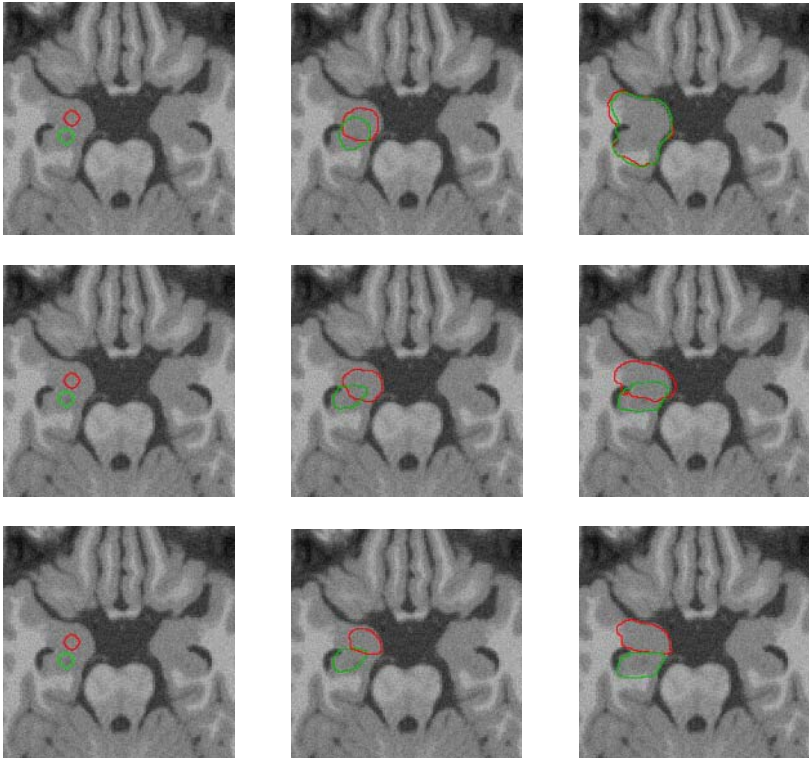


Fig. 3. Three steps in the segmentation of amygdalae and hippocampus in a brain image. (top) Results without prior knowledge. (middle) Results using only self prior shape information with $\xi_1 = \xi_2 = 0.2$. (bottom) Results using both shape priors and model competition, where $\omega_{12} = 0.2$, $\omega_{21} = 0.1$, and other parameters are the same as for the middle one. (The upper structure is object 1 and the lower one is object 2).

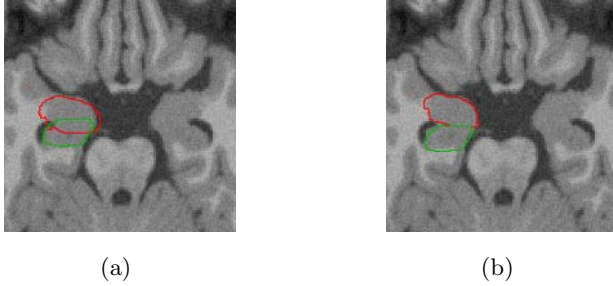


Fig. 4. MAP shape estimates of the segmentations shown at the middle and bottom rows of Fig. 3. (a) MAP shape estimate with only individual prior shape information. (b) MAP shape estimate with prior shape information and competition between two models.

$\omega_{12} = 0.2$ and $\omega_{21} = 0.1$. In the final results (column 3 of Fig. 3), the two structures are nicely segmented only when model competition is incorporated (the bottom image in column 3). We also note the overlap is avoided in the corresponding intermediate stage (the bottom image in column 2). Compared with the case of using only prior shape information (middle row), our method does not result in overlapping. Furthermore, in Fig. 4, we can see that introducing competition between models into the multiple object segmentation process will not only prevent overlapping of evolving neighboring contours but also benefit the shape estimation.

In Fig. 4, it can be seen that the final MAP shape estimates are improved after introducing interactions between models. With only shape priors, segmentation processes for each organ are treated independently. When leakage occurs, (as in the second image of the middle row in Fig. 3), an incorrect MAP shape estimate will be obtained, which tends to misguide further evolution of the contour. The final MAP shape estimates are shown in Fig. 4(a). After incorporating interactions between models, evolution of the contours and MAP shape estimates will tend to minimize the overlaps with others as well (see (1) and (6)). Each shape estimate will become less probable when it overlaps with others (see (5) and (6)). Thus, as shown in Fig. 3 and Fig. 4, allowing model competition will both benefit shape estimation and the contour evolution process to obtain good and robust segmentation results.

4 Discussions and Conclusion

In this paper, we focus on how to introduce interactions between neighboring contours and effectively incorporate them into the segmentation process. We have proposed a novel approach for segmenting multiple neighboring objects using prior shape information and competition between models, which intuitively describes the relationship between evolving contours and their own shape estimates, and the relationship between evolving contours and others' shape estimates (see (1)). Without repulsive interaction, our approach degrades to just

another formulation of active contours with prior shape influence [5]. Our approach is mainly designed for situations where prior shape information is not sufficient to achieve a good and robust estimation due to poorly defined boundaries and very similar intensities. With our method, multiple neighboring objects can be detected simultaneously without suffering from intersection problems and MAP shape estimates can be improved as well. Experiments on synthetic data and MR images show promising results. The proposed method can be applied to 3D medical image segmentation directly since implicit representations of the curves are used.

References

1. Osher, S., Sethian, J.A.: Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *J. Comp. Phys.* **79** (1988) 12–49
2. Malladi, R., Sethian, J.A., Vermuri, B.C.: Shape modeling with front propagation: A level set approach. *IEEE Trans. Pattern Anal. and Machine Intell.* **17** (1995) 158–174
3. Suri, J., Liu, K., Singh, S., Laxminarayana, S., Reden, L.: Shape recovery algorithms using level sets in 2-D/3-D medical imagery: A state-of-the-art review. *IEEE Trans. Inform. Technol. Biomed.* **6** (2002) 8–28
4. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models – their training and application. *Comput. Vision Image Understand.* **61** (1995) 38–59
5. Leventon, M.E., Grimson, W.E.L., Faugeras, O.: Statistical shape influence in geodesic active contours. In: *IEEE CVPR*. Volume 1. (2000) 316–323
6. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. *Int. J. Computer Vision* **22** (1997) 61–79
7. Qu, Y., Chen, Q., Heng, P.A., Wong, T.T.: Segmentation of left ventricle via level set method based on enriched speed term. In: *MICCAI 2004*. (2004) 435–442
8. Xie, J., Jiang, Y., Tsui, H.T.: Segmentation of kidney from ultrasound images based on texture and shape priors. *IEEE Trans. Med. Imag.* **24** (2005) 45–57
9. Zeng, X., Staib, L.H., Schultz, R.T., Duncan, J.S.: Segmentation and measurement of the cortex from 3-D MR images using coupled surfaces propagation. *IEEE Trans. Med. Imag.* **18** (1999) 927–937
10. Yang, J., Staib, L., Duncan, J.: Neighbor-constrained segmentation with level set based 3-D deformable models. *IEEE Trans. Med. Imag.* **23** (2004) 940–948
11. Pham, D.L., Bazin, P.L.: Simultaneous boundary and partial volume estimation in medical images. In: *MICCAI 2004*. (2004) 119–126
12. Chan, T.F., Vese, L.A.: Active contours without edges. *IEEE Trans. Image Processing* **10** (2001) 266–277

Point-Based Geometric Deformable Models for Medical Image Segmentation

Hon Pong Ho¹, Yunmei Chen², Huafeng Liu^{1,3}, and Pengcheng Shi¹

¹ Dept. of EEE, Hong Kong University of Science & Technology, Hong Kong

² Dept. of Mathematics, University of Florida, Gainesville, USA

³ State Key Lab of Modern Optical Instrumentation, Zhejiang University, Hangzhou, China
{garyho@ust.hk, yun@math.ufl.edu, eeliuhf@ust.hk, eeship@ust.hk}

Abstract. Conventional level set based image segmentations are performed upon certain underlying grid/mesh structures for explicit spatial discretization of the problem and evolution domains. Such computational grids, however, lead to typically expensive and difficult grid refinement/remeshing problems whenever trade-offs between time and precision are deemed necessary. In this paper, we present the idea of performing level set evolution in a point-based environment where the sampling location and density of the domains are adaptively determined by level set geometry and image information, thus rid of the needs for computational grids and the associated refinements. We have implemented the general geometric deformable models using this representation and computational strategy, including the incorporation of region-based prior information in both domain sampling and curve evolution processes, and have evaluated the performance of the method on synthetic data with ground truth and performed surface segmentation of brain structures from three-dimensional magnetic resonance images.

1 Introduction

The geometric nature of the *geometric deformable models* (GDMs) gives them several often desired properties, i.e. independence from parametrization, easy computation of geometrical characteristics such as curvatures, and natural compatibility with topological changes [1, 7]. Nevertheless, trivial level set sampling using structured image grid could significantly decrease the computational efficiency when computer memory becomes the bottleneck of the segmentation process, especially for large 3D medical images which have been becoming increasingly available. Fast marching scheme and the narrowband search have sofar provided some reliefs.

However, since most GDMs rely on the finite difference computational schemes, the resolutions of the final results are fundamentally dependent on the computational grids used to solve the discreteized level set partial differential equations (PDEs). Adaptive grid techniques to locally modify the grid structure based on certain criteria is a popular solution. The local refinement methods place additional grid points when (i.e. during narrowband refinement) and where (i.e. high gradient areas) they are needed to avoid expensive global refinement procedure [3]. More recently, moving grid methods have been proposed as simpler alternatives to the local refinement methods [4]. But the local sampling rate for such moving grid strategy is still upper-bounded by the number of initial grid lines used. Regardless of the choices of computational grids, task-specific

incorporation of image force and prior knowledge also plays essential roles in achieving accurate and attractive segmentation results, especially for medical images [2].

In the Lagrangian formulation of parametric deformable models (PDMs) and motion tracking, meshfree particle methods have been introduced as more efficient and possibly more effective object representation and computation alternatives because of their trivial $h - p$ adaptivity [6]. In addition, it has been shown that, for arbitrary shape modeling, point-sampled representation allows user-defined large deformations on the object shape [10]. Following the same spirit, it is easy to see that if general level set PDE computation can be conducted in a point-based environment for both shape and evolution (deformation) domains, we can almost freely adjust the sampling rate of the domains without many extra efforts. This can be achieved through point-wise continuous approximations of the domain and of the level set functions through polynomials fitting using local point cloud. With such inspiration, we have introduced a point-based numerical scheme for domain representation and computation [5].

In this paper, we extend the numerical scheme of [5] to point-based geometric deformable models (GDMs) for medical image segmentation. In particular, through adaptive sampling points distribution and level set formulation, our curve evolution formulation incorporates region-based prior information, which is especially useful for brain segmentation because the intensity homogeneity of the brain structures can compensate noisy or confusing boundary information. Spatial derivatives of the point-sampled level set function are obtained through moving least squares (MLS) approximation. Because of the $h - p$ adaptive nature of the sampling points, our strategy naturally possesses multiscale property and performs fast marching and local refinement concurrently.

2 Methodology

2.1 Geometric Deformable Models

Level Set Formulation: Suppose we want to model a dynamic deformable contour $C(t) \in \mathbb{R}^d$ under a known velocity field $F(\mathbf{x}, C(t))$, where $\mathbf{x} \in \mathbb{R}^d$ and the initial contour $C(0)$ is given. The dynamic equation for the moving contour is then:

$$\frac{dC(t)}{dt} = F(\mathbf{x}, C(t)) \quad (1)$$

where F is known as the speed term that determines the particular way under which $C(t)$ deforms. If we want $C(t)$ to stop at certain locations in space, such as object boundary in image segmentation cases, then those locations should afford small or zero $F(\mathbf{x})$ values. If the targeted object is complex or the image data is noisy, the actual F field could be very complicated and the resulted $C(t)|_{t \rightarrow \infty}$ requires flexible and high precision representation. Therefore, changing the representation of $C(t)$ into implicit forms had been developed, known as the level set methods [8].

The basic idea of level set methods is to implicitly represent the moving surface $C(t)$ by the zero level set of a higher dimensional hypersurface ϕ (the level set function), i.e. the signed distance transformation of $C(t)$ in \mathbb{R}^d , such that $C(t) = \{\mathbf{x} \mid \phi(\mathbf{x}, t) = 0 \forall \mathbf{x} \in \mathbb{R}^d\}$. The curve evolution is then driven by F in normal directions and we solve for convergence of the dynamic equation by time-domain finite difference:

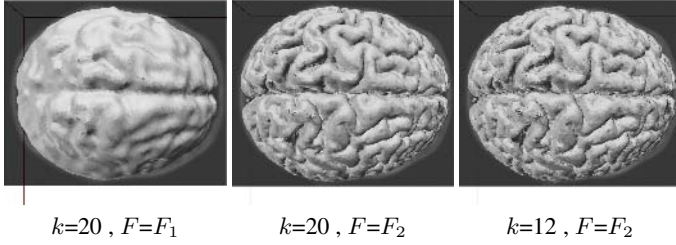


Fig. 1. Brain segmentations using same set of sampling point cloud with different forces in Eqn. (2) and samples per influence domain (k). The 1st figure was segmented without region-based prior model while the 2nd and 3rd are significantly improved results by using both boundary image information and regional priors. In the rightmost two figures, larger k makes the surface smoother while smaller k reduces computation but is more noise sensitive.

$$\phi^{t+\Delta t} = \phi^t + \Delta t F |\nabla \phi| \tag{2}$$

Construction of the Speed Term F : To achieve desirable segmentation results, task-specific incorporation of image force and other knowledge for the speed term F plays essential roles [7, 9]. One of the most popular F formulations states:

$$F_1 = \beta g \kappa - (1 - \beta) g \hat{v} \cdot \frac{\nabla \phi}{|\nabla \phi|} \tag{3}$$

where $g = \frac{1}{1+|\nabla G * I|}$ is the image gradient force, κ is the contour curvature, \hat{v} is the gradient vector flow (GVF), $\beta \in [0, 1]$ is a weighting parameter, and G denotes the gaussian smoothing kernel. This F form, however, may not produce appropriate segmentation results. In Fig. 1, the less detailed surface in the leftmost figure was segmented using F_1 . An initial guess surface was first placed inside the brain structure and then expanded outward driven. When the surface encountered sharp edges, the GVF force became dominant and attracted that portions of surface towards their nearest sharp edges. However, the image intensity gradient at certain areas of the brain image is too weak to slow down the surface expansion. Further, GVF itself is a diffusion process such that weak edges would be covered by the impact of nearby sharp edges. In the brain image, the skull layer actually has sharper boundaries than the brain surface, and thus would make the whole surface move outward beyond its proper location.

To address this problem, we further constrain the surface expansion by region-based information so that the surface would expand only if it is likely to be inside our target. The likelihood is defined by the similarity between local image intensities and a statistical mean intensity I_m with variance σ^2 , constructed from prior knowledge of the structure. With this additional information, we define our new speed term as:

$$F_2 = \alpha g \kappa - \beta g \hat{v} \cdot \frac{\nabla \phi}{|\nabla \phi|} + \gamma (F_a - e^{-\frac{1}{n\sigma^2} \sum |I_x - I_m|}) \tag{4}$$

where α, β and γ are three weighting constants. We assume our $C(t)$ surface evolves outward. F_a is an additional natural expanding speed. To avoid excessive expansion,

α and β should be smaller than the original weightings in F_1 . In the third term of F_2 , we modify the natural speed by prior information: $I_m \in [0, 1]$ and $\sigma^2 \in \mathbb{R}$. $I_x \in [0, 1]$ are image intensities around point \mathbf{x} and n is the number of neighboring points. Then, $\sum |I_x - I_m|/n$, the average inconsistency of local intensities will slow down the surface $C(t)$ motion. We would like to suggest $F_a > 1$ such that the whole third term will be positive if \mathbf{x} is inside the target and negative if outside. This would have similar effect as the GVF force in the second term.

2.2 Numerical Implementation of Point-Based GDMs

Adaptive Domain Sampling with Data- and Prior-Driven Point Cloud: Rather than using image grid or other mesh structures, we use a set of adaptively distributed points $\mathbf{x}_i \in \mathbb{R}^d$ to dynamically sample the spatial domains for Eqn. (2) [5]. The position of these sampling points can be flexible and adaptive, and we design to put more points near areas of richer features (i.e. image gradient, geometrical characteristics of the evolving contour, prior knowledge of boundary likelihood from experience or training images, etc.) in order to achieve higher precision for the ϕ function. On the other hand, for relatively featureless or uninteresting areas such as the background or homogeneous region inside the object, we only need to maintain a minimal amount of points to support proper ϕ evaluation for $C(t)$ contour propagation.

Obtaining Derivatives of ϕ : While the unstructured sampling point cloud affords greater adaptivity towards data and priors, the calculation of $\nabla\phi$ is no longer as straightforward as the conventional finite difference implementations because the relationship between the sampling points are no longer directly defined.

Assuming that the ϕ function at any location \mathbf{x} is a well-behaved polynomial surface within a small area around \mathbf{x} , we can use $\phi(\mathbf{x}_i)$ of the neighboring nodes \mathbf{x}_i in that area to recover the coefficients of ϕ surface by performing the moving least square (MLS) approximation [6]. Calling such local area the *influence domain* or *support region* of \mathbf{x} , all nodes inside the influence domain will have certain amount of effects on the recovered coefficients, weighted by their relative distance to \mathbf{x} . In our particular implementation, we actually select the proper number of neighboring nodes k for

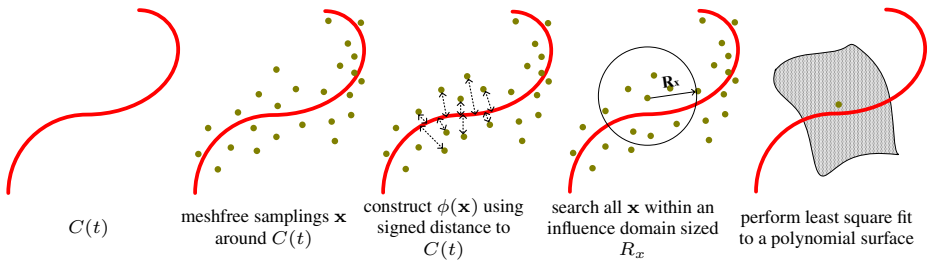


Fig. 2. From left to right: steps to construct $|\nabla\phi|$ on point cloud using moving least square fitting. The fitting is done on the $\phi(\mathbf{x})$ values. Once the polynomial coefficients have been recovered, the derivative of the fitted surface ($\nabla\phi$) can also be calculated. Then the ϕ can be updated by Eqn.(2) and $C(t + \Delta t)$ is obtained by Eqn.(7). Finally, we substitute $C(t)$ by $C(t + \Delta t)$ and the process can be repeated iteratively until C converges.

\mathbf{x} instead of determining the size of influence domain. Since the node distribution is feature-driven, choosing a fixed number of nearest neighboring nodes for \mathbf{x} effectively imply a feature-adaptive determination of the influence domain size. Fig. 2 gives an overview of the use of influence domain in the point-based GDM implementation.

Let $\tilde{\phi}$ be the MLS-approximated ϕ with the form: $\tilde{\phi}(\mathbf{x}) = \sum_j^m p_j(\mathbf{x})a_j(\mathbf{x}) \equiv \mathbf{p}^T(\mathbf{x})\mathbf{a}(\mathbf{x})$, where \mathbf{p} is the polynomial basis, m is the order of polynomial, and \mathbf{a} is the coefficient vector. To get $\tilde{\phi}(\mathbf{p})$ at arbitrary location \mathbf{x} , we determine \mathbf{a} through minimizing the weighted approximation error J of the $\phi(\mathbf{x}_i)$ values of the k -nearest neighboring nodes \mathbf{x}_i , defined as the followings:

$$J = \sum_i^k W\left(\frac{|\mathbf{x} - \mathbf{x}_i|}{R_x}\right) [\mathbf{p}^T(\mathbf{x}_i)\mathbf{a}(\mathbf{x}) - \phi(\mathbf{x}_i)]^2 \tag{5}$$

$$W(\bar{d}) = \begin{cases} \frac{2}{3} - 4\bar{d}^2 + 4\bar{d}^3 & \text{for } \bar{d} \leq \frac{1}{2} \\ \frac{4}{3} - 4\bar{d} + 4\bar{d}^2 - \frac{4}{3}\bar{d}^3 & \text{for } \frac{1}{2} \leq \bar{d} \leq 1 \\ 0 & \text{for } \bar{d} > 1 \end{cases} \tag{6}$$

where W is a cubic spline decaying weight function, $|\mathbf{x} - \mathbf{x}_i|$ is the spatial distance between \mathbf{x} and \mathbf{x}_i , and T denotes vector transpose. The derivative $\nabla\phi$ can then be approximated by the derivative of the recovered polynomial $\tilde{\phi}$.

Detecting Zero Crossings of ϕ : In any level set method, ϕ is only a medium for evolving $C(t)$ while the ultimate goal is to properly localize the $C(t)$ surface as the zero level set of ϕ . However, for the point-based GDMs, the detection of the zero level set can no longer be performed using the widely used Marching Cubes algorithm that requires a grid structure. Instead, we opt to detect the zero-crossings of evolved ϕ [8]:

$$\mathbf{z}_i = \mathbf{x}_i - \phi(\mathbf{x}_i, t) \frac{\nabla\phi(\mathbf{x}_i, t)}{|\nabla\phi(\mathbf{x}_i, t)|} \tag{7}$$

where \mathbf{z}_i is the nearest $\phi = 0$ position from point \mathbf{x}_i . The equation assumes ϕ is a signed distance function of $C(t)$ and $\nabla\phi$ denotes the normal direction towards nearest $C(t)$ position. The set of \mathbf{z}_i becomes the discrete representation of $C(t)$.

3 Experiments

3.1 Grid- and Point-Based Implementations: Comparison on Synthetic Images

A synthetic image with known boundary (Fig.3) is used to evaluate the grid- and point-based implementations. To make fairer comparison, for the point-based GDMs, we distribute sampling points according to the image gradient magnitude only, where the total number of sampling points is roughly one quarter fewer than the grid-based case. In the grid-based implementation, the Marching Cubes algorithm is used to detect the zero-crossings of ϕ and the image grid is used as the sampling grid. For both implementations, F_1 speed term is used, the image gradient $|\nabla I|$ is computed using central difference, and narrow band speedup is adopted [7].

Different levels of Gaussian noises are added to the phantom image, and the performance of the two implementations are tabulated in Fig.4. For the mean squared errors (MSE), the point-based GDMs are slightly better than the conventional finite difference

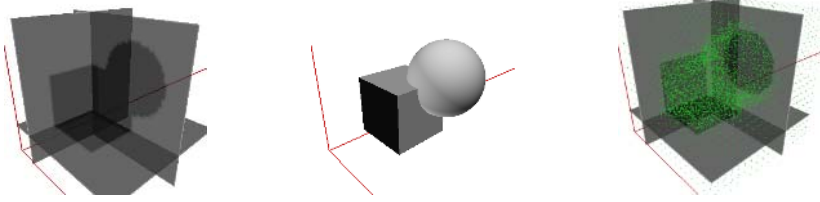
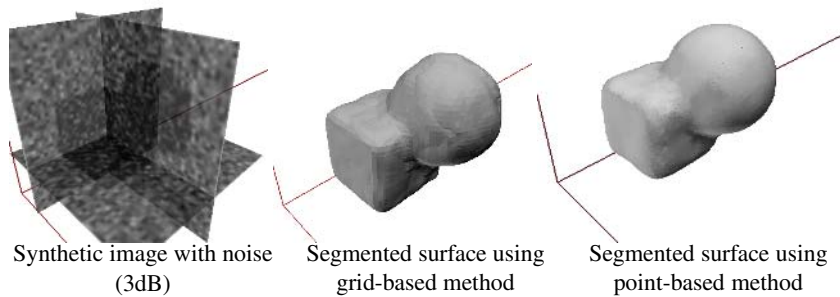


Fig. 3. From left to right: synthetic 3D image, underlying ground truth object, and sampling point cloud for experiments in Fig. 4



	SNR:	10dB	7dB	5dB	3dB
Grid-Based Implementation	Time:	90sec	98sec	86sec	171sec
	MSE:	0.1619	0.1867	0.2180	0.1970
	Std Deviation:	0.2410	0.2623	0.3438	0.3222
	Max Error:	1.3762	1.5169	1.7272	1.9957
Point-Based Implementation	Time:	100sec	84sec	102sec	100sec
	MSE:	0.1544	0.1892	0.2017	0.1951
	Std Deviation:	0.1869	0.2185	0.2392	0.2625
	Max Error:	1.2911	1.6274	1.4755	1.4260

Fig. 4. Error comparison between point-based and grid-based (finite difference) GDMs on segmentation of 3D synthetic noisy images with known ground truth

GDMs. However, due to the smoothing effect of MLS, they are much more stable in terms of the error standard deviations, especially when the noises are more significant. We also want to point out that, while the MLS procedure is computationally expensive, adaptive point sampling can efficiently limit the total amount of needed points and thus the point-based GDMs are still computationally comparable to the grid-based GDMs for practical usages, as indicating by the processing time.

3.2 Segmentation of Brain Structures

In Fig. 1, with uniformly distributed sampling points, we show the effect of the number of points k per influence domain ¹, or equivalently the sizes of the influence domain. In

¹ k is lower bounded to ensure that the MLS system of equations is invertible.

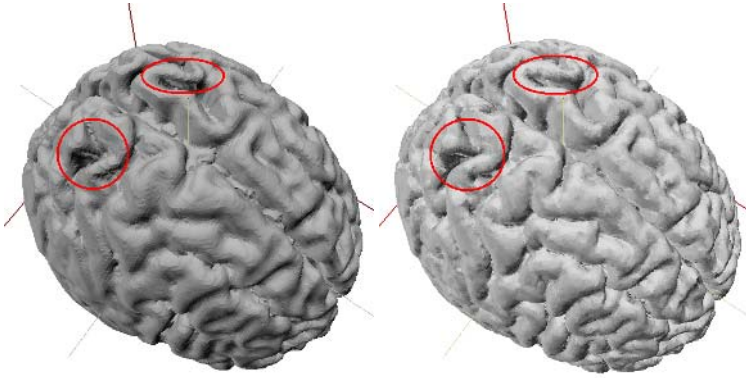


Fig. 5. Comparison between grid-based (left) and point-based (right) brain surface segmentation

general, larger k would produce smoother segmented surface (at higher computational cost), desirable for larger structures and flatter regions. On the other hand, smaller k efficiently yields *rougher* results, which is actually needed for geometrical details. Our point-based GDMs, however, adopt feature-driven point distribution scheme and thus automatically take care of the selection of k or influence domain size.

Comparison of the point- and grid-based implementations is made for brain surface segmentation (Fig. 5), both using regional priors. We first obtain a coarse segmentation using fewer sampling points which covers the whole image, and then distribute huge amount of points around the coarse segmentation results, with even more points in high image gradient areas. While the difference between the two results is not significant at smooth regions, around extremely thin structures as we have explicitly pointed out, sampling insufficiencies are uncovered by sharp changes of normal directions of the triangles in the grid-based results, extracted by the Marching Cubes algorithm.

Suppose that we have some prior knowledge about the rough locations of certain brain structures, we distribute sampling points around (inside and outside) those areas with higher point density for high probability area. Example of point distribution is

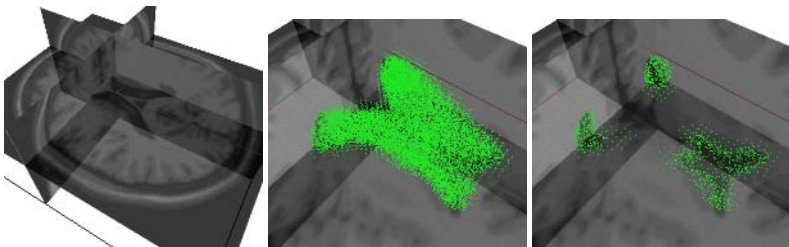


Fig. 6. Left: A Brainweb image (<http://www.bic.mni.mcgill.ca/brainweb>), T_1 weighted, 3% noise. Middle: Sampling nodes are distributed according to prior shape information obtained from training images. These nodes are used to sample the level set function ϕ during surface evolution (Fig.7). Right: Sampling points on one specific slice of the 3D image.

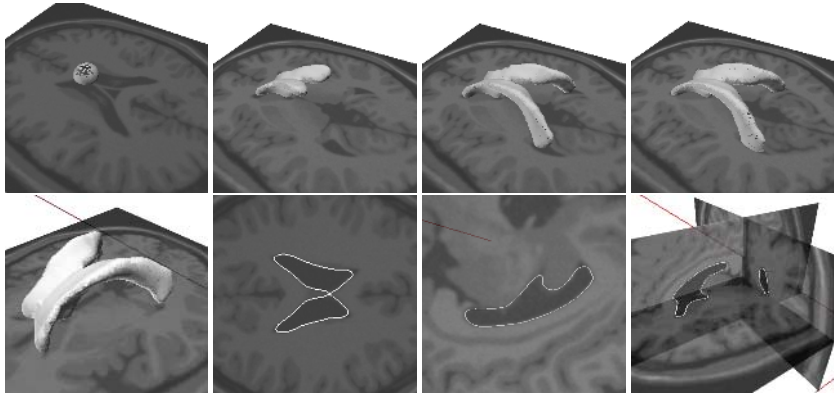


Fig. 7. Top: Brain ventricle segmentation with prior-driven node distribution. Bottom: Segmented brain ventricle surface using point-based GDM, and the intersections of surface with image slices.

shown in Fig. 6 for brain ventricle segmentation, and the point-based GDM evolution process and segmentation results are shown in Fig. 7.

Acknowledgement

This work is supported by NBRPC 2003CB716104, NSFC 60403040, HKUST6252/04E.

References

1. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *IJCV*, 22:61–79, 1997.
2. Y. Chen, H. Tagare, and et al. Using prior shapes in geometric active contours in a variational framework. *International Journal of Computer Vision*, 50:315–328, 2002.
3. M. Droske, B. Meyer, M. Rumpf, and C. Schaller. An adaptive level set method for medical image segmentation. In *Information Processing in Medical Imaging*, pages 416–422, 2001.
4. X. Han, C. Xu, and J. L. Prince. A 2d moving grid geometric deformable model. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 153–160, 2003.
5. H.P. Ho, Y. Chen, H. Liu, and P. Shi. Level set active contours on unstructured point cloud. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2005.
6. H. Liu and P. Shi. Meshfree particle method. In *ICCV*, pages 289–296, October 2003.
7. R. Malladi, J. A. Sethian, and B. C. Vemuri. Shape modeling with front propagation: a level set approach. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17:158–175, 1995.
8. S. Osher and R. Fedkiw. *Level set methods and dynamic implicit surfaces*. Springer-Verlag New York Inc., 2002.
9. N. Paragios, O. Mellina-Gottardo, and V. Ramesh. Gradient vector flow fast geometric active contours. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(3), March 2004.
10. M. Pauly, R. Keiser, L.P. Kobbelt, and M. Gross. Shape modeling with point-sampled geometry. *ACM SIGGRAPH (Trans. Graphics)*, 22(3):641–650, July 2003.

A Variational PDE Based Level Set Method for a Simultaneous Segmentation and Non-rigid Registration

Jung-ha An¹, Yunmei Chen¹, Feng Huang², David Wilson¹, and Edward Geiser³

¹ Department of Mathematics, University of Florida, USA

² Invivo Diagnostic Imaging, Florida, USA

³ Department of Medicine, University of Florida, USA

Abstract. A new variational PDE based level set method for a simultaneous image segmentation and non-rigid registration using prior shape and intensity information is presented. The segmentation is obtained by finding a non-rigid registration to the prior shape. The non-rigid registration consists of both a global rigid transformation and a local non-rigid deformation. In this model, a prior shape is used as an initial contour which leads to decrease the numerical calculation time. The model is tested against two chamber end systolic ultrasound images from thirteen human patients. The experimental results provide preliminary evidence of the effectiveness of the model in detecting the boundaries of the incompletely resolved objects which were plagued by noise, dropout, and artifact.

1 Introduction

Segmentation and registration play an important role in image processing, image analysis, and computer vision. There have been several different approaches to solve image segmentation and registration problems. In the past, solutions of these two problems have been studied separately from each other.

Segmentation techniques have been developed to capture the object boundary by several different approaches; edge-based methods mainly using active contour models, region-based methods, or the combination of the two by using Geodesic Active Region models. However, these methods may not be able to accommodate all types of imaging difficulties. The prior shape and intensity information has been incorporated into the geodesic active contour model [13, 12, 1, 4, 3, 2] to make the resulting segmentation more accurate. Recently, the prior shape and intensity information has been also incorporated into deformable models [10, 9].

Image registration is the process of overlaying two or more images of the same scene taken at different times, and/or by different sensors. Area based methods or feature based methods have been developed to match given images. For the details of the image registration techniques, please refer to [29]. Several approaches have been also made in establishing the application of mutual information for non-rigid registration [7, 16, 14, 8, 24]. Instead of intensity information, a feature based model combines region matching and segmentation with free boundary conditions in [23], but the model in [23] can be applied only in the single regions of interest.

However, segmentation or registration technique alone, dependence of segmentation on registration methods, or dependence of registration on segmentation methods

do not solve the image segmentation and registration problems completely. A joint segmentation and registration methods have been developed through active contours [27, 1, 4, 2, 3, 26]. An explicit combination of registration with segmentation has been developed in a variational framework through active contours [27]. The algorithms of [27] was extended in [15] to joint segmentation and of an object in two images. The morphing active contours algorithm is combined with the joint segmentation and registration with an application to CT scans for radiotherapy treatment planning [26].

In this paper, a new variational partial differential equation (PDE) based level set method for a simultaneous image segmentation and registration is presented with an application to two chamber human heart end systole ultra sound images. The purpose of the model is to segment the endocardial borders of the given images. The segmentation is obtained by finding a non-rigid registration to the prior shape. The model performs a simultaneous segmentation and registration in a similar way from [27, 1, 3] through active contours. But the model differs from [27, 1, 3] by using a shape prior as an initial contour. Since the shape prior is used as an initial contour, the level set form of the model is fixed which does not require the re-initialization process during the numerical calculation. The model in this paper is similar to [12, 3, 2] by utilizing both prior shape and prior intensity information to get better segmentation, but the segmentation in this model is obtained by finding the non-rigid deformation to the prior shape. This paper follows the idea from [22, 21, 25, 10, 9] for matching nonequivalent shapes by the combination of a global rigid transformation and a local non-rigid deformation. However, the model differs from [22, 21] by having an intensity matching term in the energy function which provides the difference of the intensity variation across the prior shape and the shape of interest. Minimizing the difference of the gradients of the intensity between smoothed prior image and the smoothed given image is more accurate than minimizing the gradient term in the given image, due to the loss of the information of the gradient near the edge of the desired object. This paper is organized as follows: In section 2, the model is proposed for a simultaneous segmentation and non-rigid registration and the level set form of the proposed model is described. In section 3, numerical methods pertaining to the model are explained. Experimental results of the model which were applied to human heart two chamber end systole ultrasound images are also showed in this section. In section 4, the conclusion follows and future work is stated.

2 Description of the Proposed Model

2.1 A Simultaneous Segmentation and Non-rigid Registration

A new variational PDE based level set method for a simultaneous image segmentation and non-rigid registration is presented in this section. The purpose of the model is to segment endocardial borders of the given image using a prior shape and intensity information. The segmentation is obtained by finding a non-rigid registration to the prior shape. Only small non-rigid deformation is considered in the registration. The notion of shape is independent of scaling, rotation, and translation in the model.

Let $C^*(p) = (x(p), y(p))(p \in [0, 1])$ be a differentiable parameterized curve, called shape prior, in an associated prior image I^* . The domain $N(C^*)$ is the neighborhood

of the prior shape. Let $G_\sigma(x) = \frac{1}{\sigma} * e^{-\frac{|x|^2}{4\sigma}}$ and $g(|\nabla I|)$ can be chosen as $g(|\nabla I|) = \frac{1}{1+\beta * |\nabla(G_\sigma * I)|^2}$, where $\beta > 0$ is a parameter. The model is aimed to find u, v, a, μ, R, T by minimizing the energy functional:

$$\begin{aligned}
 E(u, v, a, \mu, R, T) = & \lambda_1 \int_{N(C^*)} |\nabla(G_\sigma * I^*)(x, y) - a\nabla(G_\sigma * I)(\tilde{u}, \tilde{v})|^2 d\bar{x} \\
 & + \lambda_2 \int_{N(C^*)} (|u|^2 + |v|^2) d\bar{x} + \lambda_3 \int_{N(C^*)} (|\nabla u|^2 + |\nabla v|^2) d\bar{x} \\
 & + \lambda_4 \int_0^1 g(|\nabla I|)(\tilde{u}, \tilde{v})(C^*(p)) |(C^*)'(p)| dp,
 \end{aligned} \tag{1}$$

and the vector $\begin{bmatrix} \tilde{u}(x, y) \\ \tilde{v}(x, y) \end{bmatrix} = \mu * R * \begin{bmatrix} x \\ y \end{bmatrix} + T + \begin{bmatrix} u(x, y) \\ v(x, y) \end{bmatrix}$, where μ is a scaling, R is a rotation matrix with respect to an angle θ , and T is a translation. A prior image I^* , a novel image I , the prior shape C^* , and the neighborhood $N(C^*)$ of C^* as a domain are given and $\lambda_i > 0 (i = 1, 2, 3, 4)$ are parameters balancing the influences from four terms in the model.

The first term minimizes the difference of the intensity variation across the prior shape and the shape of interest. Minimizing the difference of gradients of the intensity between smoothed prior image and the smoothed given image is more accurate than minimizing the gradient term in the given image, due to the loss of the information of the gradient near the edge of the desired object. Since images may have different levels of intensity, the model is also minimized over the scale factor variable a . The second term is minimizing the magnitude of the non-rigid deformation term $\begin{bmatrix} u \\ v \end{bmatrix}$. Smoothing

the non-rigid term $\begin{bmatrix} u \\ v \end{bmatrix}$ is in the third term. Final term is the energy functional of active contours [5] and [11], that measures the amount of high gradient under the trace of the curve and increases the attraction of the evolving curve toward the boundary of the object. Since the prior shape C^* is used as an initial contour and the domain is taken as the neighborhood of the prior shape, it helps the numerical solutions converge faster. In addition, when the model is formulated into the level set form in Equation (2), the level set form of the active contour in this model is fixed. Therefore, re-initializing process in the calculation is not needed which decreases the numerical calculation time. In the model, a simultaneous segmentation and non-rigid registration performs. The segmentation is obtained by finding a non-rigid registration to the prior shape. A registration consists of global rigid transformation and local non-rigid deformation.

2.2 Level Set Formulation of the Proposed Model

Level set methods of [17] are extensively used in the problems of curve evolution including snakes and active contours, since they allow the curve to develop cusps, corners, and topological changes. Now, a variational level set approach is given from [28] and [6]. Let the contour C^* be the zero level set of a Lipschitz function w such that $\{z | w(z) > 0\}$ is the set inside C^* . Let $H(z)$ be the Heaviside function, that is $H(z) = 1$ if $z \geq 0$, and $H(z) = 0$ if $z < 0$, and $\delta = H'(z)$ (in the sense of distribution) be the Dirac measure. Then, the length of the zero level set of w in the conformal metric $ds = g(|\nabla I|) |(C^*)'(p)| dp$ can be measured by $\int_{N(C^*)} g(|\nabla I|) |\nabla H(w)| =$

$\int_{N(C^*)} \delta(w)g(|\nabla I|)|Dw|$, where $N(C^*)$ is the given image domain. Therefore, the level set formulation of the model is to find u, v, a, μ, R, T by minimizing the energy functional:

$$E(u, v, a, \mu, R, T) = \lambda_1 \int_{N(C^*)} |\nabla(G_\sigma * I^*)(x, y) - a\nabla(G_\sigma * I)(\tilde{u}, \tilde{v})|^2 d\bar{x} + \lambda_2 \int_{N(C^*)} (|u|^2 + |v|^2) d\bar{x} + \lambda_3 \int_{N(C^*)} (|\nabla u|^2 + |\nabla v|^2) d\bar{x} + \lambda_4 \int_{N(C^*)} \delta(d_{C^*})g(|\nabla I|)(\tilde{u}, \tilde{v})|\nabla d_{C^*}| d\bar{x}, \tag{2}$$

where d_{C^*} is the distance function from the prior shape C^* , $g(|\nabla I|)$ can be chosen as $g(|\nabla I|) = \frac{1}{1+\beta*|\nabla(G_\sigma * I)|^2}$, $\beta > 0$ is a parameter, $G_\sigma(x) = \frac{1}{\sigma} * e^{-\frac{|x|^2}{4\sigma}}$, and in a similar way, the vector $\begin{bmatrix} \tilde{u}(x, y) \\ \tilde{v}(x, y) \end{bmatrix} = \mu R \begin{bmatrix} x \\ y \end{bmatrix} + T + \begin{bmatrix} u(x, y) \\ v(x, y) \end{bmatrix}$. A prior image I^* , a novel image I , and the prior shape C^* are given and $\lambda_1, \lambda_2, \lambda_3$, and λ_4 are positive parameters.

3 Numerical Methods and Experimental Results

In this section, numerical methods to solve the Equation (2) are explained and the results of applications to the thirteen human heart two chamber end systole ultrasound images are showed. The Equation (2) was solved by finding a steady state solution of the evolution equations. The evolution equations are associated with the Euler-Lagrange equations of the Equation (2). Following the approach from [17], H and δ are replaced by the regularized versions of them during the numerical calculation. A finite difference scheme and the gradient descent method is applied to discretize the evolving equations.

Prior shape C^* and prior intensity I^* of 13 Human heart two chamber end systole ultra sound images which have $0.62mm$ resolution in each pixel were generated by the method from [2]. The prior intensity information and a prior shape are in Figure 1 (a) and (b). The Figure 1 (a) represents the average image as prior intensity information. The solid red line in Figure 1 (b) is the average curve as a prior shape.

To show the effectiveness of the local non-rigid deformation term in the model, a global rigid transformation model is created without non-rigid deformation term $\begin{bmatrix} u \\ v \end{bmatrix}$

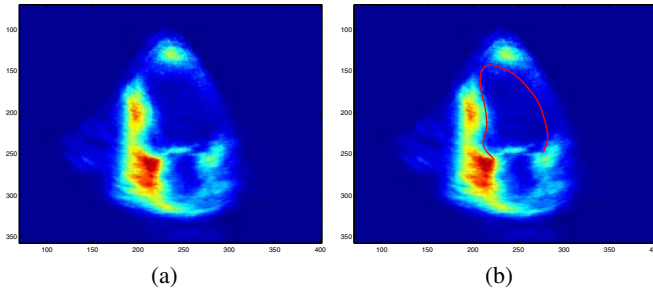


Fig. 1. (a). An average image as a prior intensity information. (b). An average curve with a solid line as a prior shape.

in Equation (2). Numerical calculation is done without non-rigid term. To validate experimental results of the model, distance measurements to the results of the model from expert's borders from Figure 2 to Figure 4 are followed in the Table 1. Table 1 provides the effectiveness of the suggested model which is better than the global rigid transformation model.

From Figure 2 to Figure 4, A solid red line in (a) is the initial contour in a novel image. A prior shape is used as an initial contour. The solid red line in (b) is the model's result with the local non-rigid deformation term. The expert endocardium border is compared as a dotted white line. The solid red line in (c) is the result of the model without a local non-rigid deformation term in the Equation (2). In a similar way, the expert endocardium border is compared as a dotted white line. From the Figures and the distance measurement results from the expert' border in Table 1, it is easily seen that the performance of the suggested model is better than the model without a local non-rigid deformation in the Equation (2). In Figure 2, the distance result from the model with non-rigid term is $1.8442(1.1434mm)$ which is better than the distance result $2.2160(1.3739mm)$ from the global rigid transformation model. In Figure 3, the distance result from the model with non-rigid term is $2.2301(1.3826mm)$ which is better than the distance result $2.8891(1.7912mm)$ from the global rigid transformation model. In Figure 4, the distance result from the model with non-rigid term is $2.0686(1.2825mm)$ which is better than the distance result $3.0470(1.8891mm)$ from

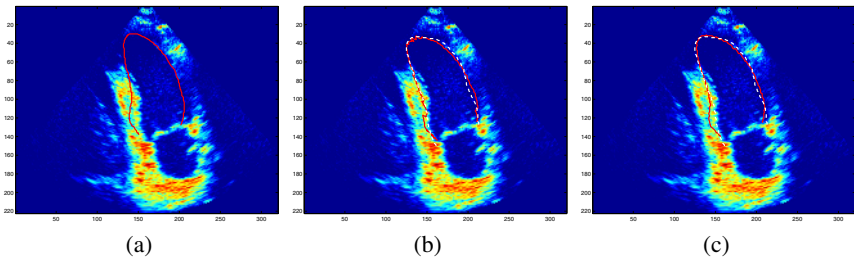


Fig. 2. (a). Solid line is an initial contour in the novel image. (b). The results from the model with non-rigid term (solid) and expert's border (dotted) in an image. (c). The results from the model without non-rigid term (solid) and expert's border (dotted) in an image.

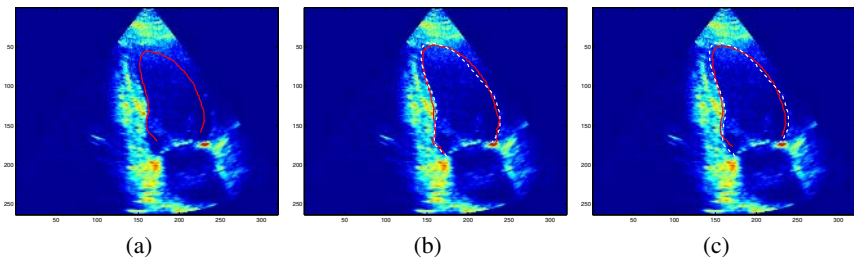


Fig. 3. (a). Solid line is an initial contour in the novel image. (b). The results from the model with non-rigid term (solid) and expert's border (dotted) in an image. (c). The results from the model without non-rigid term (solid) and expert's border (dotted) in an image.

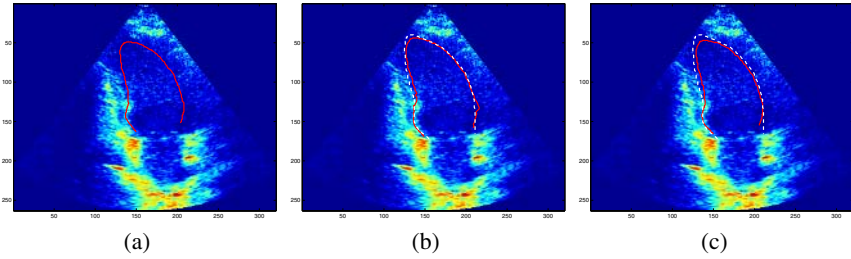


Fig. 4. (a). Solid line is an initial contour in the novel image. (b). The results from the model with non-rigid term (solid) and expert's border (dotted) in an image. (c). The results from the model without non-rigid term (solid) and expert's border (dotted) in an image.

Table 1. Distance measurements of the result of the model from expert's border from Figure 2 to Figure 4

Figures	(a)Initial	(b)Result with non-rigid term	(c)Result with rigid term only
Figure 2	25.1354 (15.5839mm)	1.8442(1.1434mm)	2.2160 (1.3739mm)
Figure 3	107.2969 (66.5240mm)	2.2301 (1.3826mm)	2.8891 (1.7912mm)
Figure 4	4.4458 (2.7563mm)	2.0686 (1.2825mm)	3.0470 (1.8891mm)

the global rigid transformation model. Since the local deformation is assumed small in the model, the distance difference between the model with non-rigid term and the global rigid transformation only model is not big. However, the Table 1 shows the effectiveness of the suggested model.

4 Conclusions and Future Work

A new variational method for a simultaneous segmentation and non-rigid registration that incorporates the shape and intensity priors is proposed. Numerical results showed the effectiveness of the model to capture the boundary of the given images in detail.

From Figure 2 to Figure 4, the model performance to the given images which is compared to expert borders is showed. The global rigid transformation model was calculated in the Equation (2) without $\begin{bmatrix} u \\ v \end{bmatrix}$ term. The distance measurements to the results of the model from the expert's borders by using the distance function are followed in the Table 1. The results show the effectiveness of the model to capture the boundary of the given image, which is better than the model with global rigid transformation term only. These results were obtained through Matlab and the computational time in each iteration was about 1 second utilizing a Pentium 4 CPU running at 2.4 GHZ with 512 MB of RAM. Windows XP Home Edition was used as the operating system.

Due to the inhomogeneity, dropout, and loss of some information near the edges in heart images, parameters are adjusted in each image to get desired results. Some of the images with a diminishing problem which depends on the parameters are observed in

the experiments. Since the shape prior is used as an initial contour, the level set form of the model is fixed which does not require the re-initialization process during the numerical calculation. The intensity matching term in the energy function of the model provides the difference of the intensity variation across the prior shape and the shape of interest. Minimizing the difference of the gradients of the intensity between smoothed prior image and the smoothed given image is more accurate than minimizing the gradient term in the given image, due to the loss of the information of the gradient near the edge of the desired object. Moreover, this model can be extended to 3-d cases and any other types of images. Generalization of the model to global non-rigid deformation is needed in the future work.

References

1. Chen, Y., Tagare, H., Thiruvenkadam, S., Huang, F., Wilson, D., Gopinath, K., Briggs, R., and Geiser, E. : Using prior shapes in geometric active contours in a variational framework. *Int. J. Comp. Vis.* Vol. **50** (3) (2002) 315–328
2. Chen, Y., Huang, F., Tagare, H., Rao, M., Wilson, D., and Geiser, E. : Using prior shapes and intensity profiles in medical image segmentation. *Proc. Int. Conf. Comp. Vis. Nice France* (2003) 1117–1124
3. Chen, Y., Huang, F., Wilson, D., and Geiser, E. : Segmentation with shape and intensity priors. *Proc. Int. Conf. Image Graph. Hefei China* (2003) 378–385
4. Chen, Y., Guo, W., Huang, F., Wilson, D., and Geiser, E. : Using prior shapes and points in medical image segmentation. *Proc. Ener. Mini. Meth. Comp. Vis. Pat. Recog. Lisbon Portugal* (2003) 291–305
5. Caselles, V., Kimmel, R., and Sapiro, G. : Geodesic active contours. *Int. J. Comp. Vis.* Vol. **22** (1) (1997) 61–79
6. Chan, T. and Vese, L. : Active contours without edges. *IEEE Trans. Image Proc.* **10** (2) (2001) 266–277
7. Gaens, T., Maes, F., Vandermeulen, D., and Suetens, P. : Non-rigid multimodal image registration using mutual information. *Proc. First Int. Conf. Med. Image Comp. Computer-Assi. Interv.* vol. **1496** (1998) 1099–1106
8. Hellier, P. and Barillot, C. : Multimodal non-rigid warping for correction of distortions in functional MRI. *Proc. Third Int. Conf. Med. Image Comp. Computer-Assi. Interv.* vol. **1935** (2000) 512–520
9. Huang, X., Li, Z., and Metaxas, D. : Learning Coupled Prior Shape and Appearance Models for Segmentation. *Proc. 7th. Ann. Int. Conf. on Med. Image Comp. Computer-Assi. Interv. MICCAI04 Vol. I LNCS-3216* (2004) 60–69.
10. Huang, X., Metaxas, D., and Chen, T. : MetaMorphs: Deformable Shape and Texture Models. *Proc. IEEE Comp. Soc. Conf. Comp. Vis. Pat. Recog. CVPR04, Vol. I* (2004) 496–503
11. Kichenassamy, S., Kumar, A., Olver, P., Tannenbaum, A., and Yezzi, A. : Gradient flows and geometric active contour models. *Proc. ICCV'95* (1995) 810–815
12. Leventon, M., Faugeras, O., Grimson, E., Wells, W. : Level Set Based Segmentation with Intensity and Curvature Priors. *Proc. Math. Meth. Biomed. Image Anal.* (2000) 4–11
13. Leventon, M., Grimson, E., and Faugeras, O. : Statistical Shape Influence in Geodesic Active Contours. *Proc. IEEE Conf. CVPR* (2000) 316–323
14. Likar, B. and Permus, F. : A Hierarchical Approach to Elastic Registration Based on Mutual Information. *Image Vis. Comp.* **19**(1-2) (2000) 33–44
15. Moelich M. and Chan, T. : Joint segmentation and registration using logic models. *UCLA CAM Report 03-06* (2003)

16. Meyer, C., Boes, J., Kim, B., and Bland, P. : Probabilistic Brain Atlas Construction: Thin-Plate Spline Warping via Maximization of Mutual Information. Proc. Second Int. Conf. Med. Image Comp. Computer-Assi. Interv. vol. **1679** (1999) 631–637
17. Osher, S. and Sethian, J. : Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulation. J. Comp. Phys. Vol. **79** (1988) 12–49
18. Paragios, N. and Deriche, R. : Geodesic active regions for supervised texture segmentation. IEEE ICCV Cofu Greece (1999) 926–932
19. Paragios, N. : Geodesic active regions and level set methods: contributions and applications in artificial vision. Ph.D. Thesis Sch. Comp. Engin. Uni. Nice/Sophia Antipolis (2000)
20. Paragios, N. : A variational approach for the segmentation of the left ventricle in MR cardiac images. Proc. 1st. IEEE Workshop Var. Level Set Met. Comp. Vis. Vancouver B.C. Canada (2001) 153–160
21. Paragios, N. and Rousson, M. : Shape prior for level set representations. Comp. Vis. ECCV2002 7th. Eur. Conf. Comp. Vis. Copenhagen Demark Proc. (2002)
22. Paragios, N., Rousson, M., and Ramesh, V. : Matching distance functions: a shape-to-area variational approach for global-to-local registration. Comp. Vis. ECCV2002 (2002) 775–789
23. Richard, F. and Cohen, L. : A New Image Registraton Technique with Free Boundary Constraints: Application to Mammography. ECCV LNCS **2353** (2002) 531–545.
24. Rueckert, D., Clarkson, M., Hill, D., and Hawkes, D. : Non-rigid registration using higher-order mutual information. Med. Imaging SPIE Proc. vol. **3979** (2000) 438–447
25. Soatto, S. and Yezzi, A. : Deformation: deforming motion, shape average and joint registration and segmentation of images. Comp. Vis. ECCV (2002)
26. Young, Y. and Levy, D. : Registration-base morphing of active contours for segmentation of CT scans. Math. Biosci. Eng. vol. **2** (2005) 79–96
27. Yezzi, A., Zollei, L., and Kapur, T. : A variaional approach to joint segmentation and registration. Proc. IEEE Conf. on Comp. Vis. Pat. Recog. (2001)
28. Zhao, H., Chan, T., Merriman, B., and Osher, S. : A variational level set approach to multi-phase motion. J. Comp. Phys. Vol. **127** (1996) 179–195
29. Zitová, B. and Flusser, J. : Image registration methods: a survey. Image Vis. Comp. **21** (2003) 977–1000

A Tracking Approach to Parcellation of the Cerebral Cortex

Chris Adamson^{1,2}, Leigh Johnston¹, Terrie Inder¹, Sandra Rees⁴, Iven Mareels², and Gary Egan^{1,3}

¹ Howard Florey Institute, University of Melbourne, Australia
Chris.Adamson@hfi.unimelb.edu.au

² Dept. of Electrical and Electronic Engineering, Univ. of Melbourne, Australia

³ Centre For Neuroscience, University of Melbourne, Australia

⁴ Department of Anatomy and Cell Biology, University of Melbourne, Australia

Abstract. The cerebral cortex is composed of regions with distinct laminar structure. Functional neuroimaging results are often reported with respect to these regions, usually by means of a brain “atlas”. Motivated by the need for more precise atlases, and the lack of model-based approaches in prior work in the field, this paper introduces a novel approach to parcellating the cortex into regions of distinct laminar structure, based on the theory of target tracking. The cortical layers are modelled by hidden Markov models and are tracked to determine the Bayesian evidence of layer hypotheses. This model-based parcellation method, evaluated here on a set of histological images of the cortex, is extensible to 3-D images.

1 Introduction

The cerebral cortex is divided into specialised regions that are likely to be characterised by their laminar structure [see Figure 1(i)]. Many brain functions are localised to individual regions, and a major challenge of functional neuroimaging research involves the correlation of these regions to specific brain functions. Functional neuroimaging studies typically report results with respect to a structural atlas, the most popular in current studies being that of Brodmann [1] [see Figure 1(ii)]. However, the embryonic and post-natal formation of these regions is influenced by a combination of personal experience (environmental) and genetic influences, leading to unknown degrees of inter-subject variability in the overall cortical structure. Thus, *individualised* structural maps have applications in understanding normal human brain function as well as in the diagnosis and treatment of neurological and psychiatric diseases.

Inconsistencies arose in early postmortem atlases due to differences in experimental conditions and subjective layering schemes [2]. More recently, analytical techniques [2–4] have focused on automatic, objective identification of homogeneous regions of cortical structure in digitised images of postmortem brain sections. However, these classification methods all use inferential statistics without directly modelling the layering structure. Therefore, these schemes are sensitive to relative layer depth changes, which occur due to factors such as curvature of the cortex and the obliqueness of histological slice cuts, and are difficult to account for.

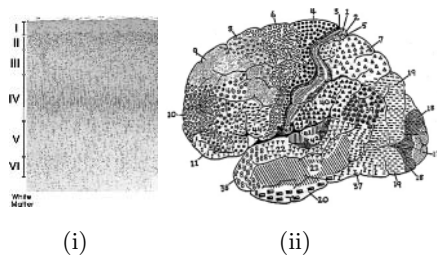


Fig. 1. (i) High-magnification image of haematoxylin and eosin (H & E) stained histological section of primary visual cortex from a baboon brain, with Roman numerals denoting the layering structure. (ii) Brodmann's cytoarchitectonic atlas for the human cortex, lateral view, from [1].

The limited number of parcellation schemes that directly model the laminar structure and are extensible to a 3D setting motivate this paper [5]. We propose a novel target tracking-based method that models the cortical layers using hidden Markov models and computes Bayesian evidence of layering to determine parcellation. Our method is grounded in the work of Xie and Evans [6], who track dominant frequencies in noisy time-varying signals using hidden Markov models. We present this paper as work that forms the basis for a method that is able to parcellate the cortex in a 3D Magnetic Resonance Image (MRI) of a living subject, as visualised in a number of recent experimental papers [7–9].

2 Method

2.1 Image Acquisition and Sampling

Coronal sections of a baboon cortex were stained with haematoxylin and eosin (H&E) and the histological images, of resolution $14\mu\text{m} \times 14\mu\text{m}$, were captured using a Nikon LS-2000 slide scanner [see Figure 2(i)].

We sample the cortex using a procedure that is directly extensible to 3-D, based on the method of Jones *et al.* [10], which solves Laplace's equation ($\nabla^2\Psi = 0$). Here Ψ is a potential function defined subject to the Dirichlet boundary conditions on the White Matter/Grey Matter (WM/GM) boundary ($\Psi(\text{WM/GM}) = 0$) and the Grey Matter/Cerebral Spinal Fluid (pial) boundary ($\Psi(\text{pial}) = 1$). Finite differences are used to determine a solution to Laplace's equation. The tangent field to this solution can be traced to define lines that span the cortex.

The raw intensity profiles resulting from this method are extremely noisy due to the high degree of texture variation in the images caused by the histological staining processes. For this reason the data is mean filtered prior to parcellation, thus smoothing the local image intensity variations. Choice of the optimal filter size is discussed in Section 3.

The inherent ordering in the set of sampled profiles implies a sequence of vectors, equivalently a matrix. When visualised as a 2D image, this ap-

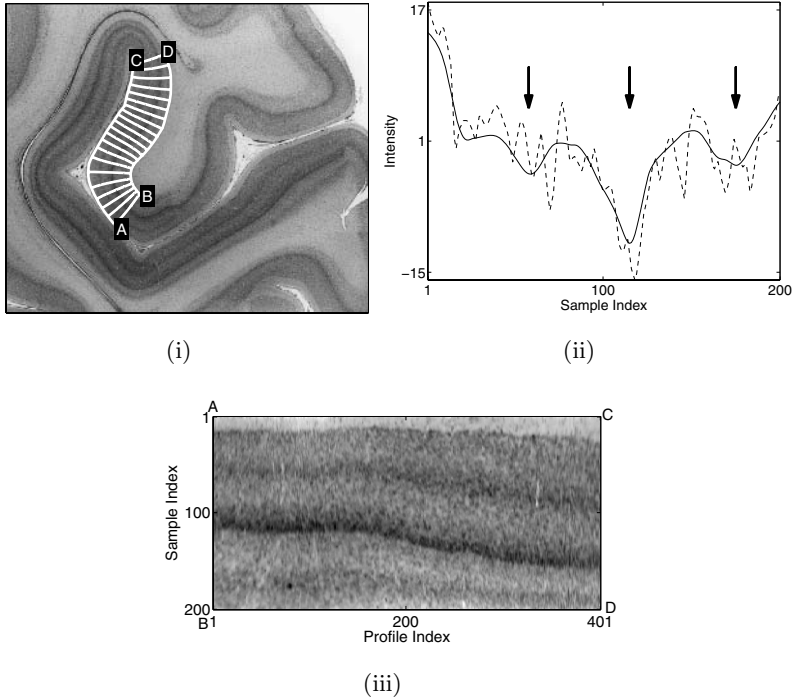


Fig. 2. (i) Histological slice of baboon cortex, with a section of the WM/GM and pial boundary contours and a sub-sample of the lines overlaid. (ii) Intensity profile 100 (dashed line = original, solid line = mean-filtered), arrows indicate centres of troughs. (iii) “Straightened” representation of section in Fig.(i).

pears to be a “straightened” representation of the sampled section of the cortex [see Figure 2(iii)]. We denote by $\mathbf{Y}_1^T = [\mathbf{y}_1, \dots, \mathbf{y}_T]$ the sequence of T filtered line profiles, where each profile is a K -vector of image intensity values, $\mathbf{y}_t = [y_{1,t}, \dots, y_{K,t}]'$, $y_{k,t} \in \mathbb{R}$.

2.2 Tracking-Based Parcellation of the Cortex

The strongest visual cue to the identification of laminar structure in the cortex is the presence of dark bands [see Figure 2(iii)]. Each distinct region in the cortex as observed in a typical histology image [see Figure 2(i)] occupies an extensive area. Thus, the dark bands that constitute these regions extend across hundreds of sampled profiles. However, the boundaries between bands in individual profiles are not always sharp and in some cases small, transient bands exist. Therefore, while a visual parcellation of the images is possible, constructing an algorithm that is able to mimic this level of non-local information processing is a difficult task.

After appropriate filtering, the dark bands manifest as troughs in the intensity profiles [see Figure 2(ii)]. The bottom of the troughs in a profile are the estimated centres of the dark bands that exist in that profile.

We have used a target tracking paradigm, since that cortical layers have predictable local dynamics well-suited to Markov modelling [11]. That is, by considering the profile index to be a temporal variable, the probabilistic centroid of a dark band is a “target” that we wish to track, taking into account its “movement” up and down through “time”.

The Markov chain state is used to denote the position of a track in the image, with the movement up and down of the bands captured by suitable choice of the transition probability matrix. The profiles are the noisy observations of these underlying states (band positions).

Introducing notation for these concepts, we define a track to be a sequence of pixel indices delineating the probabilistic centre of a dark band in the image that starts at position, (k_0, t_0) , and exists for a length of n pixels:

$$\mathbf{d}^{(k_0, t_0, n)} = \left[d_{t_0}^{(k_0, t_0, n)} = k_0, d_{t_0+1}^{(k_0, t_0, n)}, \dots, d_{t_0+n-1}^{(k_0, t_0, n)} \right] \tag{1}$$

where $d_t^{(k_0, t_0, n)} \in \{1, \dots, K\}$, $\forall t = t_0 + 1, \dots, t_0 + n - 1$. Let $\Omega = \{(k_0, t_0, n)\}$ denote the set of all track parameter 3-tuples.

We assume first-order Markov dynamics for the tracks, whereby the position of the track in profile t depends only on the position of the track in profile $t - 1$. Thus we define a transition probability matrix, A , with elements,

$$a(j, k) = p\left(d_t^{(k_0, t_0, n)} = k \mid d_{t-1}^{(k_0, t_0, n)} = j\right) \tag{2}$$

The transition matrix A is constructed to be diagonally dominant, reflecting the biological constraint that limited local “movement” is more probable than drastic “movement.”

The likelihood of a profile given a track state is constructed by assuming firstly only local dependence of the data, within a window size m ,

$$p\left(\mathbf{y}_t \mid d_t^{(k_0, t_0, n)} = k\right) = p\left([y_{k-m, t}, \dots, y_{k, t}, \dots, y_{k+m, t}] \mid d_t^{(k_0, t_0, n)} = k\right) \tag{3}$$

and secondly, that a quadratic fitted to the data $[y_{k-m, t}, \dots, y_{k, t}, \dots, y_{k+m, t}]$ will have it’s turning point, (\hat{k}, \hat{t}) , close to pixel (k, t) if this pixel is at the centre of a trough, and further away when the pixel is not at the centre of a trough. We therefore construct the likelihood function as

$$p\left(\mathbf{y} \mid d_t^{(k_0, t_0, n)} = k\right) = p\left(\rho \mid \|(k, t) - (\hat{k}, \hat{t})\| \mid d_t^{(k_0, t_0, n)} = k\right) \tag{4}$$

$$\rho \mid \|(k, t) - (\hat{k}, \hat{t})\| \sim N(0, 1) \tag{5}$$

where $\rho \in \mathbb{R}$ scales distance measure to make unit variance applicable.

Having thus formulated the hidden Markov model of the tracks, we describe the algorithm for parcellation based on this model. We calculate the evidence for the hypothesis, $H_{k, t}$, that a pixel (k, t) is situated at the probabilistic centre of a dark band. Evidence is measured by the posterior probability of the hypothesis given the data:

$$p(H_{k, t} \mid \mathbf{Y}_1^T), \quad \forall k = 1, \dots, K, \quad t = 1, \dots, T \tag{6}$$

A pixel is situated at the probabilistic centre of a dark band if and only if a track, as defined in (1), passes through that point. That is, that there exists some (k_0, t_0, n) for which $d_t^{(k_0, t_0, n)} = k$. Therefore by the law of total probability, (6) can be expressed in terms of a sum of posterior probabilities over all possible tracks:

$$p(H_{k,t} | \mathbf{Y}_1^T) = \sum_{\{k_0, t_0, n\} \in \Omega} p(d_t^{(k_0, t_0, n)} = k | \mathbf{Y}_1^T), \forall k = 1, \dots, K, t = 1, \dots, T \quad (7)$$

The posterior probability of pixel (k, t) belonging to track $\mathbf{d}^{(k_0, t_0, n)}$ is rewritten using Bayes' rule as

$$p(d_t^{(k_0, t_0, n)} = k | \mathbf{Y}_1^T) = \frac{p(\mathbf{Y}_1^T | d_t^{(k_0, t_0, n)} = k) p(d_t^{(k_0, t_0, n)} = k)}{p(\mathbf{Y}_1^T)} \quad (8)$$

The prior in (8) is considered to contain a local and a global term:

$$p(d_t^{(k_0, t_0, n)} = k) = p_{loc}(d_t^{(k_0, t_0, n)} = k) p_{glob}(d_t^{(k_0, t_0, n)} = k)$$

The local prior is the first-order Markov assumption discussed above. The global prior is a weighting term on the length of the track, such that longer tracks are more probable than shorter ones. This local/global factorisation of the prior term is used as a tool in image processing methods that attempt to capture both short and long range properties of an object of interest [12].

We make a conditional independence assumption on the image data; profiles that are not between t_0 and $(t_0 + n - 1)$ are independent of the track state $d_t^{(k_0, t_0, n)}$. Let $\mathbf{Y}_{t_0}^{t_0+n-1}$ be the sequence of profiles between t_0 and $t_0 + n - 1$. Then, by cancelling independent data in numerator and denominator, and by reverse application of Bayes' rule, (8) becomes

$$p(d_t^{(k_0, t_0, n)} = k | \mathbf{Y}_{t_0}^{t_0+n-1}) \propto p_{loc}(d_t^{(k_0, t_0, n)} = k | \mathbf{Y}_{t_0}^{t_0+n-1}) p_{glob}(d_t^{(k_0, t_0, n)} = k)$$

Here, $p_{loc}(d_t^{(k_0, t_0, n)} = k | \mathbf{Y}_{t_0}^{t_0+n-1})$ is a state occupancy probability, computed by the polynomial-time hidden Markov model (HMM) smoother, commonly known as the forward-backward algorithm [11]. The HMM smoother computes the state occupancy probability via

$$p_{loc}(d_t^{(k_0, t_0, n)} = k | \mathbf{Y}_{t_0}^{t_0+n-1}) = \alpha_t(k) \beta_t(k) / \left(\sum_{k=1}^K \alpha_t(k) \beta_t(k) \right) \quad (9)$$

where $\alpha_t(k)$ and $\beta_t(k)$ are determined in forward and backward sweeps through the data respectively (see [11] for details).

The initial condition that the track starts in pixel (k_0, t_0) is reflected by the initialisation of the forward variable:

$$\alpha_{t_0}(k) = \begin{cases} 1, & k = k_0 \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

The backward variable is initialised to reflect that the track can end in any state:

$$\beta_{t_0+n-1}(k) = 1, \forall k = 1, \dots, K \quad (11)$$

The computational requirement in running an HMM smoother for all possible track parameters, (k_0, t_0, n) , to calculate the posterior probability of the evidence (7) at each pixel is prohibitive. Instead, we limit the size of the parameter space, Ω , by considering only a set of points that are highly likely to be the start of tracks, as measured by a large change in likelihood between two horizontally-neighbouring pixels.

Having computed the posterior probability (7) for all pixels in the image, we sweep the image left-to-right and classify regions according to the appearance and disappearance of tracks.

3 Results

The results of applying the proposed method to three example histological images acquired as described in Section 2 are shown in Figures 3. The size of the mean filter was chosen to be 10×100 pixels, optimising the trade-off between

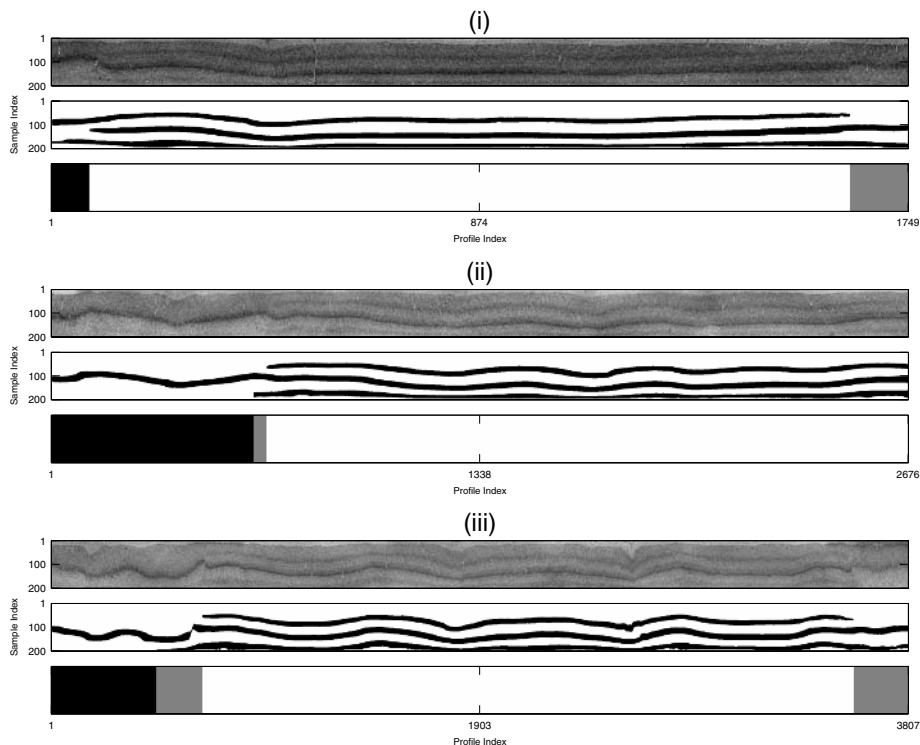


Fig. 3. (i)–(iii) Three histological sections of baboon cortex from different parts of a single brain. In each, Top: “straightened” representation of profiles, Middle: posterior probability map, dark indicates higher probability, Bottom: parcellation of line profiles; note that there is no correspondence between regions with the same colour in different figures.

unwanted image blur and noise texture smoothing. The likelihood normalising factor was set to $\rho = 200$, while the global prior was chosen to be a $N(T, 500)$ distribution, where T is the width of the image, and $\sigma^2 = 500$ was chosen to appropriately weight a range of longer tracks, i.e. those tracks more likely to within a dark band. The tuning parameters, ρ and σ^2 , while somewhat arbitrary, are analogous to tuning parameters in target tracking schemes which are in general also arbitrarily chosen.

As is evidenced in Figure 3, the parcellation into regions of distinct laminar structure has been accomplished by the proposed method. The small transition region (grey region between black and white) in Figure 3(ii) is caused by the extension of the bottom dark band past the termination of the topmost dark band. The transition region in Figure 3(iii) is caused by an erroneous weak trough attaching to a stronger dark trough. We are currently investigating likelihood models that suppress weak troughs more effectively.

4 Conclusions

The probabilistic, model-based method proposed in this paper improves on previous parcellation schemes, both in its applicability to 3D data, and its modelling of the layer structure known to be present in the data. In this latter regard, our algorithm makes use of the prior information known about the data, i.e. direct identification of the biological structure underlying the problem, and is therefore more robust and generalisable than methods based solely on inferential statistical measures. We have demonstrated that our method produces a parcellation that contains contiguous regions that correspond to discernible regions of homogeneous laminar structure.

The target tracking paradigm we employ models the probabilistic centres of the dark bands in the cortex by hidden Markov models and tracks their position through the data. A notion of causality is imposed on the data by considering the horizontal axis to be a temporal variable, which is possible due to the sequential organisation of the profiles.

Parcellation in 3D is considerably more difficult, namely that the 1D linear profile arrangement is replaced by a 2D graphical organisation. This 2D arrangement has no notion of causality, which is exploited in a target tracking paradigm. However, certain fundamental aspects of the algorithm, such as robustness to band depth changes and image artifacts will form the basis of a 3D parcellation algorithm, which is the focus of current work.

References

1. Garey, L.J.: Brodmann's Localisation in the Cerebral Cortex. Smith-Gordon, London (1994)
2. Schleicher, A., Amunts, K., Geyer, S., Morosan, P., Zilles, K.: Observer-Independent Method for Microstructural Parcellation of Cerebral Cortex: A Quantitative Approach to Cytoarchitectonics. *NeuroImage* **9** (1999) 165 – 177

3. Annese, J., Pitiot, A., Dinov, I.D., Toga, A.W.: A myelo-architectonic method for the structural classification of cortical areas. *NeuroImage* **21** (2004) 15 – 26
4. Adamson, C., Davies, R., Inder, T., Rees, S., Mareels, L., Egan, G.: Markov random field-based parcellation of the cerebral cortex: Application to histology images. In: *ISSNIP 2004. Volume 1.* (2004) 559 – 564
5. Kruggel, F., Brückner, M.K., Arendt, T., Wiggins, C.J., von Cramon, D.Y.: Analyzing the Neocortical Fine-Structure. *Medical Image Analysis* **7** (2003) 251 – 264
6. Xie, X., Evans, R.J.: Multiple target tracking and multiple frequency line tracking using hidden Markov models. *IEEE Transactions on Signal Processing* **39** (1991) 2659 – 2676
7. Walters, N., Egan, G., Kril, J., Kean, M., Waley, P., Jenkinson, M., Watson, J.: *In vivo* Identification of Human Cortical Areas using High-Resolution MRI: An Approach to Cerebral Structure-Function Correlation. *Proceedings of the National Academy of Sciences* **100** (2003) 2981 – 2986
8. Barbier, E.L., Marrett, S., Danek, A., Vortmeyer, A., van Gelderen, P., Duyn, J., Bandettini, P., Grafman, J., Koretsky, A.P.: Imaging Cortical Anatomy by High-Resolution MR at 3.0T: Detection of the Stripe of Gennari in Visual Area 17. *Magnetic Resonance in Medicine* **48** (2002) 735 – 738
9. Bridge, H., Clare, S., Jenkinson, M., Jezzard, P., Parker, A.J., Matthews, P.M.: Independent anatomical and functional measures of the V1/V2 boundary in human visual cortex. *Journal of Vision* **5** (2005) 93 – 102
10. Jones, S.E., Buchbinder, B.R., Aharon, I.: Three-Dimensional Mapping of Cortical Thickness Using Laplace’s Equation. *Human Brain Mapping* **11** (2002) 12 – 32
11. Rabiner, L.R.: A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE* **77** (1989) 257 – 286
12. Elder, J.H., Krupnik, A., Johnston, L.A.: Contour Grouping with Prior Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25** (2003) 661 – 674

Cell Segmentation, Tracking, and Mitosis Detection Using Temporal Context

Fuxing Yang¹, Michael A. Mackey², Fiorenza Ianzini^{2,3},
Greg Gallardo¹, and Milan Sonka¹

¹ Department of Electrical and Computer Engineering

² Departments of Pathology and Biomedical Engineering

³ Department of Radiation Oncology,

The University of Iowa, Iowa City, IA, 52242

fuxing-yang@uiowa.edu

Abstract. The Large Scale Digital Cell Analysis System (LSDCAS) developed at the University of Iowa provides capabilities for extended-time live cell image acquisition. This paper presents a new approach to quantitative analysis of live cell image data. By using time as an extra dimension, level set methods are employed to determine cell trajectories from 2D + time data sets. When identifying the cell trajectories, cell cluster separation and mitotic cell detection steps are performed. Each of the trajectories corresponds to the motion pattern of an individual cell in the data set. At each time frame, number of cells, cell locations, cell borders, cell areas, and cell states are determined and recorded. The proposed method can help solving cell analysis problems of general importance including cell pedigree analysis and cell tracking. The developed method was tested on cancer cell image sequences and its performance compared with manually-defined ground truth. The similarity Kappa Index is 0.84 for segmentation area and the signed border positioning segmentation error is $1.6 \pm 2.1 \mu\text{m}$.

1 Introduction

1.1 Living Cells

Study of living cells including cell death, cell motility, measurement of intracellular pro-oxidant species, and the determination of phenotypic changes observed using adenovirus-mediated gene expression systems and many other aspects will help to diagnose and assess the natural course of many diseases.

The need for investigation of cell behavior and movement led to the development of automated systems designed to quantify the reaction of cells in different environments. An automated system, the Large Scale Digital Cell Analysis System (LSDCAS) was designed to analyze large numbers of cells under a variety of experimental conditions in the Real-Time Cell Analysis Laboratory, Department of Biomedical Engineering, University of Iowa, in the previous years. The LSDCAS is capable of monitoring thousands of microscope fields over time intervals of up to one month. LSDCAS was originally designed to study stress-induced

mitotic catastrophe [1]. Current efforts concentrate on providing imaging and data modeling/visualization technologies that will be useful in many areas of biological research.

1.2 Previous Work and Associated Problems

Corresponding to changing properties of a variety of cells, many distinctive methods have been proposed and employed in the analysis of cellular structure.

A part of the methods which have been widely used in cellular imaging segmentation and tracking are model-free approaches, such as thresholding [2], median filtering [3], watershed segmentation [4], and others. However, these approaches provide no or only poor descriptions of the cellular shape. Even worse, they frequently do not produce closed-object contours and require extensive post-processing steps [5]. These approaches often fail to correctly track multiple cells in sequences exhibiting cell-cell contacts [6].

Model-based techniques, such as active contours [7], are also frequently used. These active contours may be parametric format, geometric, or may use region based models. The methods produce closed and smooth object boundaries, and may provide a first guess through the interactive initialization step. They allow keeping track of object identity in an obvious manner through automatic initialization based on the segmentation of the previous image[8, 5]. With the introduction of the Level Set Methods [9], a powerful mathematical tool to solve the problem of cusps, corners, and automatic topological changes, the active contour methods became well suited to cell image analysis.

It is not sufficient to rely on pure image information for cell analysis. Pattern recognition based segmentation method of cell nuclei in tissue section analysis was utilized in [10]. Neural network [11] and genetic algorithms [12] were applied to cell image analysis. The pattern recognition techniques helped achieve cell analysis objectives in specific cases. However, defining and computing the rules as well as performing the training steps is difficult.

Little effort was devoted to quantitative living tumor cell analysis largely due to its complicated nature: cell shape variability; contact between cells; weak cell boundaries, etc. Many existing methods are only effective in low-density situations and temporal context is frequently not considered.

2 Methods

2.1 Cell Trajectories Extraction

To identify trajectories of moving cells, time is taken as an extra dimension and the level set method-based geometric active contour is directly applied onto a 2D+time image sequence. For each frame, image boundary is taken as the initial location of the moving front. Using the Fast Marching method, the front is moved inward through the gradient defined on the cost image until it stops at the location close to the cell or cell cluster boundary. Next, Narrow Band Level Set method is applied to achieve a more accurate segmentation.

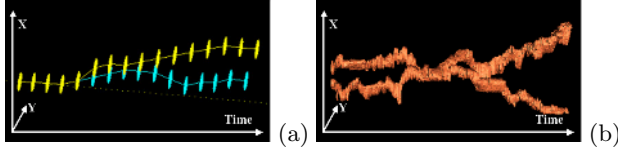


Fig. 1. Trajectory model.(a) The cell trajectory model in 2D + time space, each disk represents a cell or cell compound at a single time and the model is showing a normal cell division. (b) A cell trajectory, with 2 cells in the field, the total number of frames being 300, the time span between frames was 5 minutes and the total time span was 1500 minutes.

The cell trajectory model behind this method is shown in Fig. 1. An example of the real cell trajectory segmented from the LSDCAS data set is also shown.

In this step, the multi-channel cost image is used for level set segmentation, which is $f(intensity) + g(localvariance)$, a combination of the different kinds of information from the original image. $f()$ is the function of the image intensity and $g()$ is the function of the local variance.

A special-designed edge enhancement step is included to prevent the moving front evolving into the internal portion of the cells which have weak boundary areas. Let $\theta(x, y) : R^2 \rightarrow R$ and the arc be defined as: $d\theta = \frac{\partial\theta}{\partial x}dx + \frac{\partial\theta}{\partial y}dy$. Based on Riemannian geometry standard notation, we have:

$$\lambda_{\pm} = \frac{s_{xx} + s_{yy} \pm \sqrt{(s_{xx} - s_{yy})^2 + 4s_{xy}}}{2} \tag{1}$$

where second order derivative $s_{xx} = (\frac{\partial\theta}{\partial x}).(\frac{\partial\theta}{\partial x})$, $s_{yy} = (\frac{\partial\theta}{\partial y}).(\frac{\partial\theta}{\partial y})$, $s_{xy} = s_{yx} = (\frac{\partial\theta}{\partial x}).(\frac{\partial\theta}{\partial y})$. These derivatives are approximated using forward difference method and in 0 - 1 scale. In the image, the maximal and minimal changes at a specific location are provided by the eigenvalues λ_{\pm} . In our approach, we define the driving force for the moving front as $F = e^{-\rho*(\lambda_+ - \lambda_-)}$.

2.2 Mitosis Detection

In the life cycle of living cells, mitosis (cell division) may provide extra information. Mitosis typically involves a series of steps consisting of prophase, metaphase, anaphase, and telophase, and results in the formation of two new nuclei each having the same number of chromosomes as the parent nucleus. Mitotic cells tend to be circular and have a larger percentage of bright pixels than non-mitotic cells. Once mitosis is detected, the change of the number of the cells in current field is observed. In a long temporal sequence, the frames between mitotic states have the same number of cells (not considering cells moving in or out of the current field of view).

The mitosis detection is done according to the following procedure:

1. Apply an optimal threshold method [13] to remove the effect from non-mitotic cells in the cell compound identified in the cell trajectory and to get the Region of Interest (ROI) for the mitotic candidates.

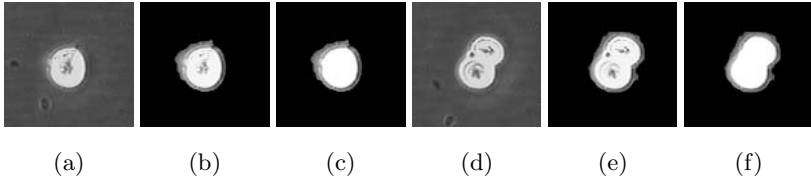


Fig. 2. Mitosis detection. (a) The original image at frame 12. (b) The profile of the cell trajectory at frame 12. (c) After threshold and movement under curvature via level set methods at frame 12, the ROI properties are $P = 144$, $A = 1540$, $AI = 237$, $C = 0.933$. (d) The original image at frame 13. (e) The profile of the cell trajectory at frame 13. (f) After threshold and curvature-based smoothing via level set methods at frame 13, the ROI properties are $P = 176$, $A = 2072$, $AI = 236$, $C = 0.841$. (P and A in pixels).

2. Apply Level Set Methods based on moving under curvature [9] to smooth the ROI boundary.
3. Compute ROI properties: area, perimeter, circularity, etc.
4. In adjacent frames t and $t + 1$, only the ROI whose area A increases, perimeter P increases, circularity C decreases, and average intensity AI remains unchanged are considered a mitosis.

The mitosis detection uses a classifier trained on an independent training set. An example of the mitosis detection procedure is shown in Fig. 2.

2.3 Cell Separation and Segmentation

The next step is extraction of the cell trajectory between the already identified mitotic frames. If there is no mitosis detected from the trajectory, the whole trajectory has a fixed number of cells (possibly more than 1 if cell compounds are present). Backward and forward adjacent frame separation and segmentation are attempted for the trajectories which have cell compounds exhibiting cell-cell contact.

A special data structure is utilized which helps to generate an effective description of the trajectory. Importantly, it sets up a connection graph through the trajectory. Each node stores the information for the individual region corresponding to each time frame. The information includes: area, perimeter, average intensity, father identifier (the connected individual region in the previous frame), father label, son number (the connected individual region at next frame), son label.

We applied marker-based watershed method for the cell separation. The watershed method is implemented via rain falling simulation [14]. The separation is directional because, in addition to the number of cells, the separating cell border in adjacent frame helps to segment and separate the cell cluster in the current frame.

For a non-mitosis example, in Fig. 3 (a) and (b), which are the connection graph and the separation result, node 13 is the first cluster to be separated using nodes 8 and 9 as the reference markers. The processing direction at node 13 is forward. Node 53 will be the second to be separated using nodes 57 and 58 as a reference and the processing direction here is backward. Basically, the separation

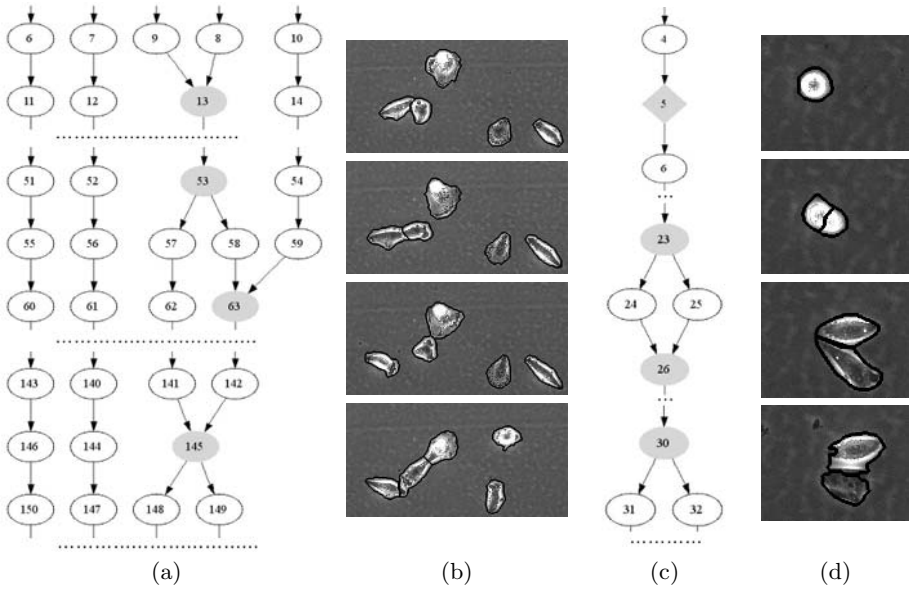


Fig. 3. Cell connection graph and cell segmentation/separation. (a) No mitosis detected. The arrow in the graph is representing the time change and links the connected node at the adjacent frame. The gray circles are the very first clusters to be separated, where cell contact happens. (b) The cell separation corresponding to (a). The first image is corresponding to frame at the first gray circle 13. The second image is for the frame at gray circle 53. The third image is for the frame at gray circle 65. The fourth image is for the frame at gray circle 145. (c) With mitosis detected. The gray diamond in this graph represent the labeled mitotic cell. The merging in this graph with gray circle is due to the cell contact and the cell trajectory generated by the level set methods was not able to separate them. (d) The cell separation corresponding to (c). The first image is for the frame at gray diamond 5, which is a mitotic cell. The second image is for the frame just after mitosis detected, at circle 6. The third image is for the frame at gray circle 23. The fourth image is for the frame at gray circle 26.

strategy is from known to unknown, from large number of cells or cell clusters per frame to a smaller number.

Similarly for a mitosis example, in Fig. 3 (c) and (d) in which a mitotic cell is detected at node 5, the whole trajectory is divided into 2 parts. The first part starts at node 1 (not shown in the image) and ends at node 5. The second part of the trajectory starts at node 6. For each new trajectory, the same method was applied as was used for the trajectory where there was no mitosis detected.

3 Experimental Methods

Image Data

The performance of our cell sequence analysis was tested in field images acquired using the LSDCAS. All the cells were U87-MG cells.

Validation

The manually identified independent standard was defined using a user-friendly border tracing program. Automated segmentation was obtained for all the cells in each of the data sets. The independent standard was used to determine the signed (bias) and unsigned mean border positioning errors as well as the maximum border positioning errors.

Positioning error: For each pixel along the automatically-identified (*AI*) border of individual cell, a nearest counterpart is found on the manually-identified (*MI*) border of the same cell and the distance between these two pixels is computed. If the counterpart is inside of the *AI* contour, we define the distance as positive. Otherwise the distance is negative. For each cell, the average value of the absolute distances was taken as the unsigned positioning error and the average value of the signed distances is taken as the signed positioning error.

Area error: Let *A* denote the *AI* segmentation area, *M* the *MI* segmentation area representing the ground truth. Ratios of overlapping areas were assessed by applying the similarity Kappa Index (*KI*)[15] and the overlap, which are defined as:

$$KI = 2 \times \frac{A \cap M}{A + M} \quad ; \quad overlap = \frac{A \cap M}{A \cup M}$$

4 Results

The reported method successfully identified the individual cells in 50 2D + time microscopy data sets collected from the LSDCAS system. A total of 3368 frames containing 6654 cells was analyzed. There are totally 26 mitoses among all the data sets and our method successfully made the detection and recorded the mitosis time and location of all of them.

The analysis errors obtained with no manual interaction or editing are shown in the following table.

Table 1. Segmentation and separation errors

validation indices	in pixel	in μm
Mean absolute positioning errors	4.2 ± 2.8	2.6 ± 1.7
Mean signed positioning errors	2.6 ± 3.4	1.6 ± 2.1

The area error *KI* was 0.84 ± 0.09 and the overlap index was 0.74 ± 0.12 .

5 Discussion and Conclusion

Quantitative analysis of living cells is a very complex task. It has strict requirements on hardware for data acquisition as well as on software for advanced quantitative analysis. All the processing modules, including preprocessing, segmentation, detection of cell splits, etc. are of great importance to the overall result. In

this paper, we introduced a new fully automated method for cell segmentation, separation and spatial cell state detection. Our approach helps solving one of the most crucial problems facilitating quantitative analysis of living cells.

The connection graph, as a by-product, gives a complete description of the cell-dividing temporal data set. Each cell trajectory can be directly used to describe behavior of a specific cell. For the cell trajectory with mitotic events, a pedigree tree can be generated by combining all nodes on each branch of the connection graph.

Under some circumstances involving abnormal cancer cells, atypical cell divisions may happen when the cell reaches the mitotic state. Compared to normal cell division when the cell begins as a single round entity and then splits into two completely separate objects, the abnormal cell division states include a multipolar state (when one cell moves into 3 or more distinct circular shapes), fuse state (which is defined as a cell which attempts to divide but is interrupted internally and re-forms as a single cell), and cell death (cell ceases to continue intracellular activities). These abnormal situations bring additional difficulty to the trajectory analysis since the a priori information used in our approach cannot be reliably extracted using the presented mitosis-detection method. A more powerful classifier combining cell trajectory analysis with cell status assessment is under development, which is based on Hidden Markov Models (HMMs) using time varying information from bright object sequences found in the cell compounds. The robustness of the HMM-based classifier shall provide reliable information helping to guide the segmentation of cancer cells in the image sequences.

Overall, the reported method is a substantial improvement in comparison with most existing cell analysis approaches. The temporal information is fully utilized, yielding more robust segmentation results compared to the previous frame-by-frame approaches [4].

References

1. Ianzini, F., Mackey, M.: Development of the Large-Scale Digital Cell Analysis System. *Radiation Protection and Dosimetry* **99** (2002) 81–94
2. Wu, H., Gil, J., Barba, J.: Optimal segmentation of cell images. In: *Proceedings of IEEE Vision, Image and Signal Processing*. Volume Vol. 145., IEEE (1998) 50–56
3. Anoraganingrum, D.: Cell segmentation with median filter and mathematical morphology operation. In: *Proceedings of 1999 International Conference on Image Analysis and Processing*. (1999) 1043–1046
4. Yang, F., Mackey, M., Ianzini, F., Gallardo, G., Sonka, M.: Segmentation and quantitative analysis of the living tumor cells using Large Scale Digital Cell Analysis System. In: *Proceedings of SPIE Conference on Medical Imaging*,. Volume 5370. (2004) 1755–1763
5. Zimmer, C., Labruyere, E., Meas-Yedid, V., Guillen, N., Olivo-Marin, J.C.: Segmentation and tracking of migrating cells in videomicroscopy with parametric active contours: A tool for cell-based drug testing. *IEEE Transactions on Medical Imaging* **Vol. 21** (2002) 1212–1221

6. Kirubarajan, T., Bar-Shalom, Y., Pattipati, K.R.: Multiassignment for tracking a large number of overlapping objects. *Multitarget-Multisensor Tracking: Applications and Advances* **Vol. III** (2000) 199–231
7. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active Contour Models. *International Journal of Computer Vision* (1988) 321–331
8. Leymarie, F., Levine, M.: Tracking deformable objects in the plane using an active contour model. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **Vol. 15** (1993) 617–634
9. Sethian, J.: *Level Set Methods and Fast Marching Methods Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*. Cambridge University Press (1999)
10. Spyridonos, P., Glotsos, D., Cavouras, D., Ravazoula, P., Zolota, V., Nikiforidis, G.: Pattern recognition based segmentation method of cell nuclei in tissue section analysis. In: *Proceedings of 14th IEEE International Conference on Digital Signal Processing (DSP2002)*. (2002) 1121–1124
11. Nattkemper, T.W., Ritter, H.J., Schubert, W.: A neural classifier enabling high-throughput topological analysis of lymphocytes in tissue sections. *IEEE Transactions on Information Technology in Biomedicine* **Vol. 5** (2001) 138–149
12. Lassouaoui, N., Hamami, L.: Genetic algorithms and multifractal segmentation of cervical cell images. In: *Proceedings of Seventh International Symposium on Signal Processing and Its Applications*. (2003) 1–4
13. Kittler, J., Illingworth, J.: Minimum error thresholding. *Pattern Recognition* **Vol. 19** (1986) 41–47
14. Stoev, S.L.: Rafsi - a fast watershed algorithm based on rainfalling simulation. In: *Proceedings of 8-th International Conference on Computer Graphics, Visualization, and Interactive Digital Media (WSCG'2000)*. (2000) 100–107
15. Zildenbos, A., Dawant, B., Margolin, R.: Morphometric analysis of white matter lesions in MR images: Method and validation. *IEEE Transactions on Medical Imaging* **Vol. 13** (1994) 716–724

A Unifying Approach to Registration, Segmentation, and Intensity Correction

Kilian M. Pohl¹, John Fisher¹, James J. Levitt², Martha E. Shenton², Ron Kikinis²,
W. Eric L. Grimson¹, and William M. Wells²

¹ Computer Science and Artificial Intelligence Lab,
Massachusetts Institute of Technology, Cambridge, MA, USA
{pohl, fisher, welg}@csail.mit.edu

² Surgical Planning Laboratory,
Harvard Medical School and Brigham and Women's Hospital, Boston, MA, USA
{jlevitt, martha, kikinis, sw}@bwh.harvard.edu

Abstract. We present a statistical framework that combines the registration of an atlas with the segmentation of magnetic resonance images. We use an Expectation Maximization-based algorithm to find a solution within the model, which simultaneously estimates image inhomogeneities, anatomical labelmap, and a mapping from the atlas to the image space. An example of the approach is given for a brain structure-dependent affine mapping approach. The algorithm produces high quality segmentations for brain tissues as well as their substructures. We demonstrate the approach on a set of 22 magnetic resonance images. In addition, we show that the approach performs better than similar methods which separate the registration and segmentation problems.

1 Introduction

With notable exceptions, segmentation and registration have been treated as two separate problems in medical imaging research. However, these techniques complement each other. For example, segmentation simplifies the registration of anatomical structures with ambiguous intensity patterns [1]. On the other hand, aligning an atlas to these anatomical structures aids the detection of indistinct boundaries and therefore simplifies the segmentation problem [2]. In this paper, we describe a simultaneous solution to both problems by combining them in a unified Bayesian framework.

The idea of the unified Bayesian framework was motivated by boundary localization techniques, such as [3, 4], which align an atlas to the subject and simultaneously estimate the shape of a structure. These methods relate both problems to each other by extending the definition of the shape to include its pose. This paper describes an integrated segmentation and registration approach for voxel-based classification methods. In contrast to boundary localization approaches, voxel-based classification methods consider the anatomical structure associated with each voxel. In addition, they often explicitly model the image inhomogeneities of Magnetic Resonance Images (MRI) to segment large data sets without manual intervention.

Voxel-based classification methods have coupled registration and segmentation of misaligned images [5, 6], however, we wish to align an atlas to MRI and separate the

images into anatomical structures. Previous voxel-based classification methods perform this task sequentially [7, 8, 2] increasing the risk of systematic biases [1]. In contrast, our new approach is based on the principle of least commitment so that an initial imperfect estimation converges to a good approximation for each problem.

Similar to [8, 9, 1], this paper is based on an instance of the Expectation Maximization Algorithm (EM) using non-stationary priors to outline structures with indistinct boundaries and to estimate image inhomogeneities. Instead of *sequentially* performing registration and segmentation [8, 9, 1], we propose in Section 2 a Bayesian framework describing the relationship between atlas registration, intensity correction, and image segmentation. This framework is based on a Maximum A posteriori Probability (MAP) estimation formulation approximating the solution to these three interrelated problems.

Section 3 applies the concept to a hierarchical registration framework modeling global- and structure-dependent deformations. The limits and benefits of the implementation are illustrated in Section 4 by presenting a study comparing the robustness of our algorithm with respect to other EM implementations. In this study, the automatic methods outline a set of 22 MRIs into the major brain tissue classes as well as the thalamus and caudate, which are structures with indistinct boundaries.

2 Deriving an EM Framework for Simultaneous Inhomogeneity Correction, Registration, and Segmentation

The accuracy of segmenting structures that have indistinct boundaries on MR images with tissue classification methods significantly depends on properly modeling the image inhomogeneities as well as correctly registering the atlas to the subject. In this section, we develop a unified framework which performs segmentation, registration and inhomogeneity correction simultaneously.

2.1 A Maximum a Posteriori Probability Estimation Problem

Due to their complex dependencies, it is very difficult to extract the inhomogeneities \mathcal{B} and the registration parameters \mathcal{R} from the MRI I without the explicit knowledge of the unknown true segmentation. However, this problem is greatly simplified within an EM framework where it is formulated as an incomplete data problem marginalizing over all possible segmentations \mathcal{T} . To determine \mathcal{B} and \mathcal{R} within this framework, we define the following MAP estimation problem over the incomplete data model:

$$(\hat{\mathcal{B}}, \hat{\mathcal{R}}) = \arg \max_{\mathcal{B}, \mathcal{R}} \log P(\mathcal{B}, \mathcal{R} | I). \quad (1)$$

In general, this results in a system of equations for which there is no analytical solution.

To estimate a solution for this problem, we propose an EM model that improves the initial estimate $(\mathcal{B}', \mathcal{R}')$ by solving the following MAP estimation problem:

$$(\mathcal{B}', \mathcal{R}') \leftarrow \arg \max_{\mathcal{B}, \mathcal{R}} E_{\mathcal{T} | I, \mathcal{B}', \mathcal{R}'} (\log P(I | \mathcal{T}, \mathcal{B}) + \log P(\mathcal{R} | \mathcal{T}) + \log P(\mathcal{B})) \quad (2)$$

The relationship between Equation (1) and Equation (2) is described in detail in our technical report [10]. This report shows that the update rule of Equation (2) yields a

better estimate of $(\mathcal{B}', \mathcal{R}')$ as measured by Equation (1). In addition, this EM framework guarantees convergence to a local maxima of the objective function [11].

Based on the derivation of [10], the solution of Equation (2) improves $(\mathcal{B}', \mathcal{R}')$ by iterating between the Expectation-Step (E-Step) and Maximization-Step (M-Step). The E-Step calculates the posterior of an anatomical structure a with respect to voxel x

$$\mathcal{W}_x(a) = \frac{P(I_x | \mathcal{T}_x = e_a, \mathcal{B}'_x) \cdot P(\mathcal{T}_x = e_a | \mathcal{R}')}{P(I_x | \mathcal{B}'_x, \mathcal{R}')} \quad (3)$$

with \mathcal{T}_x being the indicator random vector at voxel x . e_a is zero but at position a , where it is one. The M-Step updates the estimates by solving the following two MAP problems

$$\mathcal{R}' \leftarrow \arg \max_{\mathcal{R}} \sum_x \sum_a \mathcal{W}_x(a) \cdot \log P(\mathcal{T}_x = e_a | \mathcal{R}) + \log P(\mathcal{R}) \quad (4)$$

$$\mathcal{B}' \leftarrow \arg \max_{\mathcal{B}} \sum_x \sum_a \mathcal{W}_x(a) \cdot \log P(I | \mathcal{T}_x = e_a, \mathcal{B}) + \log P(\mathcal{B}) \quad (5)$$

A variety of closed-form solutions for Equation (5) have been proposed in the literature such as [8] and [12].

The remainder of this paper focuses on modeling Equation (4). Before doing so we must point the method's sensitivity towards uncommon characteristics favored by the atlas; a minor drawback of this formulation. The aligned atlas is represented in Equation (4) and Equation (3) by $P(\mathcal{T}_x = e_a | \mathcal{R})$. To match the atlas with the segmentation weights, the algorithm compensates for any biases through intensity correction and atlas realignment. For example, if the atlas does not properly capture the brain intensity distribution, the algorithm might identify the neck as part of the brain. The inhomogeneity correction will then normalize the intensity pattern of the neck to the once inside the brain and the registration might scale the atlas of the brain to cover both brain and neck. This causes the algorithm to converge to a suboptimal solution but a re-calibration of the atlas to the intensity pattern of the MRI protocol can overcome this problem.

In summary, we find a local maxima to the difficult MAP problem of Equation (1) by solving it within an EM framework. The E-Step determines \mathcal{W} of Equation (3) and the M-Step solves for the MAP estimates specified by Equation (4) and Equation (5).

2.2 Defining a Hierarchical Registration Approach

To solve the MAP estimation problem of Equation (4), we define the registration parameters \mathcal{R} , $\log P(\mathcal{T} | \mathcal{R})$, and $P(\mathcal{R})$. The parameters \mathcal{R} model a hierarchical registration framework, which distinguishes between global- and structure-dependent deformations. We then apply the registration framework to the MAP estimation problem and find a solution with another optimization algorithm.

The hierarchical registration parameters \mathcal{R} capture the correspondence between atlas, brain, and structures within the brain. The parameters \mathcal{R} can be structure-dependent or -independent. Structure-independent parameters capture the correspondence between atlas and image space. When limited to affine interpolation, the degrees of freedom of \mathcal{R} are too low to capture the characteristics of individual brain structures ([13]). The alternative is a more general non-rigid framework which often has problems aligning structures with indistinct boundaries [1].

Our Structure-dependent registration parameters treat the relationship between atlas and image space for each structure independently. Since most of the misalignment is structure-independent, e.g the patient’s head is not aligned with the atlas space, we expect small differences between structure-dependent parameters of different structures. We model this dependency with a hierarchical registration framework $\mathcal{R} = (\mathcal{R}_G, \mathcal{R}_S)$. \mathcal{R}_G are the global registration parameters, which describe the non-structure dependent deformations between atlas and image. $\mathcal{R}_S \triangleq (\mathcal{R}_1, \dots, \mathcal{R}_N)$ are the structure dependent parameters of structure $1, \dots, N$, which represent the residual structure-specific deformations not adequately explained by \mathcal{R}_G . The similarity between structure specific parameters is encoded in Equation (4) through the prior probability $P(\mathcal{R})$ enforcing tight bounds on \mathcal{R}_S and weak constraints on \mathcal{R}_G .

The mapping of the atlas to the image space is performed by an interpolation function $r(\mathcal{R}_G, \mathcal{R}_a, x)$, which maps voxel x into the coordinate system of structure a . Unlike global affine registration methods, this results in structure dependent coordinate systems represented by $(r(\mathcal{R}_G, \mathcal{R}_1, \cdot), \dots, r(\mathcal{R}_G, \mathcal{R}_N, \cdot))$ that are not aligned with each other.

Let \mathcal{R}_a of \mathcal{R}_S be the parameters specific to structure a with $a \in \{1, \dots, N\}$. If we define f_a as the probability over voxel location in the atlas space conditioned on structure a then $f_a(r(\mathcal{R}_G, \mathcal{R}_a, \cdot))$ is defined in the structure specific coordinate system of the patient. Thus, we can model the conditional structure probability:

$$P(\mathcal{T}_x = e_a | \mathcal{R}) \triangleq \frac{f_a(r(\mathcal{R}_G, \mathcal{R}_a, x))}{\sum_{a'} f_{a'}(r(\mathcal{R}_G, \mathcal{R}_{a'}, x))} \quad (6)$$

Substituting Equation (6) into Equation (4) changes the MAP problem to

$$\mathcal{R}' \leftarrow \arg \max_{\mathcal{R}} \sum_x \sum_a [\mathcal{W}_x(a) \cdot (\log f_a[r(\mathcal{R}_G, \mathcal{R}_a, x)] - \log \sum_{a'} f_{a'}[r(\mathcal{R}_G, \mathcal{R}_{a'}, x)])] + \log P(\mathcal{R}) = \arg \max_{\mathcal{R}} Q(\mathcal{R}) \quad (7)$$

where the objective function $Q(\cdot)$ of Equation (7) is defined as

$$Q(\mathcal{R}) \triangleq \sum_x \sum_a \mathcal{W}_x(a) \cdot \log(f_a[r(\mathcal{R}_G, \mathcal{R}_a, x)]) - \log(\sum_{a'} f_{a'}[r(\mathcal{R}_G, \mathcal{R}_{a'}, x)]) + \log P(\mathcal{R}) \quad (8)$$

$Q(\cdot)$ measures the disagreement between between $f_a(\cdot)$ and $\mathcal{W}_a(\cdot)$ (see Figure 1.) One can also show that $Q(\cdot)$ relates to the Kullback-Leibler divergence. The objective function is therefore maximized with respect to \mathcal{R} if $f_a(r(\mathcal{R}_G, \mathcal{R}_a, x))$ is made as close

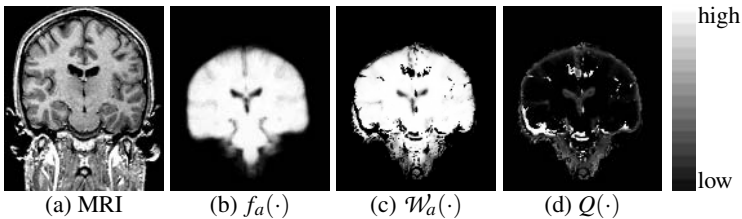


Fig. 1. Image (a) shows an MRI, which is segmented by our EM implementation using the misaligned spatial atlas of the brain $f_a(\cdot)$ in (b). \mathcal{W} in (c) is the result of E-Step and $Q(\cdot)$ in (d) is the initial cost function of the M-Step. $Q(\cdot)$ measures the disagreement between $f_a(\cdot)$ and $\mathcal{W}_a(\cdot)$ with black indicating agreement and white showing disagreement between the two instances.

as possible to $\mathcal{W}_x(a)$. For example, if voxel x is clearly assigned to structure a' with $y_a \triangleq r(\mathcal{R}_G, \mathcal{R}_a, x)$ being the coordinate of voxel x in the atlas space of structure a then $\mathcal{W}_x(a') = 1$ and $f_{a'}(y_{a'}) = \sum_a f_a(y_a)$. The value of the sum across all structures of Equation (8) at x is zero as

$$\sum_a [\mathcal{W}_x(a) \log(f_a(y_a)) - \log(f_{a'}(y_{a'}))] = \log(f_{a'}(y_{a'})) - \log(f_{a'}(y_{a'})) = 0.$$

In summary, we developed a hierarchical registration framework guided by global and structure specific deformations. We transformed Equation (4) into Equation (7), whose objective function $Q(\cdot)$ measures the disagreement between the current alignment of the atlas space and the segmentation weights.

3 Affine Registration Implementation

This section describes an implementation of the approach presented in Section 2.2. We will give an example of an interpolation function $r(\cdot, \cdot, \cdot)$, the corresponding registration parameters \mathcal{R} , a Probability Density Function (PDF) $P(\mathcal{R})$, and a maximization algorithm to solve the MAP problem defined in Equation (7).

The interpolation function $r(\cdot, \cdot, \cdot)$ of Section 2 can model various mapping approaches. For simplicity, we choose an affine interpolation so that $\mathcal{R}_z = (\vec{t}_z, \vec{r}_z, \vec{s}_z)^t$ with $z \in \{G, 1, \dots, N\}$ define displacement \vec{t}_z , rotation \vec{r}_z , and scaling \vec{s}_z . The mapping is defined as $r(\cdot, \cdot, \cdot) : \mathbb{R}^{12 \times 12 \times 3} \rightarrow \mathbb{R}, (\mathcal{R}_G, \mathcal{R}_a, x) \rightarrow A_{\mathcal{R}_G} \cdot A_{\mathcal{R}_a} \cdot (x^t, 1)^t$ where $A_{\mathcal{R}_z}$ is an affine transformation matrix based on the parameter setting \mathcal{R}_z .

We do not assume correspondence between the atlas and the image space so that we choose a uniform prior for the global registration parameter \mathcal{R}_G . As opposed to \mathcal{R}_G , the structure specific parameters $\mathcal{R}_S \triangleq (\mathcal{R}_1, \dots, \mathcal{R}_N)$ describe the residual of structure specific deformations that are not well explained by \mathcal{R}_G . In general, our model should penalize large deviations of \mathcal{R}_S from the expected mean, which is approximated by the average structure-specific registration parameters of the training data. We enforce this penalty by modeling the PDF of \mathcal{R}_S as a Gaussian distribution $N(\mu_{\mathcal{R}_S}, \Upsilon_{\mathcal{R}_S})$ with structure independent mean $\mu_{\mathcal{R}_S}$ and variance $\Upsilon_{\mathcal{R}_S}$ based on the mapping parameters of the training data. We choose a Gaussian distribution as small variance $\Upsilon_{\mathcal{R}_S}$ discourages large deformations from the mean $\mu_{\mathcal{R}_S}$. In addition, Gaussian distributions simplify the calculations in the M-Step [11].

Based on the previous modeling assumptions the objective function is defined as

$$Q(\mathcal{R}) \triangleq \sum_x \left(\sum_a (\mathcal{W}_x(a) \cdot \log [f_a(A_{\mathcal{R}_G} \cdot A_{\mathcal{R}_a} \cdot (x^t, 1)^t)]) \right. \\ \left. - \log [\sum_a f_a(A_{\mathcal{R}_G} \cdot A_{\mathcal{R}_a} \cdot (x^t, 1)^t)] \right) - \frac{1}{2} (\mathcal{R}_S - \mu_{\mathcal{R}_S})^t \Upsilon_{\mathcal{R}_S}^{-1} (\mathcal{R}_S - \mu_{\mathcal{R}_S}).$$

Determining a closed form solution to the MAP problem defined by $Q(\cdot)$ is difficult. Instead, we estimate the solutions through the Powell's method [14]. We also decouple the search for \mathcal{R}_G and \mathcal{R}_S as their dependencies can cause instability. The pseudo code for this implementation is given below.

Algorithm 1: SEGMENTATION AND REGISTRATION()**repeat****E-Step:** Update soft assignment of anatomical structures

$$\mathcal{W}'_x(a) \leftarrow \frac{1}{Z} P(I_x | \mathcal{T}_x = e_a, \mathcal{B}'_x) \cdot f_a(r(\mathcal{R}'_G, \mathcal{R}'_a, x))$$

M-Step: Update parameter space $\mathcal{B}' \leftarrow$ Estimation of the image inhomogeneities based on \mathcal{W}' $\mathcal{R}'_G \leftarrow$ Result of Powell's method with $Q((\cdot, \mathcal{R}'_S))$ $\mathcal{R}'_S \leftarrow$ Result of Powell's method with $Q((\mathcal{R}'_G, \cdot))$ **until** \mathcal{B}' and \mathcal{R}' converge**define** labelmap: $\mathcal{T}_x \leftarrow \arg \max_a \mathcal{W}'_x(a)$

4 Comparative Experiment on 22 Test Cases

The section compares three EM methods which differ in the mapping of the atlas to the patient. The first approach (EM-NonRigid) aligns the atlas using an intensity based non-rigid registration approach and then runs our EM implementation without registration parameters [1]. The second approach (EM-Affine) is similar to EM-NonRigid but uses the affine mapping method by [15] as the preprocessing step. The third approach (EM-Integrated) is our new algorithm solving registration and segmentation simultaneously. All three methods use the same atlas which was generated according to [1]. In order to compare the three methods, each of them segments 22 test cases into the three brain tissue classes and further parcellates grey matter into the subcortical structures thalamus and caudate. We then measure the agreement of the automatic segmentations of the subcortical structures to manual ones, which we view as ground-truth.

This experiment focuses on the thalamus and caudate as they are challenging structures for registration and segmentation. Purely intensity based methods, such as EM without spatial priors, cannot segment these structures because part of the boundary is invisible on MRI (Figure 3). Consequently, EM relies heavily on spatial priors making it sensitive towards misaligned priors. The registration of the priors is also a challenge. Intensity based alignment methods, such as [16], have difficulties mapping the priors to the thalamus because of the structure's similar intensity properties to the adjacent white matter; affine registration methods, such as [15], are too constrained to properly capture the bending within the horn-shaped caudate. In conclusion, a detailed analysis of the segmentation of thalamus and caudate highlights the differences between the three methods.

To measure the quality of the automatic generated results, we compare them to the manual segmentations of the thalamus and caudate using the Dice volume overlap measure. The graph in Figure 3 shows the mean and standard error of the Dice measure for the three algorithms. For the thalamus, EM-NonRigid (NRigid) performed worst because the intensity based registration method is too unreliable for structures with smooth boundaries. The method often overestimates white matter and underestimated the thalamus in this region (Figure 3(EM-NonRigid)). EM-Affine (Affine) performs much better than EM-NonRigid but the method is sensitive towards initial misalign-

ments. For example, EM-Affine cannot properly address the unusual pose of the patient in Figure 2 (Sagittal) causing a bias in the segmentation (Figure 3 (EM-Affine)).

For the caudate, EM-NonRigid performs much better than EM-affine. In contrast to the thalamus, the caudate has a different intensity profile than white matter. Only the relatively small portion of the boundary neighboring the putaman, another subcortical structure, is invisible on MRI. Thus, the intensity based registration method of EM-NonRigid correctly registers the spatial priors to this region. However, the affine registration method of EM-Affine does not have enough degrees of freedom to capture the patient specific bending of the horn-shaped caudate. This causes EM-Affine to misclassify voxels especially at the tip of the structure, which explains the lower score.

EM-Integrated (Integr.) generally outperformed the other two methods (Figure 3). For the thalamus, the method was significantly more reliable than the other two methods (mean Dice and standard error of EM-Integrated : $89.2 \pm 0.4\%$, EM-Affine: $86.9 \pm 1.2\%$, EM-NonRigid: $81.4 \pm 1.3\%$). For the caudate, EM-Integrated ($86.3 \pm 0.6\%$) was still significantly better than EM-Affine ($83.2 \pm 1.7\%$), but only slightly more robust than EM-NonRigid ($85.8 \pm 1.1\%$). However, the standard error of EM-Integrated is lower than EM-NonRigid indicating a higher predictive power of our new approach.

Our new approach performed much better than the two other method on cases, where the deformation between atlas and image space was complex. As the example of Figure 2 illustrates the accuracy of registration and segmentation greatly depend on each other. Initially, the algorithm only correctly outlines corticospinal fluid, whose disposition between the atlas and image space guides the registration (see also Figure 1(d)). As the method progresses, the overall accuracy of the registration as well as segmentation increases. In this example it took 30 iterations until the algorithm converged and correctly segmented the subcortical structures whose boundary is outlined in black.

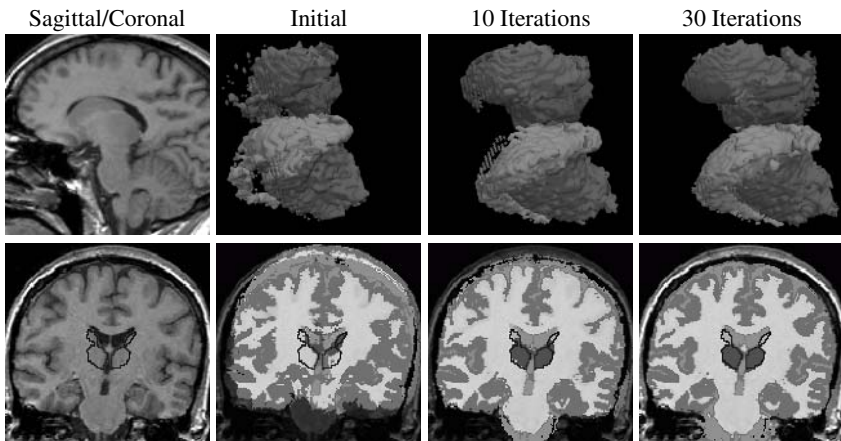


Fig. 2. The top row shows an MRI with unusual head position and the corresponding 3D models of our segmentations of the thalamus and caudate. The bottom row shows the coronal MRI with black indicating the two manually outlined structures. The segmentations to the left are produced by our method, which initial performs poorly but converges closely to the experts' results.

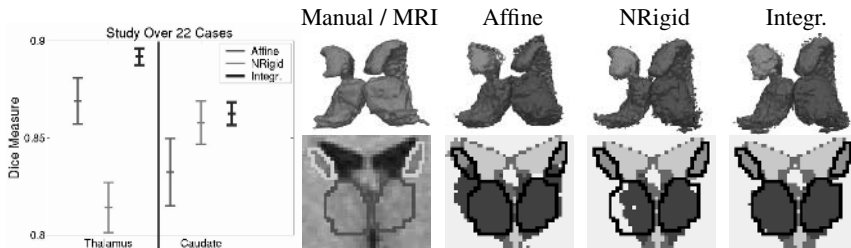


Fig. 3. The graph shows the mean and standard error of the three methods segmenting the thalamus and caudate in 22 cases. For both structures, our new approach EM-Integrated (Integr.) outperformed EM-Affine (Affine) and EM-NonRigid (NRigid). The segmentations to the right highlight deficiencies of EM-Affine and EM-NonRigid, such as the general underestimation of the thalamus by EM-NonRigid and the misalignment of the caudate by EM-Affine due to the low degree of freedom. In the MRI as well as 2D segmentations the thalamus is outlined in black.

We have demonstrated that our method performs better than EM-Affine and EM-NonRigid as a consequence that our approach directly maps the spatial priors of the structures to the segmentation model. In contrast, EM-Affine and EM-NonRigid align an MRI in the atlas space to the image of the patient, using the resulting deformation map to align the spatial priors. This inherently increases the risk of systematic biases in the model. Another explanation for the increased accuracy of our approach is the explicit modeling of dependency between segmentation and registration, which constrains the space of possible solutions and thus simplifies the segmentation problem.

5 Conclusion

We have presented a statistical framework combining inhomogeneity estimation, atlas registration, and segmentation of MRI. Unlike other voxel-based classification methods, our framework models these three problems as a single MAP estimation problem. We implemented the framework as an instance of an EM algorithm using a hierarchical affine mapping approach for anatomical structures. Our approach was validated by automatically segmenting 22 sets of MRIs into the major brain tissue classes and the sub-cortical structures thalamus and caudate that are structures with indistinct boundaries. Using manual segmentations, we then compared our results to other EM implementations which sequentially register and segment. In general, our method performed much better than the other segmentation methods. The improvement is due primarily to the seamless integration of registration into the performance estimation problem.

Acknowledgments. This investigation was supported by the Department of Veterans Affairs (VA Merit Awards and a REAP Award), and NIH grants K02 MH-01110, R01 MH-50747, R01-NS051826-01, P41 RR-13218, U24 RR021382, and U54-EB-005149. We would also like to thank Corey Kemper, Sylvain Bouix, and Polina Golland for their helpful comments.

References

1. K. Pohl, S. Bouix, R. Kikinis, and W. Grimson, "Anatomical guided segmentation with non-stationary tissue class distributions in an expectation-maximization framework," in *ISBI*, pp. 81–84, 2004.
2. B. Fischl, D. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, and A. Dale, "Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain," *Neuron*, vol. 33, 2002.
3. M. Leventon, W. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *CVPR*, pp. 1316 – 1323, 2000.
4. A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, W. Grimson, and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *TMI*, vol. 22, no. 2, pp. 137 – 154, 2003.
5. P. P. Wyatt and J. A. Noble, "MAP MRF joint segmentation and registration," in *MICCAI*, pp. 580–587, 2002.
6. C. Xiaohua, M. Brady, and D. Rueckert, "Simultaneous segmentation and registration for medical image," in *MICCAI*, pp. 663–670, 2004.
7. C. Cocosco, A. Zijdenbos, and A. Evans, "A fully automatic and robust brain MRI tissue classification method," *MIA*, vol. 7, pp. 513–527, 2003.
8. K. Van Leemput, F. Maes, D. Vanermeulen, and P. Suetens, "Automated model-based bias field correction of MR images of the brain," *TMI*, vol. 18, no. 10, pp. 885–895, 1999.
9. J. Marroquin, B. Vemuri, S. Botello, F. Calderon, and A. Fernandez-Bouzas, "An accurate and efficient bayesian method for automatic segmentation of brain MRI," *TMI*, vol. 21, 2002.
10. K. Pohl, J. Fisher, W. Grimson, and W. Wells, "An expectation maximization approach for integrated registration, segmentation, and intensity correction," AIM-2005-010, CSAIL - MIT.
11. G. McLachlan, T. Krishnan, *The EM Algorithm and Extensions*. John Wiley & Sons, 1997.
12. W. Wells, W. Grimson, R. Kikinis, and F. Jolesz, "Adaptive segmentation of MRI data," *TMI*, vol. 15, pp. 429–442, 1996.
13. S. Srivastava, F. Maes, D. Vandermeulen, W. V. Paesschen, P. Dupont, and P. Suetens, "Effects of anatomical asymmetry in spatial priors on model-based segmentation of the brain MRI: A validation study," in *MICCAI*, no. 3216 in LNCS, pp. 327–334, Springer, 2004.
14. W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes in C : The Art of Scientific Computing*. Cambridge University Press, 2 ed., 1992.
15. S. Warfield, J. Rexilius, P. Huppi, T. Inder, E. Miller, W. Wells, G. Zientara, F. Jolesz, and R. Kikinis, "A binary entropy measure to assess nonrigid registration algorithm," in *MICCAI*, pp. 266–274, 2001.
16. A. Guimond, A. Roche, N. Ayache, and J. Meunier, "Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections," *TMI*, vol. 20, 2001.

Automatic 3D Segmentation of Intravascular Ultrasound Images Using Region and Contour Information

Marie-Hélène Roy Cardinal¹, Jean Meunier², Gilles Soulez³, Roch L. Maurice¹,
Éric Thérasse³, and Guy Cloutier¹

¹ Laboratory of Biorheology and Medical Ultrasonics,
University of Montreal Hospital, Canada

roycarmh@iro.umontreal.ca, guy.cloutier@umontreal.ca

² Département d'Informatique et de Recherche Opérationnelle,
University of Montreal, Canada

³ Radiology Department, University of Montreal Hospital, Canada

Abstract. Intravascular ultrasound (IVUS) produces images of arteries that show the lumen in addition to the layered structure of the vessel wall. A new automatic 3D IVUS fast-marching segmentation model is presented. The method is based on a combination of region and contour information, namely the gray level probability density functions (PDFs) of the vessel structures and the image gradient. Accurate results were obtained on in-vivo and simulated data with average point to point distances between detected vessel wall boundaries and validation contours below 0.105 mm. Moreover, Hausdorff distances (that represent the worst point to point variations) resulted in values below 0.344 mm, which demonstrate the potential of combining region and contour information in a fast-marching scheme for 3D automatic IVUS image processing.

1 Introduction

Intravascular ultrasound (IVUS) is a medical imaging technique that produces cross-sectional images of the vascular wall as a catheter is pulled-back inside blood vessels. The ability of IVUS to characterize the vascular wall was initially proven in 1989 [1]. Since then, it has become very useful for studying atherosclerotic disease because quantitative assessment of the wall, information about the nature of atherosclerotic lesions as well as plaque shape and size are provided. IVUS is expected to play an important role in atherosclerosis research; for example, to achieve precise evaluation of the disease in progression-regression therapies [2]. Experts agree that IVUS imaging adds precious complementary information to angiography which only shows a projection of the lumen [3].

A typical intravascular ultrasound acquisition contains several hundred of images making non-automatic analysis of the data long, fastidious and subject to high intra- and inter- observer variabilities. These could be serious constraints against the clinical usage of IVUS. There is also some quality issues related to IVUS imaging due to speckle noise, imaging artifacts and shadowing of parts of

the vessel wall by calcifications. It is thus necessary to develop specific segmentation methods that take into account the nature of IVUS images.

So far, a number of segmentation techniques have been developed for IVUS data analysis. A great portion of this work was based on local properties of image pixels, namely gradient-based active surfaces [4] or multi-agents [5] and pixel intensity combined to gradient active contours [6]. Graph search was also investigated using different edge operators [7, 8, 9]. Another portion of the IVUS segmentation work was based on more global or region information. Texture-based morphological processing was considered [10]. Gray level variances were then used for the optimization of a maximum a posteriori (MAP) estimator modeling ultrasound speckle and contour geometry [11]. However, some of these methods are semi-automatic and others focus on either region or contour information without combining them.

The aim of this work was thus to achieve automatic IVUS image analysis by combining region and contour features in a segmentation model. Since pixel gray values are distributed according to Rayleigh probability density functions (PDFs) in B-mode imaging of uniform scattering tissues [12], PDF features can be of value for IVUS segmentation. In a previous semi-automatic 2D segmentation model [13], gradient- and PDFs-based methods were compared and both provided results with similar variability (PDFs parameters were calculated from manually traced regions). The combination of the region statistics and of the gray level gradient contour information, that are complementary, is thus hypothetically more suitable for IVUS image analysis, especially when the vascular wall edges are weakly defined. These features were included in a multiple interface fast-marching segmentation [14], that enables simultaneous segmentation of different parts of the vessel wall. In this paper, automatic estimation of the PDF parameters and initial contours is also presented.

2 Methods

2.1 Fast-Marching Segmentation

The fast-marching method was derived from the level-set model; they both allow the following of an interface propagating under a speed function F [15]. Fast-marching can be applied to image segmentation by interpreting an image contour as the propagating interface final position [16]. To stop the evolution of the interface on image boundaries, the speed function, that is defined in terms of image or shape features, should become close to zero when the propagating front meets with object boundaries.

In the fast-marching formulation, where an interface propagates under a unidirectional speed function, the evolution of the contour is expressed in terms of the arrival time $T(x)$ of the front at point $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$. The T function satisfies Eq. 1, stating that the arrival time difference between two adjacent pixels increases as the velocity of the contour decreases.

$$|\nabla T| F = 1 \quad . \quad (1)$$

The propagation of the interface is done in 3D via the construction of the arrival time function map (T map). The construction algorithm [17] selects the interface point having the smallest arrival time and calculates the arrival times of its neighbors. Neighbor arrival times are updated by solving a discrete 3D approximation of Eq. 1. This is repeated until the interface has propagated across the whole image or until the interface front is considered stationary (when the time gradient is sufficiently high).

In IVUS imaging, multiple contours (lumen and media) must be identified to characterize plaque development, consequently segmentation was done via a multiple interface extension of the fast-marching algorithm [14]. For this particular case, a boundary is defined by the meeting position of two contours propagating in opposite directions. A speed function must then be determined for each propagating interface; it is defined according to the PDF of the anatomical structure in which the interface is evolving. The T map is built by selecting the point with the smallest arrival time value from all interfaces¹. For the current study, the multiple interfaces evolved according to the speed function given by Eq. 2, which is based on the gray values occurring probability and gradient.

$$F_m(i, j, k) = \frac{1}{N_\nu} \sum_{s \in \nu} p_m(I_s) + 1/(1 + g(I_s)). \quad (2)$$

$$g = |\nabla G_\sigma * I| \quad (3)$$

where I_s is the gray level value of pixel s in image I ; ν is the set of the N_ν neighbors of the pixel positioned at (i, j, k) (the neighbors window size is $7 \times 7 \times 3$); $p_m(I_s)$ is the occurring probability of pixel I_s in the region m ; $g(I_s)$ is the value of the gradient g at pixel I_s and G_σ is a gaussian smoothing filter of standard deviation $\sigma = 2$. According to the first term of Eq. 2, the velocity of interface m will take higher values when it is inside a region having a grayscale distribution close to p_m ; it will decrease when approaching the boundary since the neighbors are distributed under other component PDFs.

2.2 Pre-processing

The pre-processing calculations involved the catheter artifact detection and removal; and the feature extraction (PDF mixture parameters and image gradient).

The pixels near the catheter with gray level values that correlated through the whole IVUS pullback were labeled as being part of the ring-down artifact. The catheter artifact was removed from the images by subtracting the average ring-down gray level values from each IVUS 2D frame. The catheter artifact had to be detected before each segmentation because it is specific to each pullback.

The speed function of Eq. 2 uses the gray level distributions $p_m(I_s)$ of the different vessel wall structures. The distribution of a whole IVUS pullback was thus estimated by a mixture of Rayleigh PDFs, each corresponding to the lumen, plaque structure of the intima, media, and adventitia plus surrounding tissues. The mixture PDF $p_{X|\Theta}$ with parameters $\Theta = \{(\omega_m, a_m^2)\}_{m=1}^M$ is:

¹ A detailed description of the T map construction algorithm can be found in [14].

$$p_{X|\Theta}(x | \Theta) = \sum_{m=1}^M \omega_m p_m(x | a_m^2). \quad (4)$$

$$p_m(x | a_m^2) = \frac{x}{a_m^2} \exp\left(-\frac{x^2}{2a_m^2}\right) \quad (5)$$

where X represents the pixels' gray level taking values in $[0, \dots, 255]$ (with $x > 0$); M is the number of different tissue structures; ω_m is the proportion of the m^{th} component of the mixture; and $p_m(x | a_m^2)$ is the m^{th} tissue Rayleigh probability density function with parameter a_m^2 , which is related to the variance $\sigma_m^2 = a_m^2(4 - \pi)/2$ of the distribution.

The mixture parameters Θ were estimated with the Expectation-Maximization algorithm (EM) that is an iterative computation technique of maximum likelihood estimates for incomplete data [18].

2.3 Automatic Initialization

Initial contours must be computed to initialize the fast-marching segmentation. These contours were found on one every three frames to diminish computational load without losing their 3D continuity.

An initial lumen contour was first automatically detected from the IVUS data. The log-likelihood of the lumen L_{lum} was calculated for each pixel according to Eq. 6 and by using the previously estimated lumen distribution of the PDF mixture.

$$L_{lum}(i, j, k) = \frac{1}{N_\nu} \sum_{s \in \nu} \log p_{lum}(I_s) . \quad (6)$$

where $p_{lum}(I_s)$ is the estimated occurring probability of pixel I_s in the lumen region.

The initial lumen contour was set as a smooth closed curve, not necessarily circular to preserve the irregular lumen shape, containing the catheter and the pixels maximizing L_{lum} . The initial media contour calculation was performed similarly to the lumen initialization. An elliptical contour surrounding the previously found lumen was grown radially to maximize the media log-likelihood L_{med} also calculated according to Eq. 6 but with the estimated media distribution $p_{med}(I_s)$. An elliptical contour was searched because the media boundary is generally smooth and regular (it appears as a hypoechoic ring, see Fig. 2).

These automatically computed initial contours were only rough estimates of the wall boundaries; they were used to initialize the fast-marching segmentation that allowed obtaining more precise contours. The initial contour points were spline-interpolated axially and longitudinally while being forbidden to overlap to preserve the layered structure of the vessel wall. The fast-marching requires the initial front to be inside the object to segment, the contours were thus each shifted radially (10 pixels) in the 2 opposite directions to compute the 2 pairs of 3D tube-like initial interfaces. The interfaces propagating in opposite directions (and including the exact object border) were thus separated by 20 pixels.

2.4 In-Vivo Data and Numerical Simulation

An in-vivo IVUS pullback of 600 frames from a diseased superficial femoral artery was acquired with a Volcano Therapeutics equipment (In-vision gold, California, USA) using a 20 MHz transducer. Images of size 10 x 10 mm were digitized on 384 x 384 pixel matrices and stored using the DICOM standard. The acquisition was done at a 10 images/sec frame rate and the catheter pullback velocity was set to 1 mm/sec generating 0.1 mm thick 2D slices.

In addition to the in-vivo data, realistic simulations of IVUS data were conducted. Since the exact geometry of simulated data is known, direct calculation of the detected boundary performance could be obtained. A detailed description of the image-formation model used to simulate IVUS echograms can be found in [19]. To generate a realistic vessel geometry, the vessel boundaries (lumen, plaque of the intima, media) were created from manually traced contours on an in-vivo IVUS series. The simulated pullback contained 86 IVUS 2D frames.

The automatic 3D segmentation method was applied to the in-vivo and simulated IVUS images. For validation purpose, the in-vivo IVUS segmentation results were compared with manually traced contours on 1 every 10 frames; boundary position from 60 IVUS frames were thus available. Results on simulated data were compared with the exact geometry. The average and Hausdorff point to point contour distances [20], and wall layer area differences were calculated. Hausdorff distance represents the worst case, it is the maximum distance between between corresponding contour points.

3 Results

As part of the pre-processing, the EM algorithm was applied to the IVUS series to evaluate the PDF mixture parameters. An example of an automatically detected Rayleigh PDF mixture and corresponding IVUS pullback histogram is shown in Fig. 1.

Segmentation results for the automatic 3D fast-marching method combining PDFs and gradient are shown in Fig. 2. The lumen and media detected boundaries are presented for typical cross-sectional IVUS images of the whole 3D in-vivo and simulated pullbacks. Table 1 shows the average and Hausdorff

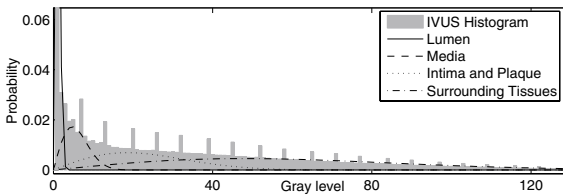


Fig. 1. Illustration of an automatically detected Rayleigh PDF mixture and corresponding IVUS gray level histogram

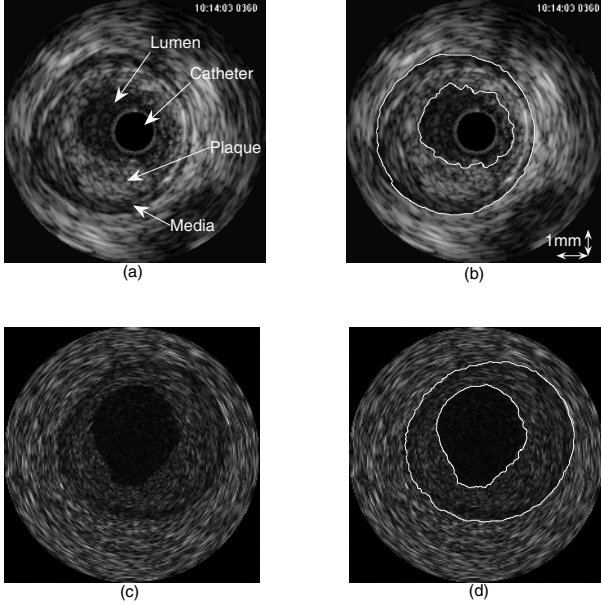


Fig. 2. Segmentation examples on in-vivo and simulated data. Typical IVUS cross-sectional image (a). Lumen and media detected contours (b). Simulated IVUS cross-sectional image (c). Lumen and media detected contours (d).

Table 1. Average distances, Hausdorff distances and area differences between detected contours on in-vivo and simulated data and the validation contours

In-vivo Data		
	Lumen	Media
Average Distances (mm)	0.083 ± 0.111	0.105 ± 0.094
Hausdorff Distances (mm)	0.344 ± 0.237	0.300 ± 0.104
Area Differences (mm^2)	-0.24 ± 1.07 ($r=0.974$)	-0.47 ± 1.14 ($r=0.890$)
Simulated Data		
Average Distances (mm)	0.059 ± 0.069	0.066 ± 0.061
Hausdorff Distances (mm)	0.235 ± 0.105	0.224 ± 0.064
Area Differences (mm^2)	0.18 ± 0.65 ($r=0.997$)	-0.65 ± 0.73 ($r=0.976$)

The pixel size is $26 \times 26 \mu\text{m}^2$.

distances and the area differences between the detected contours and the validation boundaries for the in-vivo and simulated IVUS pullbacks.

4 Discussion

The goal of this work was to demonstrate the feasibility of a fully automatic 3D fast-marching method for the segmentation of IVUS images when using PDFs and gradient in the same model.

No ground truth value was available to evaluate the precision of the calculated mixture parameters, however a rough comparison, in Fig. 1, of the pull-back gray level histogram with the estimated Rayleigh PDF mixture showed good agreement. A qualitative analysis of the automatic fast-marching segmentation combining region and contour information in Fig. 2 revealed detected contours that were very close to all vessel layers. Fig. 2 also showed that non-circular lumen contours could be precisely found. For the simulated images, Table 1 showed that the detected contours were very close to the simulated vessel boundaries with average point to point distances below 0.066 mm (2.5 pixels). Similar distances were obtained for the in-vivo IVUS data with average distances between 0.083 and 0.105 mm (3 to 4 pixels). Hausdorff distances, which represent the highest distance between the contour and the validation boundary position, resulted in values below 0.344 mm (13 pixels), this demonstrates the efficiency of the segmentation method. Area differences between detected and validation contours showed in Table 1 were slightly higher than the ones reported in [5, 8] (between -0.14 ± 1.01 and 0.27 ± 0.49 mm²), but the data was acquired on a femoral artery that is larger than the coronary arteries of these studies.

These preliminary segmentation results showed that this new IVUS segmentation approach is promising for the processing of large pullbacks without user interaction. In the near future, the segmentation method will be applied to several IVUS series acquired from different patients before and after undergoing balloon angioplasty to characterize the lumen and wall volumic changes over a whole vessel segment. The complete validation of the segmentation method will also include several experts' manually traced contours repeated in time to assess the performance in terms of intra- and inter-user variabilities.

Acknowledgments. This work was supported by grants from the Canadian Institute of Health Research (G.C., G.S., #MOP-53244) and Valorisation-Recherche Québec (group grant #2200-094), and by a National Scientist award (G.C.), a clinical research scholarship award (G.S.) and a studentship award (M.H.R.C.) from the Fonds de la Recherche en Santé du Québec. We also acknowledge the work of Zhao Qin.

References

1. Gussenhoven, E.J., Essed, C.E., Lancée, C.T., Mastik, F., Frietman, P., van Egmond, F.C., Reiber, J., Bosch, H., van Urk, H., Roelandt, J., Bom, N.: Arterial wall characteristics determined by intravascular ultrasound imaging: An in vitro study. *J. Am. Coll. Cardiol.* **14** (1989) 947–952
2. Nissen, S.E.: Application of intravascular ultrasound to characterize coronary artery disease and assess the progression or regression of atherosclerosis. *Am. J. Cardiol.* **89** (2002) 24B–31B
3. Nissen, S.E., Yock, P.: Intravascular ultrasound: Novel pathophysiological insights and current clinical applications. *Circulation* **103** (2001) 604–616

4. Klingensmith, J.D., Shekhar, R., Vince, D.G.: Evaluation of three-dimensional segmentation algorithms for the identification of luminal and medial-adventitial borders in intravascular ultrasound images. *IEEE Trans. Med. Imag.* **19** (2000) 996–1011
5. Bovenkamp, E.G.P., Dijkstra, J., Bosch, J.G., Reiber, J.H.C.: Multi-agent segmentation of IVUS images. *Pattern Recogn.* **37** (2004) 647–663
6. Kovalski, G., Beyar, R., Shofti, R., Azhari, H.: Three-dimensional automatic quantitative analysis of intravascular ultrasound images. *Ultrasound Med. Biol.* **26** (2000) 527–537
7. Zhang, X., McKay, C.R., Sonka, M.: Tissue characterization in intravascular ultrasound images. *IEEE Trans. Med. Imag.* **17** (1998) 889–899
8. Koning, G., Dijkstra, J., von Birgelen, C., Tuinenburg, J.C., Brunette, J., Tardif, J.C., Oemrawsingh, P.W., Sieling, C., Melsa, S., Reiber, J.H.C.: Advanced contour detection for three-dimensional intracoronary ultrasound: A validation - in vitro and in vivo. *Int. J. Cardiac Imag.* **18** (2002) 235–248
9. von Birgelen, C., Mario, C.D., Li, W., Schuurbiens, J.C.H., Slager, C.J., de Feyter, P.J., Serruys, P.W., Roelandt, J.R.T.C.: Morphometric analysis in three-dimensional intracoronary ultrasound : An in vitro and in vivo study using a novel system for the contour detection of lumen and plaque. *Am. Heart J.* **132** (1996) 516–527
10. Mojsilovic, A., Popovic, M., Amodaj, N., Babic, R., Ostojic, M.: Automatic segmentation of intravascular ultrasound images: A texture-based approach. *Ann. Biomed. Eng.* **25** (1997) 1059–1071
11. Haas, C., Ermert, H., Holt, S., Grewe, P., Machraoui, A., Barmeyer, J.: Segmentation of 3D intravascular ultrasonic images based on a random field model. *Ultrasound Med. Biol.* **26** (2000) 297–306
12. Wagner, R.F., Smith, S.W., Sandrik, J.M., Lopez, H.: Statistics of speckle in ultrasound B-scans. *IEEE Trans. Son. Ultrason.* **30** (1983) 156–163
13. Roy Cardinal, M.H., Meunier, J., Soulez, G., Thérasse, E., Cloutier, G.: Intravascular ultrasound image segmentation : A fast-marching method. In: *Lecture Notes in Computer Sciences. Proceedings of MICCAI 2003. Volume 2879.* (2003) 432–439
14. Sifakis, E., Garcia, C., Tziritas, G.: Bayesian level sets for image segmentation. *J. Vis. Commun. Image R.* **13** (2002) 44–64
15. Osher, S., Sethian, J.A.: Fronts propagating with curvature-dependent speed: Algorithms based on hamilton–jacobi formulations. *J. Comput. Phys.* **79** (1988) 12–49
16. Malladi, R., Sethian, J.A., Vemuri, B.C.: Shape modeling with front propagation: A level set approach. *IEEE Trans. Pattern Anal. Machine Intell.* **17** (1995) 158–175
17. Sethian, J.A.: A fast marching level set method for monotonically advancing fronts. *Proc. Nat. Acad. Sci.* **93** (1996) 1591–1595
18. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc. B* **39** (1977) 1–38
19. Maurice, R.L., Ohayon, J., Finet, G., Cloutier, G.: Adapting the Lagrangian speckle model estimator for endovascular elastography: Theory and validation with simulated radio-frequency data. *J. Acoust. Soc. Am.* **116** (2004) 1276–1286
20. Chalana, V., Kim, Y.: A methodology for evaluation of boundary detection algorithms on medical images. *IEEE Trans. Med. Imag.* **16** (1997) 642–652

Automatic Segmentation of the Articular Cartilage in Knee MRI Using a Hierarchical Multi-class Classification Scheme

Jenny Folkesson¹, Erik Dam^{1,2}, Ole Fogh Olsen¹,
Paola Pettersen², and Claus Christiansen²

¹ Image Analysis Group, IT University of Copenhagen, Denmark
jenny@itu.dk

² Center for Clinical and Basic Research, Ballerup, Denmark

Abstract. Osteoarthritis is characterized by the degeneration of the articular cartilage in joints. We have developed a fully automatic method for segmenting the articular cartilage in knee MR scans based on supervised learning. A binary approximate kNN classifier first roughly separates cartilage from background voxels, then a three-class classifier assigns one of three classes to each voxel that is classified as cartilage by the binary classifier. The resulting sensitivity and specificity are 90.0% and 99.8% respectively for the medial cartilage compartments. We show that an accurate automatic cartilage segmentation is achievable using a low-field MR scanner.

1 Introduction

Osteoarthritis (OA) is one of the major health concerns among the elderly today [1]. The main effects of OA is the degradation of the articular cartilage together with remodeling and overgrowth of bone, a process causing loss of mobility of the joints. It typically affects large weight bearing joints as hips and knees. Currently, the treatment of OA is restricted to symptom control, because as yet there are no disease-modifying drugs [2].

MRI allows for quantitative evaluation of the cartilage [3],[4], and cartilage deterioration can be detected using this technique [5]. MRI also has the advantage of being a non-invasive technique.

When assessing the cartilage, the MR scans can be manually segmented slice-by-slice by experts, but for clinical studies manual methods are too time consuming and are also prone to inter- and intra-observer variability. When automating the cartilage segmentation, the main challenges are the thin structure of the cartilage and the low contrast between the cartilage and surrounding soft tissues. The progression of OA is very often slow and it can take many years before the cartilage is reduced from its typical thickness of a few millimeters to possible total loss. It is therefore important to have high accuracy and precision of the cartilage assessment technique in order to detect statistically significant changes. This enables the correlation of the method with the effects of drugs, and the evaluation of their benefit to the joint in reducing the signs of the disease.

Several groups have developed automated methods for cartilage segmentation. 2D methods has limited continuation between slices and since they have to be converted into a 3D segmentation when finding for example thickness maps, it is advantageous to perform segmentation in 3D directly. Among the 3D techniques that have been developed, Grau et al. [6] use a semi-automatic segmentation method that is based on a watershed approach. The method is evaluated on 7 scans from 4 subjects and has an average sensitivity and specificity of 90.0% and 99.9% respectively. Pakin et al. [7] have developed an automatic segmentation method based on region growing followed by two-class clustering. It is evaluated on one scan with resulting sensitivity and specificity of 66.2% and 99.6%. The semi-automatic segmentation method of Warfield et al. [8], [9] iterates between a classification step and a template registration step, and has a lower variability compared to repeated manual segmentations on the scan it was evaluated on. Naish et al. [10] use a data set that consists of a longitudinal study of OA patients and local intensity changes over time is used as a measure of cartilage degradation. However, the cartilage is manually or semi-automatically segmented.

All of the methods mentioned (except for the one of Naish et al. but they have not focused on the segmentation part) have only been evaluated on a handful of scans, and the only fully automatic segmentation produces low sensitivity and specificity values compared to the semi-automatic methods.

In this paper, we present a method for segmenting the tibial and femoral medial cartilage in 3D MR scans of knees. The segmentation is based on an three class approximate kNN classification scheme and is improved by selecting the largest connected component from the result of the classification. The segmentation method works directly in 3D, not in 2D slices, and is fully automatic. This is an improvement of previous work [11] which was a method for locating tibial medial cartilage for the initialization of a shape model, a method based on a two class kNN classifier without any feature selection incorporated.

Our segmentation algorithm aids the automatization of cartilage assessment and is intended for clinical studies on a low-field MR scanner. Though the image quality of the scanner we are using is slightly lower compared to the conventional high-field scanners, we propose to examine if accurate automatic cartilage segmentation is achievable also on a low-field scanner. If such a scanner can be used in clinical studies it would reduce the costs significantly. It has been shown that low-field dedicated extremity MRI can provide similar information on bone erosions and synovitis as expensive high-field MRI units [12] comparing manual segmentations, but there has to our knowledge not been published any work on automatic segmentation of cartilage on low-field MRI. From the automatic segmentation, relevant quantitative measures such as the cartilage volume and thickness can be calculated either globally or locally in a point or a small area. In the latter case comparison between patients or temporal studies of the same patient will require establishing geometric or anatomical correspondence either by expert annotations or by automated modeling of landmarks. An automated approach for this is planned for future work and will not be part of this paper.

2 Methods

2.1 Image Acquisition

An Esaote C-Span low-field 0.18 T scanner dedicated to imaging of extremities acquires Turbo 3D T1 scans (40° flip angle, T_R 50 ms, T_E 16 ms). The scans are made through the sagittal plane with a voxel size in the range $0.7031 \times 0.7031 \times (0.7031/0.7813/0.8594) \text{ mm}^3$. Among the total of 71 scans, 50 have the resolution $0.7031 \times 0.7031 \times 0.7813 \text{ mm}^3$, 3 the resolution $0.7031 \times 0.7031 \times 0.7031 \text{ mm}^3$ and the remaining 18 scans have the resolution $0.7031 \times 0.7031 \times 0.8594 \text{ mm}^3$. The scans all have the size $256 \times 256 \times 104$ voxels, but we only use the central $170 \times 170 \times 104$ voxels because only they contain information.

The scans have been manually segmented on a slice-by-slice basis by a radiologist. A scan slice with the tibial and femoral medial cartilage manually segmented is shown in Figure 1.

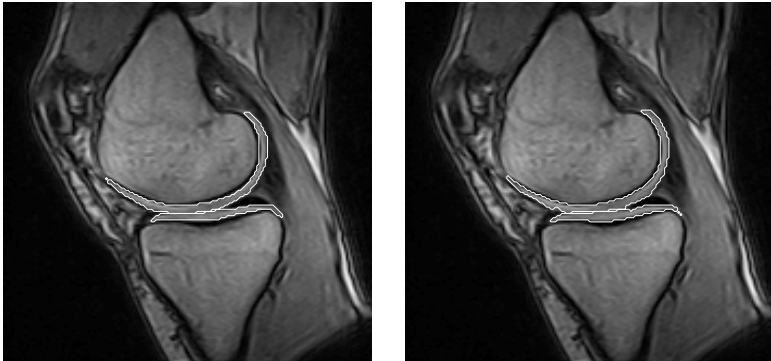


Fig. 1. To the left, a slice from a knee MR scan where the tibial medial and femoral medial cartilage is segmented manually by radiologists. The size of this slice is 170×170 pixels. To the right is the result from our automatic segmentation for the corresponding slice. The sensitivity and specificity for this scan are 92.52% and 99.82% respectively, with a dice similarity coefficient of 0.83.

The 71 scans in the data set are of both left and right knees. In order to treat all scans analogously, the right knees are reflected about the center in the sagittal plane. The test subjects are both males and females aged between 21 and 72 years. They have no or mild OA symptoms, diagnosed by radiologists as being between 0 and 3 on the Kellgren and Lawrence Index [13].

2.2 Cartilage Classification

For the segmentation of cartilage we use an approximate k NN classifier, which is implemented in an Approximate Nearest Neighbor (ANN) framework developed by Mount and colleagues [14]. The ANN classifier is in principle the same as a

k NN classifier, but with the modification that you can allow for a small amount of error in the search for nearest neighbors which may improve the run time significantly. An error bound, ϵ , is introduced, so instead of returning the k nearest neighbors from a data set, the ANN search algorithm returns k points such that the ratio of the distance between the i th reported point ($1 \leq i \leq k$) and the true i th nearest neighbor is at most $1 + \epsilon$. We have found empirically that examining the 100 nearest neighbors yields a good balance between computational complexity and accuracy, and we set $\epsilon = 2$, a value that only marginally lowers the accuracy while reducing computational time significantly.

In this work we examine the medial cartilage since OA is more often observed in this compartment [15] and in particular in the medial tibial part [16], thus these compartments are of major interest when it comes to finding disease markers for OA. In order to separate different types of cartilage from one another we use a three class classifier, where the classes are tibial medial cartilage, femoral medial cartilage and background.

The classification is hierarchical, and the first step is a two class classification where the voxels are roughly separated into cartilage or background. The k NN produces class probabilities for every voxel, and in this step we set the threshold at 0.65 yielding a sensitivity for medial cartilage close to 99%. This also results in a large amount of false positives, but since typically only a few percent of the total volume within the knee belongs to the cartilage, this first step is a way of reducing data significantly. In the second step, the voxels classified as cartilage in the first step are reconsidered. This time we use a three class classifier, where the three classes are tibial and femoral medial cartilage and background, and class membership is decided based on a majority vote. The three class classifier contains more features and the features are optimized to separate the three classes whereas the classifier in the first step has features optimized to separate cartilage from background. A sketch of the hierarchical classification scheme is illustrated in Figure 2.

We have also tested a direct partitioning into the three classes, but the hierarchical approach yields better results and is faster, since the first step has less features and thus lower computational complexity. The classifier in the first step has a set of 28 features compared to the three class classifier in the second step that contains 52 features.

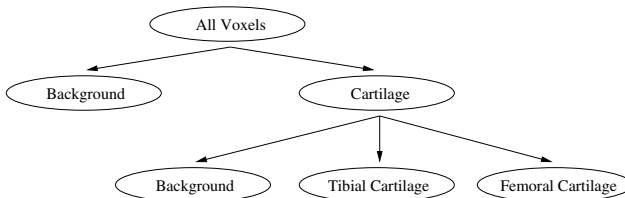


Fig. 2. Tree representation of the hierarchical classification scheme

2.3 Features and Feature Selection

In order to find a feature set that performs well for our classification scheme, we here introduce our set of candidate features and the subsets of the features that were found from our feature selection method [17], which consists of sequential forward selection followed by sequential backward selection.

When a radiologist examines an MR scan for cartilage, she or he takes the location and the intensity in the image into consideration. We therefore consider these as candidate features. Both the raw intensity and the Gaussian smoothed intensities on three different scales ($0.65mm$, $1.1mm$, $2.5mm$) are examined.

One can also consider features that are related to the geometry of the object in question. The 3-jet, which is all first, second and third order derivatives with respect to (x, y, z) forms a basis which can describe all geometric features up to third order [18] and are listed as candidate features. All the derivatives mentioned in this section are Gaussian derivatives and are defined as $I_{i_1, \dots, i_n} = \int \tilde{I}(\tilde{x}) D_{i_1, \dots, i_n} g(\tilde{x}, \sigma_1) d\tilde{x}$, where g is a Gaussian, D a differential operator and σ_1 is the scale. All features are examined on the three scales, selected to cover the range of different cartilage thicknesses, mentioned above.

Cartilage can be described as a thin curved disc in 3D. The Hessian (H), which is the symmetric matrix containing second order derivatives with respect to the coordinates (x, y, z) ,

$$H = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix},$$

is therefore considered. The eigenvectors of the Hessian points in the directions of the principal curvatures and its eigenvalues corresponds to the curvature in those directions. A thin disc such as cartilage will locally yield one large and two small eigenvalues. The eigenvalues as well as the three eigenvectors are candidate features.

A feature that has been shown to be significant in the detection of thin structures such as fingerprints is the structure tensor (ST) [19]. It is a symmetric matrix containing products of the first order derivatives convolved with a Gaussian,

$$ST = G_{\sigma_2} * \begin{pmatrix} I_x I_x & I_x I_y & I_x I_z \\ I_y I_x & I_y I_y & I_y I_z \\ I_z I_x & I_z I_y & I_z I_z \end{pmatrix},$$

where σ is not necessarily the same scale as the one used for obtaining the derivatives. The ST examines the local gradient distribution at each location (x, y, z) . The directions of the eigenvectors depend on the variation in the neighborhood. The eigenvalues and eigenvectors of the ST were considered as potential features with a combination of three scales of σ_1 and three scales of σ_2 .

The third order derivatives with respect to (x, y, z) can be conveniently represented in the third order tensor I_{ijk} . Examining the third order structure in the local gradient direction (I_x, I_y, I_z) can be described using Einstein summation as

$$L_{www} = I_{ijk} I_i I_j I_k / (I_i I_i)^{3/2}.$$

The third order tensor is examined in the gradient direction on three different scales, which were considered as possible features.

2.4 Selected Features

The features used in the two class classifier are the position in the image, the Gaussian smoothed intensities on three different scales ($0.65mm$, $1.1mm$, $2.5mm$) and the raw intensities, the first order Gaussian derivatives on scales $0.65mm$ and $2.5mm$, the eigenvalues and the eigenvector corresponding to the largest eigenvalue of the structure tensor with $\sigma_1 = 0.65mm$ and $\sigma_2 = 2.5mm$, and the eigenvalues of the Hessian on scales $1.1mm$ and $2.5mm$.

The features in the three class classifier consist of combinations of first, second and third order Gaussian derivatives on the three different scales mentioned, the Gaussian smoothed intensities on three different scales ($0.65mm$, $1.1mm$, $2.5mm$) and the raw intensities, the position, the eigenvector corresponding to the largest eigenvalue of the ST with $\sigma_1 = 0.65mm$ and $\sigma_2 = 1.1mm$, the eigenvalues of the ST with $\sigma_1 = 1.1mm$ and $\sigma_2 = 2.5mm$, the eigenvalues of the Hessian on scales $1.1mm$ and $2.5mm$.

The features selected as most significant are the Hessian and the structure tensor along with the intensity and the position in the image. The features were normalized between zero and one. Normalization for unit variance was also examined, but the normalization of values between zero and one produces slightly better results.

3 Results

From our data set of 71 scans we use 25 for training and 46 for the evaluation of our algorithm. The results of our automatic segmentation is compared to the manual segmentation made by radiologists, resulting in an average sensitivity and specificity of 90.0% ($\pm 2.6\%$ st.d.) and 99.8% ($\pm 0.06\%$ st.d.) respectively



Fig. 3. On the left is the manually segmented medial cartilage from a knee MR scan. To the right is the corresponding automatic segmentation. For this scan, the sensitivity and specificity are 94.82% and 99.79% respectively, with a dice of 0.81.

for the test set for the medial cartilage compartments. A comparison between a golden standard segmentation and an automatically segmented knee MR scan can be seen in Figure 3. A slice by slice comparison is displayed in Figure 1. The dice similarity coefficient (DSC) measures spatial volume overlap between two segmentations, A and B , and is defined as $DSC(A, B) = \frac{2 \times |A \cap B|}{|A| + |B|}$. The Dice similarity coefficient between our automatic segmentation and the golden standard segmentation is for the test set on average 0.80 (± 0.03 st.d.).

4 Discussion

The average sensitivity and specificity of our method compared to the results of methods with similar evaluation though on different data is presented in Table 1. Comparing our method with the fully automatic segmentation algorithm (Pakin et al. [7]), we get a distinctly higher sensitivity and a slightly better specificity. Though slightly worse, the sensitivity and specificity of our method are comparable to those of Grau et al. [6]. They have a higher volume overlap ($DSC=0.90$) however their method is semi-automatic. We have evaluated our segmentation algorithm on more scans than the other two methods. Some of the semi-automated

Table 1

	Our method	Method of Pakin [7]	Method of Grau [6]
Sensitivity	90.01%	66.22%	90.03%
Specificity	99.80%	99.56%	99.87%
Data set	25+46	1	7
Interaction time	0	0	5-10 min

segmentation techniques described in section 1 have been evaluated in terms of inter- and intra-observer variability of the method compared to manual segmentation. As for our method, future work will involve inter-scan variability, and we will also examine intra-user variability for the manual segmentations.

Our segmentation algorithm performs well compared to two leading cartilage segmentation schemes, which leads us to the conclusion that accurate, fully automatic cartilage segmentation is achievable in low-field MR scanners.

References

1. J.Wyngaarden, H.Smith, L., Bennett, J.: Cecil Textbook of Medicine. 19 edn. Volume 2. W. B. Saunders (1992)
2. Creamer, P., Hochberg, M.C.: Osteoarthritis. Lancet **350** (1997) 503–509
3. Muensterer, O., Eckstein, F., Hahn, D., Putz, R.: Computer-aided three dimensional assessment of knee-joint cartilage with magnetic resonance imaging. Clinical Biomechanics **11** (1996) 260–266
4. Graichen, H., Eisenhart-Rothe, R.V., Vogl, T., Englmeier, K.H., Eckstein, F.: Quantitative assessment of cartilage status in osteoarthritis by quantitative magnetic resonance imaging. Arthritis and Rheumatism **50** (2004) 811–816

5. Loeuille, D., Olivier, P., Mainard, D., Gillet, P., Netter, P., Blum, A.: Magnetic resonance imaging of normal and osteoarthritic cartilage. *Arthritis and Rheumatism* **41** (1998) 963–975
6. Grau, V., Mewes, A., Alcaiz, M., Kikinis, R., Warfield, S.: Improved watershed transform for medical image segmentation using prior information. *IEEE Transactions on Medical Imaging* **23** (2004)
7. Pakin, S.K., Tamez-Pena, J.G., Totterman, S., J.Parker, K.: Segmentation, surface extraction and thickness computation of articular cartilage. Volume 4684., *SPIE* (2002) 155–166
8. Warfield, S.K., Kaus, M., Jolesz, F.A., Kikinis, R.: Adaptive, template moderated, spatially varying statistical classification. *Medical Image Analysis* (2000) 43–55
9. Warfield, S.K., Winalski, C., Jolesz, F.A., Kikinis, R.: Automatic segmentation of mri of the knee. SPL Technical Report 91, ISMRM Sixth Scientific Meeting and Exhibition, Sydney, Australia (1998)
10. Naish, J.H., Vincent, G., Bowes, M., Kothari, M., White, D., Waterton, J.C., J.Taylor, C.: A method to monitor local changes in mr signal intensity in articular cartilage: A potential marker for cartilage degeneration in osteoarthritis. Volume 3217., *MICCAI* (2004) 959–966
11. Folkesson, J., Dam, E., Pettersen, P., Olsen, O.F., Nielsen, M., Christiansen, C.: Locating articular cartilage in mr images. Volume 5747., *SPIE* (2005) 1484–1490
12. Ejbjerg, B., Narvestad, E., adn H.S. Thomsen, S.J., Ostergaard, M.: Optimised, low cost, low field dedicated extremity mri is highly specific and sensitive for synovitis and bone erosions in rheumatoid arthritis wrist and finger joints: a comparison with conventional high-field mri and radiography. *Annals of the Rheumatic Diseases* **13** (2005)
13. Kellgren, J., Lawrence, J.: Radiological assessment of osteo-arthritis. *Annals of the Rheumatic Diseases* **16** (1957)
14. Arya, S., Mount, D., Netanyahu, N., Silverman, R., Wu, A.: An optimal algorithm for approximate nearest neighbor searching in fixed dimensions. Number 5, *ACM-SIAM. Discrete Algorithms* (1994) 573–582
15. Dunn, T., Lu, Y., Jin, H., Ries, M., Majumdar, S.: T2 relaxation time of cartilage at mr imaging: comparison with severity of knee osteoarthritis. *Radiology* **232** (2004) 592–598
16. Kamibayashi, L., Wyss, U., Cooke, T., Zee, B.: Changes in mean trabecular orientation in the medial condyle of the proximal tibia in osteoarthritis. *Calcif Tissue Int.* **57** (1995) 69–73
17. Folkesson, J., Dam, E., Olsen, O.F., Pettersen, P., Christiansen, C.: A supervised learning approach to segmenting articular tibial cartilage in knee mri. Technical report, IT University of Copenhagen, Rued Langgaards Vej 7, 2300 Copenhagen S, Denmark (2005)
18. Florack, L.: The Syntactical Structure of Scalar Images. PhD thesis, University of Utrecht (1993)
19. Weickert, J.: *Anisotropic Diffusion in Image Processing*. B. G. Teubner (1998)

Automatic Segmentation of the Left Ventricle in 3D SPECT Data by Registration with a Dynamic Anatomic Model

Lars Dornheim¹, Klaus D. Tönnies¹, and Kat Dixon²

¹ Institut für Simulation und Graphik, Fakultät für Informatik,
Otto-von-Guericke-Universität Magdeburg, Germany

² Medical Imaging Research Group, University of British Columbia, Canada

Abstract. We present a fully automatic 3D segmentation method for the left ventricle (LV) in human myocardial perfusion SPECT data. This model-based approach consists of 3 phases: 1. finding the LV in the dataset, 2. extracting its approximate shape and 3. segmenting its exact contour.

Finding of the LV is done by flexible pattern matching, whereas segmentation is achieved by registering an anatomical model to the functional data. This model is a new kind of stable 3D mass spring model using direction-weighted 3D contour sensors.

Our approach is much faster than manual segmentation, which is standard in this application up to now. By testing it on 41 LV SPECT datasets of mostly pathological data, we could show, that it is very robust and its results are comparable with those made by human experts.

1 Introduction

Heart diseases, especially heart infarcts, are one of the most common causes of death in the western world. Modern functional medical imaging techniques can contribute significantly to diagnosis and particularly to the quantitative assessment of these diseases.

Myocardial perfusion SPECT provides functional, three-dimensional images of the human left ventricle (LV) (figure 1a), where infarct-affected regions can be seen as gaps within the well-perfused LV (figure 1b). For the qualitative and quantitative assessment of the damage of the LV, anatomical knowledge is needed (figure 1c).

LV segmentation is often performed manually, which is very time-consuming. Automatic registration of an anatomical model with functional data either requires extensive training based on segmented data, or a flexible model supplementing information, if perfusion and anatomy do not coincide.

2 Related Work

A lot of literature exists on the segmentation of the LV, but only little of it deals with SPECT data. SPECT data is noisy and the level of detail is low. Anatomy may be extracted from it only if an external model is supplied.

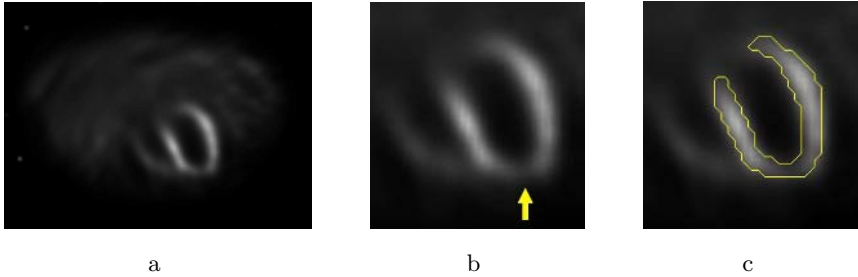


Fig. 1. (a) thorax 3D SPECT data slice. (b) enlarged LV section with gap due to infarct (arrow). (c) manual segmentation.

[1] gives an overview on previous approaches to segment the LV in SPECT data. They conclude, that many approaches (threshold-based, region-based, etc.) lead to bad results due the lack of model knowledge. Existing model-based techniques either need user interaction, or they use unflexible and simple models, so that no adequate segmentation of the LV is possible. No method exists, to cope with perfusion gaps, which are typical for infarct patients.

The approach in [2] tries to solve this problem by using superquadrics, followed by free form deformation. These implicit dynamic 3D shape models have already been used in [3] for tracking the segmented LV in 4D-SPECT data. In [2], a very simple model is used for an initial segmentation. Subsequent fine-tuning via free-form deformation does not use model knowledge, but searches the data slicewise for points supporting the interpolation. A bounding box of the LV must be provided by the user for initial placement of the implicit mode.

The most promising of these methods all use some kind of model knowledge. However, up to now, no appropriate automatic approach exists for the segmentation of the LV in 3D-SPECT data.

3 Requirements

Our approach to automatic segmentation of the LV in 3D-SPECT data registers a dynamic model with the data. We use *mass spring models*, known from the linear elasticity theory ([4]), which are a special kind of deformable models.

Unlike active contours and snakes ([5]) these models are shape models in that they model a specific shape intrinsically. They are governed by *internal forces* (from springs) and *external forces* (from sensors at the mass points). They also differ from mass spring models used in surgery simulation in that they do not try to simulate a physically correct world. Our models use forces only in order to find good segmentations.

Furthermore, the mass points do not represent exact locations with special semantic meaning like in ASMs [6]. They are (beside their function in the model dynamics) placeholders for sensors, searching for certain image features.

These mass spring models have several advantages. They can easily be applied to 3D data. No training is necessary (in contrast to stochastic models). Local

adjustment is easy, making interaction possible in a intuitive way. But also two major disadvantages exist:

- Mass spring models are already very instable in a 2D environment, which is even more so in 3D.
- A new mass spring model needs to be generated for every specific segmentation task. Its topology and shape must be adequate.

To overcome these difficulties, we have developed a new kind of stable 3D mass spring model that stabilizes itself by the use of an additional internal force, the *torsion force* ([7]), even if it is only sparsely interconnected. Stability in traditional models stems from a dense spring meshing that makes them inflexible, expensive to compute and difficult to control.

Furthermore, we developed a technique to generate complex 3D mass spring models automatically from an arbitrary sample segmentation ([8]). Models are composed of substructures representing the features of the sample segmentation; the level of abstraction can be controlled.

4 Segmentation Process

We now describe the automatic segmentation process. Its strategy is first searching the object in the dataset, then segmenting its region roughly and finally tracing its exact contour. So, the segmentation process consists of 3 phases:

1. *Initial placement of the model*: This step is crucial for the success of the following phases, since dynamic models can only adapt locally.
2. *Shape adaptation of the model*: Here, shape, size and direction of the model are fitted to the data.
3. *Contour adaptation of the model*: Now the model will be adapted to the contour of the object to segment. The shape fitting of phase 2 as a starting point will give this process a stable context.

4.1 Initial Placement

In the literature, the first and important phase of initially placing the model is usually performed interactively or discussion is omitted completely. However, shape models provide excellent possibilities to automate this step.

We generate a voxel template P from the intrinsic shape of our LV model, which we employ for a flexible pattern matching on the 3D-SPECT dataset D to find the approximate position of the LV in the data as a good initial position for our model. This approach is possible, as the size and direction of the LV never change dramatically in our datasets. Besides, there is always only one LV-shaped object in the dataset, so that no ambiguities need to be resolved. Since our model is very robust with respect to initial placement errors, a rough pattern matching should provide sufficient information.

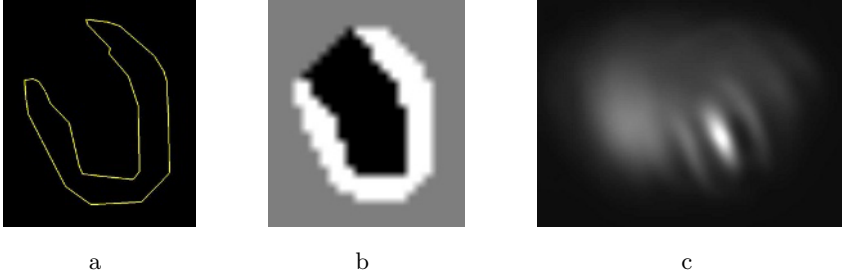


Fig. 2. (a) contour cross section of the relaxed model. (b) according slice of the generated voxel template. (c) correlation image of the slice: bright means high correlation.

The template P is constructed from the relaxed model with no external forces acting. The contour substructure of the model is treated as a polygonal mesh, which is converted to a 3D voxel template (figure 2b) by 3D polygon filling.

For the pattern matching, we use two alternative techniques. One is based on *correlation*, which can be computed very fast by FFT. The other technique is based on the *pattern vector angle*. It is reasonably fast, but much more robust to brightness artefacts in the data.

Correlation. The SPECT dataset has only non-negative values, which represent the counts in a voxel. We apply the generated LV template P to the data. P is a 3D voxel matrix with value 1 at the bright LV wall, value -1 in the dark LV interior and value 0 at unspecified outer regions (figure 2b). So, both the strong signal LV wall values and the weak LV interior values contribute to high correlation values (see figure 2c).

The similarity measure e_{CORR} we use is proportional to a 3D discrete correlation. Correlation is susceptible to intensity artefacts (figure 3a). In such cases, the second technique described in the following section is a better option.

Pattern vector angle (PVA). A fairly fast pattern matching technique that does not rely on intensity levels can be realized by the use of the *pattern vectors*. These are comprised of all the voxel values from a region in a fixed order. If one measures the similarity of two equally-sized patterns via the angle between them, it does not depend on the intensity level of the data but on the shape ([9]). The shape is represented by the direction of the pattern vectors, whereas their lengths depend on this intensity level. Equation 1 shows the calculation of this angle at position p , i.e. the similarity measure e_{PVA} .

$$e_{\text{PVA}}(p) = \cos \angle(\mathbf{D}(p), \mathbf{P}(p)) = \frac{\langle \mathbf{D}(p), \mathbf{P}(p) \rangle}{\|\mathbf{D}(p)\| \cdot \|\mathbf{P}(p)\|} \quad (1)$$

The method is less susceptible to artefactual intensity peaks (compare figures 3a and 3b) at the expense of higher computation times.

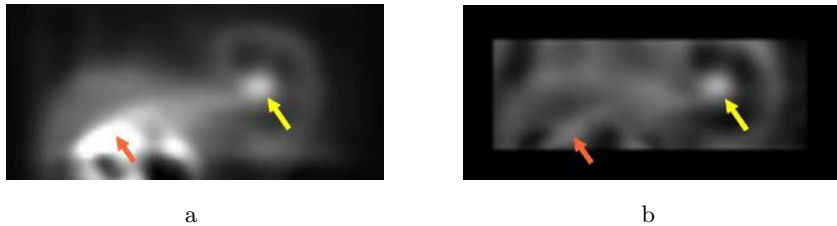


Fig. 3. (a) correlation image and (b) PVA image in Y direction. Bright regions depict high similarity values. An artefact in the dataset (red, left arrow) is only in (a) calculated with a higher similarity value than the correct position (yellow, right arrow), which has in (b) the highest similarity value.

4.2 Shape Adaptation

After initial placement, the model starts adapting to the dataset. The primary goal in this step is a rough adaptation of the model shape.

This is achieved by running the model with the external forces from its *intensity sensors* only. These sensors drive the mass points towards certain intensity levels in the data. They are evenly spread over the model and intensity information is present all over the dataset. The model therefore has good data support to adapt its rough shape quickly to it. Gaps in the LV are not critical in this phase, because they are surrounded by enough intensity information on which the model can rely. *Contour sensors*, which act on mass points towards contours, are not used in this phase. This way, a fast, flexible and robust model adaptation is achieved without disturbing the model by too many details too early.

The model movement is simulated until the model dynamics reach an equilibrium. The criterion for this is that every mass point stays within a small neighborhood around its position for a fixed number of simulation steps (equation 2, where \mathbf{m}_{i_t} is the position vector of the i th mass point at time step t).

$$\forall i \quad \forall 0 \leq j < n - 1 : \|\mathbf{m}_{i_{t-j}} - \mathbf{m}_{i_{t-(n-1)}}\| \leq \varepsilon \quad \text{for } t \geq n \quad (2)$$

4.3 Contour Adaptation

The goal of this last step is the exact model adaptation to the data. This phase is similar to the previous phase, but now contour sensors are also active. Due to the previous step they should find the correct contour very close to their starting point. This leads to a fast and robust segmentation in this final phase.

In the case of gaps in the bright LV wall, where no such contour can be found, the still active intensity sensors in this region will stabilize the model. This way, these gaps are interpolated by the model.

Because of the high complexity in 3D, there are normally many potential contours close to a sensor. We therefore developed a directionally sensitive sensor, which takes the gradient direction of the 3D contour into account ([10]).

5 Evaluation

For our tests, we developed a prototype, which was running on a standard PC (1,3 GHz Pentium M, 512 RAM) with a modern 3D graphics card (NVidia GeForce-Go5200 64MB). The concept of dynamic cached sensor cards (see [10] for that) for each sensor in the model allows for a very fast model dynamics simulation, so that the model motion can be visualized in real time. This way it is much easier to assess the model behavior on the data.

We tested our method on 41 3D-SPECT datasets of human LV's from 25 different patients (128 x 128 voxel slices, 25 to 38 slices per dataset, cubic voxels with 4.795 mm edge length). For evaluation, we used a subset of 7 datasets, which were mostly (5 of 7) affected by infarcts (up to 25 %). We used manual segmentations from a medical professional as gold standard.

For a quantitative assessment, we used several measures (relative segmentation error, Hausdorff distance and average surface distance) to calculate the similarity between the gold standard and the binary voxel segmentation.

The model was generated automatically from one of the 7 datasets. It consisted of 153 mass points, 630 springs and 124 contour faces. The average spring degree at a mass point was 8.24. This is not a very complex 3D model, but it is already too complex to be generated by hand.

We weighted external model energy in the simulation 10 times stronger than the internal energy and used a damping of 2 %. For a detailed review of all these mostly non critical parameters see [10].

5.1 Segmentation Process Evaluation

Table 1 shows the results of the two techniques for initial placement discussed in section 4.1. Correlation was fast (below 1 sec) and mostly correct. It failed only in one case, when the dataset had a bright artefact. PVA had no problems with this artefact, as expected, but was noticeably slower (35 sec). Thus, it was

Table 1. correctness of the initial model placement by means of correlation respectively pattern vector angle (PVA)

Similarity / Dataset	1	2	3	4	5	6	7
Correlation	correct	correct	correct	wrong	correct	correct	correct
PVA	correct	correct	correct	correct	correct	correct	correct

Table 2. effect of phase count on the average automatic segmentation quality on the datasets 1 to 3

Phase Count	False Seg.	Hausdorff.	Aver. Dist.
2-phases approach	27.3 %	1.79 vox	0.410 vox
3-phases approach	26.4 %	1.71 vox	0.398 vox

only selected in datasets with bright artefacts being recognizable by abnormal high count values.

Several of the parameters for the automatic segmentation process were determined on a subset of the 7 datasets (dataset 1 to 3), so that their values could be verified on the remaining 4 datasets. These tests prove the 3-phases approach to be reasonable, since without the second phase for the approximate model adaptation (section 4.2), the segmentation quality decreases (table 2).

5.2 Segmentation Results

Table 3 shows the results of the automatic segmentation of the 7 datasets. They are compared with the results of the manual segmentations by two experienced users. It can easily be seen, that the automatic technique is much faster than the manual process. Its quality is slightly better. In all cases, it lies within in the inter-observer variance. This fact is illustrated by figure 4, which gives a visual impression of the segmentation quality.

Table 3. average segmentation results of the automatic process and the manual segmentations by two experienced users (A and B)

Segmentation	Duration	False. Seg.	Hausdorff.	Aver. Dist.
Auto. Seg.	25.3 sec	29.4 %	2.21 vox	0.436 vox
Seg. by A	15.6 min	36.5 %	2.81 vox	0.667 vox
Seg. by B	21.4 min	30.6 %	2.39 vox	0.534 vox

Qualitative inspection of the results on the remaining datasets, where no gold standard existed, confirmed these results: Phase 1 always detected a correct initial model position and the next phases adapted the model visually correct to the LV (similar to figure 4).

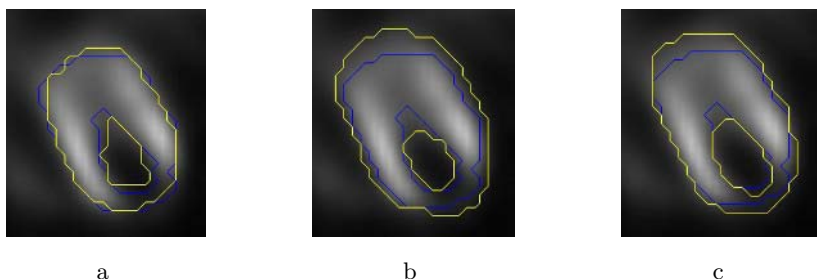


Fig. 4. results (yellow, bright) of automatic and manual segmentation in comparison to the gold standard (blue, dark). (a) automatic segmentation process. (b) manual segmentation of user A. (c) manual segmentation of user B.

6 Discussion

Our automatic segmentation method found and segmented the LV in all datasets correctly. It needs only a small fraction of the time required to achieve an equivalent result by hand. The 3-phases approach and the direction-weighted contour sensors have proven effective for the successful segmentations. Furthermore, the good results of phase 1 make the interactive initial model placement obsolete.

Our stable 3D mass spring model and the technique to generate it automatically from a sample segmentation were crucial for the good results. It enabled bridging gaps caused by reduced perfusion while being able to adapt to the individual shape of each LV.

The segmentation method was fast and robust on noisy, low-resolution data. In the next step, we will investigate its performance on more complex data.

References

1. Fernandez-Maloigne, C., Rakotobe, R.H., Langevin, F., Fauchet, M.: 3D segmentation and visualization of cardiac SPECT studies. In: 28th AIPR Workshop. Volume 3905 of Proceedings of SPIE. (2000) 222–231
2. Pohle, R., Wegner, M., Rink, K., Tönnies, K., Celler, A., Blinder, S.: Segmentation of the left ventricle in 4D-dSPECT data using free form deformation of superquadrics. In: Medical Imaging: Image Processing. Volume 5370 of Proceedings of SPIE. (2004) 1388–1394
3. Bardin, E., Cohen, L.D., Ayache, N.: Tracking and motion analysis of the left ventricle with deformable superquadrics. *Medical Image Analysis* **1** (1996) 129–149
4. Gould, P.L.: *Introduction to Linear Elasticity*. Springer (1994)
5. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. *IJCV* **1** (1988) 321–331
6. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models - their training and application. *CVIU* **61** (1995) 38–59
7. Dornheim, L., Tönnies, K.D., Dornheim, J.: Stable dynamic 3D shape models. In: *ICIP*. (2005)
8. Dornheim, L., Tönnies, K.D.: Automatische Generierung dynamischer 3D-Modelle zur Segmentierung des linken Ventrikels in 3D-SPECT-Daten. In: *Bildverarbeitung für die Medizin*. (2005)
9. Kohonen, T.: *Self-Organization and Associative Memory*. Springer (1987)
10. Dornheim, L.: *Generierung und Dynamik physikalisch basierter 3D-Modelle zur Segmentierung des linken Ventrikels in SPECT-Daten*. Diplomarbeit, Fakultät für Informatik, Otto-von-Guericke-Universität Magdeburg (2005)

Intravascular Ultrasound-Based Imaging of Vasa Vasorum for the Detection of Vulnerable Atherosclerotic Plaque

Sean M. O'Malley¹, Manolis Vavuranakis²,
Morteza Naghavi³, and Ioannis A. Kakadiaris¹

¹ Visual Computing Lab, Dept. of Computer Science, Univ. of Houston

² Dept. of Cardiology, Univ. of Athens

³ Association for Eradication of Heart Attack

Abstract. Vulnerable plaques are dangerous atherosclerotic lesions that bear a high risk of complications that can lead to heart attacks and strokes. These plaques are known to be chronically inflamed. The vasa vasorum (VV) are microvessels that nourish vessel walls. Proliferation of VV is part of the “response to injury” phenomenon in the process of plaque formation. Recent evidence has shown strong correlations between neovessel formation and macrophage infiltration in atherosclerotic plaque, suggesting VV density as a surrogate marker of plaque inflammation and vulnerability. We have developed a novel method for imaging and analyzing the density and perfusion of VV in human coronary atherosclerotic plaques using intravascular ultrasound (IVUS). Images are taken during the injection of a microbubble contrast agent and the spatiotemporal changes of the IVUS signal are monitored using enhancement-detection techniques. We present analyses of *in vivo* human coronary cases that, for the first time, demonstrate the feasibility of IVUS imaging of VV.

1 Introduction

Vulnerable plaques are subsets of atherosclerotic lesions that rupture and create blood clots resulting in acute coronary syndrome and sudden cardiac death. Plaque inflammation plays a central role in its vulnerability to future complications (e.g., rupture, hemorrhage, distal emboli, and acute stenosis). The search for an intracoronary technology capable of imaging both plaque morphology and activity (inflammation) is currently a very active topic in the cardiology community [1]. Despite major advances in the development of other intravascular imaging techniques, intravascular ultrasound (IVUS) remains the most widely-available technology to interventional cardiologists. However, the major drawback of IVUS has been its inability to gauge plaque inflammation.

The vasa vasorum (VV) are microvessels that nourish vessel walls (Fig. 1(a)). In conditions with extensive neovessel formations such as atherosclerotic plaques, tumor angiogenesis, and diabetic retinopathy, most are fragile and prone to leak or rupture. In the field of atherosclerosis, recent evidence indicates that

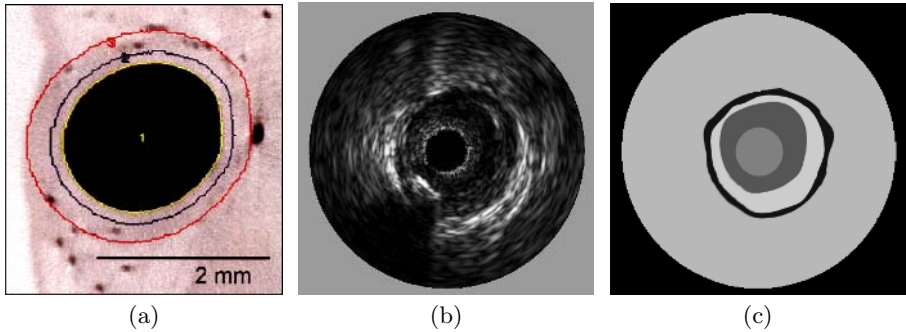


Fig. 1. (a) Vasa vasorum histology (reprinted from [9]) in relation to (b) a typical IVUS frame and (c) its labeled regions (from the center outward: the catheter; the echofree lumen; the plaque area with the intima, its membranous leading edge; the echofree media; and the adventitia)

proliferation of VV is a preceding or concomitant factor associated with plaque inflammation and instability [2, 3].

We have developed a novel method which, for the first time, enables IVUS imaging of atherosclerotic plaque inflammation based on quantification of VV density and perfusion. Our primary contributions are: 1) an IVUS acquisition protocol which utilizes a recently-developed ultrasound contrast agent to induce echogenicity in the VV, and 2) an automated algorithm for the detection, quantification, and visualization of VV in the resulting contrast-enhanced image sequences.

Related work includes plaque characterization [4, 5] and the imaging of myocardial perfusion [6] and angiogenesis [7]. Recent research into IVUS imaging has been primarily aimed at automating the segmentation task for the major IVUS features [8] (Figs. 1(b)-(c)). To the best of our knowledge, there have been no previous reports of automated detection and imaging of VV.

2 Materials and Methods

IVUS System: We use both a solid-state phased-array 20 MHz scanner (Volcano Therapeutics Inc. – InvisionTM) and a rotating single-crystal 40 MHz scanner (Boston Scientific Inc. – GalaxyTM).

Microbubbles: We use OptisonTM, a new ultrasound contrast agent composed of albumin microspheres filled with octafluoropropane gas.

VV Imaging Protocol: 1) Acquire IVUS frames for several minutes to obtain un-enhanced signal (“pre-injection” period); 2) inject contrast agent, temporarily washing out the frames due to luminal saturation (“during-injection” period); and 3) acquire frames for several minutes further to obtain signal indicating potentially enhanced areas due to perfusion into the VV (“post-injection” period).

During all periods, the IVUS catheter is held fixed; 1 min into the final period, 5 cc normal saline is injected to flush out remaining microbubbles.

Imaging Vasa Vasorum Density: Following acquisition, our VV detection, quantification, and visualization method consists of three steps.

- Step 1: Track cardiac motion throughout the sequence to allow compensation for relative catheter/vessel movement (Sect. 2.1).
- Step 2: Perform enhancement detection using difference-imaging and statistical techniques (Sect. 2.2).
- Step 3: Quantify and visualize the resulting enhancement (Sect. 2.3).

2.1 Step 1: Motion Compensation

In stationary-catheter IVUS studies, maintaining a fixed catheter position with regard to an anatomic point of reference is impossible in practice due to the periodic motion of the heart. Relative motion between the catheter and vessel produces image sequences in which each frame deviates from its predecessors; this deviation makes analysis of a specific anatomic region-of-interest (ROI) difficult. Our studies were performed on IVUS sequences which lacked associated electrocardiogram gating data. Consequently, to track cardiac motion, we have developed a method inspired by Zhu *et al.* [10], in which cardiac phase is derived from the IVUS sequence itself. This allows us, in subsequent processing, to select frames according to their position in phase (and, by extension, their physical orientation). Phase extraction is accomplished through ROI selection, intensity metasignal generation, and filtering and signal reconstruction, as follows.

ROI Selection: We select a fixed ROI in the IVUS frame to monitor the changes in intensity in this region over time. We do not analyze data from the entire frame because the lumen/catheter region has little useful signal and the adventitial region has a very low signal-to-noise ratio. Instead, the intensity study is concentrated on the region between the luminal border and the adventitia. The ROI need not be a perfect segmentation of this region, however; we may simply compute an average frame over a particular time range (e.g., one cardiac period) and create a mask which on average contains the entire ROI.¹

Intensity Metasignal Generation: To produce these metadata, we use one of two techniques: average intensity or inter-frame difference. The average intensity g at frame i over our ROI is given by $g_i = \frac{1}{n} \sum_{(x,y) \in \text{ROI}} F_i(x,y)$, where F is an IVUS frame. The inter-frame difference d between frame i and the previous frame $i-1$ is given by $d_i = \frac{1}{n} \sum_{(x,y) \in \text{ROI}} |F_i(x,y) - F_{i-1}(x,y)|$ (in both cases, n is the area in pixels of the ROI). Here we use these techniques interchangeably; typically the only difference in the methods is a phase shift between the resulting signals.

Filtering & Signal Reconstruction: Due to the spurious extrema present in the signal (Fig. 2(a)), it is necessary to apply a filter to isolate those maxima and

¹ As we rely on a fixed ROI here, we do not employ a segmentation algorithm.

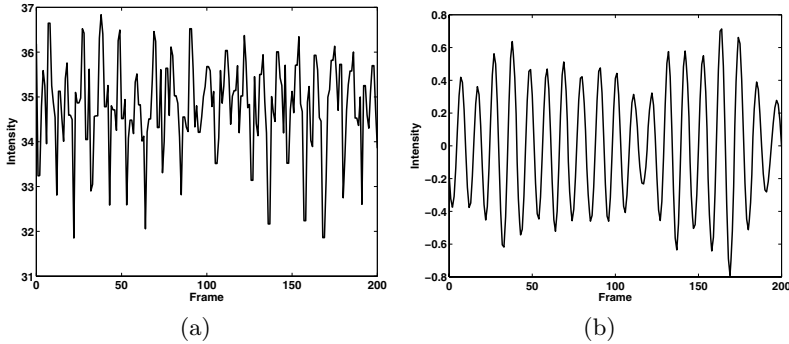


Fig. 2. (a) Plot of g_i for the first 201 frames of Case 1 and (b) the filtered result

minima which represent the motion we are interested in. A Butterworth bandpass filter is applied, $H(\omega) = 1 / \sqrt{1 + [2(\omega - \omega_c) / \Delta\omega]^{2m}}$, where the filter frequency ω_c is centered at the dominant cardiac frequency, the width $\Delta\omega = 0.6(\omega_c)$, and the order $m = 4$. We have found stationary-catheter IVUS studies to be robust with regard to automatic cardiac phase extraction as signals produced with this method exhibit a prominent peak in the frequency domain; this makes algorithmic determination of ω_c much more feasible than in previous pullback studies. While this peak is generally obvious to the human viewer, we have noted two other common sources of low-frequency noise which must be appropriately ignored: these include a prominent DC component related to the mean intensity or inter-frame difference, and other low-frequency components due to the microbubble washout during acquisition, which causes a varying mean grey-level to occur in the ROI over time. Once the signal is collected, transformed to the Fourier domain, ω_c is determined, and the signal is filtered, it is returned to the time domain (Fig. 2(b)). The set of frames associated with, e.g., the peaks in this time-domain signal are considered correlated by our definition.

2.2 Step 2: Enhancement Detection

The goal of enhancement detection is to localize and quantify the subtle intensity changes in the IVUS image sequence due to the presence of perfused VV in the field of view of the IVUS sensor following microbubble injection. For detection to be reliable, it must be invariant both to motion and to the presence of ultrasound (US) speckle noise in the image. The former issue was discussed previously (Sect. 2.1) while, to address the latter, our enhancement detection method combines difference imaging with temporal averaging to attenuate speckle.

Difference imaging, broadly, involves a “before” image (I_b) and an “after” image (I_a). Here, I_b is a baseline frame derived from a set of phase-correlated pre-injection frames which are pixel-wise averaged for noise reduction. That is, if we have a set of k phase-correlated pre-injection frames $\mathbf{F} \equiv \{F_1 \dots F_k\}$, then
$$I_b = \frac{1}{k} \sum_{i=1}^k F_i.$$
 The number of frames averaged to produce this image may be

fairly large as we are interested only in highly temporally-coherent features of the sequence for this baseline. The image I_a is either a single post-injection frame or an average of correlated post-injection frames. Single frames are useful if producing a video sequence, while temporal averaging of the post-injection set to produce I_a will suppress speckle noise along with temporally incoherent enhanced regions (e.g., passing microbubbles).

Once I_b (baseline) and I_a (post-injection) images are available, a difference image is produced by pixel-wise subtraction, $I_d = I_a - I_b$, with negative values thresholded to zero. This raw difference image shows areas of true enhancement in addition to low-valued areas of false enhancement due to US artifacts. Consequently, further thresholding must be applied to reduce the noisy appearance of the image, but this is done conservatively to avoid suppressing relevant enhancement. In this context, we have developed an automatic thresholding technique inspired by classical expectation-maximization algorithms and the work of Bruzzone and Prieto [11]. A grey-level threshold is determined for I_d under the assumption that it is a Bayesian mixture of probabilities $p(X) = p(X/v_u)P(v_u) + p(X/v_e)P(v_e)$ where $p(X)$ is the overall probability density function (PDF) of the difference image, $p(X/v_u)$ and $p(X/v_e)$ are the PDF's of the unenhanced (u) and enhanced (e) pixels respectively, and $P(v_u)$ and $P(v_e)$ are the *a priori* probabilities of these classes (unknown quantities are initially estimated from histogram statistics). Following global thresholding, a Markov modeling technique is applied to account for spatial relationships. This takes the strength of a point of detected enhancement into consideration along with its neighborhood to determine whether the enhancement is salient; this has the effect of reducing spot noise while refining the edges of enhanced regions. Note that for the purposes of estimating probabilities and enhancement levels, only the pixels in a ROI of the frame are taken into consideration: typically the intimo-medial region where we expect to find VV. As per Sect. 2.1, this ROI is fixed and need not perfectly segment the region.

Video sequences may be generated to show the changes in microbubble perfusion over time. In this case, the baseline I_b remains as defined previously, while the comparison images are a sequence of running averages of phase-correlated post-injection frames. That is, if a sequence of k correlated post-injection frames $\mathbf{C} \equiv \{C_1 \dots C_k\}$ is available and a p -frame running average is used (where p is odd), an h -frame video sequence may be produced, where $h = k - p + 1$ and the raw difference image for a particular frame $i \in [1, h]$ is $I_d = (1/p) \sum_{j=1}^p C_{i+j-1} - I_b$. In addition, by binning the frames by cardiac phase into a convenient data structure, it is trivial to extend this method to utilize *all* post-injection frames by associating each frame in a sequence with its $p - 1$ correlated neighbors.

2.3 Step 3: Quantification and Visualization of Enhancement

To quantify enhancement over time, a signal is produced composed of the average enhancement per enhanced pixel (AEPEP) of each frame. This is defined as follows. If $\mathbf{P} \equiv \{P_1, \dots, P_q\}$ is the set of intensities of all pixels labeled “en-

hanced” in the ROI of the difference image after thresholding and refinement, then the AEPEP value for the frame is given by $\varepsilon = \frac{1}{q} \sum_{p \in \mathbf{P}} p$. A useful feature of the AEPEP metric is that it remains relatively constant in spite of changes in the apparent area of the enhanced regions due to US distortion. Metrics utilizing the *entire* ROI were found to be too noisy as measures of enhancement. In addition, we fit an approximating spline to the resulting AEPEP signal to highlight trends for the human observer. Where results are presented as color-coded images, the unit of intensity is a percentage indicating an enhanced pixel’s difference level divided by the maximum grey-level difference (255).

3 Results

We conducted microbubble contrast-enhanced IVUS imaging on seven patients with coronary artery disease using the ACESTM (Analysis of Contrast-Enhanced Sequences) software developed for this project. Due to space limitations, data collected from one typical case is presented in detail and six cases presented in summary.²

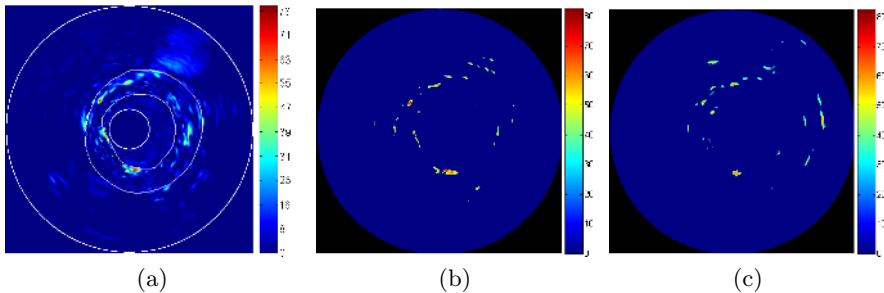


Fig. 3. Case 1: (a) raw difference image, and (b)-(c) thresholded difference images at frames 600 and 800, respectively. Unit of intensity is %, as discussed in Sect. 2.3.

Case 1: This sequence consists of 1,073 frames sampled at 10 frames/s. Contrast agent (CA) first appears in the lumen in frame 404, complete washout due to lumen echo-opacity occurs from frames 452-490, and CA reaches a minimum in the lumen around frame 530. Figure 3(a) illustrates a raw difference image showing marked enhancement in the plaque region, particularly at 6 o’clock (note relation to calcified plaque in Fig. 4(a)) and from 10-12 o’clock. Auto-thresholded difference images (Figs. 3(b)-(c)) exhibit the changes in enhancement over time due to diminution. A magnification of an area of enhancement is shown in Fig. 4; this highlights the difficulty of imaging the echo-transparent VV under normal circumstances. Enhancement is plotted over time in Figs. 5(a) and 5(b) for the intimo-medial/plaque and adventitial regions respectively.

² Additional figures are available from <http://www.vcl.uh.edu/CARDIA>.

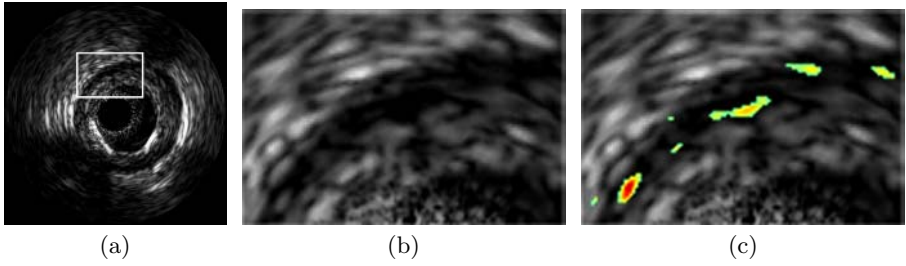


Fig. 4. (a) Unenhanced frame from Case 1, (b) a magnified ROI of the frame, and (c) the ROI with an overlay showing enhanced features invisible in (b). Observe conceptual agreement with histological observations reported in the literature (Fig. 1(a)).

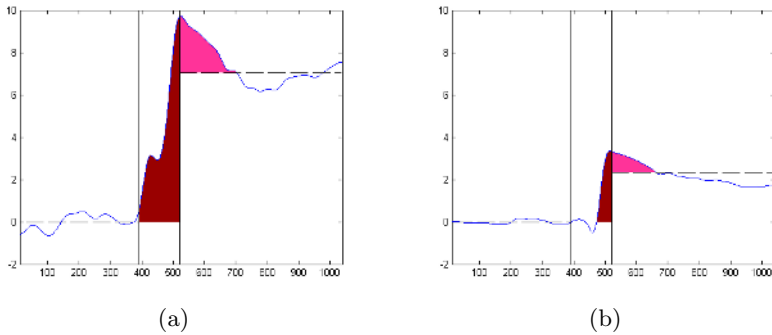


Fig. 5. Case 1: AEPEP over time for the (a) intimo-medial region and (b) adventitia. Pre- and post-injection means are shown (dotted horizontal lines; graph has been translated to 0% pre-injection mean). Injection start and peak are marked by vertical lines. Solid areas indicate the injection and return-to-baseline periods.

Summary: Enhancement statistics for six additional cases whose analyses were identical to that of Case 1 are presented in Table 1. The pre-injection mean AEPEP (μ_{pre}) scores report a positive value due to random speckle events occurring above each frame's enhancement threshold. Hence, the most meaningful

Table 1. Representative results from seven case studies.

	μ_{pre}	μ_{post}	R
Case 1	2.41%	9.13%	3.79
Case 2	1.90%	2.80%	1.47
Case 3	1.53%	3.97%	2.59
Case 4	1.88%	3.35%	1.78
Case 5	1.48%	2.27%	1.53
Case 6	2.50%	2.70%	1.08
Case 7	4.00%	5.50%	1.37

global enhancement metric we have found is the ratio of the post-injection mean (μ_{post}) to pre-injection mean, $R = \frac{\mu_{post}}{\mu_{pre}}$. This ratio appears to correlate well with the strength and quantity of visible enhancement in our studies.

4 Discussion

We have described a novel method which, for the first time, enables IVUS imaging of VV presence. Due to the inherent limitations of *in vivo* human coronary IVUS studies, we were unable to correlate our findings with histopathological evidence of VV density. However, the significant changes in the IVUS signal after microbubble passage leave no doubt as to its ability to show contrast enhancement. Knowing that almost all blood perfusion in plaques comes through VV capillaries, we hypothesize that the enhancement is related to the density of VV in the vessel wall. A comparison of our images with the μ CT and pathology images of VV reported in the literature [9] shows a conceptual agreement.

In vitro studies are currently underway to further gauge the accuracy and robustness of our technique. We are also investigating high-speed acquisition of the raw radiofrequency signal in order to more effectively measure the subtle changes in the backscatter from inside the plaque and adventitia. A clinical discussion of our results may be found in [12].

Acknowledgements. We would like to thank S. Carlier, C.J. Hartley, R.W. Metcalfe, N. Dib, R. Mehran, E. Falk, K. Gul, C. Stefanadis, T. Papaioannou, S. Vaina, M. Drakopoulou and I. Mitropoulos for valuable assistance. S. O'Malley is supported under a NSF Graduate Research Fellowship. This work supported in part by NSF Grant IIS-0431144. 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 NSF.

References

1. Naghavi, M., Libby, P., Falk, E., Casscells, S., Litovsky, S., *et al.*: From vulnerable plaque to vulnerable patient: A call for new definitions and risk assessment strategies: Parts I-II. *Circulation* **108** (2003) 1664–1672/1772–1778
2. Kolodgie, F., Gold, H., Burke, A., Fowler, D., *et al.*: Intraplaque hemorrhage and progression of coronary atheroma. *New England J of Med* **349** (2003) 2316–2325
3. Hayden, M., Tyagi, S.: Vasa vasorum in plaque angiogenesis, metabolic syndrome, type 2 diabetes mellitus, and atheroscleropathy: a malignant transformation. *Cardiovasc Diabetol* **3** (2004)
4. Tearney, G., Yabushita, H., Houser, S., Aretz, H., Jang, I.K., *et al.*: Quantification of macrophage content in atherosclerotic plaques by optical coherence tomography. *Circulation* **107** (2003) 113–119
5. Khan, T., Soller, B., Naghavi, M., Casscells, W.: Tissue pH determination for the detection of metabolically active, inflamed vulnerable plaques using near-infrared spectroscopy: An in-vitro feasibility study. *Cardiology* **103** (2005) 10–16

6. Go, V., Bhatt, M., Hendel, R.: The diagnostic and prognostic value of ECG-gated SPECT myocardial perfusion imaging. *J Nucl Med* **45** (2004) 912–921
7. Winter, P., Morawski, A., Caruthers, S., Fuhrhop, R., Zhang, H., *et al.*: Molecular imaging of angiogenesis in early-stage atherosclerosis with $\alpha_v\beta_3$ -integrin-targeted nanoparticles. *Circulation* **108** (2003) 2270–2274
8. Olszewski, M., Wahle, A., Vigmostad, S., Sonka, M.: Multidimensional segmentation of coronary intravascular ultrasound images using knowledge-based methods. In: *Proc SPIE Medical Imaging*. Volume 5747. (2005) 496–504
9. Gossel, M., Rosol, M., Malyar, N., Fitzpatrick, L., Beighley, P., Zamir, M., Ritman, E.: Functional anatomy and hemodynamic characteristics of vasa vasorum in the walls of porcine coronary arteries. *Anat Rec* **272A** (2003) 526–537
10. Zhu, H., Oakeson, K., Friedman, M.: Retrieval of cardiac phase from IVUS sequences. In: *Proc SPIE Medical Imaging*. Volume 5035. (2003) 135–146
11. Bruzzone, L., Prieto, D.: Automatic analysis of the difference image for unsupervised change detection. *IEEE Trans Geo Remote Sensing* **38** (2000) 1171–1182
12. Carlier, S., Kakadiaris, I.A., Dib, N., Vavuranakis, M., Stefanadis, C., O'Malley, S.M., *et al.*: Vasa vasorum imaging: A new window to the clinical detection of vulnerable atherosclerotic plaques. *Curr Athero Rep* **7** (2005) 164–169

Parametric Response Surface Models for Analysis of Multi-site fMRI Data

Seyoung Kim¹, Padhraic Smyth¹, Hal Stern¹, and Jessica Turner²

¹ Bren School of Information and Computer Sciences, University of California, Irvine
{sykim, smyth}@ics.uci.edu, sternh@uci.edu

² Department of Psychiatry and Human Behavior, University of California, Irvine
turnerj@uci.edu

Abstract. Analyses of fMRI brain data are often based on statistical tests applied to each voxel or use summary statistics within a region of interest (such as mean or peak activation). These approaches do not explicitly take into account spatial patterns in the activation signal. In this paper, we develop a response surface model with parameters that directly describe the spatial shapes of activation patterns. We present a stochastic search algorithm for parameter estimation. We apply our method to data from a multi-site fMRI study, and show how the estimated parameters can be used to analyze different sources of variability in image generation, both qualitatively and quantitatively, based on spatial activation patterns.

1 Introduction

Functional magnetic resonance imaging (fMRI) is one of the primary techniques for studying how the brain functions when triggered by external stimuli. Neural activity in local areas of the brain changes the oxygenated blood level, changes that in turn can be detected by MRI scanners. The data collected in an fMRI session consist of a time-series of voxel images, containing both temporal and spatial information about brain activation patterns. The temporal aspect of the data is often aggregated by fitting a linear regression model that relates the activation data at each voxel to the stimulus signal and perhaps other predictors as well. The regression coefficients of the stimulus signal at each voxel (often denoted as β) or other regression summary (e.g., t -statistic) can then be used as a statistical parametric map of activity [1]. There is a significant body of prior work using techniques such as hypothesis testing and thresholding to search these statistical voxel images for activated brain regions. Test statistics used in these techniques include the maximum activation intensity within a selected region of interest, or the spatial extent of the largest set of contiguous voxels above a threshold [2, 3].

An important aspect of the data that is typically not considered in these studies is the spatial pattern of activation across voxels. Voxels with a high activation level tend to appear in spatial clusters in the β -maps, forming what looks like a bell shape for each activation cluster. These spatial clusters can

be explained to first order by assuming that the underlying true activations are point sources that are then spatially smoothed (convolved) by Gaussian-like filters.

In this paper we characterize fMRI activation patterns in the brain by modeling the shape of each “activation cluster” by (a) its height or peak value, representing the amount of activation, and (b) the location of the cluster, modeling the center of activation in voxel-space. We represent these features using a parametric model of the functional form of the surface with free parameters for the heights and center locations of the “bumps.” These parameters can be estimated in a statistical manner from the β activation maps.

In terms of relevant prior work, Hartvig [5] also used a similar parametric function for modeling spatial activation patterns. However, the focus in this earlier work was on extracting activated voxels by thresholding after the model parameters were estimated. In the work described in this paper we are interested in the shape features themselves, and analyze the activation pattern at the feature level rather than at a voxel level. In our experiments we show how the estimated features of the local activation clusters can be used in a multi-site study of fMRI to analyze sources of variability. Evidence in this direction is suggested by Fig. 1, where the locations of the peaks of activation are more consistent in the four runs of the same subject within the same site (MRI scanner), than between the two sites.

Multi-site fMRI studies are not currently common, but are becoming more so as the need is growing to combine imaging data across sites to create larger datasets than would be possible at a single site. These datasets allow imaging experiments in rare diseases where few subjects would be available at any given site, or with very complex hypotheses where a single site study would be underpowered. In particular, in one multi-site study of signal reliability across sites [4] the activation levels and locations were similar, but the pattern of activation across sites was very different. The ability to assess such patterns of activation provides the primary motivation for the method proposed in this paper.

The paper is organized as follows. In Sect. 2, we briefly describe the data collection process and preprocessing steps used to produce the activation maps analyzed in this paper. In Sect. 3, we introduce a spatial model for activation patterns, describe inference procedures for this model and present experimental results based on fitting this model to multi-site fMRI data. In Sect. 4, we show how the shape features estimated from the surface modeling can be used to investigate multi-site variability. Sect. 5 concludes with a brief discussion on future work.

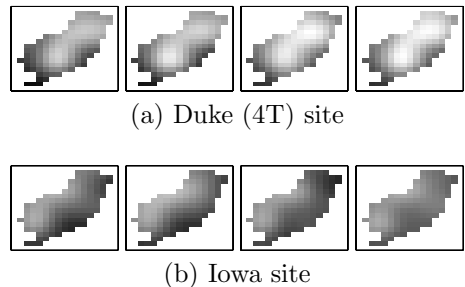


Fig. 1. Raw data (β maps) for a cross section at $z = 53$ of the right precentral gyrus of four runs within visit 2 of subject 3

2 Multi-site Data Collection and Preprocessing

fMRI scans for the same five control subjects were collected from 10 different scanners (UCSD, UCI, Stanford, Duke (1.5T), Duke (4T), New Mexico, Minnesota, Iowa, BWH, MGH) as part of a multi-site study of functional brain images, known as FIRST BIRN (Functional Imaging Research on Schizophrenia Test-bed Biomedical Informatics Research Network), also known as Function BIRN or fBIRN. For each subject there were two visits to each site, and at each visit fMRI data were collected for four runs of a sensorimotor task and two runs of breathholding, resting, and two cognitive tasks. A primary goal of this data collection is to better understand the variability of fMRI response patterns across runs, visits, scanners (sites) and subjects, so that future data collected across sites and subjects can be analyzed collectively and consistently. In this paper, we use the data from the sensorimotor task, and we focus on activation within specific regions of interest such as the right precentral gyrus and the left superior temporal gyrus that are relevant to this task.

Each run of the sensorimotor task produces a series of 85 scans that can be thought of as a large time-series of voxel images. The set of scans for each run is preprocessed in a standard manner using SPM99, with the default settings. The preprocessing steps include correction of head motion, normalization to a common brain shape (MNI template), and spatial smoothing.

A general linear model is then fit to the time-series data for each voxel. A β -map is a voxel image of the regression coefficients (β 's) that summarizes the activation across time as an activation map. Binary masks for regions of interest from the normalized atlas were then used to extract the β values for all voxels within the region.

3 Activation Surface Modeling

3.1 Model

We model the spatial activation pattern (β -map) for a region of interest as a superposition of multiple Gaussian-shaped surfaces, with a constant background term for non-activation, and with additive Gaussian measurement noise. Below we develop the model for the case of 2-dimensional slices of pixels—the 3-dimensional case can be derived directly as an extension of the 2-dimensional case, but is not pursued in this paper.

Assuming the number M of Gaussian surfaces for a 2-dimensional slice of a region is known and fixed, the model for the activation value β at position $\mathbf{x} = (x_1, x_2)'$ is defined as

$$\beta = \mu + \sum_{m=1}^M k_m \exp(-(\mathbf{x} - \mathbf{b}_m)'(\mathbf{x} - \mathbf{b}_m)/\sigma_m) + \varepsilon, \quad (1)$$

where ε is distributed as $N(0, \sigma^2)$, and μ is the background level. $\boldsymbol{\theta}_m = \{k_m, \mathbf{b}_m, \sigma_m\}$, $m = 1, \dots, M$, is the set of parameters for the m th Gaussian surface,

corresponding to activation centered around $\mathbf{b}_m = (b_{1m}, b_{2m})'$ with height of k_m , and σ_m controlling the volume under the surface. By using σ_m instead of a 2×2 covariance matrix, we assume a spherical shape with no correlation between x_1 and x_2 directions.

In the Bayesian estimation setting, we are interested in the posterior distribution of the unknown parameters $p(\mu, \sigma, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M | \boldsymbol{\beta})$. Using Bayes' rule, we write $p(\mu, \sigma, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M | \boldsymbol{\beta}) = p(\boldsymbol{\beta} | \mu, \sigma, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M) p(\mu, \sigma, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M) / p(\boldsymbol{\beta})$, where $p(\boldsymbol{\beta} | \mu, \sigma, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M)$ is the likelihood of the $\boldsymbol{\beta}$ values $\boldsymbol{\beta} = \{\beta_1, \dots, \beta_N\}$ for a region with N voxels given the model in (1). Assuming independent priors for (μ, σ^2) , and $\boldsymbol{\theta}_m$'s, we choose a uniform prior $p(\mu, \log \sigma) \propto 1$ for $(\mu, \log \sigma)$, $p(k_m) \propto 1$ for k_m 's, where $k_m > 0$ for positive activations, and assign a Gamma (a, b) prior on σ_m with hyperparameters a and b . Given that it is reasonable to believe that the center of activation \mathbf{b}_m is located inside (or close to) the region of brain under consideration, we assume $p(\mathbf{b}_m)$ is uniform inside or within one pixel of the region. It is straightforward to verify that this choice of prior distributions yields a proper posterior distribution.

3.2 Parameter Estimation Using Stochastic Search

Because of the nonlinearity of the model described in Sect. 3.1, it is not possible to determine the posterior distribution of the parameters analytically, and we need to rely on simulation methods to obtain samples from the posterior. In this work, we use a combination of the Gibbs sampler and the Metropolis algorithm.

In the Gibbs sampler, starting from initial values for the unknown parameters, at each iteration we update each parameter in turn by replacing it with a sample from its conditional posterior distribution given the data and the most recent samples of the other parameters. By repeating this update, the Gibbs sampler converges (under mild conditions) to a state where the samples can be presumed to be coming from the joint posterior distribution of interest. Direct sampling from the conditional posterior distribution is possible for μ and σ^2 , but not for the $\boldsymbol{\theta}_m$'s. The Metropolis algorithm is used to sample $\boldsymbol{\theta}_m$'s inside the Gibbs sampler. We describe the details of the sampling algorithm below.

1. Sample a new value for μ from its conditional posterior distribution $p(\mu | \sigma, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M, \boldsymbol{\beta}) = N(\hat{\mu}, \sigma^2/N)$, where $\hat{\mu} = (\sum_{i=1}^N (\beta_i - \sum_{m=1}^M k_m \exp(-(\mathbf{x}_i - \mathbf{b}_m)'(\mathbf{x}_i - \mathbf{b}_m)/\sigma_m)))/N$.
2. Sample a new value for σ^2 from its conditional posterior distribution. Given the prior $p(\log \sigma) \propto 1$, which is equivalent to $p(\sigma^2) \propto \sigma^{-2}$, the conditional posterior distribution for σ^2 can be found as $p(\sigma^2 | \mu, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M, \boldsymbol{\beta}) = \text{Inv-gamma}(\nu/2, (\nu/2)s^2)$, where the degree of freedom $\nu = N$ and $s^2 = (\sum_{i=1}^N (\beta_i - \mu - \sum_{m=1}^M k_m \exp(-(\mathbf{x}_i - \mathbf{b}_m)'(\mathbf{x}_i - \mathbf{b}_m)/\sigma_m))^2)/N$.
3. For $m = 1, \dots, M$,
 - (a) Sample k_m^* from the jumping distribution $N(k_m, (\tau_k)^2)$ and compute $r = (p(\boldsymbol{\beta} | k_m^*, \boldsymbol{\theta}_{-k_m}) p(k_m^*)) / (p(\boldsymbol{\beta} | k_m, \boldsymbol{\theta}_{-k_m}) p(k_m))$, where $\boldsymbol{\theta}_{-k_m}$ represents all parameters other than k_m . If $r > 1$, accept k_m^* as a new value for k_m , otherwise, accept it with the probability r .

- (b) Update \mathbf{b}_m as in 3a, but using the jumping distribution $N_2(\mathbf{b}_m, (\tau_{\mathbf{b}})^2 I)$.
- (c) Update σ_m as in 3a but using the jumping distribution $N(\sigma_m, (\tau_{\sigma})^2)$.

Before the start of the sampler, the values for $(\tau_k)^2$, $(\tau_{\mathbf{b}})^2$, and $(\tau_{\sigma})^2$ in Step 3 are adjusted to values that keep the acceptance rate at roughly 30-50%.

In the results described below we summarize the posterior distributions via “point estimates” (specific values of the parameters), corresponding to the mean values for each parameter over the samples (i.e., the estimated posterior mean).

3.3 Results from Surface Fitting

In this section, we discuss results of our modeling procedure for two cross sections, one at $z = 53$ of the right precentral gyrus and another at $z = 33$ of the left superior temporal gyrus, for subjects 1 and 3 in our study, for the sensorimotor task at all sites.

From visual inspection of the images, we set the number of Gaussians in the model to $M = 1$ for subject 1, and $M = 2$ for subject 3 in the case of the right precentral gyrus, and $M = 2$ for both subjects in the case of the left superior temporal gyrus. From an exploratory analysis of the data, the hyperparameters a and b of the gamma prior distributions of the σ 's were set so that the mode and variance are approximately 15 and 160 ($a = 3.0$ and $b = 0.1365$). With this setting of the model, we ran the Markov chain simulation for 20,000 iterations for each image until convergence, and output the mean of the samples from the last 10,000 iterations to provide point estimates of the parameters.

A typical example (for a particular run, visit, and subject) is shown in Fig. 2, comparing the raw β maps with the activation maps for the fitted model based on point estimates of the parameters from the same raw data. The shapes of the activation pattern in the estimated surfaces such as the locations of the peaks are consistent with what we see in the raw data. Table 1 shows an average across multiple images of the sum of squared errors (across voxels) between the fitted model and the data, divided by the voxel variance in each image and expressed as a percentage—in effect, the average amount of variability in the data not explained by the model. Both Fig. 2 and Table 1 suggest that while the functional form of

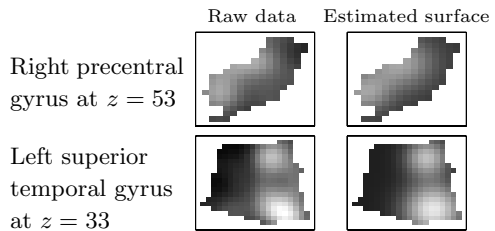


Fig. 2. Raw data vs. estimated surfaces from run 3, visit 2, and subject 3 at Iowa site

Table 1. Percentage of variance not explained by the models

	subject 1	subject 3
Right precentral gyrus ($z = 53$)	15.4%	7.1%
Left superior temporal gyrus ($z = 33$)	16.1%	5.6%

the model is a result of various simplifying assumptions it nevertheless provides a good fit to the observed spatial activation patterns.

4 Modeling Sources of Variability

We can think of the estimated parameters of the model as features that summarize the activation pattern, and analyze the statistical variation of these features rather than that of the raw data. Fig. 3 visually shows the variabilities in the estimated location parameters of the local activation centers at each of run, visit, and site levels for the cross section $z = 53$ of the right precentral gyrus of subject 3. The estimated locations are represented as o's for visit 1 and +'s for visit 2. The two bumps are connected with a line if they come from the same run. Across certain groups of sites (e.g., UCI and UCSD) there appears to be little cross-site variability in the estimated spatial locations, but between other pairs (e.g., MGH and UCI) the cross-site variation is much larger than the within-site variation. Qualitatively similar results are obtained for other cross-sections.

In this section, we attempt to quantify the contributions of site, visit, and run effects to variation in features of the local activations within each subject. In the case of $M = 2$, we first find the correspondence of the two bumps across images from different sites, visits and runs based on the estimated location parameters using a simple heuristic algorithm, and consider each bump separately in the variance component analysis. In Sect. 4.1 we discuss a Bayesian hierarchical model for a variance component analysis. In Sect. 4.2 we present the results from the model.

4.1 Bayesian Variance Component Analysis

A traditional variance components analysis can be used to decompose the observed variation in signal into portions attributable to site, visit, and run variabilities. One such model that we choose to use in our analysis is $y_{ijk} = \mu + s_i +$

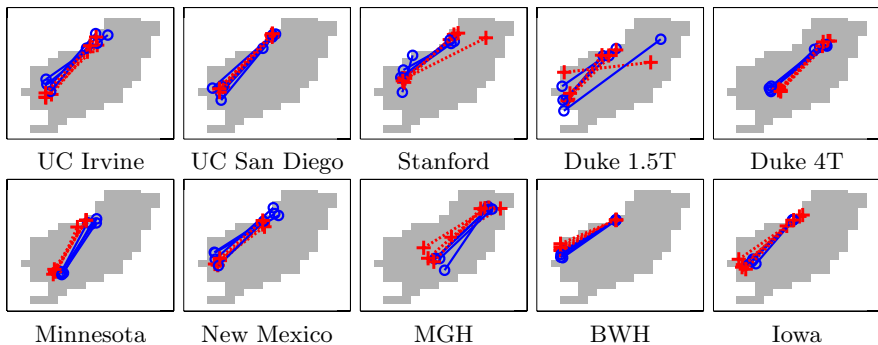


Fig. 3. Variability in locations parameters estimated from right precentral gyrus at $z = 53$ of subject 3. o's are visit 1, and +'s are visit 2. Two locations are connected with a line if they come from the same image.

$v_{ij} + r_{ijk}$, where y_{ijk} is the response measure, μ is an overall mean, s_i is a site effect, v_{ij} is a visit effect and r_{ijk} is a run effect (essentially all sources of variation not included in previous terms). The responses y_{ijk} 's considered in this analysis are the estimated locations and heights of each of the Gaussian components. Each of the effects (site, visit, run) is further modeled as a Gaussian random variable with mean zero and variance/covariance parameter that characterizes the contribution to overall response variability of the particular source. It is the variance parameters, or the variance components, that are of interest.

Parameters are estimated as means of the posterior distribution of the variance components from a Bayesian analysis with weak prior distributions on the variance components. The prior distributions for the analyses reported here are $N(0, 10^{10})$ for μ (a common vague prior for the mean of a Gaussian distribution), gamma with mean 1 and variance 100 on the one dimensional precision parameters (reciprocals of the variances), and Wishart distribution with degree of freedom 2 and scale parameters estimated empirically from the y_{ijk} 's for the covariance matrices. Samples of the variance parameters from their posterior distribution given the data are obtained using WinBugs, a freely available software implementation of the Gibbs sampler. The Gibbs sampler was run for 1,000,000 iterations by which time the draws were determined to be representative of the posterior distribution. The means of the last 200,000 draws are reported as estimates for the variance components.

4.2 Results

We compute the variance components for sites, visits, and runs of location and height parameters, and show the proportions of variance with respect to the total variation in Tables 2 and 3. For the 2×2 covariance matrices of location parameters, we use the absolute value of the determinant of the matrix as a summary of the amount of variation instead of the full matrix.

In the cross section of $z = 53$ of the right precentral gyrus the effect size due to sites is consistently larger than visit and run effects across different bumps and subjects in both locations and heights. The visit effects account for the smallest proportion of the total variabilities. Note also that the relative sizes of variance components (for sites, visits, runs) are similar for both subjects.

A slightly different pattern is observed in the cross section of $z = 33$ of the left superior temporal gyrus. The contribution from visit effects is still the smallest, but the run variability is often larger than the site variability.

Overall we see more consistent patterns of variance components in the right precentral

Table 2. Proportions of variance components for right precentral gyrus at $z = 53$

	Subject 1		Subject 3			
	Height	Location	Height		Location	
			Bump 1	Bump 2	Bump 1	Bump 2
Site	0.51	0.86	0.50	0.58	0.67	0.90
Visit	0.22	0.07	0.13	0.18	0.02	0.03
Run	0.27	0.07	0.37	0.24	0.31	0.07

Table 3. Proportions of variance components for left superior temporal gyrus at $z = 33$

	Subject 1				Subject 3			
	Height		Location		Height		Location	
	Bump 1	Bump 2	Bump 1	Bump 2	Bump 1	Bump 2	Bump 1	Bump 2
Site	0.25	0.52	0.05	0.06	0.49	0.59	0.40	0.45
Visit	0.42	0.25	0.18	0.03	0.12	0.05	0.04	0.02
Run	0.33	0.23	0.77	0.91	0.39	0.36	0.56	0.53

gyrus than in the left superior temporal gyrus. The right precentral gyrus is considerably smaller in size, therefore, is less affected by the limitation that we are looking at a 2 dimensional cross section instead of taking into account the variability in all of the three dimensions.

An alternative non-Bayesian method of moments approach provides similar results with the exception of some negative estimates of variances (a known disadvantage of this approach).

5 Conclusions

We have shown that spatial modeling of fMRI activation patterns in local brain regions can extract reliable and useful information providing (for example) a basis for statistical analysis of variability. Future work includes analyzing more regions of interest with all 5 subjects in the study, extending the model to 3-dimensional voxel data, and developing techniques for automatically selecting the number of bumps M .

Acknowledgements. This material was supported in part by the National Science Foundation under grants No. SCI-0225642 and IIS-0431085. The authors also acknowledge the support of the First BIRN (M01RR00827), the Transdisciplinary Imaging Genetics Center (P20RR020837-01) and the National Alliance for Medical Image Computing (Grant U54 EB005149), funded by NIH/NCRR.

References

1. Bullmore, E., Brammer, M., Williams, S.C., Babe-Hesketh, S., Janot, N., David, A., Mellers, J., Howard, R., & Sham, P. (1996). Statistical methods of estimation and inference for functional MR image analysis. *Magnetic Resonance in Medicine*, 35(2):261–277.
2. Cao, J., & Worsley, K. (2001). Applications of random fields in human brain mapping. In *M. Moore (Ed.) Spatial Statistics: Methodological Aspects and Applications*, Springer Lecture Notes in Statistics, 159:169–182.
3. Friston, K.J., Holmes, A.P., Price, C.j., Buchel, C., & Worsley, K.J. (1999). Multi-subject fMRI studies and conjunction analysis. *NeuroImage*, 10:385-396.
4. Casey, B. J., Cohen, J. D., O’Craven, K., Davidson, R. J., Irwin, W., Nelson, C. A. et al. (1998). Reproducibility of fMRI results across four institutions using a spatial working memory task. *Neuroimage*, 8:249-261.
5. Hartvig, N. (1999). A stochastic geometry model for fMRI data. Research Report 410, Department of Theoretical Statistics, University of Aarhus.

Subject Specific Finite Element Modelling of the Levator Ani

Su-Lin Lee^{1,2}, Ara Darzi², and Guang-Zhong Yang^{1,2}

¹ Royal Society/Wolfson Foundation Medical Image Computing Laboratory,

² Department of Surgical Oncology and Technology,

Imperial College London, London, United Kingdom

{su-lin.lee, a.darzi, g.z.yang}@imperial.ac.uk

Abstract. Understanding of the dynamic behaviour of the levator ani is important to the assessment of pelvic floor dysfunction. Whilst shape modelling allows the depiction of 3D morphological variation of the levator ani between different patient groups, it is insufficient to determine the underlying behaviour of how the muscle deforms during contraction and strain. The purpose of this study is to perform a subject specific finite element analysis of the levator ani with open access magnetic resonance imaging. The method is based on a Mooney-Rivlin hyperelastic model and permits dynamic study of subjects under natural physiological loadings. The value of the proposed modelling framework is demonstrated with dynamic 3D data from nulliparous, female subjects.

1 Introduction

The levator ani is a tripartite, striated muscle group of the pelvic floor and consists of the pubococcygeus, the iliococcygeus and the puborectalis. Childbirth has been shown to be one of the primary causes of injury to the levator ani with symptoms ranging from pain or constipation to faecal or urinary incontinence [1]. In current clinical settings, diagnoses are made either with full clinical examinations or using magnetic resonance imaging (MRI) in mid-sagittal and mid-coronal planes with measurements made from the position of organs to the pubococcygeal line [2], and an evacuation scan is popular with this method. Treatments for pelvic floor dysfunction include pelvic floor exercises and surgery. The usual surgical treatment of the pelvic floor is divided between the urologist, gynaecologist and proctologist. Existing research has shown that among women who have had surgery, approximately 20% require a second operation. This is partially due to the current surgical practice and the general lack of quantitative assessment in guiding the surgical procedures. A more unified way of understanding the levator ani in 3D, as well as the location of any injuries is essential for reducing the rate of repeated operation.

Hoyte, *et al.* [3,4] initiated the use of 3D models of the levator ani in asymptomatic and symptomatic women. Each shape was manually segmented and the groups were compared by measurements. Lee *et al.* extended 3D modelling of the levator ani with statistical shape models that allowed the identification of the main modes of variation of the shape of the levator in normal subjects at rest, on maximal contraction, and on

maximal strain [5,6]. Whilst these techniques provide the visualisation and understanding of the 3D morphology of the levator, they do not provide an insight into how the muscle strains during the performance of pelvic floor exercises.

To elucidate dynamic behaviour of the soft tissue, finite element models (FEM) have been widely used to model biological tissues such as bone, myocardium, brain deformation and breast tissue deformation. For example, Pathmanathan, *et al.* [7] have modelled breast deformation when the patient is standing, supine and compressed during mammography. d'Aulignac, *et al.*, [8] used shell elements (thin-walled structures) to build a FEM of the levator ani. The model took into account the direction of the muscle fibres and the incompressibility of the tissue, and used geometrical data obtained by Janda, *et al.*, [9-11], where morphological parameters of the levator ani, including geometry of the muscle fibres and muscle sarcomere length, are measured from cadavers. Force-length curves, which determine the maximal force in certain positions, were derived and the material properties of the muscle were investigated by using 20mm^2 samples of the cadaver tissue to determine the biaxial stress-strain [10].

In this study, a FEM based on patient specific data, rather than idealised models with a constant thickness throughout the muscle sheet, is used to predict dynamic properties of the pelvic floor under different physiological loadings. Pelvic floor exercise involves extensive tissue deformation during maximal contraction or maximal strain. Linear models are inappropriate in this case as linear constitutive laws, such as Hooke's law, do not hold. A Mooney-Rivlin hyperelastic model is used instead in this study. The value of the proposed modelling framework is demonstrated with dynamic 3D data from three nulliparous, female subjects using open access MRI.

2 Methods

2.1 Hyperelasticity

Hyperelasticity is a classical modelling framework that is applicable to biological tissues. It can account for both mechanical and geometric non-linearities in the mathematical formulation. A material is considered elastic when the stress S at a point X depends on the deformation gradient F . A material is considered hyperelastic when the stress S can be derived from both the deformation gradient F and a strain energy function W :

$$S = \frac{\partial W}{\partial E} \quad (1)$$

where E is the Lagrangian strain tensor, defined as:

$$E^{(m)} = \frac{1}{m} (U^m - I), \quad m \neq 0 \quad (2)$$

$$E^{(0)} = \ln U, \quad m = 0.$$

where U is the right stretch tensor and is obtained from the deformation gradient F by polar decomposition:

$$F = RU . \quad (3)$$

In Eq. (3), R is a proper orthogonal tensor and U is positive definite and symmetric.

As it is practically difficult to test the muscle material of a levator ani *in vivo*, a five parameter Mooney-Rivlin hyperelastic material model was adopted with the parameters from the work by [12]: $c_{10} = 2.5\text{kPa}$, $c_{20} = 0.625\text{kPa}$, $c_{01} = 0$, $c_{02} = 0$, $c_{11} = 0$. Such models have been used to model the human tongue (and in turn was originally developed for the myocardium and other general soft tissues) and facial muscles [13]. The Mooney-Rivlin model approximates the energy function W:

$$W = c_{10}(I_1 - 3) + c_{01}(I_2 - 3) + c_{20}(I_1 - 3)^2 + c_{11}(I_1 - 3)(I_2 - 3) + c_{02}(I_2 - 3)^2 \quad (4)$$

where I_1 and I_2 are the first and second invariants of the strain tensor E. An examination of a square centimetre of the material shows that stretch causes a reduction in width in the direction perpendicular to the stretch.

2.2 Magnetic Resonance Image Acquisition

The anatomical structure of the levator ani during different levels of stress were built from magnetic resonance images acquired from a 0.5T GE iMR open access scanner. Each 3D data set was acquired with a turbo spin echo sequence (TR = 4900ms, TE = 102ms, slice thickness = 5.5mm) for each of 3 nulliparous, female subjects with a total scan time of approximately 5 minutes. The ages for Subjects 1, 2 and 3 were 24, 31 and 23, respectively.

2.3 Finite Element Modelling

For this study, dynamic 3D morphological segmentation of the levator ani was based on the work by Lee *et al* [6]. The node correspondence for each levator ani surface was automatically determined by using a harmonic embedding framework developed by Horkaew *et al* [14]. The optimisation also took into account the thickness information at each node, calculated to be the minimal distance from the bottom surface of the levator to its corresponding top surface. Each shape is defined by a total of 2401 nodes (with a specific thickness at each node) and 4608 elements. Each levator ani was subsequently modelled using SHELL181 elements in ANSYS with the nodes in the areas connected to the pelvic bone, ligament and coccyx fixed in translation but left free to rotate. There are no fixed degrees of freedom on all other nodes. Thickness information was added at each node. Pressures were applied to each surface to simulate contraction and strain of the levator ani.

3 Results

The morphological change for each of the three subjects can be seen in Figure 1. The shape variation associated with contraction and strain is consistent with the expected

anatomical change. The contraction shows a lifting of the levator ani while the strain leads to a downward movement and spread of the levator ani wings. The von Mises stress for each surface is displayed in Figure 2, where the stress distribution is consistent with areas involving the most significant movement during pelvic floor exercise.

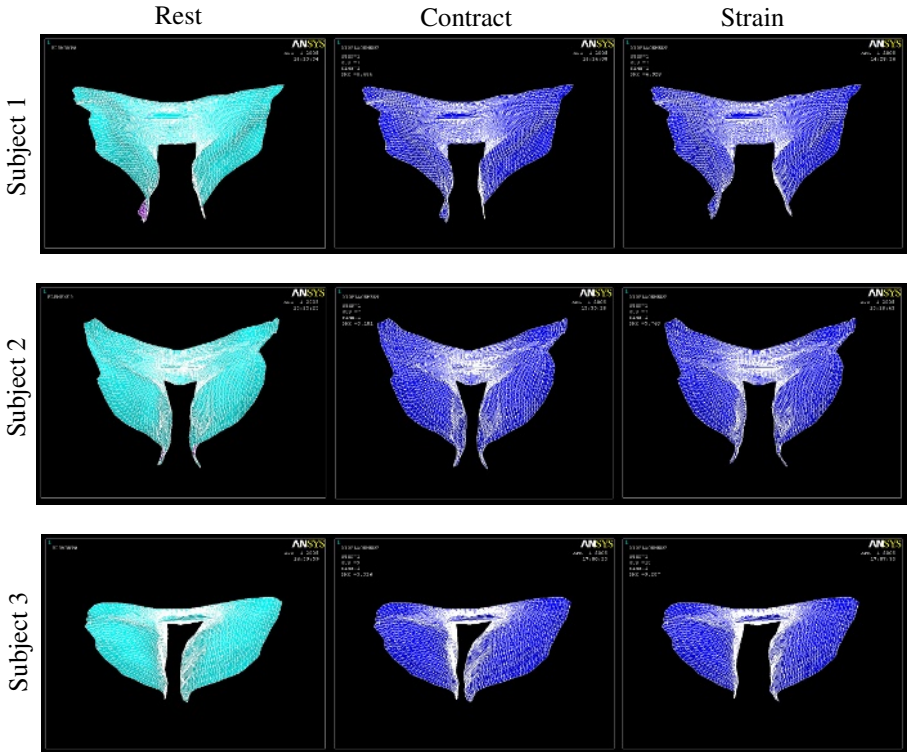


Fig. 1. The morphological changes of the levator anis for the three subjects simulated during different levels of stress: rest, contraction and strain

In Figure 3, we examine the von Mises stress in the mid-coronal plane. The graph in Figure 3b shows that greater stress is located in the anterior rather than the posterior of the levator ani. This is to be expected since this area bears most of the pressure and is responsible for straightening of the anal canal, which is required during defecation.

To assess the reliability of the model, we have evaluated the change in maximal von Mises stress when the thickness of the model is changed. The levator ani surface was constructed with equal thickness at each node. In this case, the average thickness (rounded to 3mm) was used and the results are summarised in Table 1. It is evident that small variation in thickness does not result in a large change in the stress values.

To illustrate the convergence behaviour of the model, the reaction force at each time-step is plotted in Figure 4, where the load was applied in substeps from time 0 to 1.

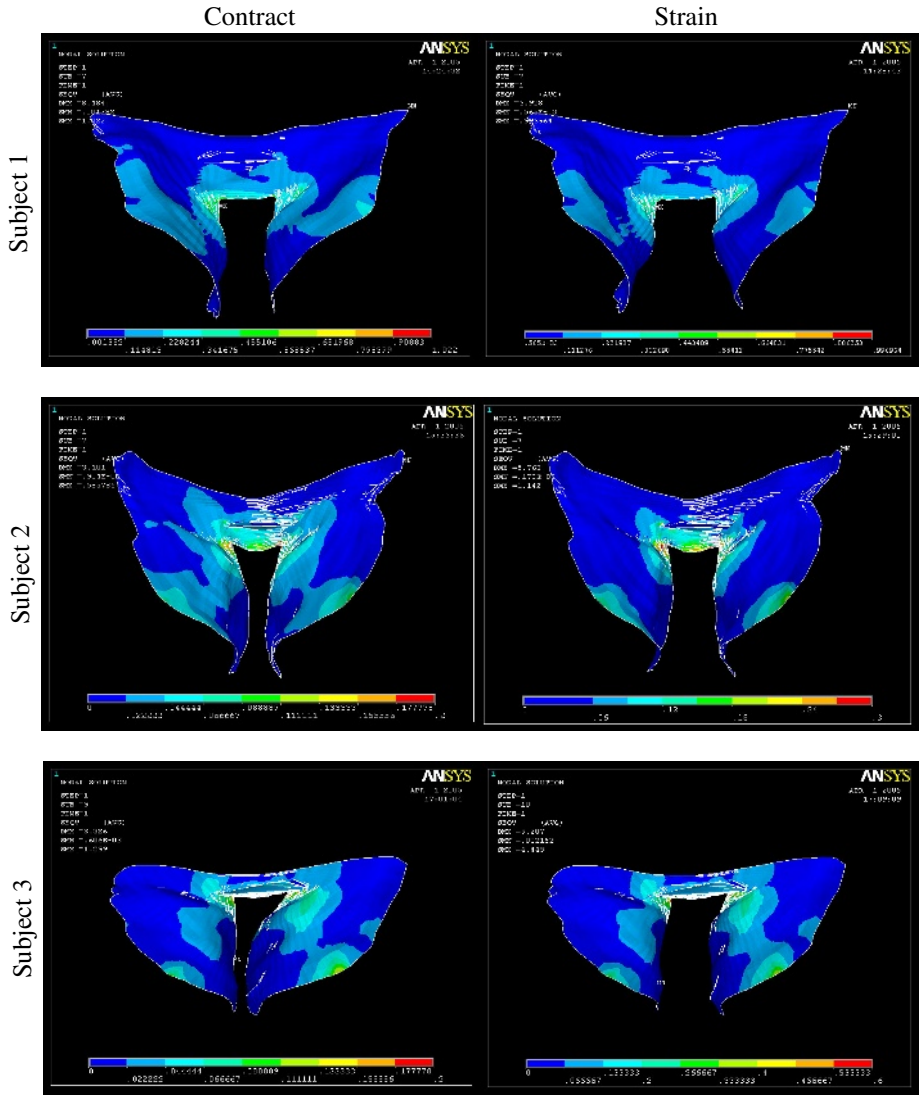
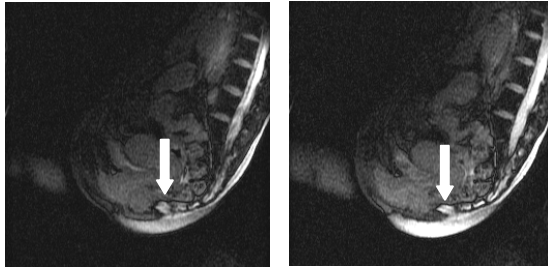


Fig. 2. The von Mises stress for each surface during contraction and strain

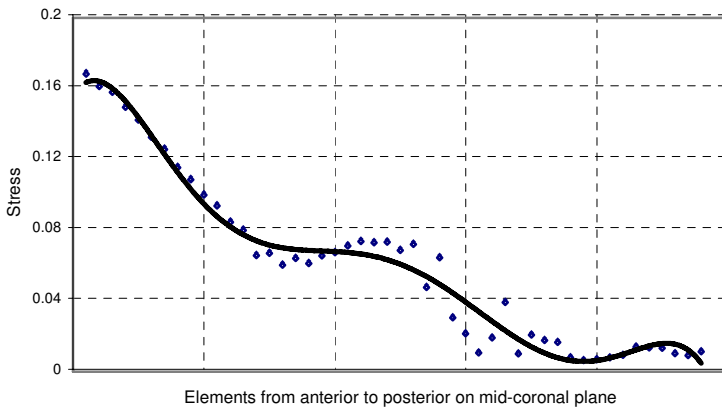
Table 1. Changes in the maximal von Mises stress when thickness is varied

Thickness	Maximal von Mises stress	Percent change
+10%	0.427622	7.68%
+5%	0.445193	3.89%
3mm	0.463207	-
-5%	0.481528	3.96%
-10%	0.499974	7.94%



(a)

von Mises Stress



(b)

Fig. 3. (a) Mid-Coronal MR images of the pelvic floor with the levator ani (white arrow) in at rest (left) and at maximal strain downwards (right) and (b) the corresponding von Mises stress in the elements of the modelled levator ani lying on the mid-coronal plane

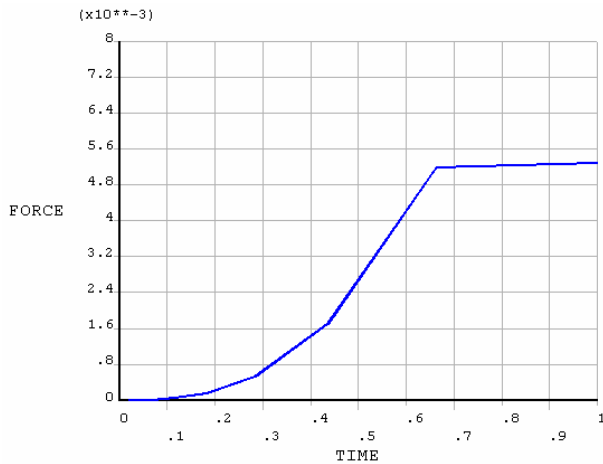


Fig. 4. Convergence behaviour of the L2 norm of the force over time

4 Discussion and Conclusions

The FEM simulation based on patient specific data has shown that stress occurs mostly in the pubococcygeus and puborectalis. The associated morphological variation corresponds well with existing clinical findings as well as current work on anatomical modelling of the levator ani. We have shown quantitatively that the support to the rectum is affected most and to our knowledge, this is the first subject-specific FEM study of the levator ani with data from dynamic MR images under natural physiological stress conditions. It is worth noting that the assessment of the accuracy of the current method is limited to evaluating the internal consistency and convergence behaviour of the model in addition to its sensitivity to potential morphological errors. In practice, *in vivo* quantitative validation of the model is difficult without the use of invasive means. Nevertheless, it is possible to indirectly gauge the strength of the levator ani by using vagina perineometer pressure measurements with a balloon catheter connected to a pressure gauge to determine the amount of squeeze by the pelvic floor muscles. The method, however, can be too intrusive to nulliparous subjects. In conclusion, the use of subject-specific FEM allows an improved understanding of the dynamics of a levator ani. Our results have shown consistent behaviour of the pelvic floor in normal subjects during both contraction and strain. This provides information that would not normally be available from traditional imaging techniques, thus permitting its practical use in real clinical settings.

Acknowledgements. The authors would like to thank W. Caspersz, C. Geddes, W. Gedroyc and the rest of the team at Interventional Magnetic Resonance in St Mary's Hospital, London, for their assistance in MR image acquisition. We also thank Dr Xiao Yun Xu for her help with finite element analysis.

References

1. Healy, J.C., Halligan, S., Rezek, R.H., Watson, S., Phillips, R.K.S., Armstrong, P.: Patterns of Prolapse in Women with Symptoms of Pelvic Floor Weakness: Assessment with MR Imaging. *Radiology*. Vol 203(1). (1997) 77-81.
2. Fielding, J.R., Hoyte, L., Schierlitz, L.: MR Imaging of Pelvic Floor Relaxation. *Journal of Women's Imaging*. Vol 2. (2000) 82-87.
3. Hoyte, L., Schierlitz, L., Zou, K., Flesh, G., Fielding, J.R.: Two- and 3-dimensional MRI comparison of levator ani structure, volume, and integrity in women with stress incontinence and prolapse. *American Journal of Obstetrics and Gynecology*. Vol 185(1). (2001) 11-19.
4. Hoyte, L., Jakab, M., Warfield, S.K., Shott, S., Flesh, G., Fielding, J.R.: Levator ani thickness variations in symptomatic and asymptomatic women using magnetic resonance-based 3-dimensional color mapping. *American Journal of Obstetrics and Gynecology*. Vol 191(3). (2004) 856-61.
5. Lee, S.-L., Horkaew, P., Darzi, A., Yang, G.-Z.: Optimal Scan Planning with Statistical Shape Modelling of the Levator Ani. *MICCAI*, Montreal, Canada (2003). *Lecture Notes in Computer Science*, Springer, Vol 2878. 714-721.

6. Lee, S.-L., Horkaew, P., Darzi, A., Yang, G.-Z.: Statistical Shape Modelling of the Levator Ani with Thickness Variation. MICCAI, St Malo, France (2004). Lecture Notes in Computer Science, Springer, Vol 3216. 258-265.
7. Pathmanathan, P., Gavaghan, D., Whiteley, J., Brady, M., Nash, M., Nielsen, P., Rajagopal, V.: Predicting Tumour Location by Simulating Large Deformations of the Breast Using a 3D Finite Element Model and Nonlinear Elasticity. MICCAI, St Malo, France (2004). Lecture Notes in Computer Science, Springer, Vol 3217. 217-224.
8. d'Aulignac, D., Martins, J., Pires, E.: Physical Modeling of the Pelvic Floor Muscles using Shell Elements. European Congress on Computational Methods in Applied Sciences and Engineering, Jyvaskyla (2004).
9. Janda, S., van der Helm, F.C.T., de Blok, S.B.: Measuring morphological parameters of the pelvic floor for finite element modelling purposes. Journal of Biomechanics. Vol 36 (2003) 749-757.
10. Janda, S., van der Helm, F.C.T., de Blok, S.B.: Measuring material parameters of the pelvic floor for FE modelling purposes. Proceedings Dutch Annual Conference on BioMedical Engineering, Papendal, NL (2002).
11. Janda, S., van der Helm, F.C.T., de Blok, S.B.: Biomechanical analysis of the pelvic floor muscles: Constitutive and FE modelling. Dutch Annual Conference on BioMedical Engineering, Papendal, NL (2003).
12. Gerard, J., Wilhelms-Tricarico, R., Perrier, P., Payan, Y.: A 3D dynamical biomechanical tongue model to study speech motor control. Recent Research Developments in Biomechanics. (2003) 49-64.
13. Chabanas, M., Payan, Y., Marecaux, C., Swider, P., Boutault, F.: Comparison of linear and non-linear soft tissue models with post-operative CT scan in maxillofacial surgery. International Symposium on Medical Simulation, Cambridge, Massachusetts (2004). Vol 3078. 19-27.
14. Horkaew, P., Yang, G.-Z.: Optimal Deformable Surface Models for 3D Medical Image Analysis. IPMI, Ambleside, UK (2003). Lecture Notes in Computer Science, Springer, Vol 2732. 13-24.

Robust Visualization of the Dental Occlusion by a Double Scan Procedure^{*}

Filip Schutyser¹, Gwen Swennen², and Paul Suetens¹

¹ Department of Medical Image Computing (Radiology - ESAT/PSI),
Faculties of Medicine and Engineering, University Hospital Gasthuisberg,
Katholieke Universiteit Leuven, BE
`filip.schutyser@uz.kuleuven.ac.be`

² Department of Plastic Surgery, University Hospital Brugmann,
Université Libre de Bruxelles, BE

Abstract. A detailed visualization of the dental occlusion in 3D image-based planning environments for oral and maxillofacial planning is important. With CT imaging however, this occlusion is often deteriorated by streak artifacts caused by amalgam fillings. Moreover, more detailed surface information at the level of the dental cuspids is often desired.

In this paper, a double scan technique is introduced to image the dental occlusion by means of a newly designed 3D splint. The patient wears this splint between the upper and lower teeth during CT-scan. In a second step, the splint is positioned between the plaster casts of the upper and lower jaw, and this setup is scanned. Based on markers in the 3D splint, both data sets are fused and a combined visualization is possible. The accuracy, robustness and applicability in clinical routine is shown.

This technology enables meticulous 3D cephalometric analysis, detailed maxillofacial planning and opens possibilities towards intraoperative support.

1 Introduction

In 3D image-based planning environments for oral and maxillofacial surgery [1], it is important to inspect the dental occlusion carefully. For total and partial edentulous patients, it is important to co-visualize the denture with the bone. In this way, oral implant planning is not only driven by bone related issues, but also by prosthetics. This improves the esthetic result. For patients with a complete dentition facing orthognatic surgery, the surgical planning implies optimizing the occlusion. Therefore, an accurate visualization of the occlusion is important.

To visualize bony structures of the head, CT (computerized tomography) is the preferred modality. However, some important drawbacks are identified: amalgam fillings generate important streak artifacts, and the acrylic resin that is used to make dentures, shows the same Hounsfield units as soft tissues.

^{*} This work partly belongs to the Flemish government IWT GBOU 020195 project on Realistic image-based facial modelling for forensic reconstruction and surgery simulation and the IWT-040310 project on fusion of medical CT image volumes with digitizations of synthetic models.

In the field of oral implant planning, we introduced the double scan procedure [2]. With this CT protocol, the patient wearing a denture equipped with gutta percha markers is CT-scanned in a first step. Then, the prosthesis alone is CT-scanned. From this scan, the prosthesis is robustly segmented. Based on the markers, visible in both scans, the prosthesis is rigidly registered towards the scan of the patient.

Nkenke et al. [3] add a digitization of plaster models, acquired by optical surface scanning or CT imaging, to the patient's CT image volume by registration of surfaces of teeth. They show that the accuracy of this method strongly reduces with the presence of artifacts.

Gateno et al. [4] explain an in vitro test of the fusion of a laser scanned dental impression with a CT scan of a patient wearing this impression. Some extra characteristic shapes are attached to the impression. Based on these shapes, the registration is performed. However, as such, this technique is not applicable in clinical routine. No patient examples are shown. Moreover, special equipment is needed for the laser scan.

In this paper, we show that the double scan procedure can be extended to robustly image the dental occlusion of patients with a full dentition with a high detail. The registration was performed fully automatically. The method has been validated on 10 cadaver skulls, and 7 patients.

2 Method

2.1 3D Splint

During orthodontic treatment, and for diagnostic work-up of orthognatic surgery, plaster casts are routinely made. These plaster casts are an accurate copy of the actual dentition of the patient. To record the relative position of the plaster casts of the dental arches, a splint is produced.

During CT-scan, the patient gently bites on this splint. In a second scan, the plaster models with that splint are CT-scanned. If a rigid registration of the splint in both scans is possible, the CT-scan of the plaster models is fused with the patient scan.

For this purpose, the splint is modified. Spherical gutta percha markers are added to the splint. This causes bright spots on the CT images of the patient and the plaster models. This will be the basis for the matching procedure. There is a trade-off for the size of the markers. To make them easily detectable, rather large balls are preferred. However, larger balls generate also CT artifacts. A diameter of 1.5 mm is found to be optimal after extensive tests.

Since the splint has an almost planar structure, an extension perpendicular to the occlusal plane is added to improve the registration accuracy. To have a comfortable solution for the patient, this extension is mounted on a bar leaving the mouth. In this way, the natural shape of the face is not disturbed.

This 3D splint (see figure 1) is the key to obtain an accurate visualization of the occlusion. This 3D splint is produced with a hard, radiolucent acrylic resin which is not toxic. It can be easily produced and it is cheap.



Fig. 1. The 3D splint is equipped with gutta percha markers and has an extraoral extension. The patient wears this splint during CT-scan. The splint is also scanned together with the plaster models.

2.2 Registration

Gutta percha is very radio-opaque. Therefore, candidate markers are extracted by thresholding. To classify these candidates into markers and other structures, two measures are computed:

1. The Normalised Shape Index (NSI), defined in equation 1, expresses the resemblance of an object to sphere. k is defined so that $NSI = 1$ for a sphere. NSI must be smaller than a threshold value to be a marker.

$$NSI = k \frac{\text{volume}}{\text{surface}} \quad (1)$$

2. The volume of the object needs to be larger than a threshold value to exclude small structures due to noise or artifacts.

To find the rigid registration matrix based on the centers of the markers in both image volumes, the correspondence between both sets is needed. Given the sparse set of points, algorithms like Iterative Closest Points (ICP) [5] often fail. Since we aim for a fully automatic procedure, a dedicated approach is needed.

The Euclidean distance between two markers of set 1, equals the distance between two markers of set 2. This holds for all distances between possible marker pairs. If markers a and b correspond with markers a' and b' , then is marker c' a corresponding marker to c if $||ac|| - ||a'c'|| < \text{threshold}$, and $||bc|| - ||b'c'|| < \text{threshold}$. Since an initial set of corresponding points is not known, the algorithm is applied for all candidate correspondences, and the solution with the largest set of corresponding markers is kept.

The translation and rotation matrix are finally computed using a least squares approach [6].

3 Results

This method was tested and validated on a set of 10 cadaver skulls. For each skull, a 3D splint and plaster models were produced. The skulls and plaster models were CT-scanned (GE LightSpeed QX/i, 120 kV, 100 mAs, spiral CT). In order to reduce the number of axial slices affected by streak artifacts, the occlusal plane was positioned perpendicular on the CT-table. For the models, the smallest available slice thickness was selected and the axial slices were reconstructed with a bone kernel and voxel size of (0.3, 0.3, 0.3) mm. The patients were scanned with a lower resolution, to reduce the radiation dose, and the axial slices were reconstructed with a bone kernel and voxel size of (0.35, 0.35, 1.0) mm.

The registration results are listed in table 1. A mean error of 0.14 mm is a fairly acceptable accuracy for the application of 3D image-based planning given

Table 1. This table lists the registration results for 10 cadaver skulls

Skull	Nr markers	Mean error (mm)	Standard deviation (mm)	Max. error (mm)	Min. error (mm)
1	10	0.13	0.05	0.19	0.08
2	9	0.14	0.04	0.17	0.09
3	10	0.14	0.07	0.28	0.03
4	9	0.17	0.07	0.25	0.08
5	9	0.12	0.03	0.16	0.06
6	11	0.18	0.06	0.30	0.10
7	8	0.11	0.04	0.20	0.07
8	11	0.08	0.03	0.13	0.03
9	8	0.12	0.04	0.21	0.08
10	12	0.18	0.08	0.35	0.07
Overall mean error			0.14 mm		
Overall standard deviation of error			0.03 mm		

Table 2. This table lists the registration results for 7 patients

Skull	Nr markers	Mean error (mm)	Standard deviation (mm)	Max error (mm)	Min error (mm)
1	5	0.13	0.04	0.19	0.09
2	6	0.17	0.06	0.28	0.09
3	4	0.15	0.07	0.24	0.10
4	4	0.15	0.04	0.20	0.11
5	8	0.16	0.05	0.24	0.10
6	9	0.21	0.10	0.41	0.09
7	14	0.15	0.07	0.34	0.04
Overall mean error			0.16 mm		
Overall standard deviation of error			0.03 mm		

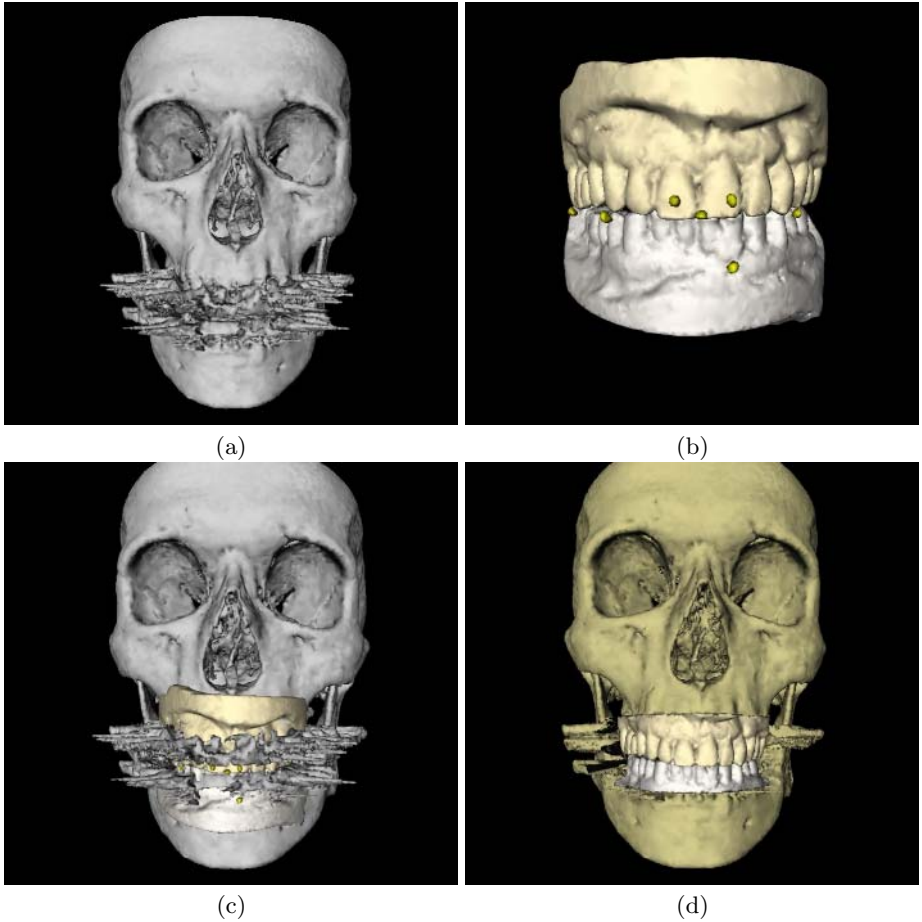


Fig. 2. This figure shows how the registration with the 3D splint enables detailed imaging of the occlusion. Streak artifacts corrupt CT images (a). Plaster models can be accurately imaged with CT-imaging (b). Based on the 3D-splint, both data sets can be fused (c) and the surface description extracted from the plaster model scan replaces the corrupted data (d).

the resolution of the CT data sets. Therefore, this technique was applied on patients with amalgam fillings or orthodontic brackets with steel wires.

Seven patients with severe amalgam fillings were selected for this test. All of them received a 3D-splint, and the double-scan technique is applied. An excessive amount of gutta percha markers is inserted in the splint because some of them will be submerged in the artifacts. All registrations are performed fully automatically with a mean registration error of 0.16 mm as shown in table 2.

The registered surfaces are co-visualized and the teeth surface from the plaster models replaces the teeth surface from the patient scan (see figure 2). In this way, a very detailed dental occlusion is visualized.

4 Discussion

The double scan procedure is a protocol that enables robust visualization of the dental occlusion. First the patient is scanned with an adapted material at the level of the occlusion. Then this adapted material is scanned alone.

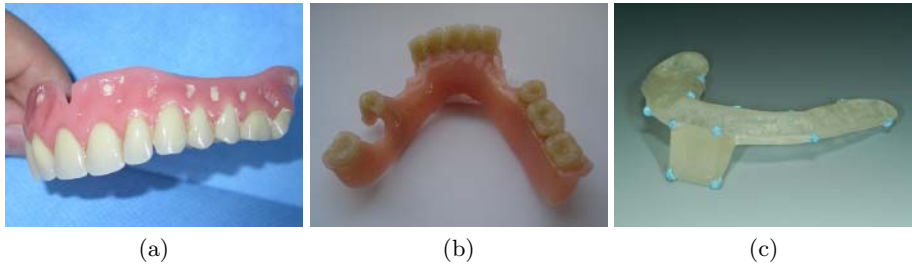


Fig. 3. This figure illustrates the evolution of a denture for a complete edentulous patient (a), over a partial denture (b) towards the 3D splint (c). Based on these materials, an accurate visualization of the dental occlusion becomes possible.

When the patient is edentulous, the prosthesis equipped with gutta percha markers is used to visualize the occlusion. For partial edentulous patients, a partial prosthesis, equipped with gutta percha markers and, if needed, support on the neighboring teeth is applied. A bite index is applied to secure the correct occlusion during CT-scan. For dentate patients, the “prosthesis” reduces to the 3D splint. This is illustrated in figure 3.

Some salient features of this approach can be enumerated:

1. The second scan can be performed with a very high resolution since the patient is not involved and thus radiation dose is not relevant.
2. Since the patient gently bites on the 3D splint, the referring clinician knows that the occlusion during CT-scan is correct. In current clinical practice, this is not guaranteed because the patient is laying on the CT table in stead of sitting on a chair and the patient might be not at ease for the CT exam.
3. Although the patient scan might have excessive streak artifacts, the second scan lacks artifacts. By this method, the clinical hindrance caused by those artifacts is reduced.
4. The markers are robustly detected and the scans are accurately registered as shown in the results.
5. This double scan technique is easy to implement in daily clinical routine.
6. It is a cheap procedure. The cost of inserting gutta percha markers and making the 3D splint is almost negligible.

This technology has several applications in the field of orthodontics, maxillofacial surgery and plastic surgery. Detailed visualization of the dental occlusion allows accurate 3D cephalometric analysis including landmarks located

on the dental occlusion. Surgical optimization of the dental occlusion can be meticulously planned, and intraoperative surgical support to obtain the planned solution can be derived.

5 Conclusion

The double scan procedure is a robust and accurate method to visualize the dental occlusion. Since the information needed to visualize the occlusion is acquired in a second scan without the patient, a highly detailed artifact free scan can be obtained. Through the use of a large number of gutta percha points, the registration is shown to be accurate and fully automatic.

This technology is useful for detailed 3D cephalometric analysis, accurate maxillofacial surgery planning and intra-operative support.

References

1. Schutyser, F., Van Cleynenbreugel, J., Ferrant, M., Schoenaers, J., Suetens, P.: Image-based 3d planning of maxillofacial distraction procedures including soft tissue implications. (2000) 999–1007
2. Verstrecken, K., Van Cleynenbreugel, J., Martens, K., Marchal, G., van Steenberghe, D., Suetens, P.: An image-guided planning system for endosseous oral implants. *IEEE Transactions on Medical Imaging* **17** (1998) 842–852
3. Nkenke, E., Zachow, S., Benz, M., Maier, T., Veit, K., Kramer, M., Benz, S., Häusler, G., Neukam, F.W., Lell, M.: Fusion of computed tomography data and optical 3d images of the dentition for streak artefact correction in the simulation of orthognathic surgery. *Dentomaxillofacial Radiology* **33** (2004) 226–232
4. Gateno, J., Xia, J., Teichgraber, J.F., Rosen, A.: A new technique for the creation of a computerized composite skull model. *Journal of Oral and Maxillofacial Surgery* **61** (2003) 222–227
5. Besl, P., McKay, N.: A method for registration of 3-d shapes. *IEEE Transaction on Pattern Analysis and Machine Intelligence* **14** (1992) 239–256
6. Arun, K., Huang, T., Blostein, S.: Least-squares fitting of two 3-d point sets. *IEEE Transaction on Pattern Analysis and Machine Intelligence* **9** (1987) 698–700

Segmentation of Focal Cortical Dysplasia Lesions Using a Feature-Based Level Set

O. Colliot, T. Mansi, N. Bernasconi, V. Naessens,
D. Klironomos, and A. Bernasconi

Montreal Neurological Institute, McGill University, Montreal, Canada

Abstract. Focal cortical dysplasia (FCD), a malformation of cortical development, is an important cause of medically intractable epilepsy. FCD lesions are difficult to distinguish from non-lesional cortex and their delineation on MRI is a challenging task. This paper presents a method to segment FCD lesions on T1-weighted MRI, based on a 3D deformable model, implemented using the level set framework. The deformable model is driven by three MRI features: cortical thickness, relative intensity and gradient. These features correspond to the visual characteristics of FCD and allow to differentiate lesions from normal tissues. The proposed method was tested on 18 patients with FCD and its performance was quantitatively evaluated by comparison with the manual tracings of two trained raters. The validation showed that the similarity between the level set segmentation and the manual labels is similar to the agreement between the two human raters. This new approach may become a useful tool for the presurgical evaluation of patients with intractable epilepsy.

1 Introduction

Malformations of cortical development (MCD) have been increasingly recognized as an important cause of medically intractable focal epilepsy. Focal cortical dysplasia (FCD) [1], a malformation due to abnormal neuroglial proliferation, is the most frequent MCD in patients with intractable extra-temporal epilepsy [2]. Epilepsy surgery, consisting in the removal of the FCD lesion, is an effective treatment for these patients. However, freedom from seizures after surgery is closely related to the resection of the whole lesion [3]. The precise delineation of lesions is thus important for surgical planning in epilepsy.

Magnetic resonance imaging (MRI) plays a pivotal role in the presurgical evaluation of patients with intractable epilepsy. Although MRI has allowed the recognition of FCD in an increased number of patients, standard radiological evaluation fails to identify lesions in a large number of cases due to their subtlety and the complexity of the cortex convolution [3]. Moreover, the spatial extension of the lesions is difficult to define on the MRI. The segmentation of FCD is thus a challenging image analysis application as the lesions are often subtle, difficult to differentiate from the normal cortex, of variable size, position and shape, and with ill-defined boundaries.

Recently, image analysis techniques have been developed to detect FCD lesions automatically on MRI, relying on different types of voxel-wise analysis [4, 5]. In particular, computational models of FCD characteristics [6] and a Bayesian classifier for lesion detection [4] were previously proposed by our group. While these approaches successfully identify the FCD in a majority of patients, they provide a very limited coverage of the lesion (about 20%) and thus cannot be considered as segmentation techniques. Therefore, to our knowledge, the question of FCD segmentation has never been addressed.

This paper presents a method for segmenting focal cortical dysplasia (FCD) lesions on T1-weighted MRI, based on a level set deformable model driven by MR features of these lesions. The method partly relies on our aforementioned detection approaches [6, 4]. However, our target application is FCD segmentation and not detection. The computational models of FCD features are used to drive a level set deformable model and the FCD classifier is used only to obtain a starting point for the segmentation procedure.

2 Methods

Our approach relies on a 3D deformable model, based on the level set method. Driving the deformable model with image gray levels would be inadequate as this attribute is insufficient to distinguish the lesion from the normal cortex. Instead, we propose a model guided by a probability map derived from FCD features. These features correspond to the visual characteristics of FCD: cortical thickening, a blurred transition between gray matter (GM) and white matter (WM), and hyperintense signal within the dysplastic lesion [7].

Additionally, it is necessary to provide a starting point for the level set evolution. To this purpose, we used our previously developed FCD classifier [4], under supervision of an expert user.

2.1 Probabilistic Modeling of FCD Features

To quantitatively evaluate the visual MR characteristics of FCD, we made use of our previous computational models (more details can be found in [6]). A cortical thickness map, denoted as Th , is computed by solving Laplace's equation over the cortical ribbon. Hyperintense signal is represented using a relative intensity index defined as $RI(x) = 1 - |B_g - I(x)|/B_g$ where $I(x)$ is the intensity at voxel x and B_g is the boundary intensity between GM and WM. Blurring of the GM/WM transition is modeled with a gradient magnitude map, denoted as Gr . These three characteristics define a vector-valued feature map $f(x) = (Th(x), RI(x), Gr(x))$ at each point x in the image space.

We then performed a supervised learning to estimate the probability of different tissue classes in the brain given the feature vector f . Four different classes, denoted as c , were considered: gray matter (GM), white matter (WM), cerebrospinal fluid (CSF) and the FCD lesion (L). Normal tissues were segmented using a histogram-based approach with automated threshold, while the FCD lesions were painted by trained observers (see Section 3). Conditional probabilities

$P(f(x)|c)$ for each class c were modeled using a trivariate normal distribution and estimated using the maximum likelihood on a learning set of patients. The posterior probabilities $P(c|f(x))$ were then obtained by Bayes' rule. As the size of FCD lesions is variable, we assumed equal prior probabilities for the different classes.

Figure 1 presents an example of the three feature maps and of the posterior probability maps in a patient with FCD.

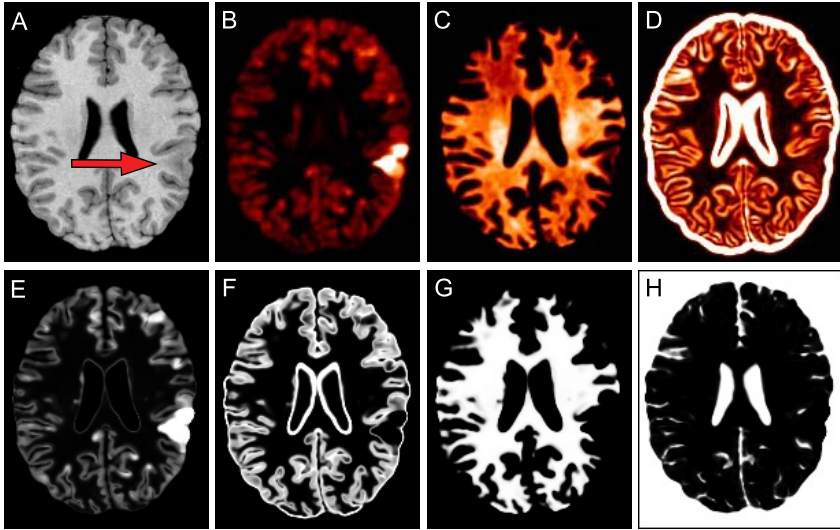


Fig. 1. Probabilistic modeling of FCD features. Upper panels: T1-weighted MRI where the FCD lesion is indicated by the arrow (A), cortical thickness map (B), relative intensity map (C), gradient map (D). The lesion is characterized by higher cortical thickness, higher relative intensity and lower gradient. Lower panels: probability maps of the lesion class (E), GM (F), WM (G) and CSF (H).

2.2 Feature-Based Deformable Model

Based on the previous features, the deformable model was designed to separate the lesion from the non-lesional regions. The region competition approach proposed by Zhu and Yuille [8] is well adapted to our purpose. It aims at segmenting an image into several regions by moving the interfaces between them. The evolution of the interfaces is driven by functions indicating the membership to each region. In our case, these functions can be derived from the FCD features.

We intended to isolate the FCD lesion from the non-lesional region, which is composed of three different classes (GM, WM, CSF). However, the boundaries between these three non-lesional classes were of no interest for our application. Thus, region competition occurred in each point between the lesion class and the most probable non-lesional class. The membership to the lesional region

was defined as $R_L(x) = P(L|f(x))$ which is the previously computed posterior probability of the lesion class. The non-lesional region was modeled by $R_{NL}(x) = \max\{P(GM|f(x)), P(WM|f(x)), P(CSF|f(x))\}$.

The feature-based deformable model describes the evolution of the interface (or surface in 3D) S of the lesional region, according to those membership functions and a regularization term. The motion of a point u belonging to S is defined as:

$$\frac{\partial u}{\partial t} = \alpha[R_{NL}(u) - R_L(u)]\mathbf{n}_u + \epsilon\kappa_u\mathbf{n}_u \quad (1)$$

where \mathbf{n}_u is the inward normal to S at point u (directed towards the interior of the lesion), κ_u is the mean curvature and α and ϵ are weighting coefficients.

In the previous equation, $\alpha[R_{NL}(u) - R_L(u)]$ is a feature-based term and $\epsilon\kappa_u$ is a regularity term producing a smooth surface. If $R_L(u) > R_{NL}(u)$, meaning that the most probable class for point u is the lesion, the surface S will be expanded, in order to include this point. On the contrary, if $R_{NL}(u) > R_L(u)$, meaning that this point should belong to one of the three non-lesional classes, the surface will be shrunk.

2.3 Level Set Evolution

The motion equation obtained for the feature-based deformable model was implemented using the level set method [9, 10]. The principle of this method is to define the surface S as the zero level set of a higher dimensional function ϕ , called the implicit function:

$$\phi(S(t), t) = 0 \quad (2)$$

As an implicit function ϕ , we chose the classical signed distance to the surface S , with negative values in the interior of S . The evolution was then performed on the function ϕ and the embedded surface S was deformed implicitly. Level set deformable models present several advantages over traditional ones: no parameterization of the surface is necessary, topology changes are handled naturally and the result is less sensitive to the initialization.

Using the derivation from curve motion to level set evolution [10], the feature-based deformable model can be described by:

$$\frac{\partial \phi}{\partial t}(u) = \alpha[R_{NL}(u) - R_L(u)]|\nabla\phi(u)| + \epsilon\kappa_u|\nabla\phi(u)| \quad (3)$$

The previous evolution equation can be seen as a particular case of the one proposed in [11].

This equation was implemented using the numerical scheme proposed in [10–chap.6]. To reduce the computational complexity, we made use of the narrow-band method [12]. A sub-voxel reinitialization of the implicit function at fixed time steps was performed to maintain the distance function [13].

3 Experiments and Results

3.1 Subjects and Image Preparation

We selected 24 patients (13 males, mean age \pm SD= 24 ± 8) with MRI-visible FCD. The Ethics Board of our Institution approved the study, and written informed consent was obtained from all participants.

3D MR images were acquired on a 1.5T scanner using a T1-fast field echo sequence ($TR = 18$, $TE = 10$, 1 acquisition average pulse sequence, flip angle= 30° , matrix size= 256×256 , FOV= 256 , thickness= $1mm$) with an isotropic voxel size of $1mm^3$. All images underwent automated correction for intensity non-uniformity and intensity standardization [14], automatic registration into stereotaxic space [15] and brain extraction [16]. Classification of brain tissue in GM, WM and CSF was done using an histogram-based method with automated threshold [6].

3.2 Manual Segmentation

Lesions were delineated independently on 3D MRI by two trained raters (VN and DK) using a software which allows painting in each of the three directions of the space. The corresponding manually labeled datasets are further denoted as M_1 and M_2 . Interrater agreement was assessed using the similarity index $S = 2 \frac{n\{M_1 \cap M_2\}}{n\{M_1\} + n\{M_2\}}$ (where $n\{M\}$ is the number of elements in set M), which is a special case of kappa statistic since the vast majority of voxels are non-lesional [17].

3.3 Results

Initialization. The FCD classifier [4] is used to initialize the deformable model. It successfully identified the lesion in 18 (18/24=75%) patients. We assessed the possibility of segmenting the six undetected lesions with a manual initialization of the procedure. However, the segmentation failed in these cases because their features were not sufficiently discriminant. The evaluation was thus done on the 18 detected lesions.

Manual Segmentation. For the 18 manual labels, the mean interrater similarity index was 0.62 ± 0.19 (range=0.22 to 0.84).

Level Set Segmentation. We compared the automated segmentations to the sets of manual labels using the similarity index S presented above. The evaluation was performed using a leave-one-out approach: for the segmentation of a given patient, this patient was excluded from the learning set (Section 2.1). This approach avoids the introduction of bias in the result. All results were obtained with $\alpha = 0.8$ and $\epsilon = 0.2$ in Equation 3. Moreover, we computed the similarity obtained with the FCD classifier [4] to evaluate the added value of the level set. Results are reported in Table 1. Figures 2 and 3 present the segmentations obtained in two patients with FCD.

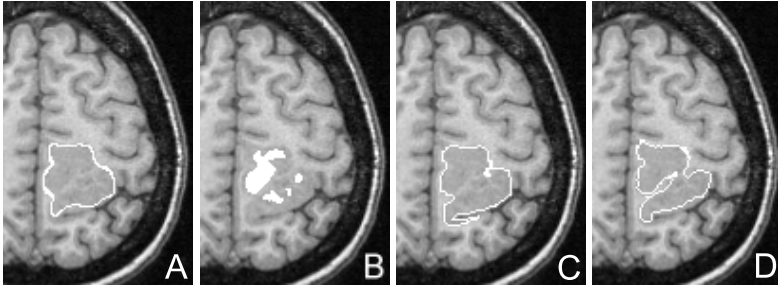


Fig. 2. Results of FCD segmentation: level set segmentation (A), initialization (B), manual tracing M_2 (C), manual tracing M_1 (D)

Table 1. The table presents the similarity indices for the level set and the FCD classifier with respect to the two manual tracings, as well as the interrater similarity. Results are reported as mean±SD with the range in parentheses.

	M_1	M_2
Level set	0.62 ± 0.16 (0.32 to 0.84)	0.63 ± 0.12 (0.43 to 0.79)
Classifier	0.30 ± 0.17 (0.11 to 0.64)	0.31 ± 0.17 (0.07 to 0.59)
Interrater (M_1 vs. M_2)		0.62 ± 0.19 (0.22 to 0.84)

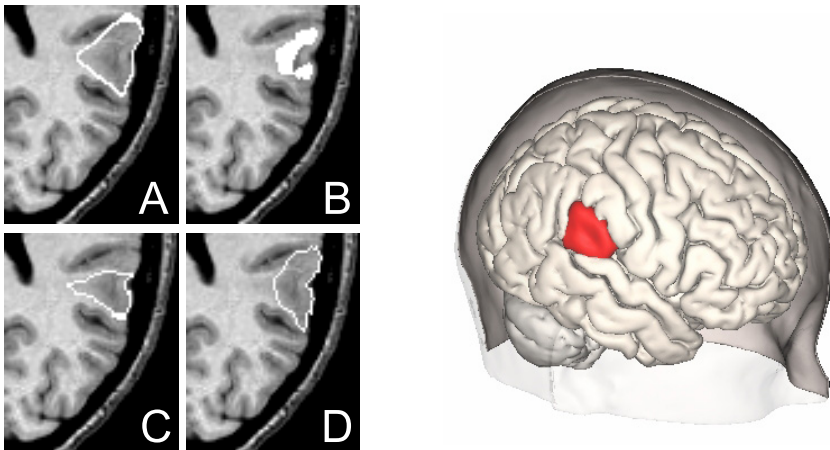


Fig. 3. Results of FCD segmentation. Left panels: level set segmentation (A), initialization (B), manual tracing M_2 (C), manual tracing M_1 (D). Right panel: 3D rendering of the FCD lesion segmentation together with the cortical surface.

To assess the robustness of our method with respect to the choice of the learning dataset, the procedure was also evaluated using a learning on the labels M_2 (in Table 1 the results are obtained using the labels M_1 as a learning set). Similarity indices for the level set (LS) were: LS vs. $M_1=0.62 \pm 0.16$ (range=0.33 to 0.82) and LS vs. $M_2=0.62 \pm 0.11$ (range=0.43 to 0.77).

4 Discussion

In this study, we proposed and evaluated a method for segmenting FCD lesions on MRI. We introduced a feature-based level set, driven by known MR characteristics of FCD. Probability maps of these features corresponding to FCD and normal tissue classes were estimated in order to guide the level set evolution.

On MRI, FCD lesions possess ill-defined contours and are not easily differentiable from normal cortex. Moreover, there is no available gold standard for evaluating the delineation of these lesions. For this reason, we compared the level set segmentation to the manual tracings of two trained observers. The interrater similarity was 0.62 which corresponds to a substantial agreement, in particular when keeping in mind the difficulty of FCD segmentation. The level set segmentations achieved a degree of similarity of 0.63 and 0.62 with the two sets of manual labels, which again constitutes a good agreement.

The similarities achieved by the level set are also very close to the interrater agreement (both were computed on the 18 detected lesions). A significant portion of the remaining differences between automated and manual labels is probably due to the interrater variability rather than to the inability of the level set to recover the full extension of lesions. This can be seen in Figure 3 where the two raters decided to exclude different parts of the lesion (Panels C and D) while these parts were included in the automated segmentation (Panel A).

To our knowledge, there is no other published work on FCD segmentation that could be used for comparison to our results. Nevertheless, compared to the FCD classifier, our method achieved a similarity twice as large and therefore constitutes a significant improvement. However, it should be noted that this classifier was designed for FCD detection and not segmentation.

The results of the automated segmentation did not depend on the manual tracings (by one rater or the other) used for the learning step. This shows the robustness of our method with respect to the learning dataset. Moreover, no fine-tuning of the level set parameters was necessary and the same parameters were used for all patients. Furthermore, the method is fast thanks to the use of the narrow-band and the subvoxel reinitialization approaches (the mean computation time for the level set evolution was 4 minutes on a PC 1.6GHz).

In conclusion, this paper demonstrates the effectiveness of a feature-based level set approach for the segmentation of FCD lesions. We do not advocate that the level set segmentation should be used in place of manual delineation but rather that it is a complementary tool. It has the potential to reduce user subjectivity and, more importantly, to unveil lesional areas that could be overlooked by visual inspection. This new method may become a useful tool for sur-

gical planning in epilepsy. Future work includes further validation using other metrics, comparison with other image segmentation techniques and a detailed study of cases that were not detected by the classifier.

References

1. Taylor, D., Falconer, M., Bruton, C., Corsellis, J.: Focal dysplasia of the cerebral cortex in epilepsy. *J Neurol Neurosurg Psychiatry* **34** (1971) 369–387
2. Sisodiya, S.: Surgery for malformations of cortical development causing epilepsy. *Brain* **123** (2000) 1075–1091
3. Tassi, L., Colombo, N., et al.: Focal cortical dysplasia: neuropathological subtypes, EEG, neuroimaging and surgical outcome. *Brain* **125** (2002) 1719–1732
4. Antel, S., Collins, D., et al.: Automated detection of focal cortical dysplasia lesions using computational models of their MRI characteristics and texture analysis. *NeuroImage* **19** (2003) 1748–59
5. Wilke, M., Kassubek, J., Ziyeh, S., Schulze-Bonhage, A., Huppertz, H.: Automated detection of gray matter malformations using optimized voxel-based morphometry: a systematic approach. *NeuroImage* **20** (2003) 330–343
6. Antel, S., Bernasconi, A., et al.: Computational models of MRI characteristics of focal cortical dysplasia improve lesion detection. *NeuroImage* **17** (2002) 1755–60
7. Barkovich, A., Kuzniecky, R.: Neuroimaging of focal malformations of cortical development. *J Clin Neurophysiol* **13** (1996) 481–494
8. Zhu, S., Yuille, A.: Region competition: Unifying snakes, region growing, and bayes/MDL for multiband image segmentation. *IEEE TPAMI* **18** (1996) 884–900
9. Osher, S., Sethian, J.: Fronts propagating with curvature-dependent speed : algorithms based on Hamilton-Jacobi formulation. *J Comp Phys* **79** (1988) 12–49
10. Sethian, J.: *Level-set methods and fast marching methods*. 2nd edn. Cambridge University Press (1999)
11. Paragios, N., Deriche, R.: Geodesic active regions and level set methods for supervised texture segmentation. *Int J Comp Vis* **46** (2002) 223–247
12. Adalsteinsson, D., Sethian, J.: A fast level set method for propagating interfaces. *J Comp Phys* (1995) 269–277
13. Krissian, K., Westin, C.F.: Fast sub-voxel re-initialization of the distance map for level set methods. *Patt Recog Letters* (2005) In Press.
14. Sled, J., Zijdenbos, A., Evans, A.: A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE TMI* **17** (1998) 87–97
15. Collins, D., et al.: Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J Comput Assist Tomogr* **18** (1994) 192–205
16. Smith, S.: Fast robust automated brain extraction. *Hum Brain Mapp* **17** (2002) 143–155
17. Zijdenbos, A., Dawant, B., et al.: Morphometric analysis of white matter lesions in MR images: method and validation. *IEEE TMI* **13** (1994) 716–724

Effects of Healthy Aging Measured By Intracranial Compartment Volumes Using a Designed MR Brain Database

B nedicte Mortamet^{1,5}, Donglin Zeng², Guido Gerig^{3,4}, Marcel Prastawa³,
and Elizabeth Bullitt¹

¹ Department of Surgery, ² Department of Biostatistics, School of Public Health,

³ Department of Computer Science, ⁴ Department of Psychiatry,

University of North Carolina, Chapel Hill, USA

⁵ ESCPE Lyon, France

benedicte.mortamet@cpe.fr, dzeng@bios.unc.edu
{gerig, prastawa}@cs.unc.edu, bullitt@med.unc.edu

Abstract. A publicly available database of high-quality, multi-modal MR brain images of carefully screened healthy subjects, equally divided by sex, and with an equal number of subjects per age decade, would be of high value to investigators interested in the statistical study of disease. This report describes initial use of an accumulating healthy database currently comprising 50 subjects aged 20-72. We examine changes by age and sex to the volumes of gray matter, white matter and cerebrospinal fluid for subjects within the database. We conclude that traditional views of healthy aging should be revised. Significant atrophy does not appear in healthy subjects 60 or 70 years old. Gray matter loss is not restricted to senility, but begins in early adulthood and is progressive. The percentage of white matter increases with age. A carefully-designed healthy database should be useful in the statistical analysis of many age- and non-age-related diseases.

1 Introduction

Any automated, computerized assessment of disease requires establishment of healthy norms against which a test subject can be compared. However, healthy aging induces brain changes [1], [2]. In young, healthy adults [3] and in nondemented older persons [4], analyses demonstrate that the aging process is a continuum with changes evident before senescence. An accurate assessment of disease by magnetic resonance imaging (MR) thus requires an image database of healthy subjects appropriately age and sex matched to the potentially diseased subject in question.

The public establishment of such a healthy database would be of value to the scientific community. We are collecting high-resolution, 3T MR images of 100 healthy subjects aged 20-60+, with 20 subjects per decade divided equally by sex. All subjects are screened for the presence of brain disease. Images include T1, T2, MRA, and diffusion tensor. Images will be made publicly available once collection is complete.

The current report describes the first use of this database at the 50% completion point (50 subjects equally divided by sex, with 5 men and 5 women included in each of five age decades). This first report does not analyze disease, but rather examines differences by age and sex within the healthy database itself. More specifically, we describe automated volumetric analysis of the three brain compartments comprising gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF). We also examine lateral ventricular volume as a percentage of total CSF volume.

Several papers have evaluated changes in ventricular volume with healthy aging [1], [5], [6], [7], [8], changes in GM-WM-CSF volumes with healthy aging [3], [4], [9], [10], [11]. These papers, however, examine different age group ranges, include variable male-female sex ratios, have not always screened healthy subjects carefully and have almost always employed low resolution MR images with large interslice spacing.

This study is based upon the assumption that careful design of a healthy subject database can provide better assessment of differences between patient populations. Consistent with previous studies, our results show significant differences between the intracranial volumes of males and females and a significant loss of GM with normal aging. However, in contrast to previous work, our results show relatively little brain atrophy in healthy subjects over 60, a continual gradation in loss of GM that begins in early adulthood, and an apparent relative increase in WM that approximately parallels the loss of GM, raising new issues about the change of imaging characteristics of GM and WM over time. Moreover, total CSF volume remains relatively constant, which has not been reported before. These results underscore the importance of healthy subject selection, the use of high quality images and the employment of standardized methods when attempting to establish group differences.

2 Clinical Material and Methods

2.1 Patient Population and MR Acquisition

The IRB-approved study included 50 volunteers (25 male, 25 female). Ten subjects, equally divided by sex, were imaged by decade (20-29, 30-39, 40-49, 50-59, and 60-72). Subjects with diabetes, hypertension, psychiatric disease, or history of brain disease were excluded. All subjects consented to make their images available via the web.

Images were obtained upon a head-only 3T MR unit (Allegra, Siemens Medical Systems Inc., Germany). A head coil was employed. T1 and T2 sequences, as well as additional image sequences, were obtained to cover the entire head. Voxel spacing was 1 x 1 x 1 mm for both T1 and T2 studies. For T1 images, acquisition parameters included TR=15, TE=7.1, fov=176x256 with a 0 gap, a matrix of 192x256, and a flip angle of 25. For T2 images, acquisition parameters included TR=7730, TE=80, fov=192x256 with a 0 gap, a matrix of 256 x 256, and a flip angle of 180.

2.2 Tissue Segmentation

The three brain compartments were defined using an expectation-maximization (EM) segmentation scheme that separated WM, GM, and CSF using both T1 and T2

images. A spatial probabilistic brain atlas [12] that represents expert prior knowledge about brain structures guided the segmentation. The registered atlas together with the set of prior probabilities for tissue were passed along with the T1 and T2 images as inputs to the segmentation algorithm. The segmentation algorithm used was based on the EM segmentation algorithm developed by van Leemput et al [13], [14] and re-written as an integrated package in ITK (Insight Toolkit). The automatic tool includes inter-modality registration, atlas registration, bias field inhomogeneity correction, and calculation of binary and probabilistic segmentation maps for CSF, GM and WM.

The lateral ventricular CSF was segmented using 3D snakes with region competition priors [15]. The snake was initialized near the ventricles and evolved based on the probability map for CSF, provided as input to the program. The GM, WM, total CSF and lateral ventricular volumes were calculated by integration over class-specific regions. Figure 1 illustrates results on an axial slice of one subject.

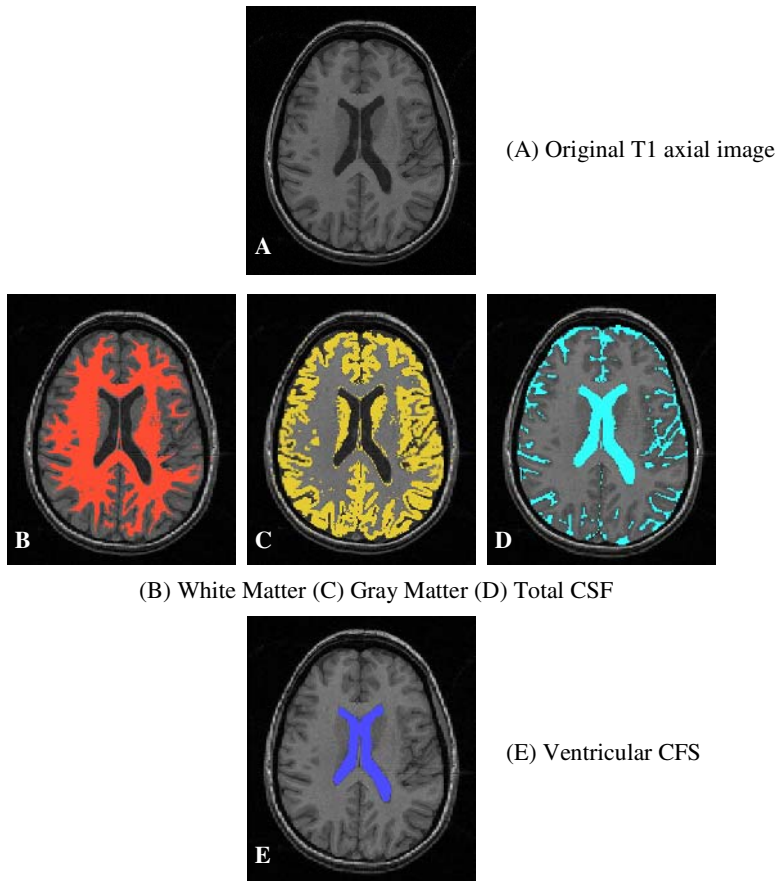


Fig. 1. Segmentation results: outputs from EM and 3D snakes segmentations

2.3 Statistical Analysis

Subjects were divided into five age groups named Agegroups 1-5, representing respectively the age ranges 20-29, 30-39, 40-49, 50-59 and 60-72. Absolute volumes were calculated for each of three compartments (WM, GM, and total CSF) as well as for the lateral ventricles. The total intracranial volume was viewed as the sum of WM+GM+CSF. Results were analyzed by age and sex not only for absolute volumes, but also for percentage of total volume occupied by each compartment. The percentage of the total CSF volume within the lateral ventricles was also examined.

Logarithmic transformations were applied to each percentage measurement and the normality of the transformed variables was tested by using the Shapiro-Wilks test and the Kolmogorov-Smirnov test. The total volume and the total CSF volume were similarly tested for the presence of normal distributions. A generalized linear model was used to fit the data. In the model, the total volume, the CSF volume and the percentages of the three components were treated as repeated measurements in the same subject so their correlations were accounted for in the analysis. The list of independent variables included the factor of age group (5 levels), the factor of gender group (2 levels) and their interactions.

Multivariable ANOVA test showed no significant interactions between age group and sex ($p=0.88$). Consequently, the final model only contained the main effects from age group and sex. The Tukey test was used to compare each measurement among the five age and two gender groups, after accounting for dependence among all the outcomes. Finally, the residue analysis post-model fitting showed that the model fits data well and that the model assumptions are thus presumably valid in this analysis.

3 Results

Results are presented in Tables 1 and 2 and illustrated in Figure 2. Table 1 shows the average total intracranial volume to be significantly smaller ($p < 0.001$) in females ($1341 \pm 100 \text{ cm}^3$) than in males ($1519 \pm 106 \text{ cm}^3$). Not surprisingly in this study of adult subjects, there is no significant difference in total intracranial volume by age ($p = 0.31$).

In terms of the ratio of WM to total intracranial volume, there is no significant difference between males and females ($p=0.07$). However, significant differences were observed among the five age groups ($p=0.002$). In particular, there appears to be a significant trend towards WM increase with age beginning at approximately age 50. Agegroup1 displays a smaller percentage of WM than Agegroups 4 and 5 ($p=0.05$ and 0.001), with Agegroups 2 and 3 displaying results similar to those of the younger group. On the average, the WM percentage for those aged 20-29 was 9.7% smaller than those 50-59, and 15% smaller than those 60 and above.¹

Males and females appear to have the same ratio of GM to total volume ($p=0.23$). However, GM volume tends to decrease with age ($p<0.001$). Agegroup1 has a consistently larger percentage of GM than their elders ($p=0.007$ for Agegroup3, <0.001 for Agegroups 4 and 5). Those aged 30-39 have a significantly larger percentage of GM than those 60 and above ($p=0.001$). On average, the percentage of GM in those 20-29 is about 7.2% larger than those 40-49, about 8.6% larger than

¹ Percentages are calculated as predicted differences of (% WM) / (% WM of compared group).

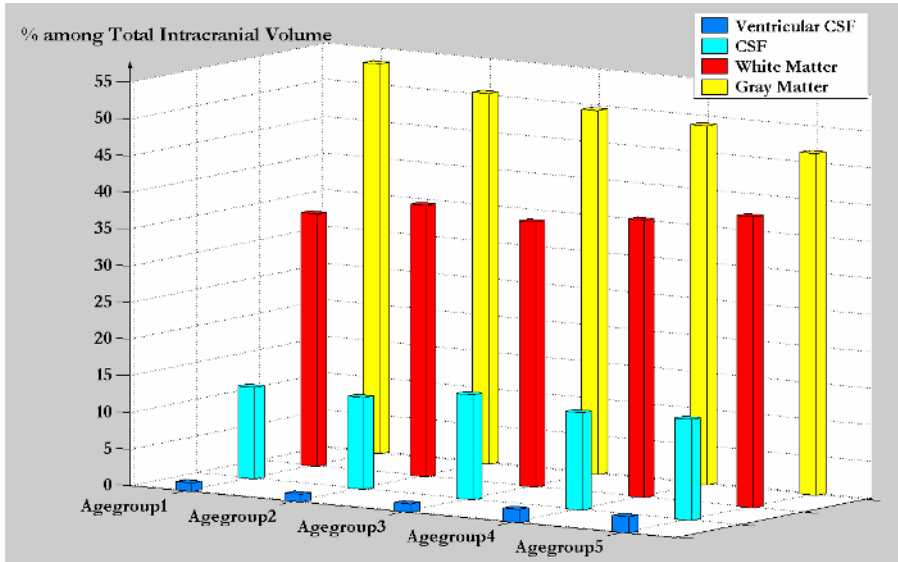


Fig. 2. Tissue volume evolution with healthy aging for overall groups

Table 1. Average values of total intracranial, WM and GM volumes in normal adults Agegroups 1-5 represent ages 20-29-...-60+ respectively, SD = Standard Deviation, % is the percentage among Total Intracranial Volume

Sex and Age Group	Total Intracranial Volume (cm ³ ±SD)	WM (cm ³ ±SD)	WM (%±SD)	GM (cm ³ ±SD)	GM (%±SD)
Both sex	1430±136	530±83	36.98±3.32	710±68	49.75±2.97
Agegroup1	1474±121	508±56	34.38±1.46	782±50	53.13±1.67
Agegroup2	1386 ±92	513 ±67	36.97±4.05	699±52	50.52±2.80
Agegroup3	1426 ±161	517 ±80	36.11±1.95	705±68	49.58±1.73
Agegroup4	1407 ±116	532 ±63	37.78±2.88	689±70	48.94±1.70
Agegroup5	1457 ±180	583±124	39.66±3.49	675±53	46.57±2.45
All Females	1341±100	488±65	36.28±3.31	672±59	50.12±3.04
Agegroup1	1387±77	472±50	33.94±1.72	751±39	54.15±1.52
Agegroup2	1326±88	498±79	37.51±5.20	665±48	50.26±3.43
Agegroup3	1310±120	461±67	35.04±2.03	650±45	49.77±2.08
Agegroup4	1326±82	496±63	37.37±3.70	644±47	48.59±1.97
Agegroup5	1358±144	512±78	37.54±1.92	648±54	47.83±1.83
All Males	1519±106	573±77	37.68±3.24	749±54	49.37±2.90
Agegroup1	1562±88	544±35	34.82±1.16	813±41	52.12±1.17
Agegroup2	1445±48	527±58	36.43±3.02	733±29	50.79±2.40
Agegroup3	1542±99	573±43	37.18±1.23	760±26	49.38±1.53
Agegroup4	1488±86	568±42	38.19±2.15	734±61	49.28±1.53
Agegroup5	1556±167	654±126	41.78±3.53	702±39	45.31±2.50

those 50-59 and about 14.2% larger than those 60 and above. The percentage of GM in people aged 30-39 is about 8.5% larger than in people 60 and over.

On the basis of literature studies, we anticipated a significant increase in CSF volume with age. The differences in CSF volume by age in the current study were relatively slight, however (Table 2), and much less marked than the age-associated

changes to GM and WM. For total CSF volume, there was no significant difference among the five age groups ($p=0.12$) or between genders ($p=0.07$). A trend towards an increased ratio of ventricular CSF to total CSF may occur with age, but our data do not provide sufficient evidence to fully support this assertion. The only statistically significant difference between groups was between those 40-49 and those 60 and above, occurring at the $p=0.04$ level.

Table2. Average values of CSF and lateral ventricles volumes in normal adults Agegroups 1-5 represent ages 20-29-...-60+ respectively, SD = Standard Deviation, % is the percentage among Total Intracranial Volume, %* is the percentage among CSF volume

Sex and Age Group	CSF (cm ³ ±SD)	CSF (%±SD)	Lateral Ventricles (cm ³ ±SD)	Lateral Ventricles (%*±SD)
Both sex	189±29	13.27±1.89	21±13	11.19±7.05
Agegroup1	184±33	12.48±1.74	18±11	9.50±4.31
Agegroup2	174±30	12.51±1.87	14±5	8.36±3.07
Agegroup3	203±29	14.31±1.63	17±10	8.10±3.94
Agegroup4	186±25	13.28±2.06	25±12	13.66±7.18
Agegroup5	199±24	13.77±1.80	32±19	16.31±10.80
All Females	182±30	13.60±2.15	17±10	9.42±5.21
Agegroup1	164±22	11.90±1.89	14±7	8.54±3.18
Agegroup2	163±36	12.23±2.24	13±3	8.41±2.47
Agegroup3	199±20	15.19±1.15	12±7	5.74±3.09
Agegroup4	186±31	14.04±2.33	22±6	12.10±4.46
Agegroup5	198±25	14.62±1.37	24±17	12.32±8.88
All Males	197±27	12.95±1.57	25±15	12.95±8.23
Agegroup1	204±30	13.06±1.55	22±13	10.46±5.42
Agegroup2	184±21	12.79±1.64	15±7	8.31±3.89
Agegroup3	208±38	13.44±1.66	22±10	10.46±3.38
Agegroup4	186±21	12.52±1.65	27±16	15.22±9.49
Agegroup5	200±26	12.91±1.90	39±19	20.31±12.00

4 Discussion

Age-associated changes in the volumes of the three intracranial compartments have interested many investigators. To our knowledge, however, the current study is the first to employ high-resolution MR, standardized protocols, careful screening of healthy subjects, and age and sex-matching of subjects by decade. Although many of our results are consistent with previous reports, some are not. We believe that these differences are most likely related to differences in image quality and subject selection.

Consistent with the results of others [1], [10], our study indicates that males have a larger total intracranial volume than females and that the mean intracranial volumes of the younger and older groups are comparable. The total intracranial volume of our subjects, however, tended to be larger than reported in other adult studies [1], [10], [16]. This difference could be attributable either to the fact that our studies covered the entire brain, whereas other studies have not, or to the high resolution of our images, allowing more accurate volumetric measurements.

Many investigators have reported brain shrinkage with age, attributed primarily to loss of GM [2], [3], [4], [10]. To our surprise, however, our study revealed relatively

little shrinkage of the total GM and WM with age. Indeed, total CSF volume remained relatively constant even for subjects in their 60s and 70s. Although the oldest subject in the current study was 72, other reports have indicated significant loss of brain tissue by this age [1], [2]. These differences are most likely related to the careful selection of healthy subjects employed in the current study, and suggest that tissue loss in these age groups may be related more to the effects of age-related diseases rather than to the aging process in isolation.

We agree with other groups that there is a loss of GM over time [2], [3], [4], [10]. Guttmann et al [2] describe a particularly nice study that examines subjects over ages, but conclude that the only significant age-related difference in GM percentage was between subjects 18-39 and those in their 50s. By contrast, our study shows a steadily progressive decline in GM percentage by decade, affecting both males and females, well-evident by middle age, and with statistically significant differences between multiple groups.

Another surprising finding was the statistically significant increase in the proportion of WM with age. To our knowledge, this finding has not been previously reported. Intuitively, it makes sense that the GM shrinks more than the WM with age-related brain atrophy. GM shrinkage is presumably due to loss of neurons. Some of these neurons have axons that project into the WM and one would expect that losing these neurons would result in a corresponding loss of WM. However, there are also neurons in the cortex that have axons that stay within the cortex or only extend centripetally as far as the subcortical U fibers. Thus, if these neurons are lost, GM shrinkage would occur without corresponding shrinkage of the WM. This explanation is consistent with our finding of relative increases on WM with age. Another interesting possibility is that age produces changes to MR imaging characteristics so that GM becomes "less gray" as seen by MR and so becomes mistakenly classified as WM. This is a potential area of future research.

The ratio of ventricular volume to total CSF volume is of particular interest to the study of diseases such as normal pressure hydrocephalus. We were unable to find another age and sex matched study of healthy controls with which to compare our results. Our results suggest that there may be slight increase in both absolute ventricular volume and in ventricular/total_CSF ratio in our oldest age group (60+), but these results were statistically significant only when compared to the 40-49 years old group and even there did not reach a high level of significance. These results are in general agreement with those of Matsumae et al. [1], who studied ventricular volume with age. Our results may establish a basis upon which to study diseases such as normal pressure hydrocephalus at a later date.

5 Conclusion

Considered together, these results lead to the suggestion that the traditional view of age-related changes to intracranial compartments should be revised. Significant atrophy does not appear to occur in genuinely healthy subjects even into the late 60s and early 70s. Loss of GM with aging appears to be a steadily progressive process that begins relatively early in adulthood, and that may be associated with a relative increase in the proportion of WM.

This study raises several questions, however. First, the current study does not address the location of GM loss (cortex? basal ganglia? thalamus?), which could be an important area of research helpful in differentiating between healthy aging and pathology. Second, there is a real question of age-related changes in the MR imaging characteristics of GM and WM. Such changes in imaging characteristics could affect results, and could provide an important area of research.

The major contribution of this study, however, is in the initiation of a high-quality, multi-modality image database of carefully screened healthy subjects, evenly divided by sex, and with an equal number of subjects per decade. This database could serve as an important resource to all investigators interested not only in normal aging but also in the statistical study of disease.

Acknowledgment

This work was supported by the grant RO1 EB000219 NIH-NIBIB.

References

1. Matsumae M, Kikinis R, Mórocz I, Lorenzo A, Sándor T, Albert MS, McL. Black P, Jolesz F: Age related changes in intracranial compartment volumes in normal adults assessed by MRI. *J Neurosurg* (1996) 84(6): 982-91
2. Guttmann CR, Jolesz FA, Kikinis R, Killiany RJ, Moss MB, Sandor T, Albert MS: White matter changes with normal aging. *Neurology* (1998) 50(4): 972-8
3. Gur RC, Gunning-Dixon FM, Turetsky BI, Bilker WB, Gur RE: Brain Region and Sex Differences in Age Association With Brain Volume: A quantitative MRI Study of Healthy Young Adults. *Am J Geriatr Psychiatry* (2002) 10:72-80
4. Resnick SM, Goldszal AF, Davatzikos C, Golski S, Kraut MA, Metter EJ, Bryan RN, Zonderman AB: One-year Age Changes in MRI Brain Volumes in Older Adults. *Cerebral Cortex* (2000) Vol. 10, No. 5, 464-472
5. Wohl MA, Mehninger CM, Lesser IM, Boone KB, Miller BL: White matter hyperintensities in healthy older adults: A longitudinal study. *International Journal of Geriatric Psychiatry* (1994) Volume 9, Issue 4, Pages 273-277
6. Condon B, Patterson J, Wyper D, Hadley D, Grant R, Teasdale G, Rowan J: Use of magnetic resonance imaging to measure intracranial cerebrospinal fluid volume. *Lancet* (1986) 1(8494):1355-7
7. Haug G: Age and sex dependence of the size of normal ventricles on computed tomography. *Neuroradiology* (1977) 14(4):201-4
8. Barron SA, Jacobs L, Kinkel WR: Changes in size of normal lateral ventricles during aging determined by computerized tomography. *Neurology* (1976) 26:1011-1013
9. Malko JA, Hoffman JC Jr, Green RC: MR measurement of intracranial CSF volume in 41 elderly normal volunteers. *AJNR Am J Neuroradiol* (1991) 12(2):371-4
10. Matsumae M, Kikinis R, Mórocz IA, Lorenzo AV, Albert MS, McL. Black P, Jolesz FA: Intracranial compartment volumes in patients with enlarged ventricles assessed by MRI based image processing. *J Neurosurg* (1996) 84(6): 972-81
11. Resnick SM, Pham DL, Kraut MA, Zonderman AB, Davatzikos C: Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. *J Neurosci* (2003) 23(8): 3295-301

12. Evans AC, Collins DL, Mills SR, Brown ED, Kelly RL, Peters TM: 3D statistical neuroanatomical models from 305 MRI volumes. Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference (1993) 1813-1817
13. Van Leemput K, Maes F, Vandermeulen D, Suetens P: Automated model based tissue classification of MR images of the brain. IEEE Trans Med Imaging (1999) 18:897-908
14. Van Leemput K, Maes F, Vandermeulen D, Suetens P: Automated model based bias field correction of MR images of the brain. IEEE Trans Med Imaging (1999) 18:885-896
15. Ho S, Bullitt E, Gerig G: Level set evolution with region competition: Automatic 3-D segmentation of brain tumors. In: Katsuri R, Laurendeau D, Suen C (Eds.), Proc. 16th International Conference on Pattern Recognition. IEEE Computer Society (2002) pp. 532-535
16. Reiss AL, Eliez S, Schmitt JE, Straus E, Lai Z, Jones W, Bellugi U: Neuroanatomy of Williams syndrome: a high-resolution MRI study. J Cogn Neurosci (2000) 12 Suppl 1:65-73

Predicting Clinical Variable from MRI Features: Application to MMSE in MCI

S. Duchesne¹, A. Caroli², C. Geroldi², G.B. Frisoni², and D. Louis Collins¹

¹ Montréal Neurological Institute (MNI), McGill University, Montréal, Canada
² IRCC San Giovanni di Dio, Fatebenefratelli, Brescia, Italy

Abstract. The ability to predict a clinical variable from automated analysis of single, cross-sectional T1-weighted (T1w) MR scans stands to improve the management of patients with neurological diseases. We present a methodology for predicting yearly Mini-Mental Score Examination (MMSE) changes in Mild Cognitive Impairment (MCI) patients. We begin by generating a non-pathological, multidimensional reference space from a group of 152 healthy volunteers by Principal Component Analyses of (i) T1w MR intensity of linearly registered Volumes of Interest (VOI); and (ii) trace of the deformation fields of nonlinearly registered VOIs. We use multiple regression to build linear models from eigenvectors where the projection eigencoordinates of patient data in the reference space are highly correlated with the clinical variable of interest. In our cohort of 47 MCI patients, composed of 16 decliners, 26 stable and 5 improvers (based on MMSE at 1 yr follow-up), there was a significant difference ($P = 0.0003$) for baseline MMSE scores between decliners and improvers, but no other differences based on age or sex. First, we classified our three groups using leave-one-out, forward step-wise linear discriminant analyses of the projection eigencoordinates with 100% accuracy. Next, we compared various linear models by computing F-statistics on the residuals of predicted *vs* actual values. The best model was based on 10 eigenvectors + baseline MMSE, with predicted yearly changes highly correlated ($r = 0.6955$) with actual data. Prospective study of an independent cohort of patients is the next logical step towards establishing this promising technique for clinical use.

Keywords: MRI, Principal Components Analysis, Intensity, Deformation, Multiple Regression, Mild Cognitive Impairment, Mini-Mental Score Examination.

1 Introduction

A number of neurological diseases exhibit pathologically-specific discriminatory information in the form of local intensity variations and shape changes when observed on magnetic resonance images (MRI). The goal of computer-aided diagnosis approaches is to focus and exploit those attributes in order to give physicians a quantitative measurement related to the disease process. Example techniques may perform analysis of the T1-weighted (T1w) signal intensity, serving as an indicator of disease progression, as subtle changes may indicate an underlying

pathological process before structure integrity is lost. Other approaches will employ co-registration, a process where individual subject images are aligned into a reference space, allowing spatial comparisons to be made between cohorts either at the voxel level, such as in voxel or deformation-based morphometry, or for surfaces. Registration further enables the comparison of individual structures once segmented; manual segmentation and volumetry are considered the gold standard for many research areas. The most promising techniques however move away from single-structure approaches to whole regions of interest, capturing the interrelations between neighboring tissues, and combine individual intensity, texture or registration information, such as in appearance-based approaches [1].

An important and valuable area of research for quantitative MRI analysis resides in the prediction of clinically measured variables. An example of the latter is the Mini-Mental State Examination (MMSE), a cognitive scale to assess normality ($MMSE = 30$), mild impairment ($23 \leq MMSE < 30$) or possible dementia ($MMSE \leq 23$) [2]. The assumption is that such neuropsychological or neurological assessments will have a morphological correlate that is detectable via MRI. The ability to categorize cognitive scores from baseline MRI may lead to increased understanding of the disease in question. Further, the ability to predict future scores (i.e., predict change) from baseline MRI is even more important, as it would improve patients management.

1.1 Mild Cognitive Impairment (MCI)

MCI is widely viewed as the transition phase between normal aging and Alzheimer's disease (AD) [3], and amnesic MCI individuals are known to be at risk for progression to AD. There is evidence that in those who will progress, measurable hippocampal and entorhinal cortex atrophy, demonstrable on T1w MRI serves as a moderate, though labor-intensive, predictor [4]. Microscopically the strongest predictor of premortem cognitive dysfunction appears to be the relative area of entorhinal cortex occupied by beta-amyloid deposition [5]. Existing MRI measures that have been developed to predict decline are longitudinal, as for example a study by Rusinek et al. [6] showing that an increased rate of atrophy in the MTL predicted future cognitive decline.

In this work we attempt to solve the challenging problem of predicting MMSE changes with a single, cross-sectional MRI measurement. An obvious advantage of having a reliable means of assessing future cognitive decline (within the limits of the MMSE) at baseline, or with one scan, resides in the potentially increased therapeutic effect that comes with earlier detection and treatment. The disadvantage of any cross-sectional approach is that the effect at hand may differ for individuals enrolled in the study or be confounded by another variable, such as aging in the case of MCI. It is important to note that not all MCI patients progress to clinically defined AD, nor show MMSE decline at identical rates [7]. However, careful design and selection of the patient population should serve to improve homogeneity of effect in the test population. Further, we believe that the rate of MMSE decline, along with MTL atrophy, will be linear for MCI patients over the short duration of the study, a reasonable assumption given long-term AD follow-up data [8].

1.2 Hypothesis and Goal

Our general hypothesis is that in the case of neurological pathologies, microscopic changes will be detected via their impact on the T1w MRI signal intensity, while macroscopic alterations in structure shape will be noticed via registration/deformation-based metrics. Consequently, we propose a prediction methodology that (i) uses a large, non-specific Volume of Interest (VOI); (ii) combines intensity and registration-based shape features; and (iii) generates a high dimensional linear model from multiple regression of highly correlated eigenvectors. Our primary goal is to see if we can create such a model, and use MCI patients as a test case of the methodology.

2 Methods

Our method can be summarized as follows. First, we generate a non-pathological eigenspace from a large training group of young subjects ($N = 152$). This multi-dimensional reference eigenspace is created by uniting results from four distinct Principal Component analyses of (i) linearly registered intensity images of the left and right VOIs; and (ii) an approximation of the determinant of the Jacobian matrix of the deformation field within those VOIs. Secondly, patients data are projected in the reference eigenspace and the correlation coefficient between the projection coordinates and the clinical variable is used to identify eigenvectors for the predictive model. The latter is generated via multiple regression against the clinical variable. We compare predictive ability by computing F-statistics based on the residuals of predicted *vs* actual values for the clinical variable. The methodological details are elaborated in the following sections.

2.1 Subjects

The Ethics Committee of the Montreal Neurological Institute (Montreal, Canada) and the IRCCS San Giovanni di Dio FBF (Brescia, Italy) approved the study and informed consent was obtained from all participants. A total of 199 subjects were included in this study. The reference group consisted in 152 young, neurologically healthy individuals from the International Consortium for Brain Mapping database (ICBM) [9], whose scans were used to create the non-pathological, reference space. The training population consisted in 47 MCI patients ($23 \leq MMSE < 30$), seen at the IRCCS San Giovanni di Dio FBF Hospital, that have been followed clinically a minimum of 12 months after their initial MR scan.

2.2 Preprocessing

MRI data for our 152 ICBM subjects was collected with a T1w MRI protocol on a 1.5 T scanner (Philips Gyroscan, Best, Netherlands) using a fast gradient echo sequence ($TR = 18ms$, $TE = 10ms$, 1 NEX pulse sequence, flip angle= 30° , matrix size= 256×256 , $FOV = 256mm$, slice thickness= $1mm$). Data for MCI patients were acquired on a 1.0T scanner (Philips Gyroscan, Best, Netherlands) using an FFE sequence ($TR = 19.7ms$, $TE = 6.9ms$, sagittal acquisition,

$0.9365 \times 0.9375 \times 1.3\text{mm}^3$). All global MRI data were processed to correct for intensity non-uniformity due to scanner variations [10]. The 152 ICBM subjects were registered in a Talairach-like stereotaxic space in the context of the ICBM project [9]. Most (33/47) of the MCI data were linearly registered (9 DoF) automatically into stereotaxic space [11] while the remaining volumes were manually registered due to high scalp brightness. All reference and training volumes were resampled onto a 1mm isotropic grid [11].

Two VOIs were selected for this study, centered on the left and right medial temporal lobe, using Talairach coordinates (start coordinates $x = [-57, +2]$ for the left and right side respectively, $y = -53$ and $z = -52$). Each VOI measured $n = 55 \times 82 \times 80 = 360800$ voxels. The VOI was selected so that its extent captured the hippocampus and neighboring MTL structures (e.g. ento and perirhinal cortex, parahippocampal gyrus), irrespective of normal inter- and intra-individual variability. After extraction, each VOI was linearly registered (9 DoF) to the reference volume to further reduce local distortions, and its mean intensity scaled to the mean intensity of the reference VOI, which serves to eliminate the first-order drift in signal measurement between patients.

2.3 Multi-dimensional Reference Space and Model Creation

Two image features at each voxel location were retained. The first feature is the *grey level intensity* consisting in the rasterized data from the intensity-scaled VOIs. The second feature is the *trace* or the first-order approximation of the determinant of the Jacobian matrix of a non-linear registration-derived deformation field. The latter is calculated to map each subject's VOI to our reference ICBM target. The trace represents an estimate of local volume change. Principal Components Analysis (PCA) is used to reduce the dimensionality of the input training data and generate linear variation models based on the $N = 152$ datasets from our ICBM normal subjects. The resulting four PC models were each $p = N - 1$ (or 151-dimensional). Most of the variation can usually be explained by a smaller number of modes, l , where $l \ll n$ and $l < p$. We proceeded in selecting 535 eigenvectors in total from our four models (left/right intensity/trace VOIs), that accounted to a per-model variance of 99.7%.

Rasterized vectors of the processed VOI intensity and trace data for each test subject are then projected into the training space, and thus form eigencoordinate vectors. While a number of possible features can be calculated on the distribution of the projected data, our predictor is based on the position along the PC axes. The distribution of eigencoordinates along any principal component for a given population is normally distributed as assessed via Shapiro-Wilke statistics. For each eigenvector the correlation of the eigencoordinate distribution with the clinical variable is then computed. We selected a number q of vectors based on an arbitrarily predefined threshold for the correlation coefficient of $r > |0.30|$. A predictive model is then built from those eigenvectors using multiple regression (JMP IN, SAS Institute, Cary, N. Carolina) and the model is then used to predict the future value of the clinical variable of interest. Residuals and correlation are computed between predicted vs actual value, and F-tests are used

to determine if the models improve the prediction of our clinical variable from MRI features.

2.4 Experiments

Four experiments were completed. **Experiment 1** served as a baseline for the classification of our patient population into 3 groups based on their MMSE changes at 1 year follow-up from clinical variables (age, sex, baseline MMSE). **Experiment 2** attempted the same 3-group classification but this time based on the projection eigencoordinates in the reference space. **Experiment 3** served as a baseline for the prediction of yearly MMSE decline by building a linear model based on clinical variables ("Clinical"). **Experiment 4** attempted the same prediction but with a model based on projected eigencoordinates, as per the methodology described above ("MRI"), while in **Experiment 5** we added baseline MMSE as an additional variable to the projected eigencoordinates ("MRI+baseline MMSE").

3 Results

When comparing MMSE results between baseline and 12 months follow-up, we can separate the 47 patients in the test population into three distinct groups: 16 decliners (> -1 point negative change in MMSE or cognitive decline), 5 improvers (> 1 point positive change in MMSE or cognitive improvement), and 26 stable individuals (MMSE change between $[-1, 1]$). Demographic information about each group can be found in fig. 1. There was no statistically significant age difference between either groups, as assessed from ANOVA and Tukey-Kramer

Groups	Decliners	Stable	Improvers
Subjects	16	26	5
Mean age (yrs)	72.4	67.6	71.8
Std dev	(4.7)	(8.5)	(5.4)
Baseline MMSE	27.5	27.8	24.8
Std dev	(1.3)	(1.4)	(1.8)
Mean MMSE Δ	-2.9	0	2.4
Std dev	(1.2)	(0.8)	(0.6)

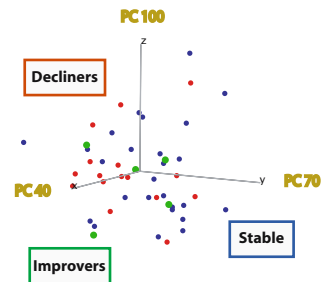


Fig. 1. (LEFT) Demographic information. **(RIGHT)** Leave-one-out, forward stepwise linear discriminant analysis of the patient eigencoordinates in the reference space was 100% accurate at classifying groups (decliners, stable, improvers). The data is shown here projected on the 3 most discriminating eigenvectors. Our goal was to find an independent basis for the classification and creation of a predictive model, rather than the optimal reference space in which to represent our population.

HSD ($P > 0.05, DF = 2$). There was a statistically significant baseline MMSE difference between the decliners and improvers ($P = 0.0003, DF = 2$), but no other significant difference between groups for baseline MMSE. The improvers had the lowest mean baseline MMSE of all three groups.

The classification is based on a leave-one-out, forward stepwise linear discriminant analyses (SYSTAT 10.2, Georgia, PA; $P - to - enter < 0.05$) of either clinical variables (age, sex and baseline MMSE) or eigencoordinates along the 535 reference space eigenvectors. The clinical classifier of **Experiment 1** was 53% accurate in separating the 47 patients into decliners, improvers and stable subjects ($DF = 2, \text{Wilk's } \lambda = 0.69$), while the 3-way classifier based on projection eigencoordinates of **Experiment 2** was 100% accurate, with 31 significantly discriminant eigenvectors ($P - to - enter < 0.05, DF = 31, \text{Wilk's } \lambda = 0$). Fig. 1 displays the data plotted along the three most discriminating eigenvectors.

While Experiments 1 and 2 classified the data into groups, in the following 3 experiments our goal was to predict the magnitude of the yearly MMSE change.

Table 1. Results from prediction models

Model	Features	r	r^2	SD	F stat to Clin.	P	F stat to MRI	P
Clinical	3	0.429	0.176	1.86	-	-	-	-
MRI	10	0.668	0.446	1.53	2.499	0.003	-	-
MRI+MMSE	11	0.696	0.484	1.48	2.691	0.002	2.585	0.002

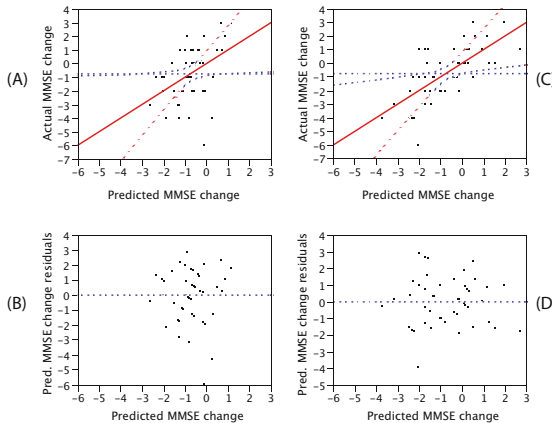


Fig. 2. (A) Clinical model built from multiple regression of age, sex and baseline MMSE against 1 year MMSE changes. (B) Residuals for the "Clinical" model. The correlation of predicted *vs* actual values was $r = 0.429$. (C) MRI+Baseline MMSE model built from multiple regression of the 10 most correlated reference space eigenvectors plus baseline MMSE. (D) Plot of residuals. With a correlation between predicted *vs* actual yearly MMSE changes of $r = 0.6955$, this model was a significant improvement over the "Clinical" one ($F - stat = 2.691, P = 0.002$).

Baseline MMSE, age and sex were all negatively and weakly correlated with 1 year MMSE change ($r = -0.25$, $r = -0.21$ and $r = 0.15$, respectively). In contrast, out of the 535 reference space eigenvectors, 10 had a correlation ratio of $r > |0.30|$. We predicted MMSE change for all patients using each linear model. The number of input features to the model, the resulting correlation (r) and squared correlation (r^2) of predicted vs. actual values, the standard deviation of the predicted score and F-test values (against "Clinical" and "ICBM") are shown in table 1. Recall that the first predictive model (**Experiment 3**) is based on the 3 clinical variables ("Clinical"), the second (**Experiment 4**) on the 10 selected eigenvectors ("MRI") and the last (**Experiment 5**) using the 10 eigenvectors plus the baseline MMSE ("MRI+baseline MMSE"). The linear fit for the "Clinical" and "MRI+baseline MMSE" models are shown in fig. 2, alongside their residual plots. The best model was the "MRI+baselineMMSE" of **Experiment 5**, with a correlation between predicted and actual value of $r = 0.6955$. It was also significantly better than either the "Clinical" model ($Fstat = 3.39$, $P = 0.0001$, $DF_1 = 43$, $DF_2 = 35$) or the "MRI" model ($Fstat = 2.59$, $P = 0.002$, $DF_1 = 36$, $DF_2 = 35$).

4 Discussion

Our goal was to demonstrate the application of a generic, multidimensional reference space from MR features for the classification and prediction of a clinical variable. We proceeded with analysis of a cohort of 47 Mild Cognitive Impairment patients and achieved 100% classification accuracy with 31 eigenvectors. We succeeded in creating an 11-variable linear model that predicted yearly MMSE changes, explaining 48.4% of the variability in the actual data.

Our approach has been to create a model eigenspace based on subjects from the ICBM database, in which we projected our MCI patients. While such a space built from young, neurologically healthy individuals may not be optimal to represent the MCI cohort, it should be noted that our primary goal was to find an independent basis for the creation of a predictive model, and not to find the best representation (mathematical, clinical or otherwise) of the patient population. As the training set is composed of ICBM subjects that are separate from the test set subjects, there is no issue of overdetermination in the creation of the reference space from PCA of the ICBM data. If we are to use this technique for a prospective or retrospective study, we will need either to use a separate test set or to use a leave-one-out technique in order for our regression model to remain independent of any and all training data.

The use of intensity features from MRI raises the question of calibration and normalization. Absolute intensities are rarely used in MRI, since they vary with machine calibration, shimming, and patient-induced variations. We have tried to limit those variations by (1) using the same scanner within groups in the study; (2) ensuring that the same quality assurance procedures were followed for each acquisition at each site; (3) by acquiring subject scans in random group order; and (4) scaling with respect to the reference image.

Not all macroscopic changes will be captured by the registration process. Point homology in nonlinear registration is of course approximate: in regions where there is complete homology, the displacement field will be nearly exact; and in regions where it is not, the result will be noisy. This uncertainty however is rejected in the PCA model, as it is uncorrelated and non-covarying, and therefore only changes associated with the pathology should remain.

While the classification procedure ought to be tested for generalizability in another set of patients, the leave-one-out classification test (**Experiment 2**) indicates that it is possible to predict 1-year MMSE changes based on a single MRI scan. This has important ramifications for patient treatment, since success for most therapies is expected to increase with earlier detection of cognitive changes and appropriate treatment. Furthermore, the "MRI + baseline MMSE" linear model (**Experiment 5**) yields an estimate of the amount of change expected in MMSE over the course of one year. Such data could be used to better tailor therapy for specific patients.

5 Conclusion

The ability to perform both classification and prediction of a clinical variable from a single, cross-sectional T1w MRI scan stands to benefit tremendously physicians in the management of patients with MCI. We have explained our methodology for model creation and demonstrated its application to the prediction of yearly MMSE changes in a cohort of MCI subjects. A prospective study on an independent cohort of patients is the next logical step towards further acceptance of this technique in the medical community.

Acknowledgments. Drs. Mazziotta and Evans (ICBM) for permission to use data.

References

- [1] T.F. Cootes *et al.*, *IEEE PAMI*, 23:6, 2001.
- [2] J.R. Petrella *et al.*, *Radiology*, 226:315-336, 2003.
- [3] C. DeCarli *et al.*, *Neurology*, 63:220-7, 2004.
- [4] L.deToledo-Morrell *et al.*, *Neurobiol. Aging*, 25(9):1197-1203,2004.
- [5] B.J. Cummings *et al.*, *Neurobiol. Aging*, 17:921-33, 1996.
- [6] R. Rucinek *et al.*, *Radiology*, 229:691-96, 2003.
- [7] M. Grundman *et al.*, *J Mol Neurosci*, 19:1-2, 2002.
- [8] D. Chan *et al.*, *Lancet*, 362:1121-2, 2003.
- [9] J. Mazziotta *et al.*, *NeuroImage*, 2:89-101, 1995.
- [10] J. Sled *et al.*, *IEEE TMI*, 17:87-97, 1998.
- [11] D. Collins *et al.*, *IJPRAI*, 11:1271-1294, 1997.

Finite Element Modeling of Brain Tumor Mass-Effect from 3D Medical Images

Ashraf Mohamed^{1,2} and Christos Davatzikos^{1,2}

¹ CISST NSF Engineering Research Center,

Department of Computer Science, Johns Hopkins University

² Section for Biomedical Image Analysis, Department of Radiology,
University of Pennsylvania School of Medicine

{ashraf, christos}@rad.upenn.edu

Abstract. Motivated by the need for methods to aid the deformable registration of brain tumor images, we present a three-dimensional (3D) mechanical model for simulating large non-linear deformations induced by tumors to the surrounding encephalic tissues. The model is initialized with 3D radiological images and is implemented using the finite element (FE) method. To simulate the widely varying behavior of brain tumors, the model is controlled by a number of parameters that are related to variables such as the bulk tumor location, size, mass-effect, and peri-tumor edema extent. Model predictions are compared to real brain tumor-induced deformations observed in serial-time MRI scans of a human subject and 3 canines with surgically transplanted gliomas. Results indicate that the model can reproduce the real deformations with an accuracy that is similar to that of manual placement of landmark points to which the model is compared.

1 Introduction

Accurate deformable registration of 3D brain tumor images into a common stereotactic space is needed for the construction of brain tumor atlases. Such atlases will be useful in planning neuro-surgical tumor operations and therapeutic approaches by linking the functional and structural information provided by multi-modality images to variables such as tumor size, grade, subsequent progression, therapeutic approach and outcome [1–4].

Currently available image registration approaches applied to a register a normal brain atlas and a tumor-bearing image have limited accuracy in and around the tumor area. This is due to the inability of these approaches to account for topological differences between the two images, severe deformation in the vicinity of the tumor, and the confounding effects of edema and tumor infiltration. Approaches that introduce a small tumor “seed” in the atlas brain and rely on image information to subsequently adapt the atlas to the subject’s images produce acceptable registration results for some tumor cases, however, patients with large tumors and substantial brain tissue deformation still present a major challenge [3, 4]. For such cases, the lack of a physically realistic model of brain tissue deformation, derails the image matching process and causes the deformable registration to fail near the tumor.

In this paper, a 3D finite element (FE) model of the mass-effect of brain tumors is presented. The model is controlled by a number of parameters that are related to variables such as the tumor size, location, and peri-tumor edema extent. The exact value of these parameters for a particular tumor patient can only be found within the context of inverse problem solving based on the patient's images. The goal behind the model presented here is not to predict tumor growth, or the mass-effect for a particular tumor patient, but rather to use this model for generating a large number of brain anatomies deformed by simulated tumors for a range of the model parameters. These simulated anatomies will act as training samples for a statistical estimator of the model parameters for particular brain tumor cases. Once the model parameters are estimated, they can be used to introduce a tumor in the atlas and simulate the associated mass-effect, prior to applying a deformable image registration approach to the two images. This paper is dedicated to presenting the details of a mechanical model for tumor mass-effect and its partial validation via a number of real tumor cases. The integrated approach for image registration is described in [5].

In contrast to previous similar models that used 2D simulations and included only bulk tumor mass-effect [1, 2, 6], the model presented here is fully 3D and includes the mass-effect of the tumor as well as that of peri-tumor edema. Most of the mass-effect of many real tumors is attributed, not to the bulk tumor itself, but to the peri-tumor edema [7]. Additionally, the model employs a non-linear constitutive material model for brain tissues and a non-linear FE formulation which allow the simulation of realistic, large-deformation tumor cases. The details of the approach are presented in Sect. 2.

Another contribution of this work is the quantitative comparison of model's predictions to deformations caused by real tumors and observed in a dataset of serial MRI scans. This dataset is composed of brain images of a brain tumor patient, and uniquely available scans of three canines with surgically transplanted tumors. Results reported in Sect. 3 show that the presented model can reproduce a large percentage of the deformations caused by the real tumors. The paper is concluded with a discussion of the results in Sect. 4.

2 Methods

The aim of the proposed model is to simulate only the mass-effect component of tumor growth via a mechanical FE model constructed from 3D medical images. Since tumor growth is not purely a mechanical process, but involves a host of interacting biological, biochemical and mechanical mechanisms, it is essential to initialize the model simulations with a configuration for the brain from which the target configuration (that deformed by the tumor at the desired stage of tumor growth) is reachable by solving a continuum mechanics problem.

The proposed approach can be understood by aid of Fig. 1. Let κ_o be the initial configuration of the brain before the tumor emergence. The stresses in κ_o are assumed negligible. Let κ_t be the configuration of the brain at the target stage of tumor development. The bulk tumor denoted by T_t , is assumed to be

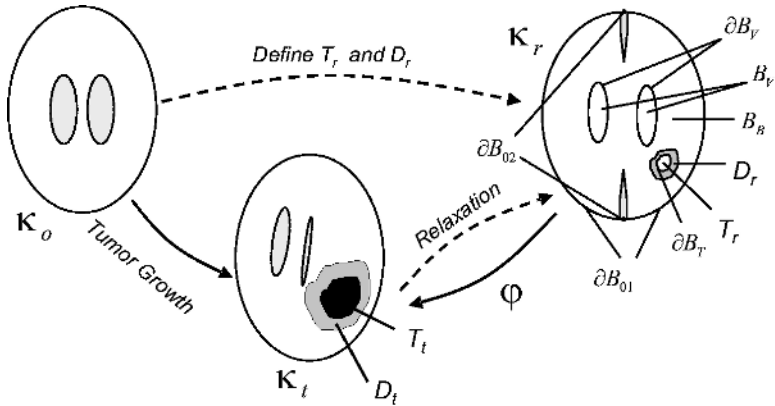


Fig. 1. A schematic showing the three configurations involved in the model. κ_o is the brain before tumor development, κ_t is brain at the desired stage of tumor growth, and κ_r is the corresponding relaxed configuration. T_t and D_t are the bulk tumor and peritumor edema regions in κ_t respectively, while T_r and D_r are the corresponding regions in κ_r . In κ_r , the ventricles are denoted by B_V , and ∂B_{01} denotes the outer surface of the brain except for ∂B_{02} where the falx meets the skull.

composed of proliferative, quiescent and necrotic tumor cells [6]. A region D_t of brain tissue swollen by edema may also be associated with the tumor.

If T_t were resected, the edema were diffused, and the stresses were allowed to relax, brain tissues will reach a relaxed configuration κ_r . There is a relaxed configuration associated with every κ_t and it is, in general, different from both κ_t and κ_o . Given κ_r , the stresses caused by the bulk tumor, and the swelling due to edema, the deformation map φ can be obtained by solving the mechanics problem described below. For real tumor cases, these parameters are not known, but they can be estimated from the measured pattern of deformed anatomy as described [5]. For simulating mass-effect of tumors starting with normal brain images, approximations of these parameters are used as explained next.

Defining κ_r involves specifying the geometry of the brain and that of T_r (which corresponds to brain tissue that is no longer present in κ_t) and D_r (which corresponds to brain tissue that is swollen by edema in κ_t). These regions are highly variable for different tumor cases and types. For tractability, herein, we will consider the approximation of T_r and D_r with two concentric spheres. Since expansion due to edema happens mainly in white matter (WM) [7, 8], T_r is restricted to WM tissues. The center and radii of T_r and D_r are treated as model parameters. It is worth noting that the shape of the final tumor depends on the generated surrounding stresses and need not be spherical [1].

To account for the mass-effect of the bulk tumor, we follow the work of Wasserman and Acharya [6] and assume that the expansive force of the neoplasm can be approximated by a constant outward pressure P acting on the tumor boundary. P is a parameter that determines the mass-effect of the bulk tumor, and therefore, to a large extent, the final tumor size.

Edema expansion in WM is mostly perpendicular to the direction of the fibers [7, 8]. Here, we assume no knowledge of WM fibers' orientation and an isotropic expansive strain e is applied to D_r by using analogy to thermal expansion. Studies of brain edema that measured a volume expansion of 200%–300% in WM [7, 8] imply $e \in [0.26, 0.44]$. For simulations starting with normal brain scans, a value of $e = 0.35$ is adopted.

Continuum Mechanics Problem Statement

Given the time scale of the tumor growth process, the deformation of brain tissues may be modeled as a quasi-static process. Additionally, if the effect of gravity is ignored, the required deformation map $\varphi : \kappa_r \rightarrow \kappa_t$ can be found by solving the static equilibrium equation $\text{Div}(\mathbf{S}) = 0$, where \mathbf{S} is the first Piola-Kirchhoff tensor which is related to strain via the material constitutive law [9].

Based on simulations using several of material constitutive laws for brain tissues suggested in the literature, we adopted the isotropic and homogeneous hyperelastic model proposed by Miller and Chinzei [10] while relaxing the perfect incompressibility assumption and ignoring viscous effects (since the time involved in tumor growth is much larger than the viscous time constants). Under these conditions, the strain energy density function of the material becomes [9]:

$$W = \frac{2\mu}{\alpha^2}(\bar{\lambda}_1^\alpha + \bar{\lambda}_2^\alpha + \bar{\lambda}_3^\alpha - 3) + \frac{1}{D_1}(J/J^{th} - 1)^2, \tag{1}$$

where $\bar{\lambda}_i = J^{-1/3}\lambda_i$, $\lambda_i, i = 1, 2, 3$ are the principal material stretches, $J = \det(\mathbf{F})$ is the volume ratio, \mathbf{F} is the deformation gradient, $J^{th} = (1 + e^{th})^3$ is the thermal volume ratio, and e^{th} is the thermal strain. The constants μ, D_1 are related to the Young's modulus at zero strain E_o , and Poisson's ratio ν by

$$\mu = \frac{E_o}{2(1 + \nu)} \quad \text{and} \quad D_1 = \frac{6(1 - 2\nu)}{E_o}. \tag{2}$$

The value $\alpha = -4.7$ determined in [10] was adopted here. Since the brain biomechanics literature includes varying accounts of brain tissue compressibility and stiffness, in the experiments described below, the effects of μ and D_1 (equivalently μ, ν) on the proposed model were investigated. The following boundary conditions (BCs) complete the statement of the problem (refer to Fig. 1):

$$\varphi(\mathbf{X}) \cdot \mathbf{N}(\mathbf{X}) = 0, \quad \mathbf{X} \in \partial B_{01} \quad \text{and} \quad \varphi(\mathbf{X}) = \mathbf{0}, \quad \mathbf{X} \in \partial B_{02} \tag{3}$$

$$e^{th}(\mathbf{X}) = e, \quad \mathbf{X} \in D_r \quad \text{and} \quad e^{th}(\mathbf{X}) = 0, \quad \mathbf{X} \in B_B \tag{4}$$

$$\mathbf{S}\mathbf{N}(\mathbf{X}) = P\mathbf{J}\mathbf{F}^{-T}\mathbf{N}(\mathbf{X}), \quad \mathbf{X} \in \partial B_T \tag{5}$$

$$\mathbf{S}\mathbf{N}(\mathbf{X}) = \mathbf{0}, \quad \mathbf{X} \in \partial B_V \tag{6}$$

where $\mathbf{N}(\mathbf{X})$ is the outward surface normal at \mathbf{X} in the relaxed configuration. Equation (3) implies a sliding BC over the brain surface except for locations where the falx meets the inner surface of the skull which are assumed pinned [11].

Equation (4) implies that the expansive strain due to edema is restricted to D_r . Equation (5) is the traction BC implied by the tumor pressure, expressed in terms on normals in κ_r . The ventricles are assumed void and eqn. (6) implies negligible intra-ventricular pressure [1].

3 Experiments and Results

Here, we provide partial validation results for the model by quantitatively comparing its predictions to the deformations observed in four brain tumor cases. The same dataset is used to guide the selection the material model parameters. Finally, a model simulation on an MRI of a normal subject is demonstrated.

Assuming that canine brain tissue properties and tumor growth process are reasonably representative of their counterparts in humans, three of the studied tumor cases were for dogs with surgically transplanted glioma cells [12] (DC1, DC2, DC3). A baseline scan was acquired before tumor growth, followed by scans on the 6th and 10th day post-implantation. Gadolinium-enhanced T1 MR images were acquired (MPRAGE for DC1, DC2). Tumors grow rapidly to a diameter of 1 – 2cm by the 10th day on which the animals were sacrificed, and prior to the presentation of significant neurological complications. The fourth dataset (HC) comes from a series of T1 MRI scans of a human with a low-grade glioma transforming into malignancy. Two scans were used with approximately 2.5 years in between. Increase in the tumor mass and significant swelling due to edema were observed. The dataset is described in Tab. 1 and example images are shown in Fig. 2.

To compare the model predictions to actual deformations in the available datasets, values of the model parameters (center, radii of T_r and D_r , and P) for these cases must be determined. To avoid optimizing all parameters for each tumor case, the first images after tumor development (6th day scan for dog cases, and the first scan for the HC) were used to approximate κ_r . This approximation involves the assumption of negligible edema spread and tumor infiltration

Table 1. Description of the image scans and results for the dog cases (DC1, DC2, DC3) and the human case (HC). Optimal values of the model parameters P and e for $\mu = 842Pa$ and $\nu = 0.485$ are provided. The number of landmarks, landmark deformation statistics (lmrk def: mean/max/std. dev.) and model residual errors (Error, mean/max/std. dev.) for the landmark points are reported.

	DC1	DC2	DC3	HC
Image dimensions	256x256x100	256x256x100	256x256x124	256x256x124
Voxel size, <i>mm</i>	0.39x0.39x0.7	0.39x0.39x0.7	0.47x0.47x1.0	0.94x0.94x1.5
P , Pa	8000	7000	15000	8000
e	0.3	0.15	0.4	0.3
Num. landmarks	25	21	20	21
lmrk def, <i>mm</i>	2.16/3.9/1.06	1.77/2.83/0.58	1.82/3.04/0.78	4.53/6.09/0.9
Error, <i>mm</i>	1.11/2.5/0.73	1.13/2.1/0.42	1.19/2.24/0.52	1.7/3.09/0.77

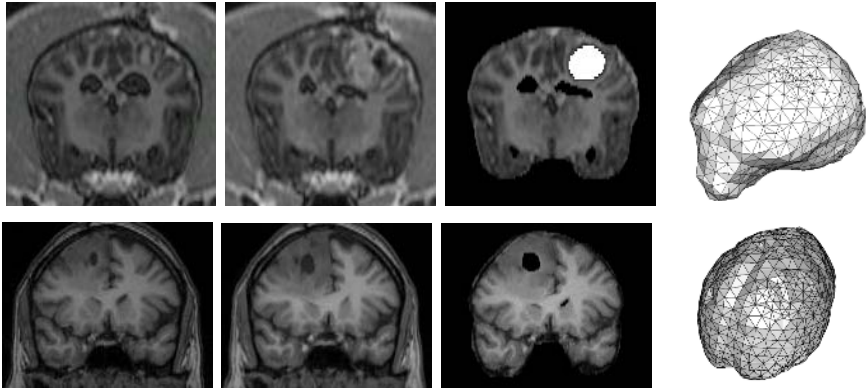


Fig. 2. Example cross sectional images from the starting (left column) and target (middle left column) 3D images for DC1 (upper row) and HC (lower row) compared to the deformed images obtained via the optimal model parameter values (middle right column). Tumors in simulated images are assigned similar intensities to the real images. The right column shows the outer surface of the used FE meshes.

between this scan and final scan, which corresponds to κ_t . Additionally, since tumors in the starting images are small, the stresses and deformation are assumed to be negligible. Under these assumptions, D_r and T_r are obtained from segmentations of the tumor and edema in the starting images. Since the starting images already had some edema, e was treated as a parameter with $e \in [0.1, 0.4]$.

Since loads in the proposed model are in the form of a pressure, P , and a prescribed strain e , the resulting deformation depends on e and the ratio P/E_o , but not P alone (confirmed by actual simulations). Given this, the value of $\mu = 842Pa$ suggested in [10] was chosen. Experiments were performed to determine ν , and P , e for each tumor case according to the following procedure.

First, rigid registration of the target (final) scan to the respective starting scan (used to approximate κ_r) was performed [13]. At least, 20 pairs of corresponding landmarks were manually identified by a human rater in the starting and target images. The landmarks were selected near the tumor, where large deformation occurs. A combination of manual and automatic segmentation of the starting images into brain, ventricles, falx, tumor and edema was then performed. A tetrahedral FE mesh was generated from the segmented images [14], and for each value of P , e , and ν , the FE simulation environment ABAQUS [9] was used to solve the continuum mechanics problem described above. The resulting deformation map φ for each simulation was used to deform the starting images and the locations of the landmark points in those images. Errors between deformed landmark coordinates and the corresponding rater’s coordinates in the target scan were computed.

Experiments with different values of ν , P , e were performed in the following sequence. With $\nu = 0.49$ [11] (which implies $E_o = 2109Pa$, $D_1 = 4.75e^{-5}Pa^{-1}$), e and $P \in [1, 16]KPa$ were varied for each case, the mean error in model predic-

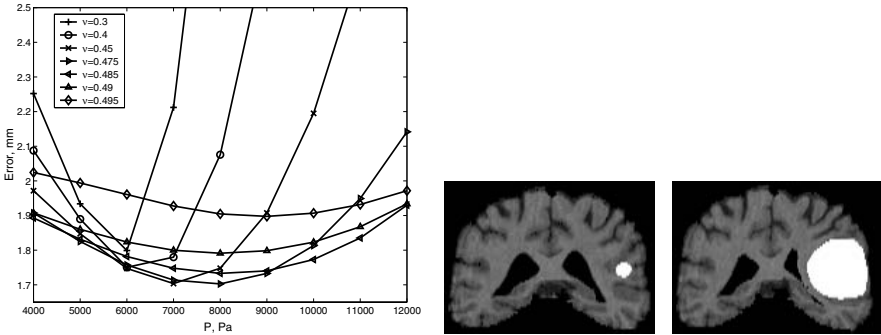


Fig. 3. Left: Mean residual errors between model predictions and deformed landmark points for HC for different values of P and ν with $e = 0.3$, $\mu = 842Pa$. Right: 2D slices through a T1-weighted MRI of a normal subject before and after simulation (radius of T_r was $5mm$, $P = 9kPa$ and no edema). The simulated tumor volume is $37cc$.

tions was computed, and the value of e for minimum error was recorded. With the value of e determined for each case, simulations were then run for $\nu \in [0.3, 0.499]$ and $P \in [1, 16]kPa$. The minimum mean error occurs for $\nu \in [0.475, 0.499]$ for all cases, which supports that the brain tissue is almost incompressible. We adopt the value $\nu = 0.485$ – near the middle of this range. With this value of ν and e determined above for each case, the optimal value of P was found. Final values of P and e and the corresponding residual errors in the model predictions are reported in Tab. 1. The residual errors for some individual landmark points were down to 15% of the deformation observed. For HC, the model is able to predict more than 62% of the deformation, on average. Simulated images using the optimal values of all parameters are compared to the real ones in Fig. 2. Due to space limitations, we provide an example of the error curves for only for HC for different values of ν in Fig. 3 where we also present the result of applying our model on a normal T1-weighted brain MRI.

To quantify the accuracy of the rater’s placement of landmarks in the target image scans, for DC1, landmarks in the target image were found by two raters. The mean inter-rater variability in this case was $1.12mm$ with a maximum of $3.29mm$ which are similar to the respective values obtained for residual model errors. In fact, the distribution of the inter-rater distances and the average residual errors (over the two raters) were statistically indistinguishable (p -value=0.95).

4 Discussion and Conclusions

A 3D FE model for simulating brain tumor mass-effect due to the bulk tumor and edema was presented. Comparison between the predictions of the model to deformations caused by four real brain tumors was performed. The results indicate that the residual errors are mostly caused by inaccuracies in the rater’s

tracking of the landmark points, as well as other errors such as modeling assumptions, errors in rigid registration and manual segmentation. In particular, in experiments on the real tumor cases, the use of segmentations of images at initial stages of the tumor to approximate the relaxed configuration makes the model unable to account for increase in infiltration and edema extent. Increase in tumor infiltration, particularly for the aggressive dog glioma tumors, may also be the reason behind perceived compressibility of brain tissues in these cases.

Future work includes investigating the relationship between model parameters and clinical variables, such as tumor type and grade. While cases studied in this paper did not involve contact between opposite walls of the ventricles, we plan to investigate the modeling of this phenomenon by using a FE contact condition, or filling the ventricles with a material that is softer than the brain tissues. The use of the presented model in establishing registration between normal brain images is presented in [5].

Acknowledgments

The authors would like to thank Dr. Nick Fox at the University College London, UK for providing the tumor patient's images, Dr. James Anderson, Carolyn Magee, Dr. Stelios Kyriacou at the Johns Hopkins University School of Medicine for preparing the dog tumor datasets. This work was supported in part by the National Science Foundation under Engineering Research Center grant EEC9731478, and by the National Institutes of Health grant R01NS42645.

References

1. Kyriacou, S.K., Davatzikos, C., Zinreich, S.J., Bryan, R.N.: Nonlinear elastic registration of brain images with tumor pathology using a biomechanical model. *IEEE Transactions on Medical Imaging* **18** (1999) 580–592
2. Mohamed, A., Kyriacou, S.K., Davatzikos, C.: A statistical approach for estimating brain tumor induced deformation. *Proc. IEEE Workshop on MMBIA* (2001) 52–59
3. Dawant, B.M., Hartmann, S.L., Pan, S., Gadamsetty, S.: Brain atlas deformation in the presence of small and large space-occupying tumors. *Computer Aided Surgery* **7** (2002) 1–10
4. Cuadra, M.B., Pollo, C., Bardera, A., Cuisenaire, O., Villemure, J.G., Thiran, J.P.: Atlas-based segmentation of pathological MR brains using a model of lesion growth. *IEEE Transactions on Medical Imaging* **23** (2004) 1301–1314
5. Mohamed, A., Shen, D., Davatzikos, C.: Deformable registration of brain tumor images via a statistical model of tumor-induced deformation. In: *Proc. of MICCAI 2005*. (2005)
6. Wasserman, R., Acharya, R.: A patient-specific in vivo tumor model. *Mathematical Biosciences* **136** (1996) 111–140
7. Kuroiwa, T., Ueki, M., Suemasu, H., Taniguchi, I., Okeda, R.: Biomechanical characteristics of brain edema: the difference between vasogenic-type and cytotoxic-type edema. *Acta Neurochir. Suppl.* **60** (1994) 158–161

8. Nagashima, T., Tamaki, N., Takada, M., Tada, Y.: Formation and resolution of brain edema associated with brain tumors. a comprehensive theoretical model and clinical analysis. *Acta Neurochir. Suppl.* **60** (1994) 165–167
9. Hibbitt, Karlsson, and Sorensen, Inc., USA: Abaqus version 6.4. (2003)
10. Miller, K., Chinzei, K.: Mechanical properties of brain tissue in tension. *Journal of Biomechanics* **35** (2002) 483–490
11. Miga, M., Paulsen, K., Kennedy, F.E., Hartov, A., Roberts, D.: Model-updated image-guided neurosurgery using the finite element method: Incorporation of the falx cerebri. In: *MICCAI 1999*. (1999)
12. Wodinsky, I., Kensler, C., Roll, D.: The induction and transplantation of brain tumors in neonatal beagles. *Proc. Am. Assoc. Cancer Res.* **10** (1969)
13. Studholme, C., Hill, D.L.G., Hawkes, D.J.: An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition* **32** (1999)
14. Mohamed, A., Davatzikos, C.: Finite element mesh generation and remeshing from segmented medical images. In: *Proc. of the 2004 IEEE ISBI, Arlington, VA* (2004)

STREM: A Robust Multidimensional Parametric Method to Segment MS Lesions in MRI

L.S. Ait-Ali¹, S. Prima¹, P. Hellier¹, B. Carsin², G. Edan², and C. Barillot¹

¹ IRISA Campus Universitaire Beaulieu 35042 Rennes Cedex, France

² CHRU Pontchaillou 2 rue Henri Le Guilloux 35000 Rennes, France
laure.ait-ali@irisa.fr

Abstract. We propose to segment Multiple Sclerosis (MS) lesions over-time in multidimensional Magnetic Resonance (MR) sequences. We use a robust algorithm that allows the segmentation of the abnormalities using the whole time series simultaneously and we propose an original rejection scheme for outliers. We validate our method using the BrainWeb simulator. To conclude, promising preliminary results on longitudinal multi-sequences of clinical data are shown.

1 Introduction

Multiple Sclerosis (MS) is an inflammatory autoimmune disorder of the central nervous system. MS is a disease with a high variability in symptoms, clinical course, neuro-radiological appearance of the lesions, response to therapy and pathological mechanisms [1].

Nowadays, MRI is widely used for disease diagnosis, patient follow-up, validation of therapies, and more generally for the understanding of MS. For instance, Gadolinium (Gd) enhanced areas in T1-weighted (T1-w) images and hyper-intense areas in T2-weighted (T2-w) and PD-weighted (PD-w) images are often considered as markers of clinical relapses, whereas brain atrophy (evaluated in T1-w or T2-w images) and hypo-intense areas in T1-w images are often seen as markers of the neuroaxonal loss [2]. These measures complement and enrich the clinical observations based on scales such as the expanded disability status scale (EDSS) or the MS functional composite (MSFC) whose drawbacks are numerous. Clinical observations are inherently subjective, show poor inter- and intra-rater reliability, have limited sensitivity, etc. These drawbacks are particularly critical when dealing with hundreds of subjects as in phase III clinical trials. On the other side, such large-scale studies, composed of multiple subjects, image modalities, time points, and acquisition centers, require automated image processing pipelines for the efficient computation of MR markers.

Many statistical approaches have been developed for the automatic segmentation of brain structures in MR data. A particularly convenient framework consists in modeling the distribution of the intensities as a finite mixture of Gaussians. Such a Gaussian Mixture Model (GMM) allows the modelization of the image intensities with a reduced number of parameters. These parameters can then be estimated within a well-defined statistical framework based on the

Maximum Likelihood Estimator (MLE) using optimization methods such as the Expectation-Maximization (EM) algorithm. Such an approach is unfortunately unadapted to pathological cases characterized by abnormal intensities in MRI. Some authors have proposed modifications of this classical approach to handle outlying intensities caused by MS lesions. Schroeter *et al.* [3] add an uniform component into the GMM to model lesions intensities. Van Leemput *et al.* [4] introduce a weight reflecting the degree of typicality of each brain voxel, the latter being consider as an outlier if its Mahalanobis distance with respect to each mixture component is greater than a threshold. Dugas-Phocion *et al.* use the Mahalanobis distance directly within the EM [5]. All these “robust” EM algorithms are computed for multi-sequences but only for one time point. On the other hand, few methods have been proposed to study the MS evolution over time [6, 7, 8, 9] using statistical tests or deformation fields. Shahar *et al.* [10] propose a combination of these two approaches for the spatio-temporal segmentation of lesions. They perform both lesions segmentation at one time point and lesions evolution over time from 2D T2-w images using a multifeature EM algorithm.

We propose an original multidimensional (multiple sequences and time points) robust algorithm using $3D + t$ MR data to segment MS lesions over time in a standardized clinical protocol. Our method consists in originally modeling the multidimensional problem. The MLE is replaced by a robust estimator, the Trimmed Likelihood Estimator (TLE). We propose an iterative scheme to compute this estimator. In the last step, we refine the segmentation using both the Mahalanobis distance and prior information coming from clinical knowledge on lesion appearance across sequences. The method is described in Section 2, and the experimental results are presented and discussed in Section 3. Finally, we conclude and address future work in Section 4.

2 Method

We first detail a classical algorithm for the segmentation of multi-sequence MR data (see Section 2.1). Then, we present several adaptations for the robust segmentation of MS lesions in multidimensional MR data (see Section 2.2).

2.1 A Classical Parametric Multi-sequence Segmentation Method

Modelization stage. MRI noise is known to follow a Rician distribution, which can be fairly approximated by a Gaussian distribution [11]. Assuming that the distribution of intensities within each brain structure is also well approximated by a Gaussian law, it is common to consider that the image intensities are generated by a Gaussian Mixture Model (GMM). When considering m different sequences simultaneously, each voxel i , $i \in \{1, \dots, n\}$, is described by a m -D intensity vector $\mathbf{y}_i = (y_i^1, \dots, y_i^m)$, which can be modeled by a m -D GMM defined by:

$$p(\mathbf{y}_i; \theta) = \sum_{j=1}^k \alpha_j (2\pi)^{-m/2} |\Sigma_j|^{-1/2} \cdot \exp\left(-\frac{1}{2}(\mathbf{y}_i - \mu_j)^T \Sigma_j^{-1}(\mathbf{y}_i - \mu_j)\right), \quad (1)$$

k is the number of mixture components and $\theta = (\alpha, \mu, \Sigma)$ is the hyper-parameter of the GMM where: (i) $(\alpha_j)_{j \in \{1, \dots, k\}}$ are the mixture proportions, (ii) $(\mu_j)_{j \in \{1, \dots, k\}}$ are the mean vectors, (iii) $(\Sigma_j)_{j \in \{1, \dots, k\}}$ are the covariance matrices.

In MRI, three structures have relatively homogeneous intensities in the intracranial cavity, and constitute the components of the GMM: White Matter (WM), Grey Matter (GM), Cerebro-Spinal Fluid (CSF). The aim is to give a label x_i to each voxel vector \mathbf{y}_i , with $x_i \in \{c_1, \dots, c_k\}$ and c_j being one of the three GMM components. The output $X = (x_1, \dots, x_n)$ is the desired image segmentation of the MR data $Y = (\mathbf{y}_1, \dots, \mathbf{y}_n)$.

Estimation stage. The computation is based on the ML estimation of the GMM hyper-parameter: $\hat{\theta} = \arg \max_{\theta} \prod_{i=1}^n p(\mathbf{y}_i; \theta)$. It can be seen as an *incomplete data problem*, since neither the voxels memberships x_i nor the Gaussian mixture parameters are known but only the voxel intensities y_i .

Optimization stage. The EM algorithm [12] is a well established method to tackle this type of problems. It consists in iterating two steps until convergence: labelization of the image based on a prior knowledge of the unknown parameters (Expectation step) and estimation of the mixture proportions and class parameters based on this labelization (Maximization step).

Practically, this EM algorithm is not directly ran on the m MR images, but rather on a m -D joint histogram built from the m images, which is computationally much more efficient.

Classification stage. The final probabilities $P(x_i = c_j | Y; \hat{\theta})$ are a natural output of the EM algorithm and are used to compute the segmentation X .

2.2 The Proposed Spatio-Temporal Robust EM: STREM

Spatio-temporal model. For the patient follow-up, the longitudinal study is essential. In this context, we propose to perform a multi-sequence segmentation with time points $(1, \dots, t)$ with a robust estimation approach. The whole dataset can then be named $Y = \{Y^{(1)}, \dots, Y^{(t)}\}$, $Y^{(l)} = (\mathbf{y}_1^{(l)}, \dots, \mathbf{y}_n^{(l)})$ being the multi-sequence MR dataset at time point l . Once the images are preprocessed (*cf.* Section 3), a m -D joint histogram is computed for each time-point and these histograms are merged into a global m -D joint histogram containing the information from all sequences and all time-points. As all time points are merged, we omit the l index for the vectors \mathbf{y}_i in the following paragraphs.

Trimmed Likelihood Estimator (TLE). In order to make the classification less sensitive to noise, registration error, partial volume effects (PVE) and last but not least pathological abnormalities (*e.g.* MS lesions) which do not fit well the GMM model, we propose to replace the MLE by the TLE proposed by Neykov *et al.* [13].

The TLE is defined by: $\tilde{\theta} = \arg \max_{\theta} \prod_{i=1}^h p(\mathbf{y}_{\nu(i)}; \theta)$, where $\nu(i)$ are the permutations of $i = \{1, \dots, n\}$ such as $r_{\nu(i)} \leq r_{\nu(i+1)}$, $r_i = -\log p(\mathbf{y}_i; \theta)$ being the residuals. The TLE and the MLE are related to each other in the same way as the LTS (Least Trimmed Square) and the LS (Least Square) criteria.

Optimization. We propose an iterative scheme to estimate $\tilde{\theta}$. We name this scheme the STREM algorithm. This scheme converges at least to a local minimum of the criterion [13]. It can be summarized in the following way:

1. computation of the MLE on the whole data set using an EM giving a first estimation of the hyper-parameter $\hat{\theta}$;
2. sorting of the residuals : $r_i = -\log p(\mathbf{y}_i; \hat{\theta})$;
3. computation of the MLE on the voxels with the h smallest residuals ($h > 50\%$ of data), giving a new estimation $\hat{\theta}'$ (still using an EM);
4. back to step 2 until convergence.

Classification. Once healthy tissue model parameters are robustly estimated using STREM, we extract MS lesions as outliers of this model using the Mahalanobis distance (STREM-1), then we refine the segmentation using MS *a priori* information (STREM-2). We term the lesions class c_{k+1} .

1. For each vector \mathbf{y}_i in the joint histogram, the Mahalanobis distance

$$d_{i,j} = (\mathbf{y}_i - \mu_j)^T \Sigma_j^{-1} (\mathbf{y}_i - \mu_j)$$

between each class c_j is computed and gives a measure of how \mathbf{y}_i fits the model. These distances follow a chi-square distribution with m degree of freedom (χ_m^2). For a vector \mathbf{y}_i , if the Mahalanobis distance between each class j is greater than the critical value of χ_m^2 distribution for a given p-value, then the vector is considered as an outlier and belongs to c_{k+1} .

2. This potential lesions set still includes partial volume effects, noise or vessels. To discriminate real lesions from these false positives, we bring some *a priori* knowledge about MS and its manifestations in MRI. This *a priori* knowledge is summarized in Tab.1, where -1 , $+1$ and 0 denote respectively hyposignal, hypersignal and normally appearing with respect to the WM. Among the outliers extracted using the Mahalanobis distance those which follow the rules in Tab.1 belong to c_{k+1} .

Finally, the segmented images $X = (X^{(1)}, \dots, X^{(t)})$, $X^{(l)} = (x_i^{(l)}, x_i^{(l)} \in \{c_1, \dots, c_{k+1}\})$ being the segmentation at time point l , are constructed.

The advantage of such a multidimensional analysis is twofold: from a computational point of view, it increases the signal-to-noise ratio of the dataset and thus gives a more precise segmentation; from a clinical point of view, we use the information from various sequences, and we study the temporal variations.

Table 1. *A priori* knowledge about MS. Inflammatory MS lesions are hyper-intense with respect to the WM in T2-w and PD-w images. Young (4 to 6 weeks) lesions are hyper-intense in Gd T1-w images. Necrosis MS lesions are hypo-intense in T1-w images.

	T1-w	Gd T1-w	T2-w	PD-w
young inflammatory lesions	0	+1	+1	+1
others inflammatory lesions	0	0	+1	+1
necrosis	-1	0	+1	+1

3 Validation and Clinical Experimentation

Before applying the STREM algorithm, we extract the intra-cranial contents using a level set method [14]. Then, a bias field correction is done to reduce the radio-frequency inhomogeneities [15]. Finally, the various sequences and time points are rigidly registered [16], and global intensity changes between consecutive scans are corrected [17].

We first quantitatively validate the method using the BrainWeb database. Then, we show preliminary results on real data.

3.1 Quantitative Validation Using BrainWeb

BrainWeb is used for validation purposes. We simulate T1-w, T2-w and PD-w synthetic 3D MR images [18] with a 1mm slice thickness and 3% noise level. We simulate a set of healthy images (set A) and a set of MS images (set B).

The Dice Similarity Coefficient (DSC) is used to compare segmentations. Given two targets R_1 and R_2 , the DSC is defined by:

$$DSC(R_1, R_2) = 2 \cdot \text{card}(R_1 \cap R_2) / \text{card}(R_1 + R_2).$$

First, we prove the impact of using multiple sequences and, second, the impact of using longitudinal data. The third paragraph shows the improvement brought by the two previously proposed refinements (STREM-1 and -2).

Impact of multi-sequence analysis. We previously stressed the importance of such a multi-sequence analysis from a clinical point of view in MS study. Here, we focus on the computational improvement of using multiple sequences even for the segmentation of healthy tissues (set A). We empirically show that using several sequences improves the segmentation doing three experiments:

- (i) mono-sequence STREM on T2-w giving X_{T2} and on PD-w giving X_{PD} ;
- (ii) multi-sequence STREM on T2-w and PD-w giving $X_{T2,PD}$;
- (iii) multi-sequence STREM on the whole set A giving $X_{T1,T2,PD}$.

The obtained DSCs are compared and show that (i) < (ii) < (iii) (see Tab. 2). Values for CSF are not mentioned because it is not well segmented using only the PD-w sequence. We use $h = 96\%$.

Impact of the longitudinal analysis. The longitudinal analysis allows to detect both MS lesions and their evolution. Our method treats in a unified manner the segmentation and tracking of MS lesions. That is to say, the segmentation at a given point is used to allow better tracking of lesions over time, and, conversely, all the information over time is used to allow better segmentation of lesions at a given time point.

As shown in the previous section, using three sequences jointly (T1-w, T2-w and PD-w) improves the segmentation. In this second experiment, we simulate a longitudinal dataset with two time points. The first time point is the set A, $Y_A = Y_{T1,T2,PD}$. The second time point is the set B, $Y_B = Y_{T1_{MS},T2_{MS},PD_{MS}}$. The following experiments are done:

- (i) multi-sequence STREM-1 on Y_A giving $X_A (= X_{T1, T2, PD})$;
- (ii) multi-sequence STREM-1 on Y_B giving X_B ;
- (iii) multidimensional STREM-1 on $Y = (Y_A, Y_B)$ giving $X = (X_{A/B}, X_{B/A})$.

MS lesions are rejected by STREM-1: $h = 96\%$ and $p_value = 0.025$.

We compare the DSCs between multi-sequence segmentation and multidimensional segmentation for Y_A and Y_B . The multidimensional (multi-sequence and longitudinal) analysis improves the segmentation (see Tab. 3).

Table 2. Set A. The DSCs increase when using several sequences simultaneously for the hyper-parameter estimation.

	WM	GM
X_{T2}	0.86	0.84
X_{PD}	0.88	0.83
$X_{T2, PD}$	0.93	0.88
$X_{T1, T2, PD}$	0.95	0.89

Table 3. Set A and B. The DSCs increase with the number of time points, particularly for all tissues in set A and MS lesions.

	WM	GM	CSF	lesions
X_A	0.95	0.89	0.56	-
$X_{A/B}$	0.96	0.95	0.65	-
X_B	0.95	0.90	0.59	0.48
$X_{B/A}$	0.95	0.90	0.58	0.75

Impact of refinements. We perform the three following experiments:

- (i) STREM on $Y = (Y_A, Y_B)$ with $h = 96\%$;
- (ii) STREM-1 on Y with $h = 96\%$ and $p_value = 0.025$;
- (iii) STREM-2 on Y with $h = 96\%$ and $p_value = 0.025$.

Table 4. DSC measures for $X_{B/A}$. The DSC of MS lesions increases with the refinements.

	WM	GM	CSF	lesions
STREM	0.95	0.90	0.58	0.58
STREM-1	0.95	0.90	0.58	0.75
STREM-2	0.95	0.90	0.58	0.76

For STREM, we consider as outliers the vectors that have the bigger residuals. STREM-1 performs better than STREM in terms of DSC because it is more specific. STREM-2 is still more specific than STREM-1 without being less sensitive and thus performs better. The increment is not obvious on this data because there are only few outliers which are not MS lesions (*cf.* Tab.4).

3.2 Results on a Real Multidimensional Data Set

To be more complete, we present a robust segmentation on a multidimensional clinical data set. This multidimensional MR dataset was acquired on a GE 1.5T scanner and corresponds to a patient follow-up every three months during one year. Each acquisition is composed of three sequences: PD-w (TR = 2740 ms, TE

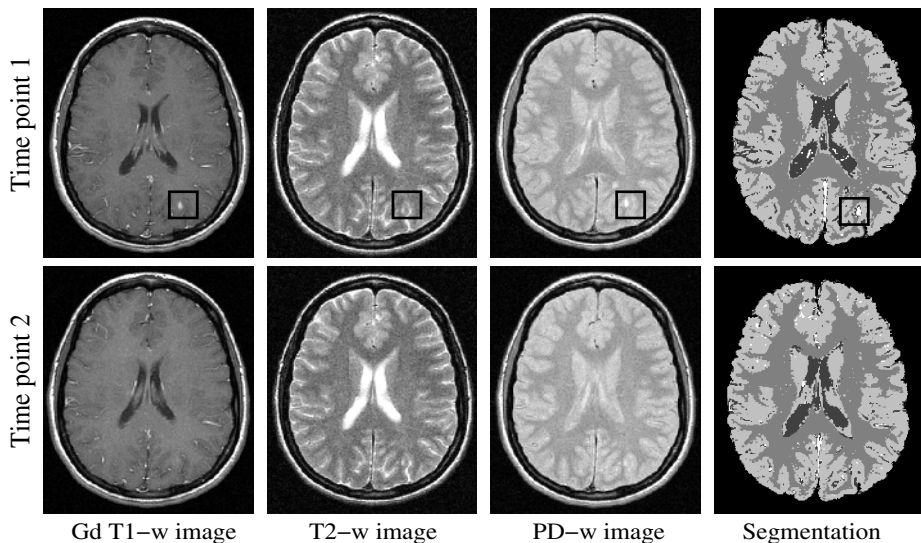


Fig. 1. STREM-1 on clinical data set ($h = 95\%$ and $p_value = 0.01$). A young lesion appears at time point 1 and disappears at time point 2. Preliminary results for multidimensional segmentation show the detection of a small and evolving lesion.

= 22.9 ms), T2-w (TR = 2740 ms, TE = 91.7 ms), and T1-w images after gadolinium injections (TR = 460 ms, TE = 20.0 ms), with a slice thickness of 3mm.

The STREM-1 is ran on the three sequences and the five time points simultaneously. Results for the first two time points are presented (there is no more evolution after). The refinement based on lesion intensity rules is not yet computed which explains the false positives (localized in the CSF). Even considering its small size, a lesion is segmented at time point 1 (see Fig. 1). This result is very promising because even very small structures are recovered. In addition, these small and evolutive lesions correspond to early stage of the pathology where MR lacks specificity. This encouraging result should even be improved by applied the lesion intensity rules to make the segmentation even more specific.

4 Conclusion

We proposed a consistent robust multidimensional iterative scheme to segment MS lesions over time. The advantages of this method are numerous. Using several sequences and time points increases the signal to noise ratio and thus improves the hyper-parameter estimation. From a clinical point of view, such a multi-sequence approach allows to take into account all types of MS lesions, while using the longitudinal data allows to handle the segmentation of static and evolving MS lesions in an unified manner. Our robust estimation scheme, coupled with refinements steps using Mahalanobis distances and a priori knowledge allows to differentiate MS lesions from other voxels showing abnormal intensities.

In the future, we will validate our algorithm on larger clinical databases, by comparing our classifications of lesions with ones manually segmented by an expert according to their type (young, inflammatory, necrosis) and their evolution characteristics over time.

References

1. C. Lucchinetti *et al.* Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination. *Annals of Neurology*, 47(6):707–717, June 2000.
2. D.H. Miller. Biomarkers and surrogate outcomes in neurodegenerative disease: lessons from multiple sclerosis. *NeuroRx*, 1(2):284–294, April 2004.
3. P. Schroeter *et al.* Robust Parameter Estimation of Intensity Distributions for Brain Magnetic Resonance Images. *IEEE TMI*, 17(2):172–186, 1998.
4. K. Van Leemput *et al.* Automated segmentation of multiple sclerosis lesions by model outlier detection. *IEEE TMI*, 20(8):677–688, 2001.
5. G. Dugas-Phocion *et al.* Hierarchical segmentation of multiple sclerosis lesions in multi-sequence MRI. In *ISBI'04*, Arlington, USA, April 2004.
6. D. Welti *et al.* Spatio-temporal segmentation of active multiple sclerosis lesions in serial MRI data. In *IPMI'2001*, Davis, USA, June 2001.
7. M. Bosc *et al.* Automatic change detection in multimodal serial MRI: application to multiple sclerosis lesion evolution. *NeuroImage*, 20(2):643–656, October 2003.
8. S. Prima *et al.* Multivariate statistics for detection of MS activity in serial multimodal MR images. In *MICCAI'2003*, Montreal, Canada, November 2003.
9. D. Rey *et al.* Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis. *Medical Image Analysis*, 6(2):163–179, June 2002.
10. A. Shahar *et al.* A probabilistic framework for the detection and tracking in time of multiple sclerosis lesions. In *ISBI'04*, Arlington, USA, April 2004.
11. J. Sijbers *et al.* Maximum Likelihood estimation of Rician distribution parameters. *IEEE TMI*, 17(3):357–361, 1998.
12. A. Dempster *et al.* Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39(1):1–38, 1977.
13. N. Neykov *et al.* Mixture of GLMs and the trimmed likelihood methodology. In *Compstat'04*, pages 1585–1592, 2004.
14. C. Baillard *et al.* Segmentation of brain 3D MR images using level sets and dense registration. *Medical Image Analysis*, 5(3):185–194, 2001.
15. S. Prima *et al.* Maximum likelihood estimation of the bias field in MR brain images: investigating different modelings of the imaging process. In *MICCAI'2001*, Utrecht, The Netherlands, October 2001.
16. F. Maes *et al.* Multimodality image registration by maximization of mutual information. *IEEE TMI*, 16(2):187–198, April 1997.
17. A. Guimond *et al.* Three-Dimensional Multimodal Brain Warping using the Demons Algorithm and Adaptive Intensity Corrections. *IEEE TMI*, 20(1):58–69, 2001.
18. C.A. Cocosco *et al.* BrainWeb: Online Interface to a 3D MRI Simulated Brain Database. *HBM'97*, 5(4), 1997.

Cross Validation of Experts Versus Registration Methods for Target Localization in Deep Brain Stimulation^{*}

F. Javier Sánchez Castro¹, Claudio Pollo^{1,2}, Reto Meuli³,
Philippe Maeder³, Meritxell Bach Cuadra¹, Olivier Cuisenaire¹,
Jean-Guy Villemure², and Jean-Philippe Thiran¹

¹ Signal Processing Institute, École Polytechnique Fédérale de Lausanne (EPFL),
CH-1015 Lausanne, Switzerland
`francisco.sanchezcastro@epfl.ch`

Departments of

² Neurosurgery and

³ Radiology, Centre Hospitalier Universitaire Vaudois (CHUV), Switzerland

Abstract. In the last five years, Deep Brain Stimulation (DBS) has become the most popular and effective surgical technique for the treatment of Parkinson's disease (PD). The Subthalamic Nucleus (STN) is the usual target involved when applying DBS. Unfortunately, the STN is in general not visible in common medical imaging modalities. Therefore, atlas-based segmentation is commonly considered to locate it in the images. In this paper, we propose a scheme that allows both, to perform a comparison between different registration algorithms and to evaluate their ability to locate the STN automatically. Using this scheme we can evaluate the expert variability against the error of the algorithms and we demonstrate that automatic STN location is possible and as accurate as the methods currently used.

1 Introduction

Deep Brain Stimulation (DBS) involves implantation of an electrode deep inside the brain. This electrode delivers electric current to specific brain cells shutting down parkinsonian symptoms. After hundreds of surgical interventions, The Subthalamic Nucleus (STN) has turned out to be the most effective target for DBS. A typical DBS procedure starts with the placement of the stereotactic head frame, fixed to the patient's skull, that will be used as a coordinate reference. Next, an imaging study is taken in order to select pre-operatively the target to be stimulated and to plan the trajectories for introducing the electrodes. Usually two kind of images are taken to be able to visualize different tissues, MR T1-weighted and MR T2-weighted images. In our state-of-the-art protocol, the selection of the STN target is performed on a coronal T2-weighted image acquired perpendicularly to the AC-PC axis and crossing the anterior limit of

^{*} This work is supported by the Swiss National Science Foundation under grant number 205320-101621.

the Red Nucleus. STN target selection depends on each institution. Common methods are the use of a stereotactic atlas of the brain [1] and the use of visible surrounding anatomical landmarks [2]. Then, the target coordinates are reported to a T1-weighted image where the trajectories are planned. Once in the operating room the head frame is fixed to the operating table, a small hole is drilled into the patient's skull according to the pre-operative trajectories. Because of the difficulty to directly see the STN on medical images, the selected pre-operative target is only an estimation of the real location. Therefore the electrode's location has to be adjusted intra-operatively using electrophysiological recordings and from stimulation tests.

This paper focuses on three main goals. First of all, the construction of a reference which is used as a ground truth for the position of the STN. Secondly, the study of the *intra* and *inter*-expert variability in identifying the STN location visually. Thirdly, to evaluate the possibility of automatically locating the STN using existing registration techniques and to compare their performance and usability. The construction of a STN location ground truth from the experts knowledge is as follows. First, a patient is chosen as atlas and each individual patient mapped back to this atlas using various registration methods. Then, an estimation of the STN location is obtained and compared to the real location given by the ground truth. Using accurate registration algorithms we demonstrate that automatic STN localization is possible and as accurate as the methods currently used. As far as we know, only one study to evaluate a standard mutual information-based non-rigid registration algorithm for automatically locating the STN has been published [3]. As opposed to us, they use post-operative coordinates as a ground truth, making the assumptions that the surgical team is able to place the electrode within the STN and that the intra-operative guidance system provides the accurate position of the electrodes.

2 Materials and Methods

2.1 Data

For each patient of our bilaterally implanted parkinsonian patient's database (37 patients, 74 STNs) two kinds of images were acquired pre-operatively: a 3D T1-weighted magnetization-prepared rapid acquisition gradient echo (MPRAGE) MRI sequence (Siemens Vision[®], 1.5T, Erlangen, Germany) TR 9.7 ms, TE 4 ms, number of slices/slice thickness: 164/1.40 mm, FOV 280x280, matrix 256x256, pixel size 1.09x1.09 mm, and few coronal slices (due to the acquisition time required for this kind of imaging sequence) of an inversion recovery (IR) T2-weighted, TR 2,560 ms, TE 4 ms, number of slices/slice thickness: 7/3 mm, FOV 300x300, matrix 512x512, pixel size 0.59x0.59 mm. Taking profit from the fact that in some rare cases the STN is visible in MR T2-weighted images, a reference is constructed and used as a ground truth. To do this, neurosurgeon experts were asked to select patients with clearly visible STN in MR T2-weighted images amongst our patient's database. After exhaustive inspection 8 patients were selected (16 STNs).

2.2 Ground Truth and Validation Scheme

Two experts, one neurosurgeon and one radiologist, both with wide experience in Parkinson’s disease (PD) surgery and targeting, have been asked to click the target point (2 STN per patient) for each selected T2 series. This process has been repeated 5 times for each patient at different days to avoid that the experts be influenced by previous targeting choices. With this data we were able to compute statistics about *intra* and *inter*-expert variability. *Inter*-expert targeting differences have turned out to not be statistically significant as shown in section 3.1. Thus, we can use the two sets of points to compute statistical mean target point coordinates, called *real targets*, and which are considered the ground truth. Amongst the 8 selected subjects, the experts have selected the one with the most clearly visible STN as a reference subject, both for the right and left sides. These data together with the real targets allow us to consider our reference subjects as an atlas. Then, by non-rigidly registering the 7 other patient’s images with the atlas, we obtain a projection of the STN of the patient in the atlas, an estimation of the STN position given by the registration method used. The procedure is as follows. First, for each patient, we perform a T1-T2 rigid registration [4] and project the real target coordinates into the T1 space. Next, we apply each of the 4 registration algorithms under study to get an estimation of the position of the STN. This estimation is obtained by registering the MR T1-weighted image of the patient under study with the MR T1-weighted image of the atlas. Finally, repeating this for the 7 datasets (14 STNs) involved in this study (leaving out the reference subjects), Euclidean distances from estimated to real targets for each STN are calculated and statistics are extracted to evaluate the performance of the different methods.

2.3 Atlas-Based Targeting and Registration Algorithms

In this study, the following registration methods have been applied and tested because of their wide use in medical image processing:

- *Atlas-based (AC-PC) targeting.* AC-PC referential together with brain atlases is one of the methods used to target the STN in medical environment, mainly when STN is not clearly visible in MR T2-weighted images which is the usual case. Neurosurgeons have to locate the anterior and posterior commissures (AC-PC points). Then, using a stereotaxic atlas and taking as the origin the midcommissural point (MCP), this procedure estimates that the STNs are located at coordinates (following Schaltenbrand-Wahren [1]): anteroposterior (AP) $-3mm$, lateral (LAT) $\pm 12mm$ (left and right side) and vertical (VERT) $-4mm$.
- *Affine registration.* We used an independent implementation based on the work of Maes et al. [4]. The 12 degrees of freedom (translation, rotation, scaling and shearing) are optimized in order to maximize the mutual information between the images to be registered [5] using a multiscale approach and a two-step optimization. First a global search using genetic algorithms [6] and

next a local optimization using the steepest descent method [7]. Affine registration is also used as a pre-alignment step for non-rigid transformations described below.

- *Demons algorithm* is an intensity-based algorithm proposed by Thirion [8] and based on the concept of optical flow. The problem of image matching is approached as a diffusion process, in which object boundaries in the reference image F are viewed as semi-permeable membranes. The other (so-called floating) image G is considered as a deformable grid, and diffuses through these interfaces driven by the action of effectors situated within the membranes. In the case of voxel-by-voxel intensity similarity, the instantaneous displacement vector for each voxel is

$$\vec{d} = -\frac{(g-f)\vec{\nabla}f}{|\vec{\nabla}f|^2 + (g-f)^2},$$

where f and g are the intensity image of F and G respectively. The deformation algorithm is applied by iterating in a hierarchical coarse-to-fine multiscale way. The smoothness of the displacement field is imposed by smoothing with a Gaussian filter of standard deviation σ (*elasticity parameter*) chosen empirically [9]. In our case, parameter σ has been chosen by exhaustive search, between 0.6 and 2.0mm by steps of 0.2mm, minimizing the distance between the estimated STN and the real targets. Finally, σ of 1mm has been chosen.

- *B-splines algorithm*. It is a mutual information-based free-form deformation algorithm whose displacement field is modelled as a linear combination of B-splines lying in a regular grid (uniformly spaced control points) similar to the method proposed by Rueckert et al. in [10]. The deformation that maximizes the mutual information between the two images involved is computed at each grid point placed on the floating image. The transformation is propagated to the rest of the image using the standard B-spline expansion with cubic splines:

$$d(x) = \sum_{k \in \mathbb{Z}} c(k)\beta^3(x-k),$$

where $\beta^3(x) = \beta^0 * \beta^0 * \beta^0 * \beta^0(x)$, $c(k)$ are the B-spline coefficients and β^0 a rectangular pulse. To speed up the optimization process the algorithm has been implemented using the communication utilities for distributed memory architectures using the MPICH implementation of the Message Passing Interface (MPI) [11]. The good interpolation properties and the suitability for multiscale processing of the B-splines are well known [12] and its deformability can be controlled by changing the spacing between the control points of the grid which we have set at 12mm.

3 Results

3.1 Target Selection and Expert Variability

In order to evaluate the *repeatability* or *intra-expert* variability of the expert targeting, we have computed the centroid of each cloud of STN points targeted

Table 1. Expert variability statistics

(a) <i>Intra</i> -expert variability.			(b) <i>Intra</i> -expert variability. References.		
$m \pm s$	Expert 1	Expert 2	$m \pm s$	Left Ref.	Right Ref.
centroid	$1.06 \pm 0.61mm$	$0.80 \pm 0.52mm$	Exp. 1	$1.10 \pm 0.32mm$	$0.75 \pm 0.38mm$
			Exp. 2	$0.79 \pm 0.30mm$	$0.38 \pm 0.25mm$

by the expert and we have calculated the Euclidean distances from the centroid to each of these points, called *centroid variability*. In table 1(a) these statistics are shown for the two experts. These quantities allow us to get an idea of the surgeon variability and its accuracy when clicking over the pre-operative target. If we only consider the 2 STNs used as a reference we obtain a centroid variability that is shown in table 1(b) for the left and right sides respectively and for the different experts.

In order to construct our ground truth a paired *T-test* of the hypothesis that the target coordinates selected by each expert come from distributions with equal means has been performed over each coordinate (x, y, z) at a 1% significance level. The results show that the hypothesis can not be rejected. Therefore our reference can be considered as the mean of two cloud of points given by each expert. The *inter*-expert variability calculated as the Euclidean distances from each expert click to the ground truth gives a mean and unbiased standard deviation of $1.61 \pm 0.29mm$ and $1.40 \pm 0.38mm$ for the expert 1 and 2 respectively. These statistics have been obtained by generating one error per

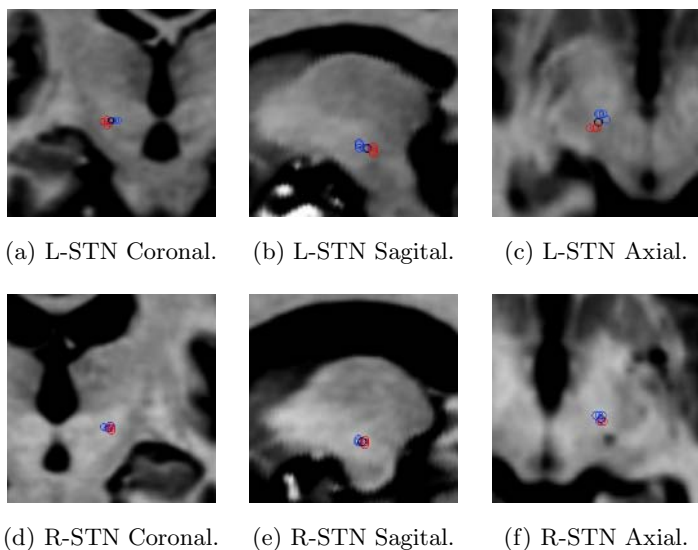
**Fig. 1.** Reference STN expert targeting

Table 2. Mean STN coordinates of the dataset and errors given by the methods

(a) STN coordinates referred to MCP.

Coordinates	<i>mean ± std</i>
AP	$-3.286 \pm 0.94mm$
LAT	$\pm 12.90 \pm 0.93mm$
VERT	$-3.23 \pm 0.78mm$
MCP-distance	$13.90 \pm 0.87mm$

(b) Estimation Errors.

Methods	<i>mean ± std</i>
Affine	$2.42 \pm 0.84mm$
Demons	$1.77 \pm 0.65mm$
B-Splines	$1.72 \pm 0.48mm$
AC-PC	$1.96 \pm 0.90mm$

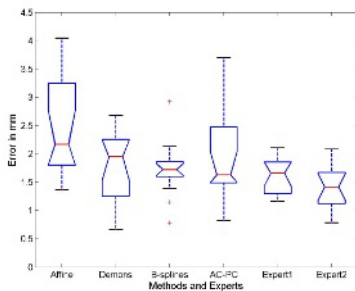
STN and per expert in the following manner. Each error is the Euclidean distance between the mean of the 5 points targeted by the expert and the ground truth.

In figure 1 the points targeted by the experts for the left (figures 1(a), 1(b), 1(c)) and right (figures 1(d), 1(e), 1(f)) STNs chosen as a reference can be seen. In blue the targeting of the expert 1, in red the targeting of the expert 2 and in black the mean point used as a reference. For visualization purposes, each point has been projected onto the three orthogonal planes passing through the centroid and showed using a circle (radius of 1mm).

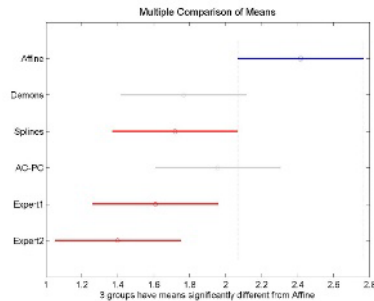
We can also decompose these distances in anteroposterior (AP), lateral (LAT) and vertical (VERT) coordinates which allows comparing directly the mean STN location for our database with usual STN coordinates from the atlases (e.g. [1]). In table 2(a) we show the mean and standard deviation of the coordinates referred to the MCP for the 16 STNs used in this study as well as their mean distance to the MCP.

3.2 Target Estimation and Evaluation of the Methods and Experts

The statistics, mean and unbiased standard deviation, of the errors committed when applying the 4 methods under study to locate the STN as described in 2.3 are shown in table 2(b). In order to compare the results statistical tests were performed. A one-way analysis of variance (ANOVA) of the hypothesis that the



(a) ANOVA statistical box plot.



(b) Multi-comparison test.

Fig. 2. Statistical tests of the errors committed using different methods and by the experts (using the *anova1* and *multcompare* functions of MATLAB[®])

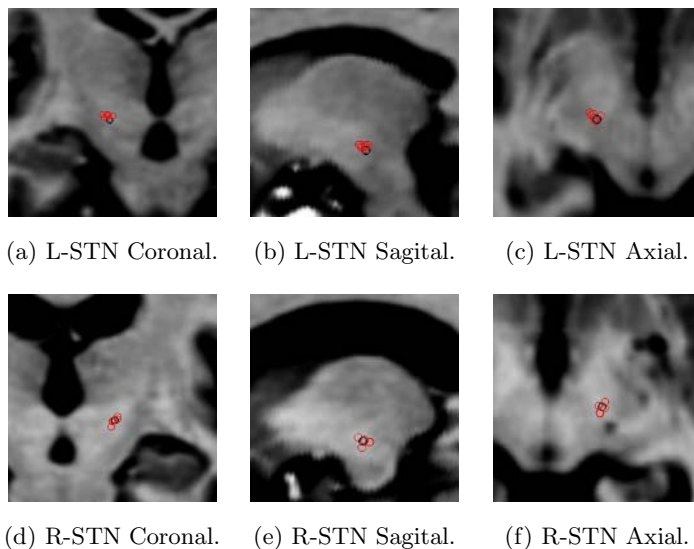


Fig. 3. STN estimation using B-splines

errors come from distributions with equal means has been performed over the errors produced by each method and by the experts at a 5% significance level. In figure 2(a) a statistical box plot produced by this test is shown as well as the result of a multi-comparison test of the means (see figure 2(b)). Two key conclusions can be drawn from these results. First that the errors committed with B-splines, demons and AC-PC based methods and the errors committed by the experts are not significantly different. Secondly that the results show that affine registration is significantly different from the B-splines method and from the experts. In figure 3 we show the projection of each STN estimation (in red) onto the reference subject (in black) using the B-splines registration algorithm. Each point is represented by a circle of $1mm$ of radius whose coordinates are projected onto the three orthogonal planes passing through the reference subject point coordinates (in black) in order to visualize the points in each view. The estimated targets are located very close to the *real target* and forming tight clouds of points showing that this kind of automatic estimation is reliable and well suited for this application.

4 Discussion and Perspectives

The main conclusion one can extract from these results is that automatic STN location is possible and accurate. As we can see by simple inspection of the numerical results, the B-splines method shows the best performance with the smallest mean error and unbiased standard deviation but closely followed by demons and AC-PC methods. The points project on tight clusters showing the

robustness of this kind of estimation method. The statistical tests have shown that global affine registration is not enough for our application while there are no significant differences between the errors committed using the B-splines, Demons or AC-PC referential-based techniques and, most importantly, that there are no significant differences between the errors committed with these three techniques and by the experts. Although the AC-PC referential-based method shows an acceptable performance its estimation needs AC and PC point's identification by an expert and does not take into account inter-patient variability, which is very important at a single patient level. Although the choice of the STN of reference can influence the results, the methods have been tested using 6 different combinations of STN-pairs (left and right references) and the results were similar. The automatic estimation of the STN can also be used as a first and fast pre-operative target estimation that can be refined by the neurosurgeon criterion. In a near future, more registration methods, mainly local ones, and experts will be added to this study. Moreover, further work will include the construction of a post-mortem atlas of the STN which will provide a ground truth without the variability produced by expert targeting over a point.

References

1. Schaltenbrand, G., Wahren, W.: Atlas for Stereotaxy of the Human Brain. 2 edn. Thieme (1977)
2. Pollo, C. et al.: Subthalamic nucleus deep brain stimulation for parkinson's disease: Magnetic resonance imaging targeting using visible anatomical landmarks. In: Proceedings of the Meeting of the American Society for Stereotactic and Functional Surgery. Volume 80. (2003) 1–4
3. DHaes, P. F. et al.: Toward the creation of an electrophysiological atlas for the pre-operative planning and intra-operative guidance of deep brain stimulators (DBS) implantation. In: MICCAI. (2004) 729–736
4. Maes, F. et al.: Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging* **16** (1997) 187–198
5. Viola, P., Wells, W.M.: Alignment by maximization of mutual information. In: Proceedings of the Fifth ICCV, IEEE Computer Society (1995) 16
6. Wall, M.: GALib 2.4.6: A C++ Library of Genetic Algorithm Components. (Massachusetts Institute of Technology) <http://lancet.mit.edu/ga/>.
7. Press, W., Teukolsky, S., Vetterling, W., Flannery, B.: Numerical Recipes in C, The art of scientific computing. (Cambridge University Press)
8. Thirion, J.P.: Image matching as a diffusion process: an analogy with maxwell's demons. *MIA* **2** (1998) 243–260
9. Bach, M.: Atlas-based Segmentation and Classification of Magnetic Resonance Brain Images. PhD thesis, EPFL, CH-1015 (2003) Thesis Num. 2875.
10. Rueckert D. et al.: Nonrigid registration using free-form deformations: Application to breast mr images. *IEEE Transactions on Medical Imaging* **18** (1999) 712–721
11. Gropp, W., Lusk, E., Skjellum, A.: Using MPI: Portable Parallel Programming With the Message-Passing Interface. MIT Press (1999)
12. Unser, M.: Splines: A perfect fit for signal and image processing. *IEEE Signal Processing Mag.* **16** (1999) 22–38

Localization of Abnormal Conduction Pathways for Tachyarrhythmia Treatment Using Tagged MRI

G.I. Sanchez-Ortiz¹, M. Sermesant², K.S. Rhode², R. Chandrashekar¹,
R. Razavi², D.L.G. Hill³, and D. Rueckert¹

¹ Imperial College London, U.K

² King's College London, U.K.

³ University College London, U.K.

Abstract. Tachyarrhythmias are pathological fast heart rhythms often caused by abnormally conducting myocardial areas (foci). Treatment by radio-frequency (RF) ablation uses electrode-catheters to monitor and destroy foci. The procedure is normally guided with x-rays (2D), and thus prone to errors in location and excessive radiation exposure. Our main goal is to provide pre- and intra-operative 3D MR guidance in XMR systems by locating the abnormal conduction pathways. We address the inverse electro-mechanical relation by using motion in order to infer electrical propagation. For this purpose we define a probabilistic measure of the onset of regional myocardial activation, derived from 3D motion fields obtained by tracking tagged MR sequences with non-rigid registration. Activation isochrones are then derived to determine activation onset.

We also compare regional motion between two different image acquisitions, thus assisting in diagnosing arrhythmia, in follow up of treatment, and in determining whether the ablation was successful. Difference maps of isochrones and other motion descriptors are computed to determine abnormal patterns. Validation was carried out using an electro-mechanical model of the heart, synthetic data, a cardiac MRI atlas of motion and geometry, MRI data from 6 healthy volunteers (one of them subjected to stress), and an MRI study on a patient with tachyarrhythmia, before and after RF ablation. A pre-operative MRI study on a second patient with tachyarrhythmia was used to test the methodology in a clinical scenario, predicting the abnormally conducting region.

1 Introduction

Supraventricular tachyarrhythmias are pathological fast heart rhythms most commonly caused by an extra electrical propagation pathway (foci) between the atria and ventricles. Treatment by radio-frequency (RF) ablation uses electrode-catheters to monitor endocardial electrical activity. An ablation electrode is then used to apply a RF current, inducing hyperthermia and thus destroying the abnormally conducting areas. The procedure is normally guided with x-rays (2D), and thus prone to errors in location and excessive radiation exposure.

Our main objective in this work is to provide pre- and intra-operative 3D MR guidance [1, 2] in XMR systems (combined X-ray and MRI room) by detecting the onset of regional motion and relating it to the electrical activation pattern. For this purpose we define a probabilistic measure of regional motion activation derived from a 3D motion field extracted by using non-rigid 3D registration of tagged MR (SPAMM) image sequences. Since we address the inverse electro-mechanical problem, trying to infer time of electrical activation by extracting information from the cardiac motion, we use an electro-mechanical model of the heart to validate these results. Isochrones computed from MR motion are compared between different image acquisitions, and also to those isochrones obtained with the model. A cardiac MR atlas of motion and geometry is also used to validate results in a realistic but relatively noise-free case.

Another objective of this work is to detect changes in regional motion patterns between two different image acquisitions. The purpose of this being the follow up of medical treatment in general, and in particular of patients that have undergone RF ablation. For these patients the method can aid in the identification and localisation of abnormal or changing motion patterns, and also can help determine whether the ablation had the desired effect of regularising cardiac contraction. For validation we use MR images of 6 healthy volunteers (one subjected to stress), synthetic data generated with a cardiac motion simulator of MR images, and pre- and post-intervention MR images on a patient with tachyarrhythmia. Difference maps of isochrones and other motion descriptors are computed on the anatomically meaningfully subdivided myocardium.

In order to test accuracy and feasibility in clinical use, a pre-operative MRI study of a second patient with tachyarrhythmia was used for predicting the abnormally conducting region and results compared against those of three experts.

2 Methods

2.1 Registration for Motion Tracking

We use a non-rigid registration algorithm [3] to track the motion and deformation of the heart in a sequence of 3D short- and long-axis tagged MR images. The goal of the non-rigid registration is to align each time frame of the tagged MR image sequence with the end-systolic (ES) time frame of the image sequence by maximising the normalised mutual information of both time frames. To model cardiac motion we use a free-form deformation based on cubic B-splines. The output of the registration is a continuous time varying 3D motion vector field, $\mathbf{F}(\mathbf{p}, t)$, where $\mathbf{F} : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ and $\mathbf{p} \in \mathbb{R}^3$ is the spatial coordinate (x, y, z) .

2.2 Coordinate System and Myocardial Segmentation

A manual segmentation of the myocardium at end-diastole (ED) is used to determine the region of interest (**myo**) for the registration at time $t = 0$. Using \mathbf{F} , the myocardial region can be automatically propagated over the entire cardiac cycle.

In order to be able to compare different image acquisitions, a common (cylindrical) coordinate system based on the left ventricle is defined for each subject, thus avoiding potential misregistration errors due to subject motion between scans. We then express the \mathbf{F} in terms of radial, circumferential and longitudinal directions, and subdivide the myocardium (**myo**) into small meaningful regions or segments s . For the purpose of comparing motion between different scans we use $S = 12$ segments (similar to those suggested by the AHA): 4 sections around the z -axis that correspond to septum, lateral, anterior and posterior walls, and 3 sections along the z -axis, corresponding to base, middle region and apex.

2.3 Differential Motion Descriptors and Changes in Motion Patterns

Some differential features derived from the motion field $\mathbf{F}(\mathbf{p}, t)$ are useful to describe myocardial non-rigid motion. We write them as the set of functions

$$F^m = F^m(\mathbf{p}, t) \text{ where } m \in \mu = \{D, R, C, Z, \dot{R}, \dot{C}, \dot{Z}, E, r, c, z, \dot{r}, \dot{c}, \dot{z}\} \quad (1)$$

and $F^m : \mathfrak{R}^4 \rightarrow \mathfrak{R}$ are defined as the total deformation or displacement $F^D = \|\mathbf{F}\|$, the radial, circumferential and longitudinal components of the deformation (F^R , F^C and F^Z) with respect to the cylindrical coordinate system and their corresponding time derivatives or velocities ($F^{\dot{R}}$, $F^{\dot{C}}$ and $F^{\dot{Z}}$), the magnitude of the strain matrix $F^E = \|E_{i,j}\|$, the radial, circumferential and longitudinal components of the strain (F^r , F^c and F^z), and their time derivatives ($F^{\dot{r}}$, $F^{\dot{c}}$ and $F^{\dot{z}}$), all with respect to the the same cylindrical coordinate system.

The values of $F^m(\mathbf{p}, t)$ are first computed for each voxel, and then averaged for each of the myocardial segments s , for all time frames during the cardiac cycle:

$$F^m(s, t) = \frac{1}{\int_{\mathbf{p} \in s} d\mathbf{p}} \int_{\mathbf{p} \in s} F^m(\mathbf{p}, t) d\mathbf{p} \quad \text{for all regions } s \in \mathbf{myo}. \quad (2)$$

In order to evaluate changes in the motion patterns between two data sets \mathbf{F}_1 and \mathbf{F}_2 , for instance those corresponding to pre- and post-ablation scans, the difference between the two functions F_1^m and F_2^m is computed for each segment, integrated over time and normalised using the maximum value of the function for the specific segment. This normalization of the values compensates for the differences in the dynamic behaviour expected in the various regions of the heart (like apex and base for instance). A statistical measure is derived from the above combined quantities [4] and each segment is assigned a measure of motion change and classified as having either no, small or significant changes.

2.4 Activation Detection

Although the study of myocardial electrical phenomena such as the excitation-contraction relation, re-entries and patterns occurring inside the myocardium remain open problems for study (see references in [5, 6]), in this work we use the underlying assumption that we can relate the onset of regional motion, derived from

the images sequences, to the electrical activation (*i.e.* using the inverse relation of electro-mechanical coupling). Ideally the onset of regional contraction could be inferred from the motion field simply from strain, but we use a more robust measure because of noise and the relatively low space and time resolution of the images and the extracted motion fields. For this purpose we investigate the subset of differential descriptors \mathbf{F}^m where $m \in M = \{R, C, Z, \dot{R}, \dot{C}, \dot{Z}, E, \dot{r}, \dot{c}, \dot{z}\}$.

The first step to characterise the regional motion of the heart during the cardiac cycle is to measure the regional ($T_{ES}(s)$) and global (T_{ES}) end-systolic times, as well as the critical times for each motion descriptor. We therefore define $T_{max}^m(s) = t^*$ such that $F^m(s, t^*) \geq F^m(s, t) \forall t \in [0, T_{ES}(s)]$ and $T_{min}^m(s) = t^*$ such that $F^m(s, t^*) \leq F^m(s, t) \forall t \in [T_{max}^m(s), T_{ES}(s)]$. Notice that for T_{min}^m the search interval begins at T_{max}^m , *i.e.* when the maximum value has been reached (it is the late minimum value of F^m that will help us define the end-systolic time, not the small values at the beginning of the cycle). Computing these values requires a first (visual) estimate of the end-systolic time, however a short iterative process rapidly provides a better estimate for $T_{ES}(s)$.

In the case of displacement and strain, the end-systolic time is linked to their maximum values, while in the case of velocity and rate of change of strain it corresponds to their minimum values (when the heart has paused its contraction). Therefore,

$$T_{ES}^m(s) = \begin{cases} T_{max}^m(s) & \text{for } m \in \{R, C, Z, E\} \\ T_{min}^m(s) & \text{for } m \in \{\dot{R}, \dot{C}, \dot{Z}, \dot{r}, \dot{c}, \dot{z}\} \end{cases} \quad (3)$$

and combining these times we obtain an estimate that corresponds to the regional time of end-systole: $T_{ES}(s) = \sum_{m \in M} w_m T_{ES}^m(s)$. The weights w_m are normalised (*i.e.* $\sum_{m \in M} w_m = 1$) and reflect the confidence we have on each of the differential motion descriptors m . At present we assign these weights manually, but a statistical measure derived from the data will be used to compute them automatically. In order to obtain a global estimate for end-systolic time for each feature we integrate those values over the entire myocardium: $T_{ES} = \int_{s \in \text{myo}} T_{ES}(s) ds$. Using the above equations we define a probabilistic measure of the activation for every voxel in the myocardium, at any time during the cardiac cycle:

$$A(s, t) = \sum_{m \in M} w_m \int_0^t \frac{F^m(s, \tau)}{\int_0^{T_{max}^m(s)} F^m(s, \tau') d\tau'} d\tau \quad (4)$$

where we impose $F^m(s, t) = 0$ if $t > T_{max}^m(s)$ in order to keep the values normalised (notice that some motion descriptors like the velocities and the time-derivatives of strain reach their maximum values before end-systole).

The value of $A(s, t)$ monotonically increases from zero to one as we expect every voxel to have been activated by the time the motion descriptors reach the maximum value at time $T_{max}^m(s)$. In order to avoid singularities in the equation we excluded from the computation, and labelled as not active, those voxels that might remain relatively static (*i.e.* those for which $F^m(s, T_{max}^m(s)) \approx 0$). By integrating over time we obtain an accumulated probability and we can therefore

set a (percentage) threshold P , between 0 and 1, to define the time t_a at which the activation of a segment s takes place. That is, if $A(s, t_a) = P$ then s becomes active for $t = t_a$. The activation *isochrones* are then defined, for a given threshold P , as the function $A(s) = t_a$, for all $s \in \mathbf{myo}$.

2.5 Cardiac Motion Simulator for Tagged MRI

In order to validate the proposed methodology with a controlled case we also implemented and modified a cardiac motion simulator for tagged MRI [7]. The motion simulator is based on a 13-parameter model of left-ventricular motion developed by Arts *et al.* [8] and is applied to a volume representing the LV that is modeled as a region between two confocal prolate spheres while the imaging process is simulated by a tagged spin-echo imaging equation [9].

A pair of sequences of synthetic tagged LV images was produced in the following manner: first, a ‘post-intervention’ (normal) sequence computed using the standard model parameters, and secondly, a ‘pre-intervention’ (abnormal) sequence computed for which the motion parameters were modified in a small region of the myocardium by bringing the contraction slightly forward in time [4].

3 Results and Discussion

3.1 Changes in Regional Motion Patterns

The detection of changes in motion patterns was evaluated on synthetic data as well as real MR data. In order to test the algorithm when the **ground truth** is available, results on the ‘pre-’ and ‘post-intervention’ sequences of **synthetic tagged LV images** were compared in two cases, with different parameters and

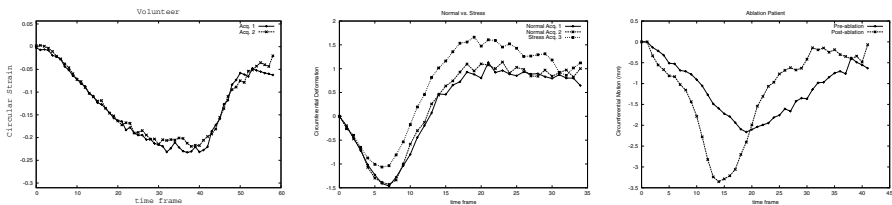


Fig. 1. Time plots of differential motion descriptors. (a) Similar F^m curves of a typical myocardial segment obtained from two independent scans of a healthy volunteer demonstrate **reproducibility**. (b) A myocardial segment of a healthy volunteer, with and without **stress**. There are no significant changes in the motion pattern between the first two image acquisitions. In the third image acquisition (under stress) a noticeable alteration was detected. (c) A myocardial segment of **patient before and after RF ablation**. Significant change can be seen in the faster and more pronounced motion of the post-intervention sequence, indicating successful regularisation of the contraction.

regions of abnormal motion (see Section 2.5). In both cases the abnormal segments were accurately located, while the remaining ones were correctly classified as having no significant change [4]. We also acquired **MR data from six volunteers**. For each of them two separate sets of image sequences were acquired with only a few minutes between the scans. Since no change is expected in these pairs of image acquisitions, this allowed us to verify the **reproducibility** of the motion fields computed by the algorithm and to test the comparison method against false positive detection. The motion patterns encountered were very similar and no region was classified as having a significant change (Figure 1a).

With another volunteer we acquired three sets of image sequences with only few minutes between the scans. The first two were normal scans as described above, but the third one was acquired while inducing **stress** on the volunteer by placing one of his feet into a bucket of ice-cold water. This experiment allowed us to compare normal motion patterns against those obtained under stress, and again, to validate the method regarding reproducibility and false positives. No segment showed a significant difference between the two normal acquisitions, but few segments did when compared to the stress acquisition (Figure 1b).

Finally, MRI data was acquired from an eight year old **patient** with acute super-ventricular tachyarrhythmia, before and after **RF ablation**. The image acquisition and catheter intervention were carried out with an XMR system [1]. Our results confirmed that the motion pattern changed in most parts of the myocardium (visual inspection of the reconstructed 3D surfaces and displacement vectors also showed pronounced changes in the overall contraction pattern), while the largest changes were found in five segments. Examples of the compared motion also show the corrective effect of the intervention (see Figure 1c).

3.2 Activation Detection

Figure 2 shows results of activation detection obtained for the MR stress study described in Section 3.1. The times of activation of different regions of the myocardium are shown as different colours over the end-diastolic myocardial surface

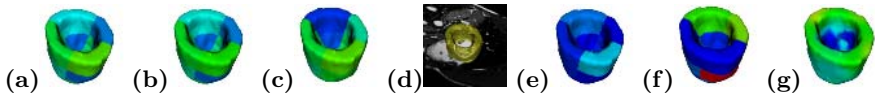


Fig. 2. Isochrones of stress data. Motion-derived activation isochrones computed from two normal MR scans, (a) and (b), and a third one acquired while the volunteer was subjected to stress (c). The anatomical MR image and LV surface skeleton shows the display orientation in (d). Isochrones subtraction maps: the difference between the two normal repetition scan in (e), and the difference between a normal and the stress scan in (f). Isochrones computed from the electro-mechanical model are shown in (g). The colour scales go from blue to red: isochrones maps from 0 to 500ms, with green approx. 200ms, and in isochrones subtraction maps from 0 to 100ms approx.

(**activation isochrones** maps). The first three images in the figure compare the isochrones obtained from the three MR data acquisitions of the same subject: two repetition scans with no changes in between them, and a third scan acquired while the volunteer was subjected to stress. Results of **subtracting pairs of isochrones maps** are also shown: the difference between the two normal repetition acquisitions, in Figure 2e, and the difference between a normal and the stress acquisition, in Figure 2f. We can see that the difference between the isochrones of the two normal acquisitions is small, thus validating the method regarding reproducibility, while on the other hand some larger changes can be appreciated between the isochrones of the normal and the stress scans, thus highlighting the regions that were most affected by stress.

Since we are addressing the problem of inverse electro-mechanical coupling, that is, trying to infer the time of electrical activation by extracting information from the cardiac motion images, we have also used a forward 3D **electro-mechanical model** of the heart [5] to validate the activation detection results. The segmentation of the myocardium of a healthy volunteer at end-diastole was used as geometric input for the model. The muscle fiber orientation and the Purkinje network location were fitted to the geometry from a-priori values of the model. Figure 2g shows the isochrones values computed using the electro-mechanical model applied to the subject of the stress study. Good correlation can be seen between these and the isochrones derived from MR motion.

We also used a **cardiac atlas** of geometry and motion, generated from 3D MR images sequences of 14 volunteers, to test our activation measure in a real-

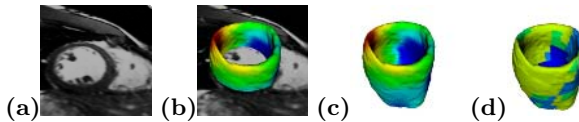


Fig. 3. Isochrones of cardiac atlas. Isochrones were computed for the atlas using both, the electro-mechanical model ((b) and (c)), and the proposed activation measure derived from the motion field ((d)). The colour scale goes from blue to red (earliest to latest activation time). The orientation of the left and right ventricle can be seen on the MR images of the subject used as a reference for the atlas ((a) and (b)).

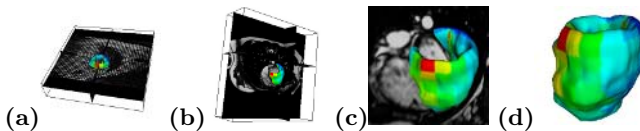


Fig. 4. LV surface with activation times derived from the motion field of a patient with WPW syndrome. The orientation of the left and right ventricles can be seen on the tagged image (a), and in two views ((b) and (c)) of the high resolution anatomical image. In order to highlight the area of earliest motion ((d)), fed by the abnormal conduction pathway, the colour scale in this figure goes from red (earliest) to blue (latest).

istic but smooth and virtually noise-free data set [10]. For the purpose of comparing activation detection results to those obtained with the high-resolution electro-mechanical model, a larger number of smaller segments was used (also, segments can be very small in this case since there is little noise in the data). Figure 3 compares the isochrones for the atlas computed by both, the electro-mechanical model, and the proposed activation measure derived from the motion field. Promising agreement can be seen on these results of activation detection.

In order to test the **accuracy and feasibility in clinical use**, a pre-operative MRI study of a second patient with tachyarrhythmia (Wolff-Parkinson-White (WPW) syndrome) was processed before the RF ablation, and the location of the abnormal conduction pathway automatically estimated as the region of earliest activation. To increase the accuracy of the estimated position, instead of using 12, we used 120 segments for the LV: 24 subdivisions around the z-axis and 5 from apex to base. The geometric centre of the earliest activated segment was used as the automated estimated position of the pathway (Figure 4).

For the purpose of validation, two experts involved with the patient's RF ablation were separately asked to estimate the location of the abnormal pathway based on careful visual inspection of a 3D anatomical MR image sequence of the patient. This anatomical scan, acquired immediately after the tagged one, had higher time resolution, thus facilitating the visual assessment of the earliest site of motion. With respect to the expert's estimations, the distances (errors) of the automatically estimated position were 5.01mm and 5.27mm. The distance between the expert's positions was under 2mm. Further confirmation of this result came from the analysis of the patient's ECG recordings by a third expert who estimated the location of the pathway to be in the posterior-septum.

4 Conclusions and Future Work

Despite current limitations such as distinguishing between epi- and endo-cardial activation patterns, the methodology seems promising for the assessment of intervention results and could also be used for the detection of arrhythmia, ischaemia, regional dysfunction, and follow-up studies in general. Because acquisition of tagged images can be carried out in less than 20 minutes, either immediately before the RF ablation or the day before the intervention, the proposed analysis is suitable for clinical practice in guiding and monitoring the effects of ablation on ventricular arrhythmias, with little extra discomfort added to the patient. As has been shown, the error in the location of the abnormal pathway can be as small as 5mm, as independently confirmed by two experts. In order to account for possible changes in the heart rate between the pre- and post-intervention acquisitions (or for instance, in the case of the stress study where there was a small change in the heart rate), we intend to re-scale one of the image sequences in the time domain, by using the 4D registration technique described in [10].

References

- [1] K.S. Rhode, D.L.G. Hill, P.J. Edwards, J. Hipwell, D. Rueckert, G.I. Sanchez-Ortiz, S. Hegde, V. Rahunathan, and R. Razavi. Registration and tracking to integrate X-ray and MR images in an XMR facility. *IEEE Trans Med Imag*, 22(11):1369–78, 2003.
- [2] G.I. Sanchez-Ortiz, M. Sermesant, R. Chandrashekhara, K.S. Rhode, R. Razavi, D.L.G. Hill, and D. Rueckert. Detecting the onset of myocardial contraction for establishing inverse electro-mechanical coupling in XMR guided RF ablation. In *IEEE Int Symp Biomed Imag (ISBI'04)*, pages 1055–8, Arlington, USA, Apr 2004.
- [3] R. Chandrashekhara, R. Mohiaddin, and D. Rueckert. Analysis of 3D myocardial motion in tagged MR images using nonrigid image registration. *IEEE Trans Med Imag*, 23(10):1245–1250, 2004.
- [4] G.I. Sanchez-Ortiz, M. Sermesant, K.S. Rhode, R. Chandrashekhara, R. Razavi, D.L.G. Hill, and D. Rueckert. Detecting and comparing the onset of myocardial activation and regional motion changes in tagged MR for XMR-guided RF ablation. In *Functional Imaging and Modeling of the Heart*, LNCS 3504, pages 348–58, 2005.
- [5] M. Sermesant, K. Rhode, A. Anjorin, S. Hedge, G. Sanchez-Ortiz, D. Rueckert, P. Lambiase, C. Bucknall, D. Hill, and R. Razavi. Simulation of the electromechanical activity of the heart using XMR interventional imaging. In *MICCAI'04*, LNCS 3217, pages 786–94, France, 2004.
- [6] E. McVeigh, O. Faris, D. Ennis, P. Helm, and F. Evans. Measurement of ventricular wall motion, epicardial electrical mapping, and myocardial fiber angles in the same heart. In *FIMH'01*, LNCS 2230, pages 76–82, 2001.
- [7] E. Waks, J. L. Prince, and A. S. Douglas. Cardiac motion simulator for tagged MRI. In *IEEE Worksh. Math. Meth. Biomed. Imag. Anal.*, pages 182–91, 1996.
- [8] T. Arts, W. Hunter, A. Douglas, A. Muijtjens, and R. Reneman. Description of the deformation of the LV by a kinematic model. *Biomech.*, 25(10):1119–27, 1992.
- [9] J. L. Prince and E. R. McVeigh. Motion estimation from tagged MR images. *IEEE Transactions on Medical Imaging*, 11(2):238–249, June 1992.
- [10] D. Perperidis, M. Lorenzo-Valdes, R. Chandrashekhara, R. Mohiaddin, G.I. Sanchez-Ortiz, and D. Rueckert. Building a 4D atlas of the cardiac anatomy and motion using MR imaging. In *IEEE Int. Symp. Biomed. Imag.*, pages 412–5, USA, 2004.

Automatic Mammary Duct Detection in 3D Ultrasound

Mark J. Gooding¹, Matthew Mellor², Jacqueline A. Shipley¹,
Kathy A. Broadbent³, and Dorothy A. Goddard⁴

¹ Medical Physics Dept., Royal United Hospital, Bath, BA1 3NG, UK
`mark.gooding@ruh-bath.swest.nhs.uk`

² Wolfson Medical Vision Laboratory, Dept. of Engineering Science,
University of Oxford, OX1 3PJ, UK

³ Dept. of Medical Physics and Bioengineering,
Bristol General Hospital, Bristol, BS1 6SY, UK

⁴ Breast Unit, Royal United Hospital, Bath, BA1 3NG, UK

Abstract. This paper presents a method for the initial detection of ductal structures within 3D ultrasound images using second-order shape measurements. The desire to detect ducts is motivated in a number of way, principally as step in the detection and assessment of ductal carcinoma in-situ. Detection is performed by measuring the variation of the local second-order shape from a prototype shape corresponding to a perfect tube. We believe this work is the first demonstration of the ability to detect sections of duct automatically in ultrasound images. The detection is performed with a view to employing vessel tracking method to delineate the full ductal structure.

1 Introduction

Most research into breast ultrasound analysis concerns the detection and diagnosis of lesions, therefore the detection of mammary ducts may seem somewhat peripheral. However, this is not the case; the role of the duct is central to the function of the breast, and consequently the ductal structures are an important area for analysis. Most malignancies originate in the epithelial tissues of the ducts, and it has been suggested that ductal echography, a process whereby the breast volume is systematically scanned along the line of duct, is the best way to manually find and diagnose lesions [1]. The automatic detection of the ductal structure from 3D ultrasound could direct the reslicing and visualisation of the ultrasound data along the axis of each duct for clinical analysis, or enable automatic detection of anomalous ductal structure as early indicator of possible tumor development [2]. The analysis of the duct may also have a role in the analysis of detected lesions, since the detection of duct-size features has been found to be an indicator of malignancy [3]. To this end, comparison with the normal appearance of ducts in ultrasound images is required for diagnosis by ductal echography, although it should be noted that this appearance changes with age [4,1].

The extent of each lobe of the ductal system is also of importance. Typically, a carcinoma will spread along a duct, as ductal carcinoma in situ (DCIS), prior to any invasive spreading. As observed by Going and Moffat [5]:

If duct systems are independent of each other, then defining their boundaries prior to surgery for DCIS could assist in achieving complete removal of the affected ‘lobe’ while sparing the unaffected tissue.

Even in cases where the tumor is invasive, the ductal spread will be extensive and the lobe must be considered the minimum excision margin. However, the presence of any ductal anastomoses must be detected if the complete extent of possible spread is to be correctly assessed [6].

Once a ductal anomaly has been identified, ductal endoscopy and lavage provide possible routes for minimally invasive diagnosis. However, some lobular systems are not openly accessible on the nipple surface [5]. A complete tree of the ductal system would facilitate assessment of whether the duct can be accessed in such a way. Furthermore, a map of the 3D ductal structure would also facilitate registration based on anatomy [7], which would have application in assessing tumor changes following of treatment. Applications may also be found in assessment of problems associated with lactation [8], and in research into mammary development [9].

Thus, the reasons motivating the automatic detection and segmentation the ductal structure of the breast are:

- A method for the detection and diagnosis of lesions by comparison with normal ductal appearance.
- A Guide to excision margins by considering the possible intraductal spread.
- An aid for ductal endoscopy to assess access and navigation.
- Facilitation of anatomy-based registration for treatment assessment.
- An aid to the assessment and diagnosis of problems associated with lactation.
- Facilitation of mammary development research and other temporal studies of the breast.

In this paper, we review previous tube (mostly vascular) detection and tracking methods to motivate shape-based detection. A simple method for the initial detection of ductal structures is proposed and demonstrated. We believe this is the first demonstration of the ability to detect sections of duct automatically in ultrasound images. The paper concludes with a brief discussion of the direction of future research.

2 Method

2.1 Previous Tube Detection and Tracking Approaches

The detection of tubular structures has mostly been focused on the detection of vessels with MRA and CT and of bronchial tubes in CT. The approaches taken can be broadly divided into intensity-based segmentation and shape-based

detection and tracking. Intensity-based detection (e.g. [10]) is not suitable for small ultrasound structures, such as ducts, since at a fine scale the image intensity is not directly representative of the tissue type. (At a coarse scale some tissue/intensity relationship may be noticed.) However, shape-based detection may show promise since the approximate structure of the tubes is still evident in the images, although significantly corrupted by speckle.

The second-order structure (the Hessian) of image intensity is often considered to be a good measure of the local shape. Frangi et al. [11] and Danielsson et al. [12] develop measures which describe the local 3D shape and “curvedness” based on the principal curvatures. Their work can be considered an extension of the shape index (SI), introduced by Koenderink and van Doorn [13], into 3D. The measures in [12,14] are sufficient to uniquely describe the local second-order shape of an image. Subsequently, both authors develop further specific measures to detect vascular structure based on their local shape measures. Similarly, Sato et al. [15] use the principal curvatures of the image intensity to calculate measures of second-order structure corresponding to a sheets, tubes and spheres. These are used heuristically to segment image data from a range of modalities. All three authors demonstrate the detection of tube-like structure for the purpose of vessel segmentation in MRI volumes.

A number of effective methods have been developed for tracking of tubular structures. These can be roughly divided into those which track second-order structures (e.g. [16,17]) and those based on medial node tracking (e.g. [18,19]). The use of medial-nodes has been demonstrated for ultrasound data at a large scale [20], however the small size and poorly defined boundaries of ducts mean that this approach is not suited to the application of ductal tracking. Both of these approaches require initialisation. For the MRI and CT data, on which the methods have been demonstrated, the majority of initialisations have been achieved using an intensity threshold of the image data. However, intensity thresholding for the initial detection of ducts would be ineffective within ultrasound images. Section 2.2 addresses this initialisation problem, proposing a method for the initial detection of ducts within 3D ultrasound images, with a view to employing a tracking method to find the complete ductal structure.

2.2 Simplified Detection Method

In this section we propose a simple method for the detection of specific second-order shape features directly from the principal curvatures, which could be used as an initialisation to the centerline tracking methods discussed previously. Although Danielsson’s angular measures of shape provide a way to express the complete range of second-order shapes from volumes, the tubular shape measures introduced in [11,12,14,15] to detect vessels are unnecessarily complicated. We propose a simplified approach to tube detection. Any shape of interest within the second-order space can be uniquely expressed in terms of the three principal curvatures (the eigenvalues of the Hessian). Consequently, it is desirable to be able to measure the variation from a specific shape description expressed in terms of these curvatures. Therefore, we opt to detect tubular shapes directly from the

principal curvatures, rather than introducing further measures of shape. The desired tube-shape prototype and the local shape can be expressed as vectors of the form

$$\mathbf{shape\text{-}vector} = \begin{pmatrix} \kappa_1 \\ \kappa_2 \\ \kappa_3 \end{pmatrix} \quad (1)$$

where κ_1 , κ_2 and κ_3 are the three principal curvatures. The orientation of this vector describes the local shape, while its length describes the degree of local curvedness. For a tubular shape the prototype is

$$\mathbf{tube} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}. \quad (2)$$

We propose to measure the deviation of the local shape from the prototype shape as the normalised dot product between their shape vectors, as shown in equation 3, which can be thought of as an angular distance measure. This gives a result of 1, for a perfect match, -1 for a perfect match of the opposite sign (i.e. same shape but opposite intensity structure), and zero where the local shape can be considered to be as far away as possible from the desired shape.

$$\mathbf{tubeness} = \begin{pmatrix} \frac{\kappa_1}{\sqrt{\kappa_1 + \kappa_2 + \kappa_3}} \\ \frac{\kappa_2}{\sqrt{\kappa_1 + \kappa_2 + \kappa_3}} \\ \frac{\kappa_3}{\sqrt{\kappa_1 + \kappa_2 + \kappa_3}} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = \cos \theta_{tube} \quad (3)$$

Normalisation of the shape vectors ensures that this shape measure is invariant to contrast. Tubes are detected at multiple scales using scale normalised Gaussian convolution to generate progressively coarser scales [21]. The shape detected at a particular location may vary over the scales, as finer features are removed, e.g. A duct may be detected as tube like at a fine scale, but the more spherical shape of the local glandular tissue may be more prominent at a larger scale. Therefore, it is necessary to choose an appropriate scale to detect features of interest. We achieve this in a number of ways. First, we only process the scales over which we know the features exist, i.e. in looking for ducts we only search smaller scales. Second, we opt choose the scale at which each location looks most similar to the prototype shape vector. Detection of tubular features is achieved by accepting areas which do not deviate too far from the desired shape, and which have a high curvedness at the scale chosen. As noted in [12], curvedness, measured as $\sqrt{\sum \kappa_i^2}$ is not shape invariant; a spherical shape having higher curvedness than a tube-like one. However, we have chosen to detect features matching a single shape prototype, therefore using this measure is not a problem since we are comparing the curvedness of features with similar shapes.

2.3 3D Ultrasound Acquisition

In-vivo acquisitions were carried out with consenting patients in accordance with an ethically approved protocol, using the radial scanning system described in

[22,23], a Technos ultrasound scanner (Esaote S.p.A, Genova, Italy), and linear array probe (LA532 10-5) at 9MHz. The pathologies present were assessed by a clinician at the time of scanning, and subsequently by reviewing the full 3D data available. The 3D images were reconstructed on a 1mm^3 voxel array. The radial scanning system is appropriate for the detection of ducts for a number of reasons: ducts are oriented approximately radially from the nipple, therefore scanning radially allows sections of duct to be viewed along their length [1]; the scanning system also uses low compression reducing the possibility that the ducts will collapse under the pressure of the probe during scanning; and the majority of the breast tissue volume is imaged, allowing the ductal structure to be viewed for most of the breast.

3 Results

For the examples given, features were detected by accepting voxels whose shape vectors were within 32° of the prototype shape vector ($\cos\theta_{tube} = 0.85$). Empirically this was found to give good results, however the optimum distance has not been investigated. The curvedness used for feature acceptance was set for each 3D image as a percentage (typically in the range 40% to 60%) of the maximum curvedness found for the tube feature over all scales. The exact value was found qualitatively in these examples, depending on the dynamic range and maximum curvedness of the particular scan. The scales used for detecting ducts and vessels had a Gaussian standard deviation ranging $\sigma = 1$ to 3 voxels. This equivalent to a one standard deviation width of between 2 and 6mm, similar to the size of enlarged ducts [8]. Detection was performed on a voxel-by-voxel basis.

Figures 1 and 2 demonstrate the potential of this system (both the radial scanner and detection method) to detect the ductal system in the breast. Figure

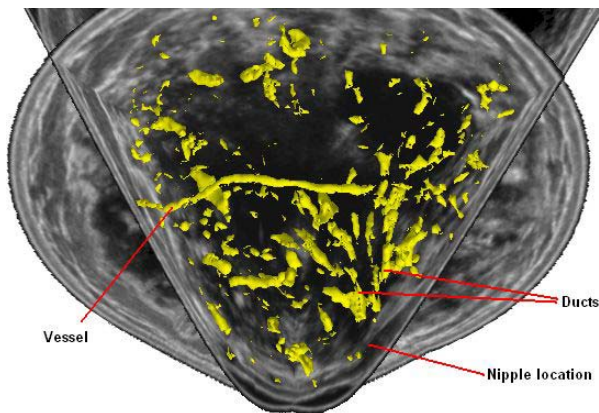


Fig. 1. The detection of ducts and vessels in a paracystic breast. Additional less interesting tube structures have been detected at cyst edges. However, these features still appear tube-like at the scale selected.

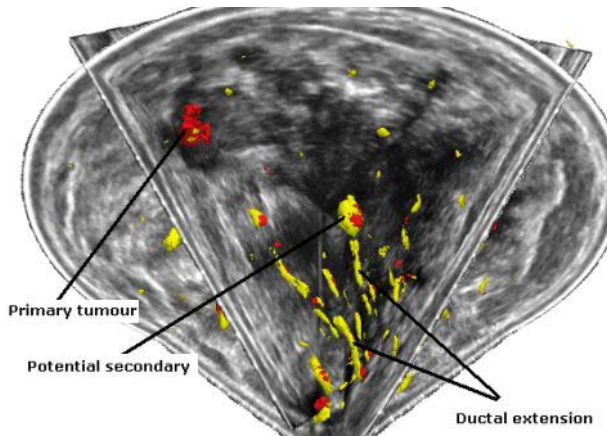


Fig. 2. The detection of ductal extension. A duct has been detected extending from the primary tumor. A possible secondary has also been detected within the same ductal tree, illustrating the importance of ductal tracking in assessing tumor invasiveness. Sphere-like structures (shown in red in the online proceedings) were also detected using a different shape prototype.

1 is a scan from a breast displaying polycystic changes and some ductal ectasia. Both vessels and ducts have been detected, although a number of less interesting detections have occurred at the edge of cysts. This is expected, since the edge of spherical objects appear tube-like in shape at small scales, and the cysts have high contrast resulting in high curvedness. Figure 2 shows the detection of ductal extension from a tumor in a breast containing a malignant ductal carcinoma. The extension of the tumor along the duct affects patient management, making mastectomy, rather than excision, the preferred option. The figure indicate areas which have been detected using a sphere-like structure prototype, rather than a tube-like one (shown in red in the online proceedings).

4 Discussion, Conclusion and Future Work

In this paper we have presented a method of tube-like feature detection for 3D ultrasound. Several methods of shape analysis and shape-based feature detection from the existing literature were considered. The method presented here simplifies these by detecting specific shapes directly from the principal curvatures, via comparison with a prototype feature, rather than introducing intermediate steps. The performance of this feature detector was demonstrated on 3D ultrasound data of breast volumes from two patients, showing the detection of ducts and vessels.

Several observations can be made from these results. First, that the detected objects correspond to perceptually significant image structures. However, interesting image structure does not necessarily equate to interesting pathology. This

is not unexpected, since we are seeking a low-level shape feature rather than an anatomical feature. However, some knowledge-based constraints (e.g. we expect ducts to be long connected tubes) must be introduced to remove non-anatomical features for clinical use. Second, it is also noted that the shape of the detected volume does not correspond to the shape of the object itself. This is expected since a perfect tube will only be detected at its centre (see [18] for a discussion). Although the shape-measure itself cannot be used for accurate segmentation, it does give a pointer to scale and approximate tube position. Therefore, this detection method could be used as an initialisation for a centreline tracking method, such as in [16,17].

The challenge presented in [2] to provide a method for automatically detecting the entire ductal tree structure is significant. We believe this work is the first demonstration of the ability to detect sections of duct automatically in ultrasound images. Future work will consider the use of centerline tracking, by finding the connecting line which best corresponds to a tube as in [16], to form a continuous ductal tree structure. Quantitative evaluation of this tree structure will be performed by comparison with manual delineations. The extension of this ductal tree delineation to the normal breast, with non-enlarged ducts, is a long term goal.

Acknowledgements

We wish to acknowledge Francis Duck, Alison Noble, and Mike Halliwell for their help and advice in this work. Mark Gooding is supported by the research and development committee of the Royal United Hospital, Bath, NHS Trust.

References

1. Teboul, M., Halliwell, M.: Atlas of ultrasound and ductal echography of the breast. Blackwell Science Ltd, Oxford (1995)
2. Moskalik, A., Carson, P., Roubidoux, M.: 3D tracking and display of mammary ducts. In: Proc. IEEE Ultrasonics symposium 1995. Volume 1636. (1995) 1167–1170
3. Donohue, K., Huang, L., Genis, V., Forsberg, F.: Duct size detection and estimation in breast tissue. 1999 IEEE ultrasonics symposium (1999) 1353–1356
4. Nani, M., Piperno, G.: Il sistema duttale mammario. aspetti ecografici. *Radiol Med* **83** (1992) 745–750
5. Going, J., Moffat, D.: Escaping from flatland: clinical and biological aspects of human mammary duct anatomy in three dimensions. *Journal of Pathology* **203** (2004) 538–544
6. Ohtake, T., Kimijima, I., Fukushima, T., Yasuda, M., Sekikawa, K., Takenoshita, S., Abe, R.: Computer-assisted complete three-dimensional reconstruction of the mammary ductal/lobular systems. *Cancer* **19** (2001) 2263–2272
7. Aylward, S., Jomier, J., Weeks, S., Bullitt, E.: Registration and analysis of vascular images. *Int. J. Comp. Vision* **55** (2003) 123–138
8. Ramsay, D., Kent, J., Owens, R., Hartmann, P.: Ultrasound imaging of milk ejection in the breast of lactating women. *Pediatrics* **113** (2004) 361–367

9. Moffat, D., Going, J.: Three dimensional anatomy of complete duct systems in human breast: pathological and developmental implications. *Journal of Clinical Pathology* **49** (1996) 48–52
10. Wilson, D., Noble, J.: An adaptive segmentation algorithm for time-of-flight mra data. *IEEE Trans. Med. Imag.* **18** (1999) 938–945
11. Frangi, A., Niessen, W., Vincken, K., Viergever, M.: Multiscale vessel enhancement filtering. *MICCAI 1998, LNCS 1496* (1998) 130–137
12. Danielsson, P.E., Lin, Q., Ye, Q.Z.: Efficient detection of second-degree variations in 2D and 3D images. *Journal of Visual Communication and Image Representation* **12** (2001) 255–305
13. Koenderink, J., van Doorn, A.: Surface shape and curvature scales. *Image Vision Comput.* **10** (1992) 557–564
14. Lin, Q., Danielsson, P.E.: Stenosis detection using a new shape space for second order 3D-variations. *IMPI 2001, LNCS 2082* (2001) 388–394
15. Sato, Y., Bhalarao, A., Nakajima, S., Shiraga, N., Tamura, S., Kikinis, R.: Tissue classification based on 3D local intensity structure for volume rendering. *IEEE Transactions on visualization and computer graphics* **6** (2000) 160–180
16. Aylward, S., Bullitt, E.: Initialization, noise, singularities and scale in height ridge traversal for tubular object centerline extraction. *IEEE Trans. Med. Imag.* **21** (2002) 61–75
17. Lorigo, L., Faugeras, O., Grimson, W., Keriven, R., Kikinis, R., Nabavi, A., Westin, C.: Curves: Curve evolution for vessel segmentation. *Medical Image Analysis* **5** (2001) 195–206
18. Krissian, K., Malandain, G., Ayache, N., Vaillant, R., Troussset, Y.: Model based detection of tubular structures in 3D images. *Computer Vision and Image Understanding.* **80** (2000) 130–171
19. Fridman, Y., Pizer, S., Aylward, S., Bullitt, E.: Segmenting 3D branching tubular structures using cores. *MICCAI 2003, LNCS 2879* (2003) 570–577
20. Stetten, G., Pizer, S.: Medial-node models to identify and measure objects in real-time 3-D echocardiography. *IEEE Trans. Med. Imag.* **18** (1999) 1025–1034
21. ter Haar Romeny, B.: 6. In: *Front-end vision and multi-scale image analysis*. Kluwer Academic Publishers, Dordrecht (2003) 132–133
22. Shipley, J., Duck, F., Thomas, B.: Registration of ultrasound breast images acquired from a conical geometry. In: *Proc Medical Image Understanding and Analysis 2003*. (2003) 45–48
23. Shipley, J., Duck, F., Goddard, D., Hillman, M., Halliwell, M., Jones, M., Thomas, B.: Automated quantitative volumetric breast ultrasound data acquisition system. *Ultrasound in Medicine and Biology* **31** (2005) 905–917

Automatic Segmentation of Intra-treatment CT Images for Adaptive Radiation Therapy of the Prostate

B.C. Davis^{1,2}, M. Foskey^{1,2}, J. Rosenman², L. Goyal²,
S. Chang², and S. Joshi^{1,2}

¹ Department of Computer Science, University of North Carolina, USA
{davisb, joshi}@cs.unc.edu

² Department of Radiation Oncology, University of North Carolina, USA

Abstract. We have been developing an approach for automatically quantifying organ motion for adaptive radiation therapy of the prostate. Our approach is based on deformable image registration, which makes it possible to establish a correspondence between points in images taken on different days. This correspondence can be used to study organ motion and to accumulate inter-fraction dose. In prostate images, however, the presence of bowel gas can cause significant correspondence errors. To account for this problem, we have developed a novel method that combines large deformation image registration with a bowel gas segmentation and deflation algorithm. In this paper, we describe our approach and present a study of its accuracy for adaptive radiation therapy of the prostate. All experiments are carried out on 3-dimensional CT images.

1 Introduction

One major treatment method for prostate cancer is external beam radiation therapy, which uses high energy x-rays that are delivered in a series of 40 or more daily treatments. To be safe and effective, the radiation dose to the cancer-containing prostate should be as high as possible while the dose to surrounding organs such as the rectum and bladder must be limited. This effect is achieved by using multiple radiation beams that overlap on the tumor and are shaped to exclude normal tissue as much as possible. However, internal organ motion and patient setup errors present a serious challenge to this approach. The prostate, rectum, bladder and other organs move in essentially unpredictable ways, and even small changes in their position can result in either tumor under-dosing, normal tissue over-dosing, or both.

Adaptive radiation therapy (ART), which uses periodic intra-treatment CT images for localization of the tumor and radiosensitive normal structures, is being investigated to meet this challenge. In this method a feedback control strategy [1] is used to correct for differences in the planned and delivered dose distributions due to spatial changes in the treatment volume early in the treatment period.

Although in-treatment-room CT scanners provide the enabling imaging hardware to implement ART, no software methods or tools for automatic image

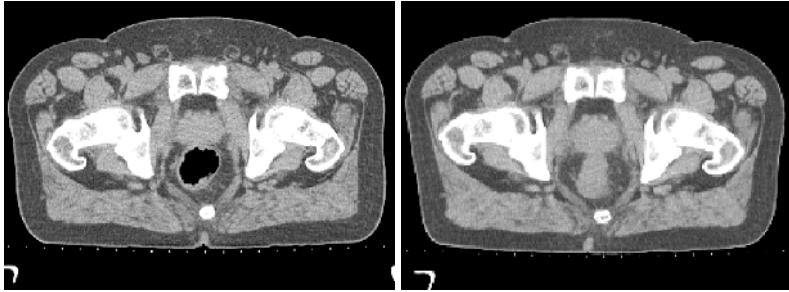


Fig. 1. Axial CT slice of the same patient acquired on different days, showing the effect of bowel gas

processing exist to enable the incorporation of these images in the adaptive treatment of prostate or other cancer. As a result, all such work must be done manually. However, manual segmentation of the tumor and neighboring organs places an impractical burden on highly skilled and already overburdened personnel. Moreover, clinically significant inter- and intra-user variability of manual segmentations introduces a source of treatment uncertainty that current adaptive radiation therapy techniques do not address [2, 3].

We have been developing an approach for automatically quantifying organ motion over the course of treatment. Our approach is based on deformable image registration, which makes it possible to establish a correspondence between points in images taken on different days. This correspondence can be used to study organ motion and to accumulate inter-fraction dose.

In prostate images, however, the presence of bowel gas can cause significant correspondence errors as no correspondence exists for pockets of gas across different days. Shown in Figure 1 are two rigidly aligned axial images of a patient taken on two different days. Due to the transient nature of bowel gas, it is present in one of the days but absent in the other. To account for this problem, we have developed a novel method that combines large deformation image registration with a bowel gas segmentation and deflation algorithm. In this paper, we describe our approach and present a study of its accuracy for adaptive radiation therapy of the prostate.

Several deformable image registration methods are currently being investigated for alignment of serial CT data [4, 5, 6, 7]. However, none of these studies address the problem of bowel gas for deformable registration of CT images. Also, while some authors present validation studies based on known transformations or phantoms, to our knowledge none have presented an analysis of the accuracy of their methods for automatic segmentation of a large number of treatment images based on physician drawn contours.

2 Methods

We use the CT taken at planning time, the *planning image*, as a reference. On each treatment day, the patient is positioned and then, prior to treatment, a

new CT scan is acquired using an in-treatment-room CT scanner that shares a table with the linear accelerator. Each *treatment image* characterizes the patient configuration at that treatment time.

If there were absolutely no organ motion then the planning and treatment images should all be the same, except for noise from the imaging device. However, because there is organ motion, these images will differ, and the difference characterizes the organ motion. We have understood the motion when we can tell, for each point in the planning image, which point in the treatment image it corresponds to. In this way organ motion and image registration are linked—we can understand organ motion if we can estimate image correspondence.

We can view an image as a function I from the spatial domain $\Omega \subset \mathbb{R}^3$ to an intensity value in \mathbb{R} . Image correspondence is expressed as a function $h: \Omega \rightarrow \Omega$, called a *deformation field*. For $x \in \Omega$, $h(x)$ is the point in the treatment image, I_T , that corresponds to the point x in the planning image, I_P .

The transformation h is estimated as follows. First, the planning and treatment CT data sets are rigidly registered. This quantifies the rigid patient setup error. In order to accommodate bowel gas we apply our algorithm for segmenting and deflating bowel gas to produce deflated images I_{P_d} and I_{T_d} . Finally, I_{P_d} and I_{T_d} are registered using a high dimensional large-deformation image registration algorithm. h is defined as the composition of these transformations.

Rigid Registration. The planning and treatment images are thresholded so that only bone is visible. The region of interest is restricted to the pelvis as it remains fixed while the femurs and spine can rotate or bend. The rigid transformation, r , is estimated using an intensity based gradient descent algorithm [8].

Accommodating Bowel Gas. As the contrast between gas and surrounding tissue is very high in CT images, we create a binary segmentation of the gas in an image using a simple thresholding operation. We refine this binary segmentation using a morphological close operation, which eliminates small pockets of gas. Next, we construct a deflation transformation s based on a flow induced by the gradient of the binary image. Points along the gas-tissue border, where the

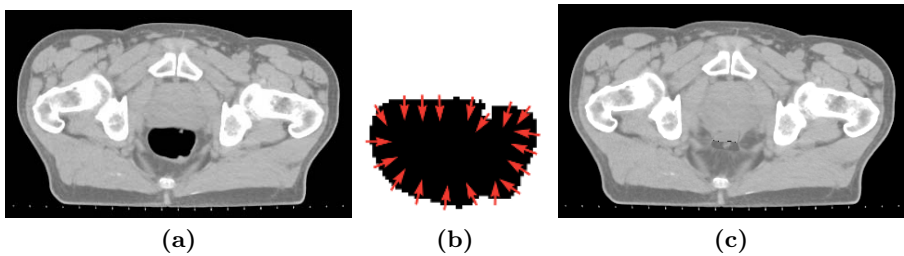


Fig. 2. Gas Deflation Algorithm. (a) Axial slice CT image with large pocket of bowel gas. (b) Zoomed in on the gas pocket. The gas is segmented using simple thresholding. Gas is deflated by a flow induced by the gradient of the binary image. (c) The image after application of the deflation transformation.

gradient is non-zero, flow in the direction of the gradient. As a result, gas filled regions collapse toward their medial skeletons—deflating like a balloon. Importantly, we do not aim to simulate the true motion of the tissue but to deflate the gas so that the image may be accurately registered.

More precisely, we construct a non-diffeomorphic deflation transformation $s: \Omega \rightarrow \Omega$ such that $I(s(x))$ is the image $I(x)$ after a deformation that deflates gas. The transformation s is constructed by integrating velocity fields $v(x, t)$ forward in time, i.e. $s(x) = x + \int_0^1 v(s(x, t), t) dt$. These velocity fields are induced by a force function $F(x, t) = \nabla(I \circ s_t)(x)$ that is the gradient of the binary image. The force function and velocity fields are related by the modified Navier-Stokes operator $(\alpha \nabla^2 + \beta \nabla(\nabla \cdot) + \gamma)v(x, t) = F(x, t)$. We solve for s using an iterative greedy method.

Figure 2 shows the result of our gas deflation algorithm. The large pocket of gas present in the image has been deflated, resulting in an image that can be accurately registered using deformable image registration.

Deformable Image Registration. We apply the theory of large deformation diffeomorphisms [9, 10] to generate a deformation $h_{\text{def}}: \Omega_{P_d} \rightarrow \Omega_{T_d}$ that defines a voxel-to-voxel correspondence between the two gas deflated images I_{P_d} and I_{T_d} . The registration is determined by finding the deformation field h_{def} that minimizes the mean squared error between I_{P_d} and the deformed image $I_{T_d} \circ h_{\text{def}}$,

$$D(h) = \int_{x \in \Omega} |I_{P_d}(x) - I_{T_d}(h_{\text{def}}(x))|^2 dx.$$

The transformation is constrained to be diffeomorphic by enforcing that it satisfy laws of continuum mechanics derived from visco-elastic fluid modeling [10, 9].

Composite Transformation. Correspondence between the original images I_P and I_T is estimated by concatenating the rigid, deflation, and deformable registration transformations, i.e.

$$h_{P \rightarrow T} = r(s_T(h_{\text{def}}(s_P^{-1}(x)))).$$

That is, the point x in the planning image corresponds to the point $h_{P \rightarrow T}(x)$ in the treatment image. This composite transformation is not guaranteed to be diffeomorphic. However, the non-diffeomorphic part of the transformation is restricted to the region of the rectum that contains gas—where no correspondence exists.

Figure 3 shows an example of the application of method described above. Panel (b) shows the result of automatic segmentation using only large deformation image registration. Manually drawn contours of the prostate and rectum are mapped, using this correspondence, from the reference image (a) onto the daily image. Manual contours are drawn in red while mapped contours are drawn in yellow. Notice the misalignment of the manual and automatically generated contours in the daily image; the presence of bowel gas has caused correspondence errors around the rectum. A more accurate correspondence between the reference and daily images is established by concatenating registration and deflation



Fig. 3. Automatic segmentation of the prostate and rectum. Manually segmented structures in the planning image (a) are mapped to the daily image (b) *before* accounting for bowel gas, and (c) *after* accounting for bowel gas with our gas deflation algorithm. Manually drawn contours are shown in red and mapped contours are shown in yellow.

transformations as shown in panel (c). Notice the close alignment between the manual contours and the contours generated by our method.

3 Results

We now present detailed statistical analysis of the application of our methods to a set of 40 CT images from 3 patients undergoing ART in our clinic. Each CT scan was collected on a Siemens Primatom CT-on-rails scanner with resolution $0.098 \times 0.098 \times 0.3$ cm. We analyze the accuracy of our method by comparing automatically generated segmentations to manual, hand-drawn, segmentations. Because of inter-rater variability, however, there is no ground truth manual segmentation to compare against. We therefore compare our automatically generated segmentations with the segmentations from two different manual raters, and then make the same comparisons between the segmentations from the manual raters.

The experimental setup is as follows. The planning image for each patient is manually segmented by rater A. Each treatment image is manually segmented twice, once by rater A and once by rater B. For each patient, our method is used to compute the transformations h_i that map the planning image onto the treatment image for each day of treatment i . An automatic segmentation is generated for each treatment image by applying h_i to the segmentation in the planning image. We can consider our automatic method for producing treatment image segmentations as rater C (for “computer”).

Each segmentation is represented by a triangulated surface. For manual segmentations, the surface is constructed by applying the power crust algorithm [11] to a set of contours drawn in the axial plane by the manual raters. For automatic segmentations, the surface is generated by applying a transformation h to the vertices of the surface given by the manual segmentation in the planning image.

For each patient and for each treatment day, we make three comparisons: **CA**, automatic segmentation versus manual segmentation by rater A; **CB**, automatic segmentation versus manual segmentation by rater B; and **BA**, manual segmentation by rater B versus manual segmentation by rater A. It should be

emphasized that the automatic segmentations are produced by transforming manual planning segmentations produced by rater A, not rater B. Thus, we expect the CA comparisons to be more favorable than the CB comparisons.

In the rest of this section, we present the results of this experiment when comparing centroid differences and relative volume overlap of segmentations.

Centroid Analysis. The centroid of the prostate is especially important for radiation treatment planning and therapy because it is the origin, or isocenter, for the treatment plan. To measure the accuracy of our automatic segmentations with respect to centroid measurement, we compare the centroid of each automatic segmentation with the centroids of the corresponding manual segmentations. The differences in the lateral (X), anterior-posterior (Y), and superior-inferior (Z) directions are measured separately.

Figure 4 shows box and whisker plots of these differences for CA, CB, and BA comparisons. All measurements are made in centimeters. Additional summary statistics are presented in table 1.

Shown in Table 1 are the 99% confidence intervals for the true mean of each distribution of centroid differences. The confidence intervals for the means of the CA and CB differences both overlap with the confidence interval of the differ-

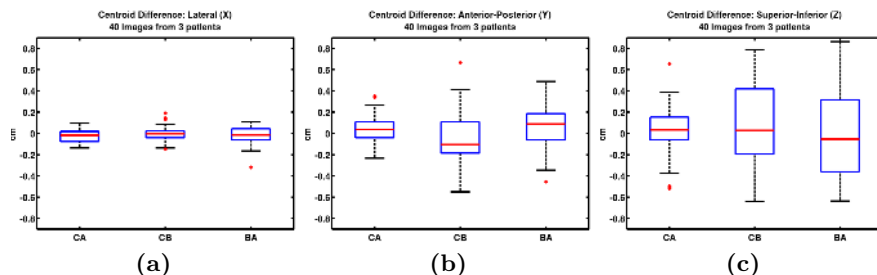


Fig. 4. Centroid differences in the lateral (X), anterior-posterior (Y), and superior-inferior (Z) directions (cm). The horizontal lines on the box plots represent the lower quartile, median, and upper quartile values. The whiskers show the extent of the rest of the data. Outliers, which fall outside 1.5 times the interquartile range, are denoted with the ‘+’ symbol.

Table 1. Summary statistics showing mean, median, standard deviation, and 99% confidence interval of the mean for centroid differences

Centroid Difference Summary (cm)									
	Lateral (X)			A-P (Y)			Sup-Inf (Z)		
	CA	CB	BA	CA	CB	BA	CA	CB	BA
mean	-0.026	-0.007	-0.022	0.035	-0.052	0.070	0.022	0.065	-0.046
median	-0.018	-0.004	-0.015	0.040	-0.104	0.089	0.030	0.028	-0.054
std. dev.	0.06	0.07	0.08	0.14	0.23	0.20	0.24	0.38	0.38
99% CI min	-0.047	-0.030	-0.049	-0.010	-0.129	0.007	-0.054	-0.058	-0.167
99% CI max	-0.006	0.016	0.004	0.081	0.023	0.133	0.10	0.189	0.073

ences between human raters (AB), and are on the order of one voxel. Note that the superior-inferior (Z) direction has a slice thickness of 0.3 cm. We conclude that the automatic segmentation method is as accurate for estimating centroids as human raters and, as seen by the standard deviations, just as reliable.

Relative Volume Overlap Analysis. A measure often reported for comparison of segmentations is relative volume overlap. This measure has been defined in several ways. For this study, we use the Dice Similarity Coefficient (DSC) [12] which is defined for two segmentations S_1 and S_2 as

$$DSC(S_1, S_2) = \frac{\text{Volume}(S_1 \cap S_2)}{\left(\frac{\text{Volume}(S_1 \cup S_2) + \text{Volume}(S_1 \cap S_2)}{2}\right)}. \tag{1}$$

Figure 5 (a) shows a box and whisker plot of the relative volume overlap for the CA, CB, and BA comparisons. To statistically quantify the difference between the relative volume overlaps of the three segmentations A, B, and C, we performed right sided t-tests with the alternative hypothesis $X > Y$. Figure 5, panel (c), reports the P-values of these tests. It can be seen from the table that the volume overlap measures for the CA comparisons are significantly higher than the volume overlap measure for the manual rater comparison BA. There is also no statistically significant difference between the relative volume overlaps from the CB comparison with the two manual raters. Also note that the automatic segmentations have a significantly better overlap with rater A than with rater B. This is expected as the planning image was segmented by rater A.

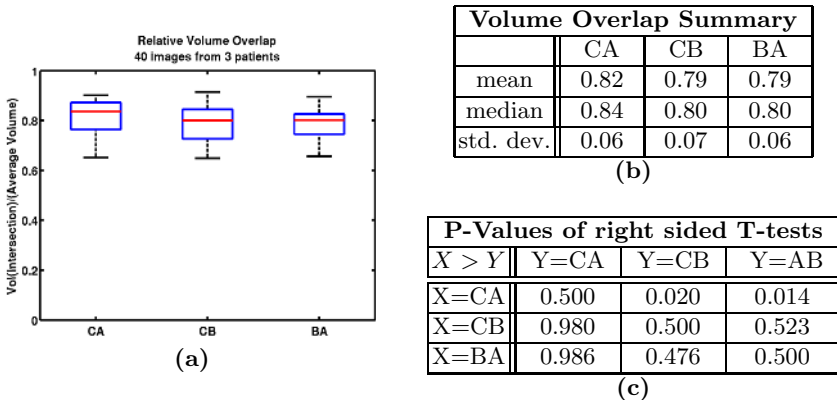


Fig. 5. (a) Relative volume overlap as measured by Equation 1. (b) Volume overlap summary statistics. (c) P-value results of right sided t-test comparing the relative volume overlaps between the various raters.

4 Conclusion

We have presented an approach for automatically quantifying organ motion for adaptive radiation therapy of the prostate. This method extends deformable image registration to accommodate bowel gas, which creates image regions where no correspondence exists. We statistically analyzed the accuracy of our automatic method against the standard of manual inter-rater variation. We showed that for centroid and volume overlap of the prostate, the automatic method is statistically indistinguishable from human raters. We are currently working on applying our method to a larger number of patients and evaluating the clinical effect of organ motion by measuring effective delivered dose and biological effect.

Acknowledgments

We thank Ed Chaney, Ph.D., for his help with manual contouring of the CT data sets. We also thank Gregg Tracton, Nathalie Strehl, and Sandrine Tomei for their help organizing and processing the CT data sets. This work was supported by the DOD Prostate Cancer Research Program DAMD17-03-1-0134.

References

1. Yan, D., Lockman, D., Brabbins, D., Tyburski, L., Martinez, A.: An off-line strategy for constructing a patient-specific planning target volume in adaptive treatment process for prostate cancer. *International Journal of Radiation Oncology*Biography*Physics* **48** (2000) 289–302
2. van Herk, M., Bruce, A., Guus Kroes, A.P., Shouman, T., Touw, A., Lebesque, J.V.: Quantification of organ motion during conformal radiotherapy of the prostate by three dimensional image registration. *International Journal of Radiation Oncology*Biography*Physics* **33** (1995) 1311–1320
3. Ketting, C.H., Austin-Seymour, M., Kalet, I., Unger, J., Hummel, S., Jacky, J.: Consistency of three-dimensional planning target volumes across physicians and institutions. *International Journal of Radiation Oncology*Biography*Physics* **37** (1997) 445–453
4. Christensen, G.E., Carlson, B., Chao, K.S.C., Yin, P., Grigsby, P.W., N, K., Dempsey, J.F., Lerma, F.A., Bae, K.T., Vannier, M.W., Williamson, J.F.: Image-based dose planning of intracavitary brachytherapy: registration of serial-imaging studies using deformable anatomic templates. *International Journal of Radiation Oncology*Biography*Physics* **51** (2001) 227–243
5. Schaly, B., Kempe, J.A., Bauman, G.S., Battista, J.J., Dyk, J.V.: Tracking the dose distribution in radiation therapy by accounting for variable anatomy. *Physics in Medicine and Biology* **49** (2004) 791–805
6. Lu, W., Chen, M., Olivera, G.H., Ruchala, K.J., Mackie, T.R.: Fast free-form deformable registration via calculus of variations. *Physics in Medicine and Biology* **49** (2004) 3067–3087
7. Wang, H., Dong, L., O’Daniel, J., Mohan, R., Garden, A.S., Ang, K.K., Kuban, D.A., Bonnen, M., Chang, J.Y., Cheung, R.: Validation of an accelerated ‘demons’ algorithm for deformable image registration in radiation therapy. *Physics in Medicine and Biology* **50** (2005) 2887–2905

8. Joshi, S., Lorenzen, P., Gerig, G., Bullitt, E.: Structural and radiometric asymmetry in brain images. *Medical Image Analysis* **7** (2003) 155–170
9. Miller, M.I., Joshi, S.C., Christensen, G.E.: Large deformation fluid diffeomorphisms for landmark and image matching. In Toga, A.W., ed.: *Brain Warping*. Academic Press (1999)
10. Christensen, G.E., Rabbitt, R.D., Miller, M.I.: Deformable templates using large deformation kinematics. *IEEE Transactions On Image Processing* **5** (1996) 1435–1447
11. Amenta, N., Choi, S., Kolluri, R.K.: The power crust. In: *ACM Symposium on Solid Modeling and Applications*. (2001) 249–260
12. Dice, L.R.: Measures of the amount of ecologic association between species. *Ecology* **26** (1945) 297–302

Inter-Operator Variability in Perfusion Assessment of Tumors in MRI Using Automated AIF Detection

Edward Ashton¹, Teresa McShane², and Jeffrey Evelhoch³

¹ VirtualScopics, LLC, 350 Linden Oaks, Rochester, NY, 14580, USA
ashton@virtualscopics.com

² Pfizer Global Research and Development, Eastern Point Road, Groton, CT, 06340, USA
teresa.m.mcshane@pfizer.com

³ Amgen, One Amgen Center Drive, Thousand Oaks, CA, 91320, USA
evelhoch@amgen.com

Abstract. A method is presented for the calculation of perfusion parameters in dynamic contrast enhanced MRI. This method requires identification of enhancement curves for both tumor tissue and plasma. Inter-operator variability in the derived rate constant between plasma and extra-cellular extra-vascular space is assessed in both canine and human subjects using semi-automated tumor margin identification with both manual and automated arterial input function (AIF) identification. Experimental results show a median coefficient of variability (CV) for parameter measurement with manual AIF identification of 21.5% in canines and 11% in humans, with a median CV for parameter measurement with automated AIF identification of 6.7% in canines and 6% in humans.

1 Introduction

Dynamic contrast enhanced MRI (DCE-MRI) has demonstrated considerable utility in both diagnosing and evaluating the progression and response to treatment of malignant tumors.[1,2] DCE-MRI involves the periodic acquisition of T_1 -weighted images before, during and after injection of a gadolinium labeled tracer such as gadopentetate dimeglumine. The change over time in signal intensity in a voxel or region of interest in this time series can then be related to tracer concentration. By making use of a two-compartment model, with one compartment representing blood plasma and the other extra-vascular extra-cellular space (EES), the observed enhancement curves in tissue and plasma can be used to estimate various physiological parameters [3,4,5,6].

The parameter of primary interest in this work is the volume transfer constant between blood plasma and EES, commonly referred to as K^{trans} [7]. This parameter is related to both blood flow and endothelial permeability-surface area product, and is therefore a good endpoint for assessing the blood supply to a target malignancy.

One of the primary challenges in estimating perfusion parameters is identifying an accurate arterial input function (AIF). One common approach is to avoid this problem by making use of a general concentration-time curve such as that utilized by Weinmann *et al.* [8], as described by Tofts and Kermonde [9]. However, using a

theoretical AIF ignores differences in injection rate and cardiac output, which may vary from subject to subject and even with a single subject over time [10]. These differences can greatly reduce measurement reproducibility. Galbraith *et al.* have reported thresholds for statistically significant change over time in K^{trans} in a single tumor using this method of -45% and +83% [11]. A second option is for an analyst to draw a manual ROI within an artery, and use the mean enhancement curve within that ROI as the subject-specific AIF, as described by Vonken *et al.* [12]. This approach is complicated by the fact that the MR signal in arteries is frequently corrupted by flow artifacts, with the result that regions of interest at different points in the same artery or in other nearby vessels can provide grossly different enhancement curves. It should be noted that these artifacts are sequence dependent, and are worse for 2D methods than for central regions of 3D methods. This problem is illustrated in Fig. 1.

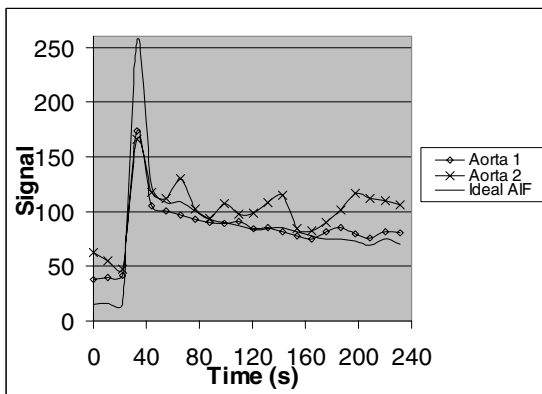


Fig. 1. Raw enhancement curves for two ROIs taken from the aorta of one human subject, with an automatically identified ideal AIF. Note that aorta 1 and aorta 2 show significantly different enhancement curves. Moreover, neither vessel shows the distinctive sharp peak and subsequent plateau characteristic of plasma enhancement following injection of a tracer bolus.

We have developed a method for the identification of an optimized AIF, described below, which is intended to eliminate this source of measurement variability and thereby increase the sensitivity to change of perfusion parameter measurements. This method is similar in aim to those presented previously by Rijpkema *et al.* [13] and Murase *et al.* [14]. Our method differs from the former in that it is completely automated, whereas the Rijpkema method requires some operator interaction, primarily to eliminate voxels consistent with veins rather than arteries. Our method is able to accomplish this without human interaction (see Fig. 2). Our method differs from the Murase technique in that it searches the entire volume for possible arterial voxels, whereas the Murase method begins with a user-defined arterial ROI, then divides that region into two or more sub-regions and averages the voxels in the optimal sub-region to generate its AIF. The AIF generated by our method for one human subject is shown in Fig. 1. Note that this curve shows a greater enhancement peak than that of either manually identified ROI, a smoother plateau, and a more characteristic shape. The primary purpose of this study was to assess the reproducibility of vascular perfusion

parameter calculation using this automated AIF identification technique with respect to that which is achievable using manual AIF identification. As a secondary point, we have also provided an example of the use of this technique to assess response to treatment of naturally occurring canine mammary tumors with an anti-vascular agent.

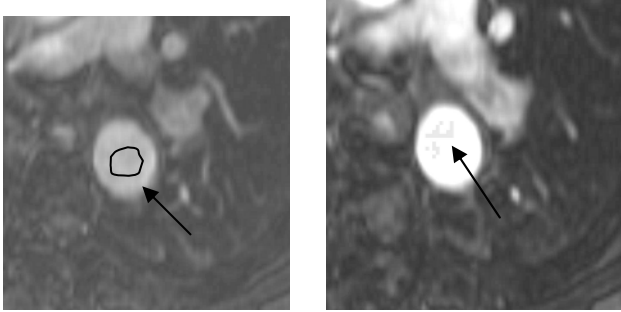


Fig. 2. (l) Manually drawn arterial ROI used to derive “Aorta 1” AIF shown in Fig. 1. (r) Light gray shaded voxels show the automatically identified voxel set used to derive “Ideal AIF” shown in Fig. 1.

2 Measurement Techniques

Tumor margins were identified in this study using a computer-aided boundary finding technique [15]. The resulting contour was then converted into a *snake* [16] which could be interactively corrected by the analyst if the initial result was sub-optimal.

After identifying the tumor margins, the analysts were required to identify a region of interest within an artery, preferably in close proximity to the tumor. This was done using manual tracing with a computer mouse. At this point, the identified arterial region was used for parameter calculation, as described below. In addition, the identified arterial region was used to initialize an automated search algorithm whose intent was to identify an optimized AIF for the data set under consideration. Each voxel in the data set was assigned a score based on time point of maximum enhancement, slope at maximum enhancement, peak value, and conformance to a gamma variate curve. This was accomplished by sorting all voxels in the data set based on the first three inter-related parameters. The highest scoring fifty voxels were then re-sorted based on the fourth parameter. The highest scoring twenty-five voxels in this sub-set were then assigned to the ideal AIF region of interest, and the automated AIF was generated by averaging the enhancement curves observed in these voxels (see Fig. 2).

After an arterial region of interest had been identified by either manual or automated means, enhancement curves were generated for both tumor and arterial plasma. These were designated $C_t(t)$ and $C_p(t)$, respectively. In the interests of noise reduction, both plasma and tumor data were fit to gamma-variate curves (17). The vascular bed was modeled as a linear system, such that:

$$C_t(t) = C_p(t) * h(t) \quad (1)$$

with impulse response $h(t)$ given by:

$$h(t) = K^{trans} e^{-k_{ep}t}, \quad (2)$$

where k_{ep} is the rate constant between the EES and blood plasma. Given $C_t(t)$ and $C_p(t)$, K^{trans} and k_{ep} were estimated using a gradient-descent energy minimization scheme in which local minima were avoided through the use of multiple instantiations with different initial parameter settings.

3 Experimental Procedure

The experiments involved in this study were intended to assess the reproducibility of perfusion measurements using manual and automated plasma identification, and to determine the percentage of measurement variability due to differences in tumor margin and plasma region of interest, respectively. Experimental data for the animal model were derived from three dogs with naturally occurring mammary tumors. These data were collected as part of an experiment testing an anti-angiogenic drug with vascular endothelial growth factor (VEGF) receptor tyrosine kinase inhibitory activity at a maximally tolerated dose in dogs. This drug was later discontinued from further development. Each animal was imaged three times over a period of 12 weeks. At each time point all animals were imaged in a single 48 hour period. Images for this study were acquired using a GE 1.5T LX/CV scanner. Three slices through each tumor were acquired using a cardiac coil. Perfusion images used a GRE pulse sequence with a repetition time of 20ms, echo time of 1ms, and a flip angle of 40 degrees. Imaging time for each image set was seven seconds, with a two second scanner delay, yielding temporal resolution for the data set of nine seconds. The reconstruction matrix was 256x192, FOV was 140mm, and slice thickness was 4mm.

Experimental data from humans were taken from a subset of images collected as part of a Phase I clinical trial. Twelve subjects with tumors of the lungs, liver, head or neck were imaged twice, at baseline (within 3 days before day 1 treatment with a VEGF/PDGFR/c-Kit tyrosine kinase inhibitor) and on day 2 (cycle 1) after the morning dose, at the estimated C_{max} . Images for this study were acquired using a GE 1.5T Signa LX scanner. Ten slices through each tumor were acquired using appropriate coils with respect to the area being imaged. Dynamic contrast enhanced images used a 3DFSPGR pulse sequence with a repetition time set by the scanner of 6-7ms, echo time of min full (the minimum TE at which a full echo can be acquired – in this case 2ms), and a flip angle of 25 degrees. Imaging time for the dynamic contrast enhanced scan consisted of 11 breath-holds at approximately 30 seconds per breath-hold, with two phases being acquired per breath-hold. The reconstruction matrix was 256x128, FOV was kept to the smallest allowable so that no wrap in the images was present, and slice thickness was 5mm.

Because a primary aim of this study was the assessment of inter-operator variability, four analysts were trained in the use of the analysis software. Two analysts were experienced MR technicians, and two were radiologists with extensive oncology experience. All analysts were trained in the appearance of canine mammary

tumors and the selection of appropriate plasma regions using images from animals not included in this study. Each analyst was then asked to identify and delineate both tumor and plasma in each of the nine included canine data sets as well as each of the 24 human subject data sets. When identifying arterial plasma, the analysts had the option to view the enhancement curve for the manually selected region at any time, and to erase, modify or replace the currently selected region. In this way each analyst was able to manually select a reasonably optimized arterial plasma region.

Once all regions of interest were delineated, K^{trans} values were calculated first using the regions of interest as identified by the analyst, and then using the analyst's tumor identification with the automatically identified plasma enhancement curve. By comparing the variance seen between analysts using manually identified AIF with that seen between analysts using the automatically identified AIF, which was identical across analysts, it was possible to isolate variability related to AIF selection from that related to differences in tumor margin identification.

4 Results and Discussion

Coefficients of variability in measurement of K^{trans} among the four analysts, defined as measurement standard deviation divided by measurement mean, were calculated separately for manual and automatic AIF identification, and for each of the nine canine and 24 human cases examined. For the nine canine manual AIF identifications, coefficients of variability ranged from 3.1% to 39.2%, with a root mean square average of 23.7% and a median value of 21.5%. For the nine canine automated AIF identifications, coefficients of variability ranged from 3.1% to 11.8%, with a root mean square average of 7.0% and a median value of 6.2%. For the 24 human manual AIF identifications, coefficients of variability ranged from 1% to 32%, with a root mean square average of 14.1% and a median of 11%. For the 24 human automated AIF identifications, coefficients of variability ranged from 1% to 24%, with a root mean square average of 11.2% and a median of 6%. It should be noted that in the human data there was a single case where the analysts disagreed significantly over the correct tumor boundaries. This case was the source of the high maximum variation in both the manual and automated AIF results.

Bearing in mind that the same tumor margins were used for both the automated and manual AIF calculations, in the canine experiment it can be generally surmised that approximately two thirds of the variability seen in the manual measurements was a result of differing AIF identifications, with the remaining one third attributable to differing tumor margin identifications. In the human experiments the median variability due to tumor margin identification was similar – approximately 6% -- while the variability due to AIF identification was significantly lower. This is reasonable, as the smaller vessels and higher blood velocity present in the canines makes manual identification of an uncorrupted arterial region much more difficult than in the human case. Additionally, the 2D imaging method used in the canine study is more sensitive to in-flow effects than the 3D method used in the human study.

An examination of scatter-plots of K^{trans} measurements using manual vs. automatic AIF identification for both canine and human data (see Fig. 3) shows that the correlation between the two measures is reasonable given the high variability of

the manual measurements. It also shows a slope of 0.58 for the canine data and 0.85 for the human data, indicating that on average the manual measurement gives a somewhat higher estimation of K^{trans} than the automatic measurement, and that this effect is more pronounced in the canine data than in the human data. This is as expected, since the general effect of flow artifacts will be to reduce the apparent plasma enhancement, thereby exaggerating the proportion of tracer apparently passing into the EES.

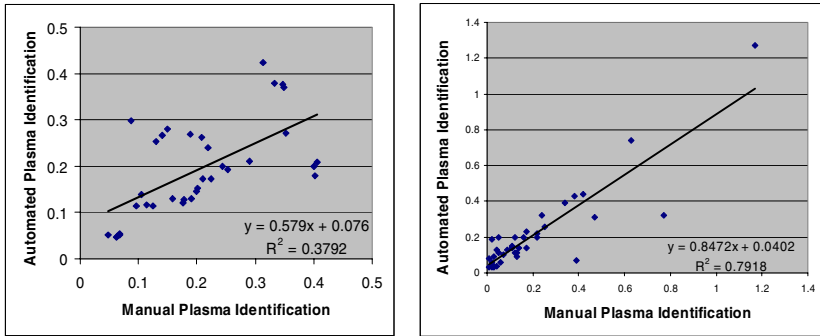


Fig. 3. Scatterplots of K^{trans} values using manual and automatic plasma identification for canine (l) and human (r) data sets. Note that in both cases the trend line slope is less than 1, due to the fact that manual values are generally over-estimated due to poor plasma identification. This effect is less pronounced in the human data, and the correlation between manual and automated measurements is higher due to the improved manual reproducibility in the human data.

An examination of the trend over time in K^{trans} for canine subject 1 (see Fig. 4) using manual and automated AIF identification highlights the value of the reduced measurement variability afforded by the automated process. Although both trend lines indicate that K^{trans} for this tumor is declining over time, higher variability makes that assumption statistically insupportable at time two for manual AIF identification, and marginally supportable after time three. Using automated AIF identification, however, this subject may be confidently classified as declining after time 2.

5 Conclusions

The two most commonly used methods for AIF definition in current clinical and experimental practice are manual AIF identification and use of a theoretical or model-based AIF. The results of this study indicate that increased accuracy and sensitivity to change could be achieved by making use of an automated method for AIF identification such as the one described here. Use of such a method allows AIF identification to be both consistent and repeatable among operators, and also patient and time-point specific. Variability due to partial volume effects and artifacts is reduced with respect to that of manual AIF identification, while variability due to

differences in injection and cardiac output is reduced with respect to that seen with the use of a model AIF.

In considering the differences between human and canine results given here, it should be noted again that the difficulty of identifying a suitable AIF is typically greater in smaller animals such as the canines used in this study than in humans. This is due to small animals' higher blood velocity, which exaggerates flow artifacts in the arteries, as well as to the lower signal to noise ratio that is achievable when imaging smaller anatomy.

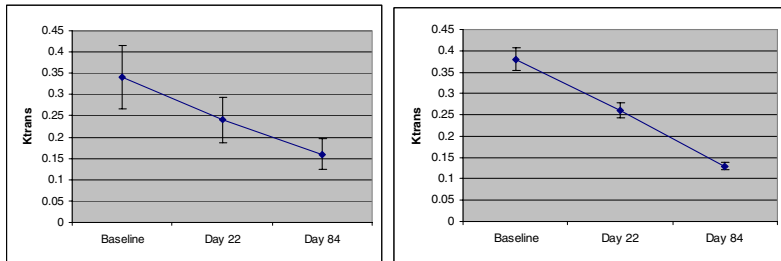


Fig. 4. Trend line using manual (l) and automated (r) plasma identification for canine subject 1. Note that the subject cannot be confidently classified as declining until time 3 using manual plasma identification, while the reduced variability of the automated measurements allows a confident classification after time 2.

References

1. Orel SG, Differentiating benign from malignant enhancing lesions identified at MR imaging of the breast: are time-signal intensity curves an accurate predictor? *Radiology* 1999; 211:5 – 7.
2. Taylor J, Tofts P, Port R, Evelhoch J, Knopp M, Reddick W, Runge V, Mayr N. MR imaging of tumor microcirculation: Promise for the new millennium. *J Magn Reson Imag* 1999; 10:903 – 907.
3. Jackson A, Haroon H, Zhu X, Li K, Thacker N, Jayson G. Breath-hold perfusion and permeability mapping of hepatic malignancies using magnetic resonance imaging and a first-pass leakage profile model. *NMR Biomed* 2002; 15:164 – 173.
4. Tofts P. Modeling tracer kinetics in dynamic Gd-DTPA MR Imaging. *J Magn Reson Imag* 1997; 7:91 – 101.
5. Evelhoch J. Key factors in the acquisition of contrast kinetic data for oncology. *J Magn Reson Imag* 1999; 10:254 – 259.
6. Murase K. Efficient method for calculating kinetic parameters using T1-weighted dynamic contrast-enhanced magnetic resonance imaging. *Magn Reson Med* 2004; 51:858 – 862.
7. Tofts P, Brix G, Buckley D, Evelhoch J, Henderson E, Knopp M, Larsson H, Lee T, Mayr N, Parker G, Port R, Taylor J, Weisskoff R. Estimating kinetic parameters from dynamic contrast enhanced T1 weighted MRI of a diffusible tracer: standardized quantities and symbols. *J Magn Reson Imag* 1999; 10:223 – 232.
8. Weinmann HJ, Laniado M, Mutzel W. Pharmacokinetics of Gd-DTPA/dimeglumine after intravenous injection into healthy volunteers. *Physiol Chem Phys Med NMR* 1984; 16:167 – 172.

9. Tofts P, Kermode A. Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental Concepts. *Magn Reson Med* 1991; 17:357 – 367.
10. Padhani A, Hayes C, Landau S, Leach M. Reproducibility of quantitative dynamic MRI of normal human tissues. *NMR Biomed* 2002; 15:143 – 153.
11. Galbraith S, Lodge M, Taylor N, Rustin G, Bentzen S, Stirling J, Padhani A. Reproducibility of dynamic contrast-enhanced MRI in human muscle and tumours: comparison of quantitative and semi-quantitative analysis. *NMR Biomed* 2002; 15:132 – 142.
12. Vonken E, Osch M, Bakker C, Viergever M. Measurement of cerebral perfusion with dual-echo multi-slice quantitative dynamic susceptibility contrast MRI. *J Magn Reson Imag* 1999; 10:109 – 117.
13. Rijpkema M, Kaanders J, Joosten F, van der Kogel A, Heerschap A. Method for quantitative mapping of dynamic MRI contrast agent enhancement in human tumors. *J Magn Reson Imag* 2001; 14:457 – 463.
14. Murase K, Kikuchi K, Miki H, Shimizu T, Ikezoe J. Determination of arterial input function using fuzzy clustering for quantification of cerebral blood flow with dynamic susceptibility contrast-enhanced MR imaging. *J Magn Reson Imag* 2001; 13:797 – 806.
15. Ashton E, Takahashi C, Berg M, Goodman A, Totterman S, Ekholm S. Accuracy and reproducibility of manual and semiautomated quantification of MS lesions by MRI. *J Magn Reson Imag* 2003; 17:300 – 308.
16. Kass M, Witkin A, Terzopolous D. Snakes: active contour models. *Int J Comput Vision* 1988; 321 – 331.
17. Moran G, Prato F. Modeling tissue contrast agent concentration: a solution to the tissue homogeneity model using a simulated arterial input function. *Magn Reson Med* 2001; 45: 42 – 45.

Computer-Assisted Deformity Correction Using the Ilizarov Method

A.L. Simpson¹, B. Ma¹, D.P. Borschneck², and R.E. Ellis^{1,2,3}

¹ School of Computing

² Department of Surgery

³ Department of Mechanical Engineering,

Queen's University, Kingston, Ontario, Canada, K7L 3N6

{simpson, mab, ellis}@cs.queensu.ca

Abstract. The Taylor spatial frame is a fixation device used to implement the Ilizarov method of bone deformity correction to gradually distract an osteotomized bone at regular intervals, according to a prescribed schedule. We improve the accuracy of Ilizarov's method of osteogenesis by preoperatively planning the correction, intraoperatively measuring the location of the frame relative to the patient, and computing the final shape of the frame. In four of five tibial phantom experiments, we were able to achieve correction errors of less than 2 degrees of total rotation. We also demonstrate how registration uncertainty can be propagated through the planned transformation to visualize the range of possible correction outcomes. Our method is an improvement over an existing computer-assisted technique (Iyun *et al.* [3]) in that the surgeon has the same flexibility as in the conventional technique when fixating the frame to the patient.

1 Introduction

Rotational and translational deformities in the long bones are commonly corrected using a method of osteogenesis developed by Russian orthopaedic surgeon Gavril Ilizarov. Ilizarov's method is based on the biological principle of inducing new bone growth by gradually distracting a fracture at regular intervals. More specifically, the method is achieved by performing a corticotomy on the deformed bone, fixating the distressed bone with a mechanical fixator, and distracting the bone according to a set schedule of corrections. This technique has been successfully applied to treating malunions, nonunions, bone defects, limb elongation, fractures, and angular deformities, to name a few (Feldman *et al.* [2]).

Conventional surgical technique for the Ilizarov fixator is plagued by two sources of error: (1) preoperatively planning the required correction requires the precise measurement of 13 parameters from anteroposterior and lateral radiographs, and (2) angular and translational errors may be present once the frame is mounted to the patient.

1.1 The Taylor Spatial Frame

The Taylor spatial frame (Smith & Nephew, Memphis, TN) is an external fixator that combines gradual distraction principles of the Ilizarov method with deformity analysis provided by a computer program. The frame consists of six telescopic rods (called

struts) that connect two circular bases (or *rings*), in a symmetric configuration of a Stewart–Gough platform [7]. By simply adjusting strut lengths, one ring moves with respect to the other and can be positioned to mimic any deformity (Taylor and Paley [9]).

Preoperatively, the surgeon determines the nature of the deformity, the desired correction, and the final (or *neutral*) height of the frame. Six specific strut lengths are calculated by a computer program using the initial and desired final frame configurations. Intraoperatively, the frame is attached to the bone by placing each ring substantially perpendicular to the bone, lengthening the six struts according to the preoperative plan, and fixating the frame with a combination of Kirschner wires, Steinman pins, and Rancho cubes (for attaching the wires or pins to the rings). The correction schedule is prescribed postoperatively by the surgeon. Once the correction schedule is complete and the frame is in its neutral position, any residual deformity is corrected by applying a secondary correction schedule; this residual correction phase is usually required (Taylor and Paley [9]).

1.2 Related Work

Early work by Kochs [4] attempted to reduce complications due to incorrect preoperative planning and inaccurate application of the frame by simulating the planned correction. Optimal joint positions and ring locations were obtained by simulation on images acquired from hand–tracing radiographs and scanning these images. Postoperatively, a radiograph was compared to the preoperative plan to determine the necessary residual corrections. Lin *et al.* [5] proposed a preoperative planning system for the Ilizarov method by (1) creating a bone template using an ultrasonic digitizer, (2) manually characterizing the deformity from radiographs and patient examination, (3) determining the weight–bearing axis, (4) performing virtual osteotomies on the computer, (5) aligning the bone fragments in the preoperative plan, (6) constructing the frame based on a physical examination of the patient, and (7) assembling the frame using a life–size diagram of the fixator assembly output by the computer.

More recently, Iyun *et al.* [3] proposed a method to apply the inverse kinematics of the Taylor spatial frame to calculate the initial position of the frame and fixation pins, the strut lengths, and the daily schedule of corrections. This research combined preoperative planning of the strut lengths with intraoperative guidance of the placement of the Kirschner wires and Steinman pins. Their methodology has two impractical assumptions. The first and most significant assumption was that the frame is always mounted using rigid pins or wires, which is not the case when a ring is mounted close to a joint line (where weaker bone can result in deviations from the rigid–fixation assumption). The second assumption was that the location and direction of the fixation pins could be determined during the planning phase. In practice, the configuration of the Rancho cubes, pins, and wires are best chosen intraoperatively because of anatomical constraints that may not be apparent preoperatively. The laboratory study was further limited by a learning effect present in the results; once the surgeon mounted the first frame in the conventional manner, subsequent frames were mounted without error.

1.3 Our Approach

Rajacich *et al.* [6] observed that a single point of failure in applying the Ilizarov method is planning the procedure. In the case of the Taylor spatial frame, 13 frame parameters

must be measured from the patient and radiographs. Since the Taylor frame is tightly coupled (Stewart [7]), errors in any one parameter propagate through the entire preoperative plan. A second source of error is the misapplication of the frame such that translational and angular problems are introduced during surgery.

We aim to improve the efficacy of the Ilizarov method by (1) eliminating the need to measure parameters preoperatively, (2) allowing the surgeon complete flexibility in frame placement and configuration, and (3) improving the accuracy of the correction. Our method relies on the idea that it is not necessary for the frame to end in a neutral, or highly symmetric, configuration. The surgeon can simply mount the frame on the patient and, based on the actual position of the frame relative to the patient, we can compute the strut lengths and required schedule of adjustments to achieve the desired correction.

2 Methodology

In this section, we describe the conventional surgical technique for applying Ilizarov's method using the Taylor spatial frame as well as the proposed computer-assisted technique.

2.1 Traditional Surgical Technique

In the conventional technique, the surgeon measures the deformity, plans the necessary correction, and specifies the 13 mechanical parameters of the Taylor spatial frame. The three rotational parameters are measured from plain X-ray images using anteroposterior angulation, lateral angulation, and axial angulation views. The translational parameters are measured in a similar way. The frame itself is described by the proximal and distal ring diameters and the neutral frame height. The four remaining parameters are obtained in clinic and measure the location and axial rotation of the reference bone fragment with respect to the reference ring; the measurement of these four frame offset parameters is described by Taylor [8]. The Taylor spatial frame is kinematically equivalent to the Stewart platform, which is fully coupled (Stewart [7]); hence, any changes in the length of any one strut results in changes to all six strut lengths. The 13 parameters are used by the computer program supplied by the manufacturer of the frame to generate the six specific strut lengths and the daily schedule of adjustments that must be made to the struts by the patient (Taylor and Paley [9]).

There are three standard methods of surgically mounting the spatial frame. The chronic deformity-correction method requires that the surgeon attach the deformed frame (which would mimic the deformity) to the patient; the deformity is fully corrected once the frame reaches its neutral shape with all six struts having equal length. Alternatively, the rings-first method of deformity correction mounts the rings to the patient prior to attaching the struts. Finally, the residual deformity-correction method simply compensates for any residual deformity which may exist after either of the first two methods is used. In each of these methods, the surgeon mounts the rings perpendicular to the weight-bearing axis of the limb under fluoroscopic guidance using either Steinman pins or flexible tensioned Kirschner wires and centers the frame on the bone (Taylor and Paley [9]).

2.2 Computer-Assisted Technique

Our proposed technique modifies conventional approaches to deformity correction using the Ilizarov method in four fundamental ways:

1. The need for the surgeon to preoperatively measure the 13 frame parameters is removed.
2. The performed correction is based on the actual location of the frame with respect to the anatomy; any translational or angular problems that occur while mounting the frame are compensated for immediately thereby potentially removing the need for residual correction.
3. The correction is calculated based on 3D coordinates from CT data rather than measured from multi-planar radiographs.
4. The rings do not have to mimic the deformity; we essentially bypass the “chronic deformity correction” step to the “total residual correction” phase.

Unlike the computer-assisted method described by Iyun *et al.* [3], our method allows the surgeon to use any configuration of Rancho cubes, pins and wires.

We performed a small number of experiments using tibial phantoms (Sawbones, Pacific Research Laboratories, Inc., Vashon, Washington, USA). Our apparatus consisted of an Optotrak optical tracking system (Northern Digital Inc., Waterloo, Ontario, Canada) dynamic reference bodies (DRBs) attached to the proximal and distal ends of the tibia phantom, a tracked surgical probe, and 5 tibia phantoms¹.

For each of the tibia phantoms, a 3D surface model was constructed from CT data. Planning software was used to plan the necessary correction (a normal bone phantom was deformed in an arbitrary way or a deformed phantom was corrected). The Taylor spatial frame was mounted to the bone phantom using tensioned Kirschner wires and Steinman pins. Note that we could have chosen to intraoperatively guide the placement of the rings; however, for the purposes of this study, we chose to mount the rings arbitrarily in order to demonstrate that we can compensate for errors in frame placement. A DRB was attached to the proximal and distal ends of the tibia phantom. We registered the bone phantom to the 3D model using a registration algorithm based on Besl and McKay’s [1] ICP algorithm. Approximately 20 registration points were collected from the osteotomy region as well as from other surfaces that could be easily digitized percutaneously, such as the shaft and medial malleolus of the tibia. We then digitized three well-defined points on each ring of the frame. Using the ring measurements and the registration transformation, we computed the location of the rings in CT coordinates. Using the planned transformation of the mobile proximal fragment, we transformed the location of the proximal ring to its planned corrected location and calculated the necessary strut lengths. The tibia phantom was then cut and distracted by changing the strut lengths to those calculated by our model. Figure 1 demonstrates the Taylor spatial frame after correction is achieved. We chose three types of corrections which are visualized in Figure 2.

¹ We were limited to a small phantom study because our apparatus is used clinically by our affiliated hospital.

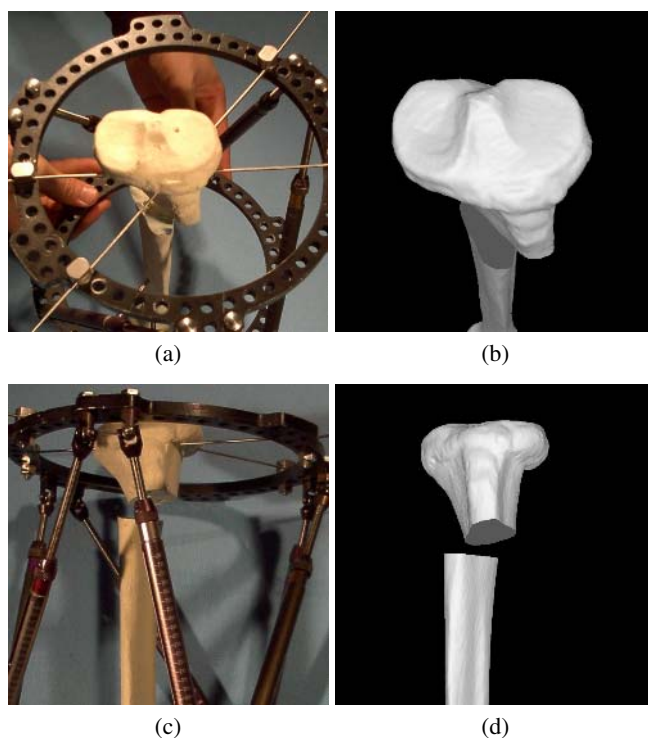


Fig. 1. Frame after correction in (a) axial and (c) lateral views and corresponding computer model of the planned correction in (b) axial and (d) lateral views

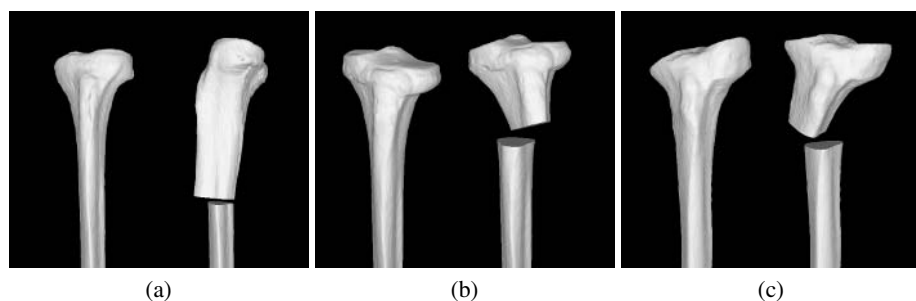


Fig. 2. The three six-degrees-of-freedom planned corrections used in our experiments. The primary modes of correction were (a) axial rotation, (b) lateral opening wedge, and (c) medial opening wedge.

3 Results

We originally attempted to track the motion of the distracted proximal end with respect to the distal end by tracking both ends of the phantom, but we could not reliably

Table 1. Alignment errors between planned and performed corrections

Correction	Rotational Errors (deg)				Translational Errors (mm)			
	θ	ϕ	δ	Total	x	y	z	Total
axial rotation	0.7	1.2	-4.2	4.4	-1.4	-1.2	1.3	2.23
lateral wedge	0.3	-0.2	0.7	0.78	-0.1	0.8	0.1	0.84
lateral wedge	-1.1	-0.2	1.4	1.82	7.1	-2.2	0.7	7.50
medial wedge	1.9	0.8	1.1	1.32	-0.9	-3.5	-2.5	4.37
medial wedge	-0.4	0.1	1.1	1.16	2.8	6.5	1.3	7.22

maintain fixation of the reference targets during the cutting and distraction processes because of the intense vibration associated with the cutting process. Instead, we digitized anatomic landmarks along with widely spaced registration points P and D separately from the proximal and distal ends, respectively, not restricting ourselves to surgically accessible surfaces. The registration transformation $\mathbf{T}_{D,CT}$ of the distal end to CT coordinates was calculated and applied to the proximal registration points. The transformed proximal points $P' = \mathbf{T}_{D,CT}P$ were registered to CT coordinates to obtain the transformation $\mathbf{T}_{P',CT}$, and the displacement of the proximal end with respect to the distal end in CT coordinates was the inverse of this transformation $\mathbf{T}_{PD,CT} = \mathbf{T}_{P',CT}^{-1}$.

The rotation component $\mathbf{R}_{PD,CT}$ of $\mathbf{T}_{PD,CT}$ was compared to the rotation component \mathbf{R}_{plan} of the planned correction \mathbf{T}_{plan} by computing the difference in rotation Δ where $\mathbf{R}_{PD,CT} = \Delta\mathbf{R}_{plan}$ and converting Δ to its screw representation; the rotation about the screw axis was the total angular error. We also calculated errors using Taylor's rotation matrix (Iyun *et al.* [3]) $\mathbf{R}_{\theta,\phi,\delta}$ where θ , ϕ , and δ were the projected angles of rotation measured in the anteroposterior, lateral, and axial views. The error in translation was taken to be the difference between the center of the proximal ring under $\mathbf{T}_{PD,CT}$ and \mathbf{T}_{plan} . All error measurements are tabulated in Table 1.

Retrospectively, we computed the uncertainty of the registration parameters for the axial deformity case by using a particle filter as part of the registration algorithm, de-

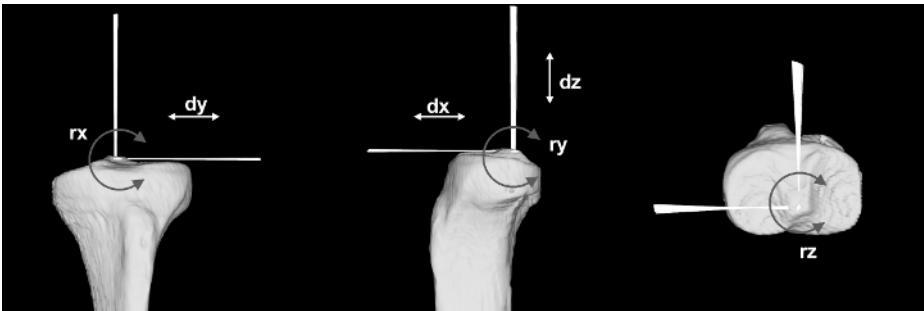


Fig. 3. Uncertainty in the location and orientation of the mobile proximal fragment for the axial rotation case. The total range of rotational uncertainty was 1.3° about the anteroposterior (AP) axis (rx), 1.3° about the mediolateral (ML) axis (ry), and 5.4° about the long axis (rz) of the fragment. The total range of translational uncertainty was 3.0mm along the AP axis (dx), 2.5mm along the ML axis (dy), and 3.9mm along the long axis (dz).

scribed by Ma *et al.* [10]. This algorithm produced a sampled distribution of the registration parameters rather than a single point estimate. The distribution of registration transformations was used to predict the range of the expected location and orientation of the mobile fragment. Figure 3 illustrates what happens to the three anatomical axes of the mobile fragment when the distribution of registration transformations is propagated through the planned correction transformation.

4 Discussion

It appears that the proposed method successfully implements the Ilizarov method of deformity correction using the Taylor spatial frame with computer assistance. The accuracy of our computer-assisted technique is better than the previous computer-assisted technique by Iyun *et al.* [3], which reported mean rotational and translational error of 3.2 degrees and 5.4 degrees, respectively. The major source of error in our method lies in the mechanical loading of the bone phantom when mounting the frame, which resulted in some strain of the phantom; the stress was released when the phantom was cut, displacing the bone fragments in the transverse (XY) plane. This phenomenon does not occur in a clinical setting. In the case of the large axial rotation correction, the final configuration of the frame was very unusual. Consequently, we found that there was significant rotational laxity about the vertical axis greater than our reported correction error of -4.2 degrees. Since we were simulating a deformity rather than a correction in this case, the laxity we observed would not occur clinically after a correction.

The primary disadvantage of our method is that a preoperative CT scan of the patient is needed, which is not generally required by the conventional surgical technique, in order to construct the 3D computer model. Furthermore, generating computer models and preoperative plans can be labor intensive. However, manually computing the 13 frame parameters in the traditional method requires approximately one hour to complete. Indeed, surgeons are conditioned to planning corrections based on bi-planar radiographs rather than 3D models; hence, it is unclear whether surgeons would be willing to visualize the necessary corrections in 3D. Finally, our method does not rely on the existing clinically used software provided by the frame manufacturer; therefore, significant testing is required to ensure the correctness of our system. A weakness of this study is the small sample size used in our experiments; we are attempting to perform more experiments as clinical conditions permit.

In clinical practice, the preferred method of registration would probably be from fluoroscopic images to the CT volume or anatomical atlas. We note that the registration algorithm would need to cope with occlusion artifacts caused by the presence of the metal rings of the frame. If we were able to register to an atlas, we would be able to use this method for trauma cases as this would eliminate the need for a CT scan. Registration remains the single point of failure in this method.

Practitioners of Ilizarov's method for deformity correction using the Taylor spatial frame admit that there is a steep learning curve in using the frame. It has been shown by Feldman *et al.* [2] that increased surgical experience with the system decreases the complication rate and increases the accuracy of correction using the frame. This is in part due to the fact that it is difficult to accurately measure the 13 frame parameters

and to mount the frame without some residual rotational and translational errors. Our technique aims to reduce complications due to these factors by preoperatively planning the desired correction directly and reducing the possibility of errors introduced during surgery by calculating the correction based on the actual location of the frame with respect to the anatomy. Moreover, the surgeon has greater flexibility in choosing the position of the rings since this technique does not depend on placing the rings in a particular orientation. This relatively new approach of bypassing the “chronic deformity correction” stage directly to the “total residual correction” step is reported by Feldman *et al.* [2].

References

1. P. J. Besl and N. D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.
2. D. S. Feldman, S. S. Shin, S. Madan, and K. J. Koval. Correction of tibial malunion and nonunion with six-axis analysis deformity correction using the Taylor spatial frame. *Journal of Orthopaedic Trauma*, 17:549–554, 2003.
3. O. Iyun, D. P. Borschneck, and R. E. Ellis. Computer-assisted correction of bone deformities using a 6-DOF parallel spatial mechanism. In *MICCAI 2001*, pages 232–240. Springer Lecture Notes in Computer Science 2488, 2001.
4. A. Kochs. Computer-assisted assembly and correction simulation for complex axis deviations using the Ilizarov fixator. *Archives of Orthopaedic and Trauma Surgery*, 114:287–291, 1995.
5. H. Lin, J. G. Birch, M. L. Samchukov, and R. B. Ashman. Computer-assisted surgery planning for lower extremity deformity correction by the Ilizarov method. *Journal of Image Guided Surgery*, 1:103–108, 1995.
6. N. Rajacich, D. F. Bell, and P. F. Armstrong. Pediatric applications of the Ilizarov method. *Clinical Orthopaedics and Related Research*, 280:72–80, 1992.
7. D. Stewart. A platform with six degrees of freedom. *Proceedings of the Institute for Mechanical Engineers*, 180(1):371–386, 1965.
8. J. C. Taylor. The correction of general deformities with the Taylor spatial frame fixator. <http://www.jcharlestaylor.com>, 2005.
9. J. C. Taylor and D. Paley. Six axis deformity analysis and correction. In *Principles of Deformity Correction*, Springer-Verlag, New York, 2002.
10. B. Ma and R. E. Ellis. Surface-based registration with a particle filter. In *MICCAI 2004*, pages 566–573, Springer Lecture Notes in Computer Science 3216, 2004.

Real-Time Interactive Viewing of 4D Kinematic MR Joint Studies

Heinrich Schulz¹, Kirsten Meetz¹, Clemens Bos², Daniel Bystrov¹, and Thomas Netsch¹

¹ Philips Research Laboratories, Sector Technical Systems,
Roentgenstrasse 24-26, D-22335 Hamburg, Germany
heinrich.schulz@philips.com

² Philips Medical Systems
Veenpluis 4-6, 5680 DA Best, The Netherlands

Abstract. Assessment of soft tissue in normal and abnormal joint motion today gets feasible by acquiring time series of 3D MRI images. However, slice-by-slice viewing of such 4D kinematic images is cumbersome, and does not allow appreciating the movement in a convenient way. Simply presenting slice data in a cine-loop will be compromised by through-plane displacements of anatomy and “jerks” between frames, both of which hamper visual analysis of the movement. To overcome these limitations, we have implemented a demonstrator for viewing 4D kinematic MRI datasets. It allows to view any user defined anatomical structure from any viewing perspective in real-time. Smoothly displaying the movement in a cine-loop is realized by image post processing, fixing any user defined anatomical structure after image acquisition.

1 Introduction

Starting in the late eighties, MRI has been used for imaging of moving joints [1,2]. Most approaches were based on fast 2D imaging, which limited the study to a single predefined view of a few anatomical structures. In addition, in order to keep the anatomy of interest stable in the imaging plane, devices are needed to restrain a part of the joint [3], thereby limiting the freedom of movement. As a result pathological behavior in the kinematics of the investigated structures might be alleviated or even not show up at all.

MRI is capable of acquiring time series of 3D images, which gives a 4D examination that can be used for kinematic joint imaging of an unrestricted movement. Furthermore, 4D data acquisition greatly simplifies scan planning, because it obviates the need for defining and tracking the diagnostically relevant 2D view during image acquisition. However, slice-by-slice viewing of the 4D images is cumbersome, and does not allow to appreciate the movement. Simply presenting slice data in a cine-loop will be compromised by through-plane displacements of anatomy and “jerks” between frames, both of which hamper visual analysis of the movement.

We have implemented a demonstrator for viewing 4D kinematic MRI datasets that addresses the following requirements: 1) The motion of any anatomy can be viewed from any perspective. 2) The user may define that object or region of interest, e.g. the distal femur in kinematic joint imaging of the knee, which should remain fixed in the

viewing plane during the movement. 3) The user interaction is reflected immediately in the viewing plane.

2 Method

2.1 Viewing Workstation

The proposed viewing workstation for 4D kinematic studies comprises two main steps of user interaction: the definition of the anatomical structure to be stabilized in visualization and real-time interaction with the video component in order to refine the selection of the anatomical structure of choice as well as to control imaging parameters like the zoom factor. Fig 1 shows a typical configuration of the user interface.

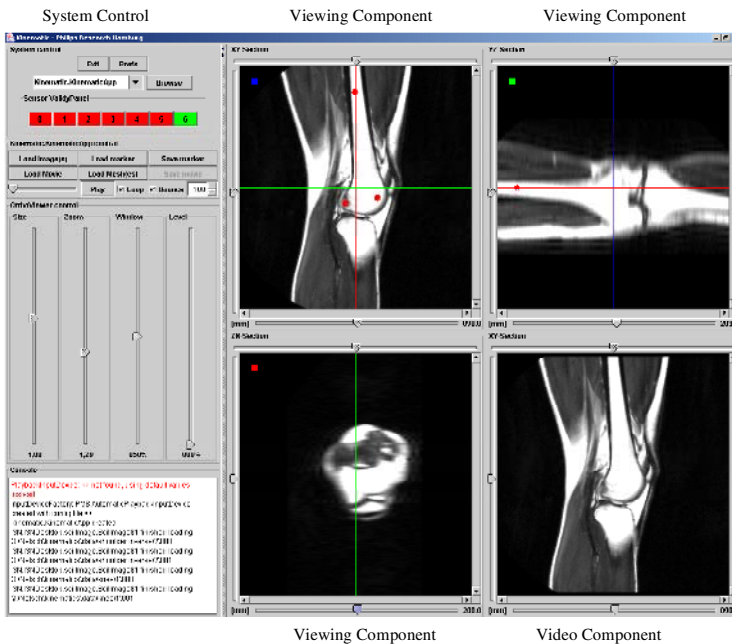


Fig. 1. User interface of the kinematic viewing station

2.1.1 Definition of Anatomical Structure

The viewing workstation allows the inspection of 4D data sets in three orthogonal viewing planes. It comprises the well-known 3D Orthoviewer functionalities depicting one 3D data set. In addition it allows switching between n 3D data sets retaining a previously defined viewing perspective with respect to the scanner coordinate system. Viewing functionality is implemented to enable the user to navigate through the data sets by translation and rotation, to scale the data set, and to manipulate window and level. Thus the entire 4D data set can be inspected from any user selected perspective. The definition of an anatomical structure of interest is done in order to fix this

structure for visualization in cine mode. The fixation can be applied to virtually arbitrary regions and comprises two steps:

First, the movement of an anatomical structure through all 3D data sets is calculated by motion tracking. Therefore, the object or region of interest is defined interactively by setting reference points in one of the 3D data sets. The user has to set a minimum of 3 non-collinear reference points by clicking at any appropriate point within the viewing component. These points are propagated to all other data sets by using a motion estimation method. The propagation of the selected points for all data sets is calculated in advance by an elastic registration approach. The tracking of the object or region of interest is calculated by rigid point-based registration of the n sets of reference points using singular value decomposition [4]. Both methods are described in Section 2.2 in detail.

Secondly, the inverse of the transformation defined by the tracking is used to align the 3D data sets, such that the defined objects or regions of interest remain stable when presented in cine loop.

2.1.2 Real-Time Interaction

The anatomical structures to be inspected for medical examination do not necessarily match those being fixated. When the kinematics data set is presented in cine mode the ability to manipulate the visualized part in real-time is provided to the user. For that purpose, the previously described user-interactions of defining a viewing perspective and defining an anatomical structure to be stabilized are combined appropriately. The combination of image transformations is applied to each of n 3D data sets, showing the stabilized anatomy in the defined viewing perspective in cine mode. Every user interaction concerning viewing aspects like the selection of reference points or the change of the viewing perspective is reflected immediately in the video component.

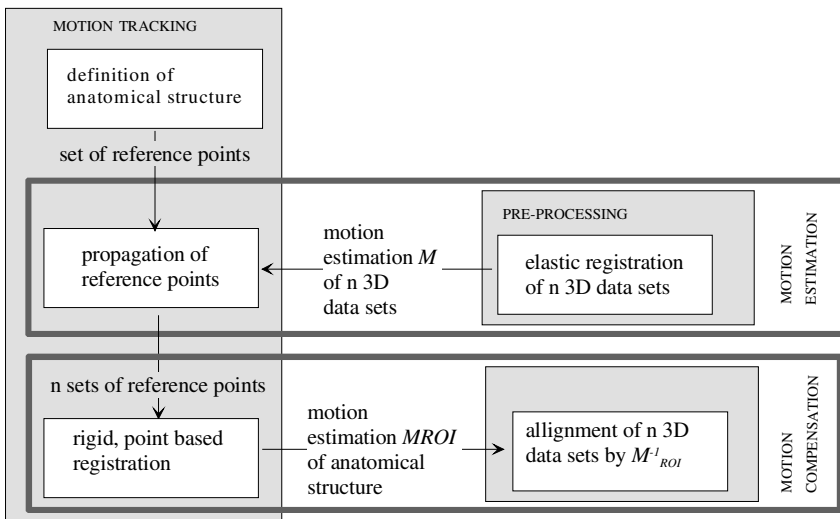


Fig. 2. Methods of image post-processing for fixation of an anatomical structure

2.2 Image Post-processing

The essential part of the proposed kinematic workstation is motion tracking and motion compensation, which is realized by two different registration approaches (see Fig. 2).

2.2.1 Motion Estimation

Motion estimation of the entire kinematic 3D images is used to propagate the user defined reference points (see Section 2.1.2). The motion estimation is calculated by an elastic volumetric grey-value registration method based on a B-spline representation [5,6]. A mesh of control points with uniform spacing in each direction defines the transformation of the image. As cost function the sum of squared differences (SSD) over the total number of voxels is selected because the image contrast does not significantly change in the series. The optimization routine then determines from a set of admissible displacements a transformation, which minimizes this cost function. The optimization needs special attention, since the large number of parameters involved may result in high computational costs. We therefore choose an iterative Levenberg-Marquardt method. Furthermore the B-spline registration is embedded into a multi-scale approach employing both an image pyramid and a parameter pyramid of three to four levels. The multi-scale approach provides additional regularization and an increased capture range, and speeds up the overall algorithm.

2.2.2 Motion Compensation

Motion compensation is used to keep the user-defined anatomical region fixed in the viewing plane. Therefore the rotational and translational component of the region's motion has to be estimated, based on the selected and propagated set of reference points. Due to the propagation, a correspondence between all sets of reference points has been installed. Thus a rigid, point-based registration approach can be used, with the aim to estimate the motion of consecutive sets of reference points. Following Arun et.al. [4] the motion can be estimated by minimizing the least square distance of corresponding points of two 3D point sets. Using a singular value decomposition (SVD), the rotational matrix and based on it the translational vector can be calculated.

Since the user is not restricted in the selection of reference points these do not necessarily follow the same motion trajectories through the kinematic data sets. This might be desired in order to stabilize soft tissue structures like ligaments, but for reference point sets completely placed on solid structures like bones the point set should move in parallel. Therefore the SVD error provides an easy control value to verify quantitatively the reliability of the registration results as well as to check for misplaced reference points.

3 Result

3.1 Viewing Component

Our approach has been tested for different 4D kinematic MR data sets of the knee, the shoulder and the foot. The MR datasets were obtained using T1-weighted 3D gradient echo sequences for 7 different positions of the studied joint. Various anatomical objects within these data sets have been selected by (clinical) users, like the femur, the tibia, and the patella of the knee images, the scapula, and the tendon of the M. su-

praspinalis of the shoulder images, and the tibia, fibula and talus in foot images. In all cases, the defined objects are fixed sufficiently, preventing in-plane and through-plane motion. The fixation of the femur is depicted in Fig. 3. The position of the distal part of the femur and the orientation of the femur's shaft remain the same during the movement from the fully stretched (Fig. 3 A) to the maximally inflected knee (Fig. 3 C).

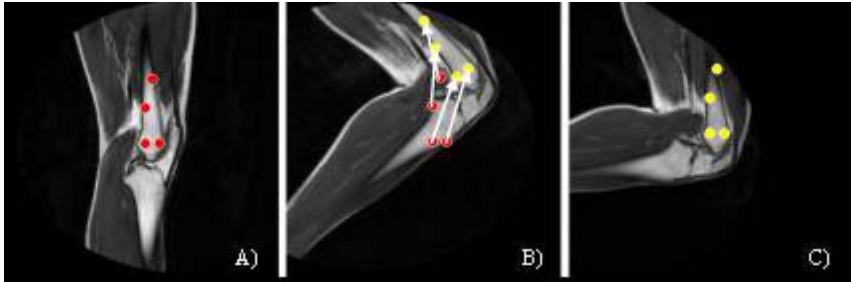


Fig. 3. A) Reference points are placed in the first frame. B) The motion of reference points is estimated from pre-computed elastic registration. C) The viewing plane is adapted to remove motion of reference points.

Following any user interaction, e.g. changing the viewing perspective or redefining the anatomy of interest, a tracked kinematic sequence is calculated instantaneously from the 4D MR data set, the user-defined reference points, and the tracking of these points via the stored elastic registration. This enables the clinical user to explore the kinematic data in real time, and allows the inspection of complex anatomical structures that have never been seen in (stabilized) motion before.

3.2 Image Post-processing

The quality of the automatic propagation of reference points by the B-spline registration algorithm is evaluated. Ideally, a reference point that has been propagated by motion estimation $M_{i,i+1}$ from image i to image $i+1$ and subsequently has been propagated back from image $i+1$ to i by motion estimation $M_{i+1,i}$ will be located at its initial starting position. Thus the Euclidian distance of the original position of each reference point and its position can determine the consistency of the propagation after its propagation.

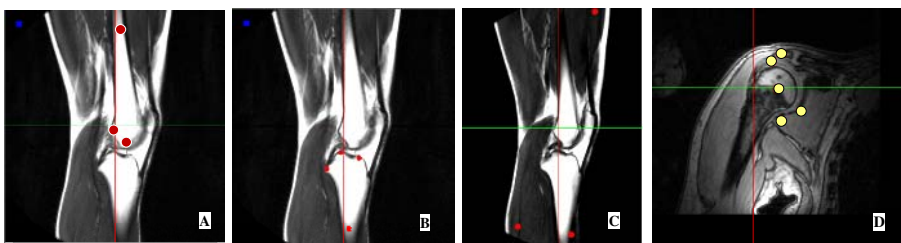


Fig. 4. Reference points of A) the femur, B) tibia, C) a non-rigid structure and D) the shoulder

Table 1 shows the mean distances of different sets of reference points propagated between consecutive 3D images. The sets of reference points used in the knee and shoulder images are shown in Figure 4.

Table 1. Mean distance of sets of reference points in consecutive 3D images defining the femur, the tibia and the shoulder

Mean Distance	1 - 2	2 - 3	3 - 4	4 - 5	5 - 6	6 - 7
Femur	1.9 mm	0.7 mm	0.4 mm	0.8 mm	0.9 mm	2.1 mm
Tibia	1.2 mm	0.3 mm	0.4 mm	0.5 mm	1.2 mm	0.8 mm
Shoulder	0.8 mm	0.3 mm	0.2 mm	1.9 mm	1.0 mm	0.5 mm

Another crucial step is the definition of the anatomical object or region of interest by the user. The residual error of the motion compensation is minimal at the centre and linearly grows with the distance from the centre of the mass of all reference points. At least 3 non-collinear points are required to define an object, where the points should be distributed evenly at its borders. However, we have experienced that (re-) defining an object by clicking is very intuitive for clinical experts, especially if feedback on the result is provided immediately.

In addition to the immediate visual feedback provided to the user table 3 shows the mean distance of reference points propagated using the SVD result versus using the registration parameter at their individual location. This value gives a quantitative measurement of the rigidity of the selected structure given a robust image registration. As expected from visual inspection table 3 shows small mean distances for rigid structures, whereas fixation of non-rigid regions leads to significant distances.

Table 2. Singular value decomposition error

Mean Distance	1 - 2	2 - 3	3 - 4	4 - 5	5 - 6	6 - 7
Knee						
femur	0.8 mm	0.4 mm	0.3 mm	0.3 mm	0.3 mm	0.7 mm
tibia	0.6 mm	0.4 mm	0.3 mm	0.5 mm	0.5 mm	0.4 mm
patella	0.4 mm	0.4 mm	0.3 mm	0.2 mm	0.4 mm	0.1 mm
non-rigid region	4.4 mm	2.6 mm	2.9 mm	5.5 mm	14.2 mm	12.1 mm
Shoulder						
ligament	1.2 mm	0.5 mm	0.8 mm	3.7 mm	2.2 mm	1.8 mm

The B-spline registration has been embedded into a multi-scale approach. The image and mesh resolution is given by Table 2. The original image size of the femur and tibia is $400 \times 400 \times 180 \text{mm}^3$ and the image size of the shoulder is $256 \times 256 \times 150 \text{mm}^3$.

Table 3. Multiscale resolution of image and parameter space

Resolution	Image [voxel ³]	Mesh [control point ³]
Femur and Tibia	36 x 36 x 16	2 x 2 x 2
	73 x 73 x 32	4 x 4 x 4
	256 x 256 x 36	8 x 8 x 8
Shoulder	62 x 62 x 36	2 x 2 x 2
	124 x 124 x 72	4 x 4 x 4
	256 x 256 x 100	8 x 8 x 8
		16 x 16 x 16

4 Conclusion

In conclusion, we have developed a workstation that facilitates viewing of 4D kinematic data sets. It allows to view in real-time any user defined anatomical structure from any viewing perspective. Smoothly displaying the movement in a cine-loop is realized by image post processing, fixing any user defined anatomical structures after image acquisition. Unrestricted from any pre-defined viewing perspective the clinical expert is enabled to examine any (fixed) anatomical structure during the movement of the entire joint in a 4D kinematic study.

References

1. F.G. Shellock, J.H. Mink, A.L. Deutsch, T.K.F. Foo. Kinematic MR imaging of patella-femoral joint: comparison of passive positioning and active movement techniques. *Radiol.* vol.184, pp.574-577, 1992.
2. U.H. Melchert, C. Schröder, J. BrossMann, C. Muhle. Motion-Triggered Cine MR Imaging of Active Joint Movement. *MRI* vol.10, pp.457-460, 1992.
3. D.K. Hodge, et.al. Dynamic MR imaging and testing in glenohumeral instability. *JMRI* vol.13, pp.748-756, 2001.
4. 4.K.S. Arun, T.S. Huang, S.D. Blostein. Last squares fitting of two 3D point sets. *IEEE Trans. Pat. Anal. Mach. Intell.*, vol 9, pp. 698-700, 1987.
5. S. Kabus, et.al., B-spline registration of 3D images with Levenberg-Marquardt optimization. *SPIE Medical Imaging*, 2004.
6. D. Rueckert, L.I. Sonoda. D.L. Hill, M.O. Leach, and D.J. Hawkes. Nonrigid registration using free-form deformations: Application to breast MR. *IEEE Transactions on Medical Image Processing*, 18(8): 712-721, 1999.

Computer-Assisted Ankle Joint Arthroplasty Using Bio-engineered Autografts

R. Sidler¹, W. Köstler², T. Bardyn¹, M.A. Styner³, N. Südkamp²,
L. Nolte¹, and M. Á. González Ballester¹

¹ MEM Research Center, Institute for Surgical Technology
and Biomechanics, University of Bern, Switzerland
`rudolf.sidler@MEMcenter.unibe.ch`

² Universitätsklinikum Freiburg,

Dept. of Orthopaedics and Traumatology, Freiburg i.B., Germany

³ Departments of Computer Science and Psychiatry,
University of North Carolina at Chapel Hill, U.S.A.

Abstract. Bio-engineered cartilage has made substantial progress over the last years. Preciously few cases, however, are known where patients were actually able to benefit from these developments. In orthopaedic surgery, there are two major obstacles between in-vitro cartilage engineering and its clinical application: successful integration of an autologous graft into a joint and the high cost of individually manufactured implants. Computer Assisted Surgery techniques can potentially address both issues at once by simplifying the therapy, allowing pre-fabrication of bone grafts according to a shape model, individual operation planning based on CT images and providing optimal accuracy during the intervention. A pilot study was conducted for the ankle joint, comprising a simplified rotational symmetric bone surface model, a dedicated planning software and a complete cycle of treatment on one cadaveric human foot. The outcome was analysed using CT and MRI images; the post-operative CT was further segmented and registered with the implant shape to prove the feasibility of computer assisted arthroplasty using bio-engineered autografts.

1 Introduction

Tissue-engineered articular cartilage has been a subject of research for a number of years (eg. [1], [2]). Although difficulties with cartilage structure and integration still persist, techniques using combined grafts using cancellous bone and autologous cartilage are approaching clinical application ([3], [4], [5]). Common to all therapies with in vitro engineered autografts is the need for optimal fit of the implant, which is imperative for screwless implantation and successful integration. Computer Assisted Surgery (CAS) techniques as used in orthopaedic surgery (e.g. [6]) have the potential both to ensure the required accuracy and simplify the therapy. A pilot study was hence conducted aiming at assembling a set of methods to realise and prove the feasibility of computer assisted arthroplasty using bioengineered autografts. The ankle joint was chosen as a first target

because of the lack of suitable alternatives: post-traumatic osteoarthritis can be diagnosed in patients as young as 20 years, where classic therapies like total ankle joint arthroplasty using an artificial prosthesis or arthrodesis with fixation screws have considerable drawbacks (loss of mobility, poor long-term outcome expectation, difficult revision).

At the Universitätsklinikum Freiburg, Germany, one case of post - traumatic osteoarthritis has been treated with a bioengineered implant. The intervention was conducted in two steps: one for arthrotomy and defect moulding, a second for implanting the bio-engineered construct. Between the two operations several weeks were needed to proliferate autologous chondrocytes and let them integrate into a cancellous bone construct shaped after the defect mould. Albeit clinically successful, this procedure does not lend itself well to routine application: the two-step operation, the long period of treatment and the high cost of individually constructed autografts make it a time-consuming and costly alternative to classical therapies. Using CAS technologies, the procedure can be significantly simplified and generalised to allow pre-fabrication of implant parts. The revised procedure consists of planning based on CT image data, harvesting mesenchymal stem cells by needle biopsy, constructing the autograft according to the planning and conducting one single intervention for the arthrotomy and construct implantation. The defect debridement has to be accurate enough to make the pre-constructed graft fit without a screw; proving this accuracy in the context of a complete cycle of treatment was a main goal of this initial study.

2 Methods

2.1 Rotational Symmetric Ankle Joint Model

Based on the hinge-like articulation of the upper ankle joint, a rotational symmetric joint was assumed in the region of interest for arthroplasty. This approach allows shape determination using a small number of points on the joint surface, effectively circumventing the need for CT segmentation. A new software was developed to define the ankle joint shape model interactively on this basis. It consists of two steps: determine the joint axis and define the rotational profile. To determine the joint axis, arbitrary joint surface points are identified on sagittal planes in the region of interest. On each plane the software performs a least-mean-square fit of a circle to the points, storing the circle center as one point on the rotation axis. A second least-mean-square fit is performed in 3D to find the optimal fit of a line through all circle centers, yielding the joint axis. The accuracy of this axis calculation depends on the number of points selected and on the anatomy of the individual joint. Selecting 40 to 80 surface points in a rotational symmetric region of Talus and/or Tibia usually gives good results; in this study, about 100 points were used per axis calculation.

Once the joint axis is established, a model for the joint follows directly from rotating a joint profile around the axis. Accordingly, the software allows interactive checking whether the target bone is sufficiently rotational symmetric in the area of interest.

2.2 Preoperative Planning and Construct Manufacturing

The axis determination software was further extended to allow the interactive planning of ankle joint arthroplasty prior to the operation. The planning consists of four steps:

1. Determine the joint axis

As described in chapter 2.1, arbitrary points are selected on sagittal planes in the region of interest. The software performs an optimal fit of a rotational symmetric body to the points selected to determine the joint axis.

2. Determine the lateral graft profile

A "hub view" is defined along the joint axis. In this view, a "U" profile can be defined determining the front, bottom and rear face of the implant (Fig. 1a). The hub view can be shifted along the joint axis to determine the optimal section of the joint to be replaced.

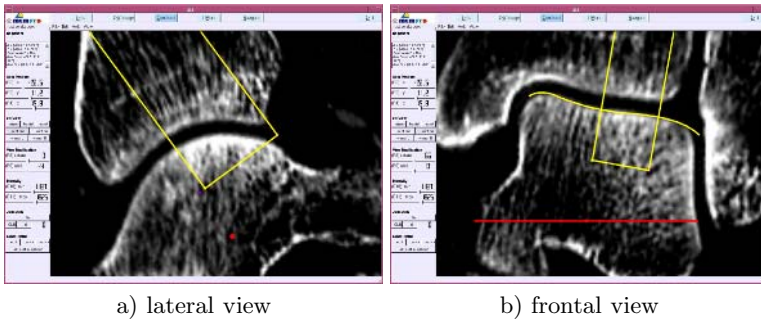


Fig. 1. Determine the implant profiles in the two characteristic views

3. Determine the frontal graft profile

A "cut view" is defined through the joint axis. The axis is always shown horizontally, with the current view rotating around it. In this view, a profile can interactively be defined consisting of an "U" shape and a spline-interpolated curve following the joint surface (Fig. 1b). A landmark is set at an arbitrary position to localise the U shape in 3D.

4. Visualize the resulting construct shape

Having defined the implant shape from hub view and from cut view, the visualisation is done using the CAD software SolidWorks® (SolidWorks Corporation, Concord, U.S.A.). The output of the shape is stored in a CAM environment (construct manufacturing) as well as in our institute's CAS environment.

The CAD part description of the planned implant shape was used to program a CAM device to manufacture the dummy implant. The implants were custom milled according this plan in PU plastic "ep-Dur" (Emaform AG, Gontenschwil, Switzerland).

2.3 Preparing the CAS Environment

At our institute a modular CAS platform has been developed allowing efficient and re-usable implementation of applications for computer assisted surgery. This study was conducted using an active Optotrak® high resolution optical tracking system from Northern Digital Inc., Waterloo, Canada. Tests have been carried out with angled chisels to study their behaviour in human bone and the feasibility limits of possible construct shapes. For navigated ankle joint arthroplasty, a set of three chisels has been designed and tested: one chisel with a blade angled 90° , one with a straight blade, 10mm tip bent by 45° and one chisel with a straight blade, 9.5mm tip. Experiments on a human cadaver showed that the chisels tested are well suited for generating the planned defects. Infrared markers were attached to the bone under treatment and to every surgical tool used. In-house image processing algorithms were used to establish the correspondence between bone, tools and the visualised CT image by pair-point matching of landmarks on the bone. The result was improved by matching arbitrary surface points. The result of the preoperative planning, the implant shape and position, was imported into the CAS system. The contours of the implant were visualised in the CT image to allow navigated operation according to the plan. Surgical tools were tracked by the Optotrak® Camera and visualised at their proper position in the CT volume. Instead of showing the actual physical appearance of the chisels, aiming models were displayed as shown in Fig. 2a, allowing a precise navigation of the tools (Fig. 2b).

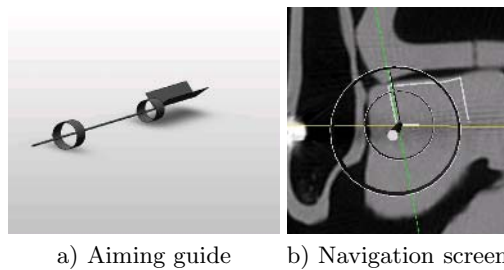


Fig. 2. Navigated Arthroplasty

2.4 Pre-trial Using a Plastic Model and First Accuracy Evaluation

A first trial combining the main elements of the target therapy was made using a plastic foot model from 3B Scientific GmbH, Hamburg, Germany. One talar implant shape was planned based on a preoperative CT. The operation was conducted using the chisels and analysed with postoperative CT. For this preliminary trial, no dummy implant was manufactured. The defect was analysed comparing distances on pre-op and post-op CT images (see Fig. 3a) and double-checking them with measures on the plastic model. The error measures are shown in Fig. 3b and listed in Table 1.

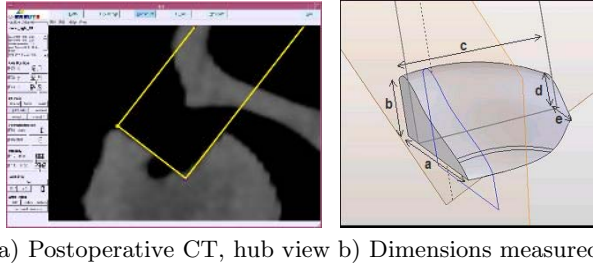


Fig. 3. Postoperative analysis of the pre-trial

Table 1. Postoperative measurements of the pre-trial operation

Distance		Planned Value [mm]	Achieved Value [mm]	Error [mm]
a	medial depth	14.58	14.10	0.48
b	medial height	8.87	8.40	0.47
c	rear width	15.45	15.20	0.35
d	lateral height	11.02	9.50	1.52
e	lateral depth	17.17	14.10	3.07

This analysis revealed a software error displaying the implant shape which affected the lateral face of the planned defect. Measurement of the other distances showed results well below 1mm.

3 Results

A trial operation was conducted on one cadaveric human foot. For the first time a closed cycle of treatment was simulated from pre-operative planning to implanting a custom graft shaped using the rotational symmetric model.

A first intuitive result is how well the implant fits into the defect made. In the experiment, proper fixation of the graft was ensured by driving it into the defect using a hammer and pestle. Visual inspection showed a good restoration of the joint surface, with some locations where the surrounding bone had broken off due to the chiseling. Live bone, being less brittle as cadaveric one, can be expected to allow generally better results.

Second, the defect was measured manually in the post-operative CT image; the results are shown in Fig. 4a, b (dorsal view) and Fig. 5a, b (lateral view). The "U" shape of the graft profile is displayed on the post-operative image exactly as planned preoperatively. Table 2 gives an overview over the dimensions as observed in the postoperative image.

The outcome was further analysed by segmenting both pre- and post-operative images, performing a normalized mutual information registration ([7]) and calculating the difference to obtain a surface model of the defect. A graphical view of

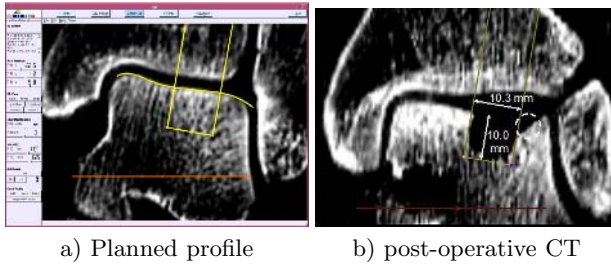


Fig. 4. Dorsal view of the operative result

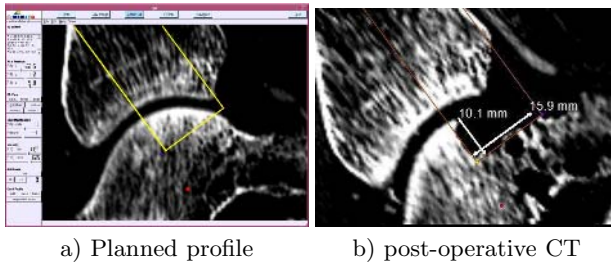


Fig. 5. Lateral view of the operative result

Table 2. Accuracy measurement using the post-operative CT image

Distance	Planned Value [mm]	Achieved Value [mm]	Error [mm]
a) medial depth	15.80	15.1	0.7
b) medial height	9.64	9.5	0.1
c) rear width	10.27	9.3	1.0
d) lateral height	9.11	10.3	1.2
e) lateral depth	15.26	16.4	1.1

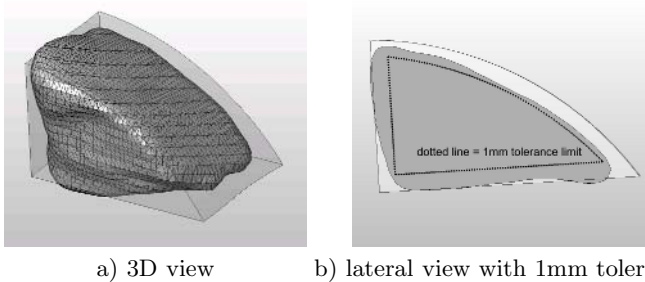


Fig. 6. Analysis of the segmented defect

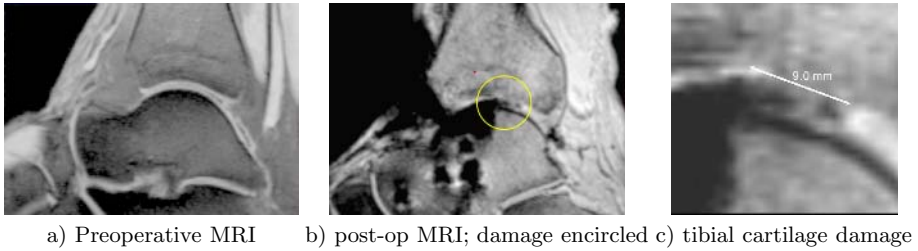


Fig. 7. Medial view of the operative result in MRI

this analysis is shown in Fig. 6a. The lateral profile (Fig. 6b) shows the defect surface to be well within an error limit of 1mm, which can be considered clinically successful.

A last analysis was conducted using post-operative MRI imaging (Fig. 7). The dummy autograft, manufactured in PU plastic, is not visible. While showing intact cartilage on the Talus side, the Tibia side has been damaged during the operation; for clinical application, the tools used need to be further optimised to avoid these injuries.

4 Discussion and Conclusion

The presented study does successfully prove that CAS techniques can be applied to support ankle joint arthroplasty using bio-engineered autografts. It is possible to make a custom-built bio-implant in a parameterisable shape planned on the basis of CT images and to implant it successfully at the planned site.

The proof of this hypothesis was done assuming a rotational symmetric shape model of the ankle joint surface, which is only correct for a limited part of talus and tibia. A more general model (e.g. [8]) could improve the matching of implant surface and surrounding joint surface, while at the same time making the technique applicable to other joints with a different kind of articulation. A big advantage of the rotational symmetric model is that no preoperative segmentation of the CT volume is required - a task which is difficult to achieve automatically, especially in joints with a narrow cavity. Manual segmentation is a very time-consuming alternative and would make the method impractical for clinical application. The determination of the ankle joint axis from the joint surface as developed in this project could be an interesting base for further studies: for example, the axis derived from the tibial or talar surface could be compared with each other and used for diagnosis. For that purpose, a closer investigation of its accuracy is required, which was not part of this study. By computing a segmented profile of the difference between pre- and post-operative images, the study tried to extract objective measures for surface errors of the resulting defect. While being sufficient to show the general feasibility, further work is needed to objectively analyse a series of interventions. This is particularly desirable for the actual joint surface, which is the main clinically relevant parameter, provided

the implant is properly fixed in the bone. Such a surface quality measure could also be helpful to optimise the conditions for graft integration.

Acknowledgements

This study was funded by the AO Foundation, Davos, Switzerland. We would like to thank Mrs. E. Spielvogel and Dr. M. Bonél of Insel Hospital, Bern, Switzerland, for providing CT and MRI images and for their radiological expertise; Dr. J. Kowal et al for building MEM center's CAS platform used to perform the operation; U. Rohrer and T. Gerber for manufacturing implants and operating tools.

References

1. Caplan, A.e.a.: Principles of cartilage repair and regeneration. *Clinical Orthopaedics and Related Research* (1997) 254–269
2. Klemm, C.: Tissue engineered cartilage with cultured chondrocytes and a collagen sponge matrix. In G.B., S., Horch, R., Tanczos, E., eds.: *Biological Matrices and Tissue Reconstruction*. Number 1682 in LNCS, Springer (1998) 169–172
3. Oakes, B.: Orthopaedic tissue engineering: from laboratory to the clinic. *Medical Journal of Australia* **180** (2004) S35–S38
4. Lynn, A., Brooks, R., Bonfield, W., Rushton, N.: Repair of defects in articular joints. *Journal of Bone and Joint Surgery (Br)* **86** (2004) 1093–9
5. Landis, W.e.a.: The potential of tissue engineering in orthopedics. *Orthopedic Clinics of North America* **36** (2005) 97–104
6. Nolte, L., Ganz, R., eds.: *Computer Assisted Orthopedic Surgery (CAOS)*. Hogrefe and Huber (1999)
7. Studholme, C., Hill, D., Hawkes, D.: An overlap invariant entropy measure of 3d medical image alignment. *Pattern Recognition* **32** (1999) 71–86
8. Gerig, G., Styner, M., Szekely, G.: Statistical shape models for segmentation and structural analysis. *IEEE International Symposium on Biomedical Imaging* (2002)

Prospective Head Motion Compensation for MRI by Updating the Gradients and Radio Frequency During Data Acquisition

Christian Dold¹, Maxim Zaitsev², Oliver Speck², Evelyn A. Firlre¹,
Jürgen Hennig², and Georgios Sakas¹

¹ Fraunhofer Institute for Computer Graphics,
Dept. Cognitive Computing & Medical Imaging, Darmstadt, Germany

² University Hospital Freiburg, Dept. of Diagnostic Radiology,
Medical Physics, Freiburg, Germany
`Christian.Dold@igd.fraunhofer.de`

Abstract. Subject motion appears to be a limiting factor in numerous magnetic resonance imaging (MRI) applications. For head imaging the subject's ability to maintain the same head position for a considerable period of time places restrictions on the total acquisition time. For healthy individuals this time typically does not exceed 10 minutes and may be considerably reduced in case of pathology. In particular, head tremor, which often accompanies stroke, may render certain high-resolution 2D and 3D techniques inapplicable. Several navigator techniques have been proposed to circumvent the subject motion problem. The most suitable for head imaging appears to be the orbital or spherical navigator methods. Navigators, however, not only lengthen the measurement because of the time required for acquisition of the position information, but also require additional excitation radio frequency (RF) pulses to be incorporated into the sequence timing, which disturbs the steady state. Here we demonstrate the possibility of interfacing the MR scanner with an external optical motion tracking system, capable of determining the object's position with sub-millimeter accuracy and an update rate of 60Hz. The movement information on the object position (head) is used to compensate the motion in real time. This is done by updating the field of view (FOV) by recalculating the gradients and the RF-parameter of the MRI tomograph during the acquisition of k-space data based on the tracking data. Results of rotation phantom, in vivo experiments and the implementation in two different MRI sequences are presented.

1 Introduction

Patient motion remains a significant problem in many MRI applications, including fMRI [1, 2], cardiac- and abdominal imaging as well as conventional acquisitions. Many techniques are available to reduce or to compensate for bulk motion effects, such as physiological gating, phase-encode reordering, fiducial markers [3], special algorithms [4], fast acquisitions, image volume registration, or alternative data acquisition strategies such as projection reconstruction, spiral and PROPELLER [5, 6]. Navigator echoes are used to measure motion with one or

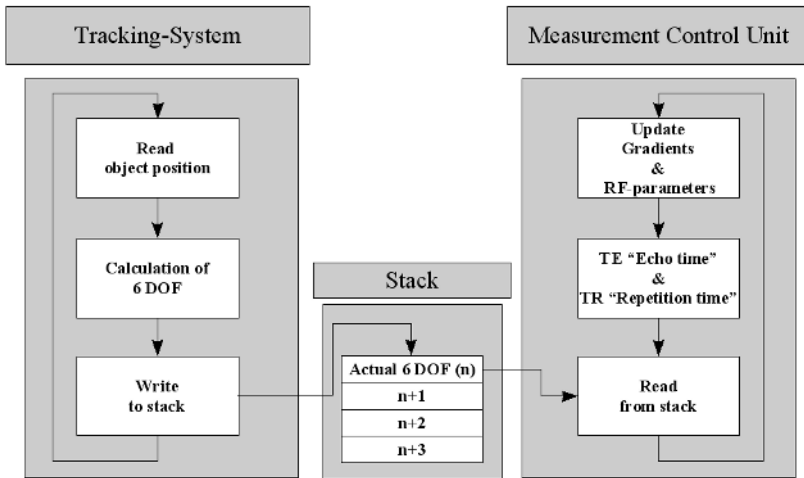


Fig. 1. The setup of the data handling to update the field of view depending on the subject motion

more degrees of freedom [7]. The motion is then compensated for either retrospectively or prospectively. An orbital navigator (ONAV) echo captures data in a circle in some plane of k -space, centered at the origin [8–10]. This data can be used to detect rotational and translational motion in this plane, and to correct for this motion. However, multiple orthogonal ONAVs are required for general 3D motion determination, and the accuracy of a given ONAV is adversely affected by motion out of its plane. Methods capable of correcting for head motion in all six degrees of freedom have been proposed for human Positron Emission Tomography (PET) brain imaging [11]. These methods rely on the accurate measurement of head motion in relation to the reconstruction coordinate frame. Implementing a similar technique in MRI presents additional challenges. Foremost the tracking system and the MRI system have to be compatible. High magnetic fields ≥ 1.5 Tesla in MRI systems require that the tracking camera system be positioned at a sufficient distance from the MRI system to ensure proper function and safety. Additionally, efforts need to be taken to ensure radiofrequency screening on the tracking hardware. Functional MRI also proves challenging because of the high spatial accuracy (root mean square (RMS) < 0.3 mm) required by the complete measurement chain with a small latency time of the tracking system. A precise relationship between the spatially varying magnetic field gradients and the spatial tracking information is necessary to compensate for motion artifacts. Our initial trials using an external tracking system to compensate for movement artifacts in MRI are published in [12].

2 Material and Methods

The technique was implemented on a Siemens Magnetom Trio 3T whole-body system (Siemens Medical Systems GmbH) at the University Hospital Freiburg.



Fig. 2. A volunteer with a mouthpiece inside the head-coil but outside of the scanner

The cameras of the optical motion tracking system (ARTrack1) together with an own developed software has been used to create a stereoscopic reconstruction of rigid bodies from gray scale images. The tracking system was capable of reporting positions and orientations of rigid targets fitted with “passive” retro reflective markers in six degrees of freedom (6DOF) using two progressive scan cameras synchronized by a sync card [13]. The passive targets consisted of at least four coplanar retro reflective markers. All markers were filled with doped water to be detectable by both MR and the optical tracking system. The cameras were equipped with infrared flashes to illuminate the scene with infrared light. Several targets could be tracked simultaneously at a sampling rate of up to 60 Hz, with a quoted positional accuracy less than 0.1 mm (RMS).

Our goal was to prospectively correct artifacts by tracking the head motion by means of the camera immediately before performing the acquisition and then use the motion data in order to internally re-position the coordinate system of the scanner on-line for performing the next scan. A small latency between measuring, scanner coordinates updating and scan performance is essential for the success of the method. Communication with the measurement control unit (MCU) of the MRI instrument took place over a TCP/IP connection. During the calibration step (performed once) the point of origin of the tracking system was transformed to the physical center of the gradient system and the corresponding coordinate transformation matrix has been calculated to overlap both coordinate systems. The information on the target position and orientation was used by the MRI system to update the position of the imaging volume in real time by means of respective transformations of gradients and radiofrequency shifts. For this purpose the scanner software is changed as well. As a result, scanning of sub-subsequent planes or lines are aligned to the first acquisition and, thus, the target appears to be virtually immobilised and without motion artifacts.

Phantom scans as well as in vivo imaging experiments were performed with healthy volunteers. All experiments with human subjects were performed in ac-

cordance with the guidance of the local institutional review board (IRB) regulations and the informed consents were obtained prior to measurements.

2.1 Phantom Tests with Small and Large Rotations

The rotation of the phantom was realized manually in the MRI. No rotation, small rotation up to 5 degree and large rotation up to 20 degree in each direction were done by motion correction (Mo-Co) enabled/ disabled as can be seen in the figure: 3. Every translation and rotation value (6DOF) is stored in a log file. The motion correction was performed prospectively by updating the MR-gradients (x,y,z) and the RF-parameter based on the tracking information.

2.2 Slice by Slice Correction for Echo Planar Imaging (EPI)

The EPI sequence was modified to enable real-time *slice-by-slice* feedback from the MCU. The latter was communicating with the optical motion tracking system and generating the feedback information for the measurement. Imaging parameters were as follows: FOV=256 mm, image matrix of 64^2 , 32 slices, 4 mm slice thickness, 1.2 mm slice gap, interleaved multi-slice acquisition, echo delay time (TE) was 16 ms, repetition time (TR) was 1660 ms. In order to track position of the subject's head a mouthpiece with 4 retro-reflective markers was used. Subjects were instructed to bite the mouthpiece tightly to make sure it

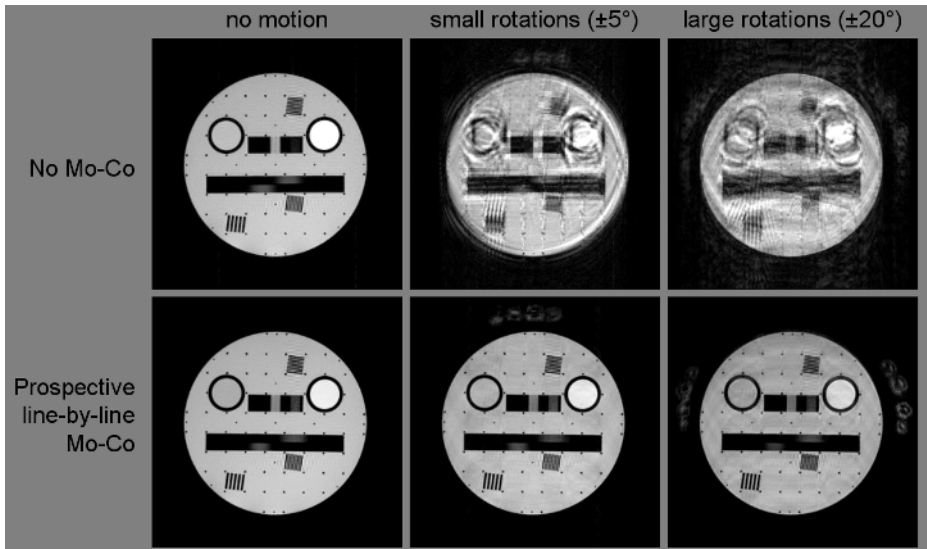


Fig. 3. 2D spin echo images in the top row are shown with motion correction (Mo-Co) disabled and in the bottom row the ones with Mo-Co enabled. Motion compensation was done in real-time by using the data delivered by the tracking system to update orientation and position of the imaging volume.

remains in contact with the upper jaw in order to make the system of 4 markers and the skull a rigid body. No head fixation pads were used in order to enable exaggerated motion during the imaging experiments.

2.3 The Line by line Correction for Turbo Spin Echo (TSE) Sequence

The TSE sequence is a pulse sequence characterized by a series of rapidly applied 180° rephasing pulses and multiple echoes which resulting in shorter scan times. The product TSE sequence was modified to enable real-time *line-by-line* motion correction. Imaging parameters were: FOV=224 mm, image matrix of 256^2 , TE=12 ms, TR=500 ms, phase-encoding dimension was vertical from anterior to posterior (AP).

3 Result

The positional accuracy of the tracking system with an FOV of 50 cm^3 was better than 0.1 mm (RMS). The mean value of the measured reproducibility of the rotation and translation was 0.008 degree/0.067 mm. No artifacts were detected in the MR images originating from the possible interactions of the MR system and the tracking system. The latency time of the whole measurement chain and correction was 30 ms. The prospective *slice-by-slice* motion correction avoids unrecoverable volume distortions. The remaining modulation may be attributed to EPI geometric distortions, which are known to be position-dependent (figure: 4).

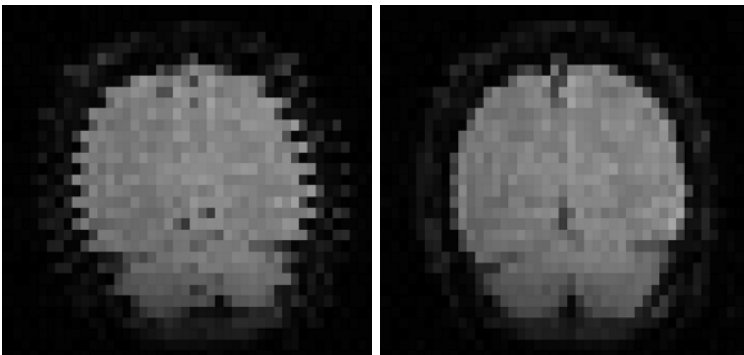


Fig. 4. Representative coronal images produced by re-slicing transversal EPI data from the interleaved multi-slice acquisition. Severe displacements between the neighboring slices are apparent when motion occurs during the acquisition of a single volume (left). Adjusting the position of each slice circumvents the intra-volume distortion problem (right). The remaining modulation and/or step-like structures at the surface of the brain are due to geometric distortions of EPI, which can change with the orientation of the head. The motion corresponds to a single continuous head rotation of yaw axis.

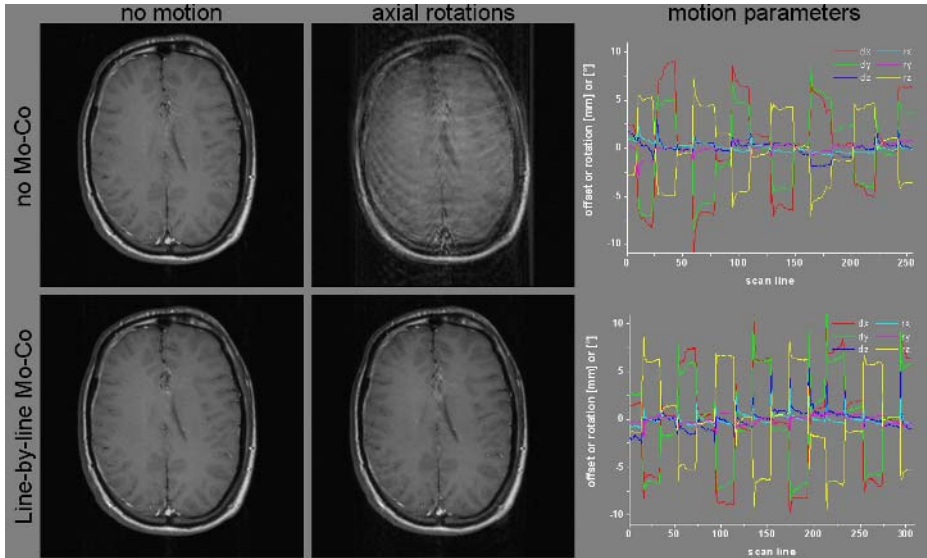


Fig. 5. Brain spin-echo images in the absence of subject motion, motion correction disabled (top left) and enabled (bottom left). Strong subject motion with Mo-Co disabled (top middle) and enabled (bottom middle). Corresponding motion patterns are presented (top/ bottom right). They are shown similar motion values to each other, but the number of scanned lines was in the corrected one higher because of the rejected k-lines including too much motion. Significant reduction of artifacts by the motion correction is apparent. In fact the corrected image shows effectively the same quality as the reference one, while the uncorrected one is useless for medical purposes!

The result of the method demonstrates the advantages of optical motion tracking to correct for motion in echo-planar time series by not only enabling *volume-to-volume* correction without additional computational overhead, but also *slice-by-slice* correction without increasing the scanning time.

In figure: 5 a slice through the brain of a normal volunteer is presented, acquired in 4 measurements. In the first two experiments the volunteer was instructed not to move the head. The figure: 5 top left was acquired with the motion correction disabled and serves as a reference for the maximum achievable image quality. Figure:5 bottom left was acquired with motion correction enabled but still with an unmoved patient in the gantry. Given the absence of subject motion, the experiment has been done in order to visualise the effects of “camera noise” or “jittering effects” visible as an increasing of image noise and slight artifacts in phase encoding dimension due to the uncertainty of the position determination generated by the noising of the optical motion tracking system. In our case these effects were not only negligible compared to the usual physiological motion of the head of living persons, but above this the acquired image shows even improved sharpness and contrast as compared to each other with the software IQM (Image Power Spectrum) [14] and 5 radiologists in an independent questioning.

In the further two experiments the volunteer was instructed to move the head during the acquisition. In figure: 5 top middle the resulting image quality showing significant artifacts with the motion correction disabled. The figure: 5 bottom middle presents the image acquired with imaging coordinates updated for each *k-space line*. The corresponding motion patterns are presented in figure: 5 top right and top left. As can be seen, for comparable motion patterns, the motion artifacts are much reduced, when prospective motion correction is performed.

4 Discussion

We presented a motion correction method which is working in real-time only by updating the FOV of the MRI tomograph. The movement term, acquired through the tracking system, is transferred to the MRI tomograph and used for adapting the gradients and RF-parameter to compensate the complete motion. The results show the feasibility of using an external optical motion tracking system to correct for subject motion during the acquisition of a single image on a line by-line basis (*k-space line*).

Even without any intended motion the image in figure: 5 bottom left is improved in comparison to the reference in figure: 5 top left. This signifies that the noise in the position data generated by the tracking system, introduces less artifacts than the minimal, physiologically unavoidable movement of the head during data acquisition. As a result, even such minimal movements can be compensated resulting in visible quality improvements also for co-operative patients. In case of moving subject, the corrected image acquired in the presence of motion (figure: 5 bottom middle) shows considerable quality improvement in comparison to the image acquired in the presence of comparable motion without correction (figure: 5 top middle). The latency time of the whole measurement chain was now reduced to approximately 30 ms. Doing movement correction within echo- and repetition time to minimize phase effects appears to be possible. This further improvement needs a movement detection and a adjustment of the gradients and RF-parameter in less than 10 ms. One of the limitation, correcting movement artifacts, is always the non linearity of the gradients in the magnetic field. As further limitation can be seen the small inaccuracy of the cross-calibration between the MRI system and the tracking device and the lag in the motion data can also contribute to the reduction of image quality.

Expected benefits of our technique include a significant reduction in imaging time since we perform prospective correction, no oversampling and we avoid repeating acquisitions corrupted by motion. Therefore, our technique helps to increase the clinical efficiency, the patient throughput and reduces measurement redundancy. In the near future the proposed method may be used to examine non-cooperative patients, e.g. the difficult and dangerous tranquillisation of children with sedatives is to become redundant.

Acknowledgement

The research work is funded by the European Union IST (Information Society Technologies) program. The project is named **MRI-MARCB**.

References

1. Babak A. Ardekani et al.: A quantitative comparison of motion detection algorithms in fMRI. *Magn. Reson. Imag.* 19:959-963, 2001.
2. Lee CC et al.: Real-time adaptive motion correction in functional MRI. *Magn. Reson. Med.* 36:436-444, 1996.
3. Korin HW, Felmlee JP, Riederer SJ, Ehman RL: Spatial-frequency-tuned markers and adaptive correction for rotational motion. *Magn. Reson. Med.* 33:663-669, 1995.
4. Shankaranarayanan A et al.: Two-step navigatorless correction algorithm for radial k-space MRI acquisitions. *Magn. Reson. Med.* 45:277-288, 2001.
5. Pipe JG et al.: Motion correction with PROPELLER MRI: Application to head motion and free-breathing cardiac imaging. *Magn. Reson. Med.* 42:963-969, 1999.
6. Starty GE et al.: Single Trajectory Radial (STAR) Imaging. *Magn. Reson. Med.* 51:445-451, 2004.
7. Thesen S et al.: Prospective acquisition correction for head motion with image-based tracking for real-time MRI. *Magn. Reson. Med.* 44:457-465, 2000.
8. Lee CC et al.: A prospective approach to correct for inter-image head rotation in fMRI. *Magn. Reson. Med.* 39:234-243, 1998.
9. Fu ZW et al.: Orbital navigator echoes for motion measurements in magnetic resonance imaging. *Magn. Reson. Med.* 34:746-753, 1995.
10. Welch EB et al.: Spherical Navigator Echoes for Full 3D Rigid Body Motion Measurement in MRI. *Magn. Res. Med.* 47:32-41, 2002.
11. R. Fulton et al.: Correction for Head Movements in Positron Emission Tomography Using an Optical Motion-Tracking System. *IEEE Trans. Nucl. Sci.* 49:116-123, 2002.
12. C. Dold et al.: The Compensation of Head Motion Artifacts using an Infrared Tracking System and a new Algorithm for fMRI. *Medicine Meets Virtual Reality* 12:75-81, 2004.
13. A. Wagner et al.: Quantitative analysis of factors affecting intraoperative precision and stability of optoelectronic and electromagnetic tracking systems. *Medical Physics* Vol. 29(5):905-912, 2002.
14. N.B. Nill and B.H.Bouzas et al.: Objective Image Quality Measure Derived from Digital Image Power Spectra. *Optical Engineering* 31:813-825, 1992.

Harmonic Skeleton Guided Evaluation of Stenoses in Human Coronary Arteries

Yan Yang¹, Lei Zhu¹, Steven Haker²,
Allen R. Tannenbaum¹, and Don P. Giddens¹

¹ Department of Biomedical Engineering,
Georgia Institute of Technology, Atlanta, GA, 30332, USA
gtg929i@mail.gatech.edu, zlzl@ece.gatech.edu, tannenba@ece.gatech.edu,
don.giddens@coe.gatech.edu

² Surgical Planning Lab, Brigham and Women's Hospital,
Harvard Medical School, Boston, MA 02115, USA
haker@bwh.harvard.edu

Abstract. This paper presents a novel approach that three-dimensionally visualizes and evaluates stenoses in human coronary arteries by using harmonic skeletons. A harmonic skeleton is the center line of a multi-branched tubular surface extracted based on a harmonic function, which is the solution of the Laplace equation. This skeletonization method guarantees smoothness and connectivity and provides a fast and straightforward way to calculate local cross-sectional areas of the arteries, and thus provides the possibility to localize and evaluate coronary artery stenosis, which is a commonly seen pathology in coronary artery disease.

1 Introduction

Stenosis is a commonly seen pathology resulting from atherosclerosis, which is a systemic disease of the vessel wall that occurs in the aorta, carotid, coronary, and peripheral arteries [1]. Atherosclerotic plaques that develop in vessel walls can intrude into the vessel lumen and result in a narrowing (stenosis) of the lumen. When stenosis occurs in coronary arteries, it will cause insufficient supply of blood to the heart tissue, and sometimes leads to serious results such as heart attacks. The detection of stenoses *in vivo* can greatly assist the diagnoses and treatments of coronary heart diseases. The conventional way of identifying stenoses is by using coronary angiography imaging, but this method is invasive, and it only provides 2D projections of the vessel lumen, which may yield biased or inaccurate evaluation of the stenosis. Thanks to the development of multislice-CT (MS-CT), we are now able to acquire 3D volumetric data of the entire coronary artery tree, and obtain the complete information of its geometry so that we can perform more accurate measurements based on these data.

Numerous methods have been proposed to extract skeletons of tubular structures in the human body such as blood vessels, airways and colons. Most methods in literature fall into the following two classes [4]: *boundary peeling* (also called thinning, erosion, etc.) [5], and *distance coding* (distance transform) [2, 3]. Although both classes of methods have yielded promising results, these methods

usually have problems of preserving connectivity and being smooth. Also, direct extension of these methods to 3D is usually difficult and may not guarantee a unique solution for a single structure.

In this paper, we explore a skeletonization method based on the solution of the Laplace equation with boundary conditions of the first kind (Dirichlet problem) on the surface of a tubular structure. Since the solution of the Laplace equation is a harmonic function, the skeleton extracted using this method is called the *harmonic skeleton* [8]. The harmonic skeleton is easy to compute, is guaranteed to be smooth, and can automatically give a viewing vector when used as a guidance for fly-throughs for virtual endoscopy. Moreover the space of harmonic functions is linear (as opposed to that of the distance transform on which the ordinary skeleton is founded). The distance transform is commonly used in endoscopy; see [9] and the references therein.

Further, we develop an effective way to determine the boundary conditions for solving the appropriate Laplace equation, so that the points that form the final skeleton are evenly distributed, without being too dense or too sparse, thus guaranteeing the connectivity of the skeleton. Most importantly for our applications, the cross-sectional area at a certain location on the artery can be conveniently measured using the least squares plane that fits all the points on the iso-value contour of the solution of the Laplace equation. This means that we can acquire the cross-sectional areas throughout the artery in order to identify stenoses.

We tested our method on CT data sets of both normal and diseased coronary arteries and developed different ways to visualize and analyze stenoses of the arteries. The method is described in detail in Section 2, and we show the results in Section 3.

2 Methodology

2.1 Harmonic Skeleton

The harmonic skeleton can be obtained by solving the Laplace equation on the surface of the artery. Let Σ denote the surface, which is topologically a branched tube. Each branch has a boundary σ_i ($i = 0, \dots, N - 1$), which is a closed contour, as shown in Figure 1. Here N denotes the total number of boundaries including the root σ_0 , which can be selected arbitrarily or selected based on physiological considerations. We have chosen the root σ_0 to be at the proximal section of the left main (LM) coronary artery. Note that we are working on a triangulated representation of Σ , which makes it convenient for performing numerical computations. The surface can be obtained by segmenting the MS-CT images. In this case, we have used a segmentation approach that combines Bayesian classification and active contours [6, 7] to get the required surface for the skeletonization algorithm.

Next we can calculate the harmonic function u by solving the Laplace equation on $\Sigma \setminus \sigma_i$ ($i = 0, \dots, N - 1$):

$$\Delta u = 0 \tag{1}$$

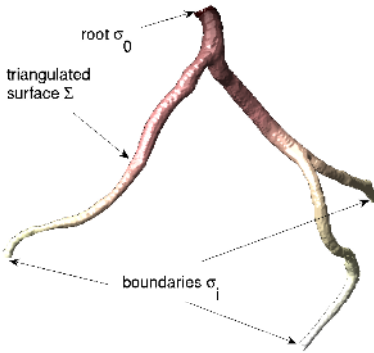


Fig. 1. The segmented and opened surface of left coronary arteries with boundary contours at each end; the surface is color-painted with solved harmonic function u , where dark color means lower u , and light color means higher u

with proper boundary conditions. To find these boundary conditions, we assign the values at the boundaries in the following manner:

1. Set the boundary value of u at the root σ_0 to be zero, and set $StepNumber = 0$;
2. Label all the points on σ_0 as “visited”;
3. Label all the neighbors of the “visited” points as “visited”;
4. $StepNumber ++$;
5. Repeat steps 3. and 4. until “visited” points reached another boundary σ_i ;
6. Set the value at the current boundary σ_i to be the current value of $StepNumber$.

In other words, we set the boundary value of σ_i to be the number of triangle strips between σ_i and the root boundary σ_0 . In coronary arteries, there are several branches, some of which are long and some of which are short. Note that if we used standard Dirichlet boundary conditions in which $u(\sigma_0) = 0$ and $u(\sigma_i) = 1$ for all i 's, as for example, is done in [13], the resulting points that form the skeleton will be denser for the short branches, and sparser for the longer ones. Hence we employ the boundary conditions stated above to get more evenly distributed skeleton points.

Having obtained the boundary conditions for the Dirichlet problem, we can solve the Laplace equation $\Delta u = 0$ on the surface using finite element techniques [6, 12]. From the standard theory [10], u exists and is unique, with the maximum and minimum values on the boundaries. In Figure 1, the surface of the artery is painted with the solved u , where brighter color indicates higher u , and darker color indicates lower u .

Using the harmonic function u , we can now build the harmonic skeleton. We first find the level sets of u on the surface, i.e., sets $\{x|u(x) = \nu\}$ for values of ν ranging from 0 to $max(u)$. Each level set corresponds to a group of points on the surface that have the value of a certain ν . These points are not necessarily located on the vertices of our original triangulation of the surface, so we obtain these level points by linearly interpolating the coordinates of the vertices according to the values of u .

We also check the change in the topology of the level contours as we proceed, and further partition these contours into sub-contours if not all the points are connected to form a single contour. This means there is a bifurcation and one contour splits into two at the current level of u . We assume one contour cannot split into more than two contours, which is usually true for coronary arteries. The centroid of each contour or sub-contour is then calculated: $\frac{1}{M} \sum_{i=1}^M \mathbf{x}_i$, where $\mathbf{x}_i = (x_i, y_i, z_i)'$ is a point on the contour, and M is the total number of points on the current contour. Each centroid contributes a point to the final harmonic skeleton. By increasing ν from 0 to the maximum value of u , we can find a series of centroids that can finally be connected to build a structured tree, which is the harmonic skeleton.

For the sake of accuracy, an additional step can be performed to refine the harmonic skeleton by replacing the previous boundary conditions with new values obtained as the distance along the skeleton from the root boundary to the other boundaries. We then solve the Dirichlet problem with the new boundary conditions and further obtain a refined harmonic skeleton.

2.2 Cross-Sectional Area Measurement

The harmonic skeleton obtained using the above method can be used as a guide in calculating cross-sectional areas at specific locations along the vessels, and the cross-sectional areas can in turn be used to precisely evaluate stenoses in the coronaries. Cross-sectional areas are better than diameters because atherosclerotic plaques are usually eccentric and asymmetric, and the cross-sections of the narrowed vessel lumen may not be perfectly circular. The stenosis may be underestimated if one uses only diameters for stenosis evaluation. Cross-sectional areas, on the other hand, provide more information about the stenosis, and are directly related to the flow rate of blood through the cross-section.

One straightforward way of obtaining the cross-sections of the artery is by looking for a plane that is perpendicular to the skeleton to cut the tube at a certain location. But this strongly depends on the local properties of the skeleton, and a plane that is perpendicular to the skeleton locally may not be perpendicular to the vessel itself. What we do here is to start with the contours or sub-contours obtained in Section 2.1, and measure the cross-sectional area at the same time as we calculate the centroid.

The first thing to do is to find the cross-sectional plane. We have many more points (about 50 points, somewhat more for wider contours and somewhat less for narrower contours) on a level contour than what is necessary (3 points) for determining a plane. Hence we need to use least squares to solve this over-determined problem. Suppose we have a set of points $X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)'$, where $\mathbf{x}_i = (x_i, y_i, z_i)'$, $i = 1, \dots, M$ is a point on the level contour, and we build a matrix $A = X - \mathbf{x}_0$, where $\mathbf{x}_0(x_0, y_0, z_0)'$ is the centroid of all the points on the contour. Calculating the Singular Value Decomposition (SVD)[14] of matrix A , we get $A = USV^T$, where U and V are orthogonal matrices and S is a diagonal matrix with non-negative diagonal entries in decreasing order. The singular vector in V that corresponds to the smallest singular value in S is the

cosine of the direction of the normal to the best-fit plane. The centroid and the normal vector can now be used to uniquely determine the least squares plane.

We now use this plane to cut the surface of the vessel and get a series of points on the surface that also belong to the least squares plane. Again, these points may not coincide with the vertices of the triangulated surface, so we find the points on the triangle edges, and connect them to form a polygon. We then calculate the area of the polygon, and the result is the required cross-sectional area at the current location on the vessel.

3 Results

3.1 Clinical Data

We tested our method on a healthy volunteer without any clinical symptoms and a patient with plaques in left anterior descending (LAD) coronary artery. The healthy volunteer was imaged using a GE LightSpeed16 CT scanner with a slice spacing of 0.625mm and an in-plane resolution of 0.60mm, while the patient with LAD plaques was imaged using a Siemens Sensation64 CT scanner with a slice spacing of 0.75mm and a in-plane resolution of 0.39mm. Contrast agents were used in both cases. The coronary arteries were extracted by image segmentation using the techniques of [6, 7], and the triangulated surfaces were generated using the Visualization Toolkit (VTK) [15]. We then performed our skeletonization and cross-sectional area measurement methods on the triangulated surfaces.

3.2 Generation of Harmonic Skeletons

The harmonic skeletons of both cases were generated using the method described in Section 2.1. Figure 2 (Left) and (Middle) show the surfaces in half-transparent mode with the skeletons inside the surfaces, and Figure 2 (Right) shows the skeleton of the diseased (Middle) coronary arteries. The skeletons provide the

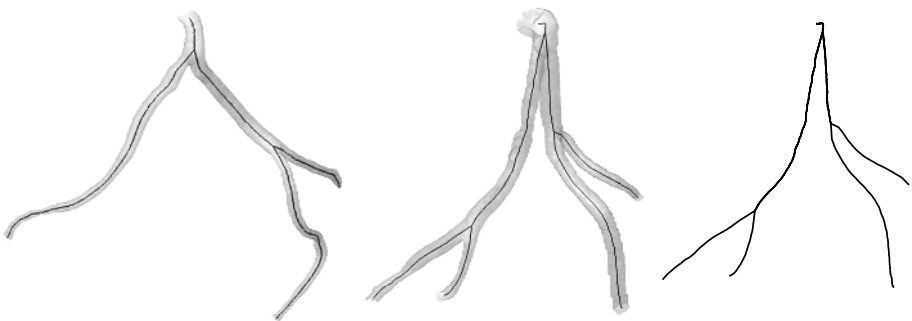


Fig. 2. Left: Coronaries and the skeleton of a healthy volunteer. Middle: Coronaries and the skeleton of a patient with plaques in LAD. Right: The skeleton of the middle coronaries.

basic information of the arteries, such as bifurcation sites, bifurcation angles and curvatures along the vessels, but they themselves alone are not able to provide much information about the stenosis.

At the bifurcation sites, some post-processing was performed to keep the skeleton smooth and natural. Increasing the value of u and tracking the topology change of the level contours, one contour will split into two at a certain u level, and there will be a jump between the centroid of the single parent contour and the centroids of the two daughter contours. If we connect the centroids directly, the resulting skeleton may not be smooth where it bifurcates. To solve this problem, we start from the daughter branches, and trace back until they reach the bifurcation site, and then extend the branches further until they meet the parent branch. After the post-processing, the skeleton is visually more natural at the bifurcation sites.

3.3 Stenosis Evaluation

We calculated cross-sectional areas along the vessel with a resolution of $\delta u = 0.5$, which means we measure the cross-sectional area once we increase the harmonic function u by 0.5. For each coronary artery, there are roughly 400 levels of cross-sectional areas, and at each level, there are one or several cross-sectional areas measured according to the number of branches at this level.

We can visualize the quantitative measurements in two different ways. The first way is to specify a u value, and show the cross-sectional areas at this u level, as illustrated in Figure 3 (Left). This is equivalent to showing the areas by specifying a location on the vessel. The second way is to create a color map of the cross-sectional area values, and paint the surface with the color of these values. This provides an intuitive way to localize the stenosis, by showing different colors where the cross-sectional area has an apparent change along the vessels, as indicated in Figure 3 (Right). In this figure, lighter color indicates normal sized vessel lumen, and darker color means the lumen is narrowed, thus implying stenosis.

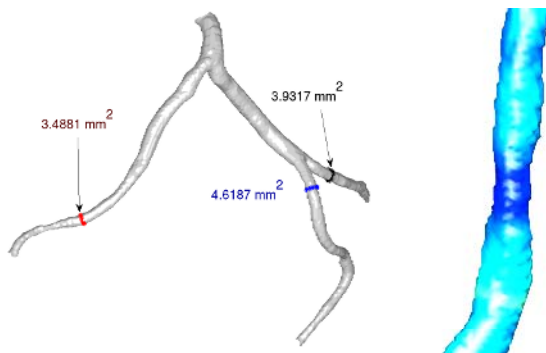


Fig. 3. Cross-sectional area visualization. Left: The cross-sectional area of each branch of the coronaries for a certain u . Right: Color-painted LAD indicates a stenosis site in the vessel.

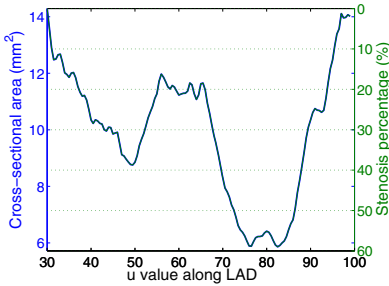


Fig. 4. Cross-sectional area along the LAD shown in Fig. 3 (Right). The left y-axis shows the cross-sectional area in mm^2 , and the right y-axis shows the stenosis percentage.

Figure 4 is a plot of u versus cross-sectional area along the LAD of Figure 3 (Right). The left y-axis shows the cross-sectional areas in mm^2 , and the right y-axis shows the stenosis percentage, which is an important measure for the evaluation of stenosis. We can see from this plot that this LAD has two stenosis sites, one mild (a little less than 40% blockage) and another more severe (60% blockage). This is in consistent with what is shown in Figure 3 (Right).

4 Discussion and Conclusions

In this note, we have proposed an approach to extract harmonic skeletons of tubular structures and applied it to both healthy and diseased coronary arteries. The skeleton generated can be used as a guide for cross-sectional area measurements in order to localize and evaluate stenosis of the arteries. The harmonic skeleton is smooth and unique, and can also serve as a guide in other applications such as virtual endoscopy.

Our algorithm is reasonably fast, taking 2-3 minutes in Matlab to generate the harmonic skeleton from a surface consisting of about 30,000 triangles and 12,000 vertices, on a 1G Hz personal computer running Linux. If the skeleton extraction and cross-sectional area measurement are performed at the same time, it takes 6-7 minutes for the same sized data. This makes it possible to use this algorithm in a real clinical setting.

We should also point out that this skeletonization and cross-sectional area measuring approach is performed on a segmented surface of a tubular structure, so the measurement results highly rely on the correctness of the image segmentation. On the other hand, the cross-sectional area can also serve a measure for the validation of segmentation to compare the values of a ground truth model (such as a phantom or a vessel cast) and the segmentation result.

Acknowledgements

This work is part of the National Alliance for Medical Image Computing (NAMIC), funded by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant U54 EB005149. This work was also supported by a grant from NIH (NAC P41 RR-13218 through Brigham and Women's Hospital). S. Haker's

work is supported by NIH grants R01CA109246, R01LM007861, R01CA1029246 and P41RR019703.

References

1. V. Fuster, Z.A. Fayad, J. J. Badimon, "Acute coronary syndromes: biology", *Lancet* 1999; 353 (suppl 2): SII5-SII9
2. M. Maddah, H. Soltanian-Zadeh, and A. Afzali-Kusha, "Snake modeling and distance transform approach to vascular centerline extraction and quantification", *Computerized Medical Imaging and Graphics*, **27**, 2003, pp. 503-512
3. M. Wan, Z. Liang, Q. Ke, L. Hong, I. Bitter, and A. Kaufman, "Automatic Centerline Extraction for Virtual Colonoscopy", *IEEE Trans. on Med. Imag.* **21**, 2002, pp. 1451-1460
4. Y. Zhou, and A. W. Toga "Efficient Skeletonization of Volumetric Objects", *IEEE Tran. on Visualization and Computer Graphics*, Vol. 5, NO. 3, 1999, pp. 196-209
5. L. Hong, A. Kaufman, Y. Wei, A. Viswambharan, M. Wax and Z. Liang, "3D Virtual Colonoscopy", *Proc. of IEEE Biomedical Visualization*, 1995, pp. 26-32
6. S. Haker, G. Sapiro, A. Tannenbaum, "Knowledge-Based Segmentation of SAR Data with learned Priors", *IEEE Tran. on Image Processing*, Vol. 9, No. 2, Feb. 2000, pp. 299-301
7. Y. Yang, A. Tannenbaum, and D. Giddens. "Knowledge-Based 3D Segmentation and Reconstruction of Coronary Arteries Using CT Images", *In Proceedings of the 26th Annual International Conference of the IEEE EMBS*, pages 1664-1666, 2004
8. L. Zhu, S. Haker, and A. Tannenbaum. "Flattening Maps for the Visualization of Multi-branched Vessels", *IEEE Trans. on Med. Imag.*, 24(2):191 - 198, 2005
9. S. Bouix, K. Siddiqi, A. Tannenbaum, "Flux Driven Automatic Centerline Extraction," *Medical Image Analysis* **9**, 2005, pp. 209-221.
10. Stanley J. Farlow "Partial Differential Equations for Scientists and Engineers", *Dover Publications, Inc. New York*, 1993
11. S. Haker, S. Angenent, A. Tannenbaum, and R.Kikinis, "Nondistorting flattening maps and the 3D visualization of colon CT images", *IEEE Trans. Med. Imag.*, Vol. 19, 2000, pp. 665-670
12. T. Hughes, "The Finite Element Method", *Upper Saddle River, NJ: Prentice-Hall*, 1987
13. A. J. Yezzi, Jr., J. L. Prince, "An Eulerian PDE Approach for Computing Tissue Thickness", *IEEE Trans. Med. Imag.*, Vol. 22, No. 10, 2003, pp. 1332-1339
14. W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C, 2nd ed.* Cambridge: Cambridge Univ. Press, 1992
15. Kitware Inc. <http://www.vtk.org>

Acquisition-Related Limitations in MRI Based Morphometry

Arne Littmann^{1,2}, Jens Guehring¹, Christian Buechel³,
and Hans-Siegfried Stiehl²

¹ Siemens Medical Solutions, Magnetic Resonance, Erlangen, Germany

² Department of Informatics, University of Hamburg, Germany

³ NeuroImage Nord, University Hospital Eppendorf, Hamburg, Germany
arne.littmann@siemens.com

Abstract. Although significant effort has been spent over the past decades to develop innovative image processing algorithms and to improve existing methods in terms of precision, reproducibility and computational efficiency, relatively few research was undertaken to find out to what extent the validity of results obtained with these methods is limited by inherent imperfections of the input images. This observation is especially true for MRI based morphometry, which aims at precise and highly reproducible determination of geometrical properties of anatomical structures despite the fact that MR images are geometrically distorted. We here present (a) a method for characterization of site-specific geometrical distortions and (b) the results of a long term study designed to find out how precisely geometrical properties and morphological changes of brain structures can, in principle, be detected in images acquired with MRI scanners. Due to the long-term character of our study, our findings include effects resulting from limited hardware stability as well as from variations in patient positioning. Our results show that these effects can be strong enough to substantially confound MRI studies of small morphological changes.

1 Introduction

Over the last decade MRI based morphometry has experienced a remarkable gain in attention in medical research, one of the most prominent areas of application being the detection of statistically significant morphological differences in the brain between a population known to be affected by a certain disease and supposedly healthy controls. Recently, serial studies examining the progress of diseases over time have attracted attention, and it can be expected that in the future single subject applications of morphometry will become more and more important in the context of computer aided detection and staging of diseases.

Even though there exists a great variety of methods that allow for these demanding morphological analyses [1-3], few research was undertaken to find out to what extent the validity of results obtained with these methods is limited by inherent imperfections in the data acquisition process. For a profound characterization of small phenomena like beginning gray matter (GM) atrophy in early Alzheimer's disease, how-

ever, it is inevitable to exactly know to what extent the results obtained with state-of-the-art morphometric methods might be confounded by these limitations.

Only very little information concerning these effects could be found in the literature: Jovicich et al. either characterize distortions by the change and variance of a phantom's total diameter [4] or evaluate the influence of gradient nonlinearity correction on hippocampus volumetry in single session repositioning experiments [5]. To our knowledge no publications are available concerning the precise quantification of the crucial long-term reproducibility of geometrical distortions on a local basis. We therefore developed a new method for characterization of site-specific geometrical distortions and used it for quantitative analyses of the previously mentioned limitations in a long-term study carried out on state-of-the-art MR scanners. Our evaluation method is capable of determining positional and volumetric deviations on a dense 3D grid both in an absolute and in a relative way. We additionally compared our findings to results obtained from the analysis of head images of a healthy volunteer acquired in the same imaging sessions with identical protocols.

1.1 Sources of Geometrical Distortion

A detailed description of the different sources of geometrical distortions in MRI is beyond the scope of this paper, which is why we restrict our delineation of these error sources to a basic level. More information on this topic can be found in [6]. In short, there are four main sources of geometrical distortions in MRI:

- **Magnetic susceptibility** differs between different materials which is why their larmor-frequency varies slightly. This leads to geometrical distortions mainly in frequency-encoding direction whose magnitude in image space is also depending on the imaging protocol's bandwidth.
- **Static field inhomogeneities** cause distortions in the same way as differences in magnetic susceptibility. Typically these errors are compensated for the most part by shimming.
- **Gradient coil nonlinearities** cause distortions in all directions. Since gradient coils are designed to minimize these effects around the magnet's isocenter, nonlinearities are stronger in outer regions.
- **Eddy currents** are induced in conductive material within the scanner every time a gradient is changed in the course of the scanning process. As eddy currents cause dynamic magnetic fields, geometrical distortions arise. Quite obviously, errors caused by eddy currents are dependent on the actual imaging sequence and its parameterization.

Whereas static field inhomogeneities, gradient nonlinearities and eddy currents are independent of the object being scanned (as long as it is not conductive), and thus, stationary in the coordinate system of the magnet, magnetic susceptibility artifacts are a characteristic of the object and can be thought of as approximately stationary in the object's coordinate system. This distinction is important to take into consideration when analyzing relative distortions between two (or more) scans caused by slight differences of the object's position in the scanner.

2 Methods

2.1 Data Acquisition

Based on our literature research, MPRAGE was identified as the most commonly used sequence in brain morphometry today. In cooperation with sequence developers at Siemens Medical Solutions, Magnetic Resonance, Erlangen, the following protocol was defined for the Avanto 1.5T and the Trio 3T systems: TR 2300 ms, TI 1100 ms, TE 3.93 ms, flip angle 12° , matrix 224×256 , bandwidth 130 Hz/pixel, asymmetric echoes allowed. All datasets were acquired in sagittal orientation and with an isotropic resolution of 1.0 mm.

Using this protocol, a phantom and a healthy volunteer (m, 28y) were regularly scanned over 6 months alternating between an Avanto and a Trio. The phantom is cylindrically shaped (diameter 250 mm) with an integrated 3D grid of small spherical holes (4 mm radius) each having a distance of 20 mm to its nearest neighbor in x- and y-, and 19 mm in z-direction (979 spheres in total). These holes are connected in z-direction by cylindrical bars (2 mm radius), and the complete system of cavities is filled with fluid.

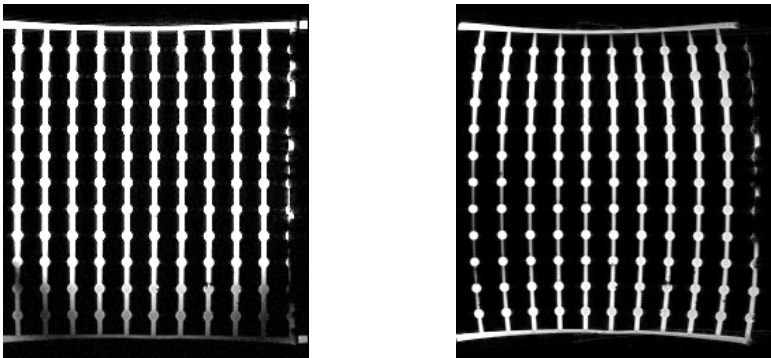


Fig. 1. Two slices of a phantom dataset acquired with the Avanto (left) and Trio (right). The interruptions in the bright bars are caused by inclusions of air.

2.2 Phantom Data Evaluation

The method for evaluating each of the phantom datasets is a two-stage procedure: First, the ellipsoids' center points are determined, and then this set of center points is compared to a reference phantom model G_0 obtained from the analysis of a high-resolution CT-scan (0.5 mm isotropic resolution).

The identification of the ellipsoids, again, can be divided into two passes: In a 2D operation, they are first separated from the interconnecting bars by eliminating all those pixels that belong to an object whose bounding box is smaller than a user-defined minimum. With the subsequent connected component analysis the remaining 2D objects are combined into 3D blobs again:

```

smooth image;
for each slice{
  binarize via Otsu separation [7];
  identify objects belonging to an ellipsoid;
}
do 3D connected component analysis;

```

These blobs are then filtered by position (to eliminate irregular ellipsoids occurring as a result of wrap-around artifacts), and by shape (to eliminate degenerated objects resulting from air inclusions).

The second pass consists of a computationally more demanding determination of the remaining ellipsoids' center points. Its principle idea is to determine the point-symmetry center for each ellipsoid in the original image. To ease the optimization strategy, candidate center points need to be located on the interconnecting bar's center line, which is determined via regression analysis. As a metric for symmetry the negative summed squared difference between each voxel and its mirrored counterpart (resulting from 3rd order B-spline interpolation) was used:

```

for each remaining blob{
  extract a small subimage (bounding box + margin);
  for each slice in subimage{
    determine intensity-weighted center (centerX,
      centerY) using the smoothed original image;
  }
  do regression analysis of centerX, centerY against z;
  determine that point on the regression axes for which
    the point-symmetry of the subimage is maximal;
  store this point as ellipsoid's center;
}

```

The resulting point set GI is then compared to the previously mentioned reference phantom model $G0$:

- The positional comparison is done straightforward and merely consists of a calculation of the absolute deviation in position for each sphere compared to its position according to the reference model $G0$.
- For volumetric analysis GI is subdivided into polyhedral cells by iteratively grouping eight adjacent center points together in order to form one cell, the shape of which can be intuitively described as a "degenerated cuboid". For determination of the detected cells' volumes, the polyhedrons are further divided into tetrahedrons, the volumes of which can be easily calculated; finally, each cell's volume is compared to the volume of the corresponding cell in $G0$.

2.3 Head Image Evaluation

Since this project's main focus was the precise quantification of limitations in serial MRI based morphometry resulting from data acquisition alone, it was put more emphasis on the phantom experiments than on the head image evaluations, which were expected to be further confounded by movements during scan, morphological changes due to differences in the volunteer's water balance and by a presumably less precise

segmentation procedure. The head image evaluation therefore was merely done for verification purposes. It was almost completely carried out with SPM2 [3].

After correcting each dataset for intensity inhomogeneities, a template image was created for each data series (one template for each scanner) by rigidly aligning all datasets to the series' baseline scan. Having rigidly aligned each dataset to its template, gray matter (GM) was automatically segmented in each image with SPM2.

We then defined three regions of interest (ROIs) in the template images (see Fig. 2) and determined the GM volume as the sum of GM density within these ROIs. We chose large box-shaped ROIs that stretch across the complete slice stack in order to minimize effects resulting from remaining imperfections of the rigid registration.

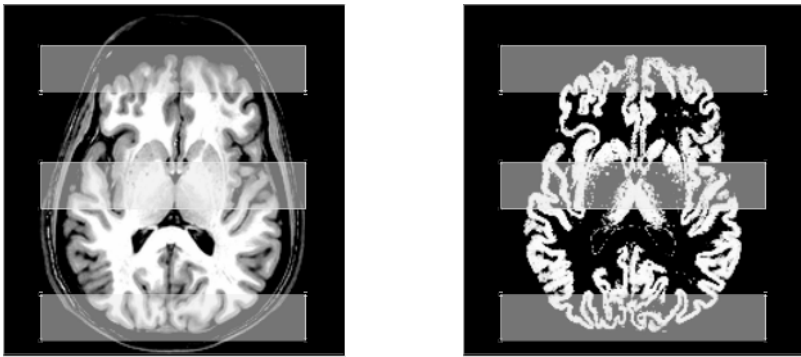


Fig. 2. ROI locations (anterior, medial, posterior) in the Avanto template image (left), and in a GM segmentation image (right)

3 Results

3.1 Validation

Whereas the precision of the proposed phantom evaluation method is limited by noise and partial volume effects, the reproducibility of the detection of the ellipsoids' center points is additionally limited by temporal drifts in the scanners' resonance frequency during data acquisition and in between two scans. This error in frequency encoding results in a slightly varying discretization scheme in readout direction even within one scanning process. For two consecutive scans this effect looks similar to a slight shift in readout direction in image space. Practical test-retest experiments consisting of evaluations of datasets that were acquired directly one after the other with identical protocols without repositioning resulted in maximal relative positional differences of less than 0.15 mm for the Avanto, and less than 0.11 mm for the Trio in a subvolume relevant for brain morphometry (spherical subvolume around the isocenter with a radius of 100 mm). The corresponding numbers for relative volume differences are 0.35 % respectively 0.26 % (relative to the volume of a standard cell in *GO* of 7600 mm³). The slightly better results for the Trio most likely are a result of a higher signal-to-noise ratio at 3T.

Analogous test-retest experiments for validation of the head image evaluation method resulted in the following volumetric changes for the Avanto: anterior ROI 0.50 %, medial ROI 0.16 %, and posterior ROI 0.35 %. The results for the Trio are: anterior ROI 0.25 %, medial ROI 0.39 %, and posterior ROI 0.12 %. Note that the test-retest datasets, too, were rigidly aligned to the template datasets before segmentation.

3.2 Repositioning Influences

An important influence on the long-term study's results were expected to be slight differences in positioning. Due to the nonlinear character of the distortion field, a different position of an unaltered object in the scanner is accompanied by relative morphological changes in image space. To get insight into the magnitude of relative geometrical distortions that are purely caused by minor errors in repositioning, the outlined evaluation was performed on two datasets, that, again, were acquired directly one after the other with a slight difference in the phantom's position in the scanner (rotation less than 5°, translation less than 2 mm, new selection of field of view, new shim). These changes produced positional deviations of up to 0.6 mm (for both scanners), and relative volumetric changes of up to 1.2 % for the Avanto and 1.8 % for the Trio. Relative morphological changes were smaller for the Avanto due to the smaller spatial gradients of the absolute distortion fields (see Sect. 3.3).

Again, analogous repositioning experiments were carried out with the human volunteer, resulting in the following volumetric changes for the Avanto: anterior ROI 0.78 %, medial ROI 0.57 %, and posterior ROI 1.32 %. The corresponding numbers for the Trio are: anterior ROI 0.95 %, medial ROI 0.43 %, and posterior ROI 0.16 %.

3.3 Long-Term Absolute and Relative Distortions

Absolute positional and volumetric distortions from the phantom's reference model *G0* were computed for the complete data collection for both scanners. It turned out that absolute geometrical distortions are smaller for the Avanto than for the Trio. As expected, geometrical distortions are depending on the object's distance to the magnet's isocenter. In summary, the maximum errors obtained were 2.0 mm positional deviation and -3.9 % volume mismatch for the Avanto, and 4.8 mm respectively -16.5 % for the Trio within a volume relevant for brain morphometry (spherical subvolume around the isocenter with a radius of 100 mm). The difference between the scanners is mainly a result of different gradient coils being integrated in them.

Relative positional and volumetric changes were determined by comparing positions of corresponding spheres, and volumes of corresponding cells with each other rather than against their reference values in *G0*. Not surprisingly, the magnitude of the distortions respectively volume changes is smaller for the relative than for the absolute evaluations (see Fig. 3 and Fig. 4).

In detail, the maximum relative morphological changes are: 1.0 mm positional, and 2.0 % volumetric change (relative to a standard cell in *G0* of 7600 mm³) for the

Avanto, and 0.6 mm and 2.0 % for the Trio. The slightly better results for the Trio most likely are a result of higher reproducibility in phantom positioning.

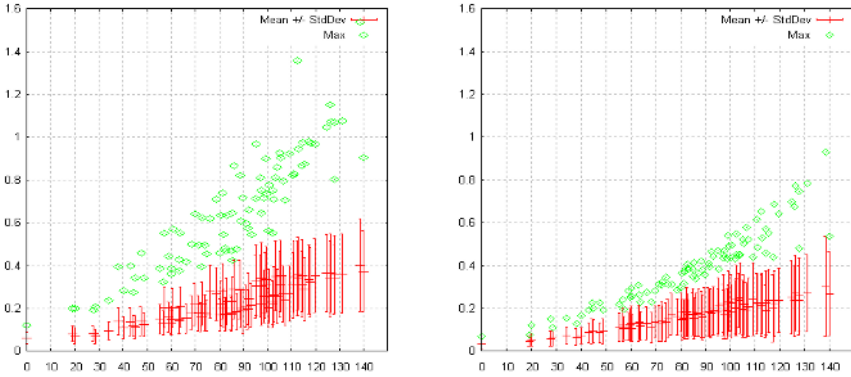


Fig. 3. Distribution of relative positional differences (in mm) of detected spheres in dependence of their distance (in mm) to the isocenter: Avanto (left), Trio (right).

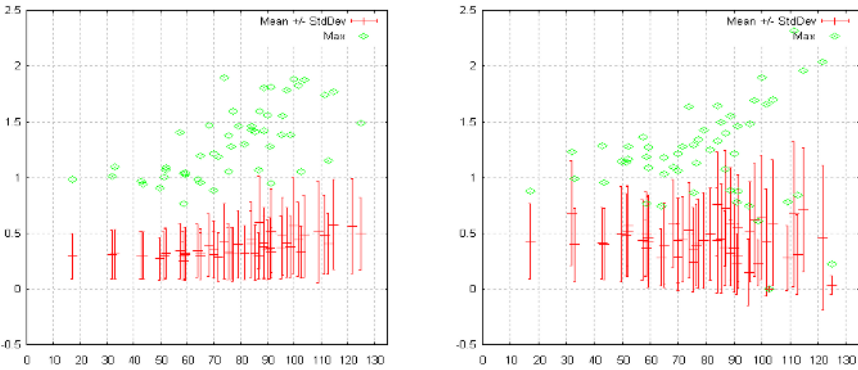


Fig. 4. Distribution of relative volume change (in % of a standard cell's volume in G_0 of 7600 mm^3) in dependence of the cells' distances (in mm) to the isocenter: Avanto (left), Trio (right)

The results of the head image evaluation are given in table 1.

Table 1. Results of head image evaluation: long-term variability in local GM volumetry estimated by range divided by mean (given in brackets)

	Avanto	Trio
anterior ROI	2.88 % (of 40773.8 mm^3)	4.74 % (of 29693.4 mm^3)
medial ROI	2.69 % (of 187359.2 mm^3)	2.60 % (of 184959.6 mm^3)
posterior ROI	5.24 % (of 68907.8 mm^3)	4.78 % (of 86514.3 mm^3)

4 Conclusions

In this paper, the crucial long-term reproducibility of geometrical distortions in MRI was examined for the first time on a local basis. We quantified both practical limitations in serial MRI based morphometry by analyzing head image series of a healthy volunteer acquired over a 6-months period of time on identical hardware, as well as purely acquisition-related limitations by analyzing analogous phantom data series.

As a result of the high reproducibility of our phantom evaluation method (0.15 mm maximal relative positional difference and 0.35 % maximal volume difference including effects caused by the scanners' resonance frequency drift), the results obtained from our analyses can be considered a precise quantification of limitations imposed on serial MRI based morphometry by imperfections in the data acquisition process and by patient repositioning errors. To our understanding, the analysis of these effects is an indispensable prerequisite for a profound validation of image processing methods in the field of morphometry.

In detail, the Avanto and the Trio proved to be equally suitable for serial morphometric studies, as their long-term reproducibility of geometrical distortions was equally high: The maximal relative morphological changes were 1.0 mm positional and 2.0 % volumetric deviation on a 7600 mm³ cuboid. Any morphological changes detected in longitudinal studies carried out with these scanner types that are not larger than the given values have to be interpreted as not significant.

The results of the phantom experiments are well reflected in the evaluations of the head images: Especially cortical areas that are relatively far away from the magnet's isocenter are affected by geometrical distortions. As our volunteer was positioned with great care, and since he was mentally and physically fit and very cooperative, the numbers given in table 1 (up to 5.24 % insecurity in local gray matter volume determination) have to be considered a realistic estimation of serial reproducibility in state-of-the-art MRI based morphometry of cortical areas in terms of a lower limit.

References

1. Thompson. P.M. et al.: Mapping Cortical Change in Alzheimer's Disease, Brain Development, and Schizophrenia; *Neuroimage*, 23 Suppl 1, S2-18, 2004
2. Fischl, B. et al. : Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain; *Neuron*, 33, 341-355, 2002
3. Ashburner, J.T., Friston K.J.: Voxel-Based Morphometry – The Methods; *NeuroImage*, 11, 805-821, 2000
4. Jovicich, J. et al.: Reliability in Multi-Site Structural MRI Studies: Effects of Gradient Non-linearity Correction on Volume and Displacement of Brain Subcortical Structures; *Human Brain Mapping*, Budapest, 2004
5. Jovicich, J. et al.: Multi-site Structural MRI Studies: An Evaluation of Image Distortions and Image Intensity Reproducibility; *ISMRM 2004*
6. Haacke, E.M.: *Magnetic Resonance Imaging – Physical Principles and Sequence Design*; Wiley, 1999
7. Otsu, N.: A Threshold Selection Method from Grey-Level Histograms; *IEEE Transactions on System, Man, and Cybernetics*, 9, 62-66, 1979

Combining Classifiers Using Their Receiver Operating Characteristics and Maximum Likelihood Estimation*

Steven Haker, William M. Wells III, Simon K. Warfield, Ion-Florin Talos,
Jui G. Bhagwat, Daniel Goldberg-Zimring, Asim Mian,
Lucila Ohno-Machado, and Kelly H. Zou

Surgical Planning Lab, Brigham and Women's Hospital,
Harvard Medical School, Boston, MA, 02115, USA
haker@bwh.harvard.edu

Abstract. In any medical domain, it is common to have more than one test (classifier) to diagnose a disease. In image analysis, for example, there is often more than one reader or more than one algorithm applied to a certain data set. Combining of classifiers is often helpful, but determining the way in which classifiers should be combined is not trivial. Standard strategies are based on learning classifier combination functions from data. We describe a simple strategy to combine results from classifiers that have not been applied to a common data set, and therefore can not undergo this type of joint training. The strategy, which assumes conditional independence of classifiers, is based on the calculation of a combined Receiver Operating Characteristic (ROC) curve, using maximum likelihood analysis to determine a combination rule for each ROC operating point. We offer some insights into the use of ROC analysis in the field of medical imaging.

1 Introduction

It is often desirable in clinical practice to combine the results of two or more diagnostic tests or classifiers in order to obtain a more accurate and certain diagnosis. In the field of medical imaging, combinations of independent assessments based on multiple imaging modalities can be combined to create a joint classifier. See [2] for example. Results from segmentation or recognition algorithms can also be combined [8, 3] to produce an improved estimate of ground truth. Ideally, combination of classifiers would be done by joint training and analysis on a common dataset to which all classifiers can be applied. Standard methods in machine learning (logistic regression, PCA, SVMs, *etc.*) could then be used to find an optimized combined classification scheme [5, 6]. In practice, however, it is often the case that joint training data is not available, or is of insufficient quantity. Indeed, there is a “power rule” involved: if it takes roughly

* This work is supported by NIH grants R01CA109246, R01LM007861, R01CA1029246 and P41RR01970.

N data points to estimate a distribution in order to train a single classifier, it is reasonable to expect the need for on the order of N^c data points to estimate the joint distribution needed to train c classifiers. In light of initiatives established to encourage the sharing of algorithms, such as the ITK project (www.itk.org), the lack of sufficient quantities of data for joint training has become more apparent. Accordingly, we have developed the following simple algorithm, used to combine multiple classifiers without the need for joint training. It is based on the maximum likelihood analysis of ROC curves of classifiers.

Although ROC analysis is widespread and standard in the medical field wherever diagnostic tests are analyzed, it is far less common within the field of medical image analysis [9]. We feel this is unfortunate, and that a wider use of these techniques would help lead to general acceptance of image analysis algorithms, *e.g.* algorithms for detection, segmentation and registration, within the clinical community.

2 Background on ROC Analysis

We begin with some basic notions from the standard ROC theory. See [4] for a review of its use in biomedicine. Let I be an image, depending on a binary random variable $T \in \{0, 1\}$ representing unknown “truth” and suppose we have a classification process, or test, A estimating T and depending on a vector of parameters k_A , so that $A(I, k_A) \in \{0, 1\}$. A simple example would be where I is a pixel in a CT image, $k_A = (I_{low}, I_{high})$ consists of a range for Hounsfield units used to segment some structure, and A is then either 1 or 0, indicating the absence or presence of said structure, *i.e.* whether or not the intensity lies in the range given by k_A . A more sophisticated example might be where A is a segmentation algorithm depending upon several parameters.

For each setting of the parameter k_A we define two probabilities, the *true positive rate* $tp_A = \Pr(A = 1|T = 1)$ and the *false positive rate* $fp_A = \Pr(A = 1|T = 0)$. The true positive rate is also known as the *sensitivity* of the classifier, while $1 - fp_A$ is known as A 's *specificity*. We would generally like a classifier to be specific and sensitive. Thus, these notions give us a partial ordering of the unit square $[0, 1]^2$: an operating point (fp_A^1, tp_A^1) is *superior* to (fp_A^0, tp_A^0) if $fp_A^1 \leq fp_A^0$ and $tp_A^1 \geq tp_A^0$.

The *Receiver Operator Characteristic* or ROC for A is the set of points $\{(fp_A(k_A), tp_A(k_A))\} \subset [0, 1]^2$, as k_A ranges over all of its possible values. When k_A is a single scalar value, the ROC is a curve in the unit square parameterized by k_A . We will assume that our ROC curves are concave, and that $tp \geq fp$ for each point on the curve. Concavity is a standard and mild assumption, for any ROC can be made concave by adding a stochastic component to the classifier [7]. Given concavity, $tp \geq fp$ on the ROC curve as long as it contains some points which are superior to $(0, 0)$ and $(1, 1)$. Our work is related to that of [7], who used stochastic methods to create a combined classifier having an ROC equal to the convex hull of the ROCs of the individual classifiers. Our method can pro-

duce superior classifiers, in the sense of having an ROC superior to this convex hull, but requires a conditional independence assumption.

3 Combining Classification Processes

3.1 Model Assumptions

Our model assumes that the classifiers A and B are conditionally independent. This means that given some unknown truth, positive ($T = 1$) for example, we assume that the output of A and B can be modeled as independent Bernoulli processes with respective probability of success tp_A and tp_B , *i.e.* the true-positive rates for the two processes. Note that we do *not* assume the independence of A and B ; only the much weaker assumption of independence conditioned on the true underlying value is required. Conditional independence assumptions are common in machine learning and statistical and information theoretic image processing, especially in relation to maximum likelihood estimation. In the area of ROC analysis, and application to combinations of classifiers, the role of conditional independence is investigated in [1]. This work is related to our own, but differs in the combination technique, estimation of priors, and derivation of a joint statistic.

3.2 Maximum Likelihood Estimation

Let us assume we have two classifiers A and B , and that they are operating according to respective parameters k_A and k_B . We assume we know the ROC curves of the two processes, and the true positive and false positive rates for every value of the parameters k_A and k_B . Given some input, processes A and B will output either 0 (false) or 1 (true), giving us a total of 4 possible cases. For each case we have an expression for the maximum likelihood estimate (MLE) of the unknown truth T :

Table 1. Binary Output for Classifiers A, B and the Maximum Likelihood Combination

$A \setminus B$	Combined MLE of Truth T
1 1	$\Pr(A = 1, B = 1 T = 1) \geq \Pr(A = 1, B = 1 T = 0)$
1 0	$\Pr(A = 1, B = 0 T = 1) \geq \Pr(A = 1, B = 0 T = 0)$
0 1	$\Pr(A = 0, B = 1 T = 1) \geq \Pr(A = 0, B = 1 T = 0)$
0 0	$\Pr(A = 0, B = 0 T = 1) \geq \Pr(A = 0, B = 0 T = 0)$

Each inequality (logical expression) in the rightmost column evaluates either to 0 or 1, and the resulting value is the maximum likelihood estimate of the truth T . If conditional independence is assumed, then $\Pr(A = 1, B = 1 | T = 1) = \Pr(A = 1 | T = 1) \Pr(B = 1 | T = 1) = tp_A tp_B$. See [1] for more details. Proceeding similarly for the other terms in the rightmost column above, we get the following table:

Table 2. Binary Output for Classifiers A, B and the Maximum Likelihood Combination

A	B	Combined MLE of Truth T
1	1	$tp_A tp_B \geq fp_A fp_B$
1	0	$tp_A(1 - tp_B) \geq fp_A(1 - fp_B)$
0	1	$(1 - tp_A)tp_B \geq (1 - fp_A)fp_B$
0	0	$(1 - tp_A)(1 - tp_B) \geq (1 - fp_A)(1 - fp_B)$

From our assumptions detailed above, $tp_A tp_B \geq fp_A fp_B$ and $(1 - tp_A)(1 - tp_B) \geq (1 - fp_A)(1 - fp_B)$, so the first and last rows of Table 2 are determined, and whenever A and B are in agreement their common output is the maximum likelihood estimate of T . Thus, only the middle two rows of the table above need to be determined, resulting in one of 4 possible MLE combination schemes, which we mnemonically name scheme “ A and B ,” scheme “ A ,” scheme “ B ,” and scheme “ A or B .” These are summarized in the following table:

Table 3. Schemes for Combining Processes A and B

A	B	Scheme “ A and B ”	Scheme “ A ”	Scheme “ B ”	Scheme “ A or B ”
1	1	1	1	1	1
1	0	0	1	0	1
0	1	0	0	1	1
0	0	0	0	0	0

It’s easy to calculate the false positive fp and true positive tp rates for these schemes, again using the assumption of conditional independence:

Table 4. False (fp) and True (tp) Positive Rates by Combination Scheme

Scheme	fp	tp
“ A and B ”	$fp_A fp_B$	$tp_A tp_B$
“ A ”	fp_A	tp_A
“ B ”	fp_B	tp_B
“ A or B ”	$fp_A + fp_B - fp_A fp_B$	$tp_A + tp_B - tp_A tp_B$

Thus, under the assumption of conditional independence, these rates can be calculated from information contained in the ROCs for A and B alone. In practice, this means that decision processes can be combined without retraining, since there is no need to estimate joint distributions for the output of A and B , nor the need to know the distribution of the underlying truth T .

3.3 Effect of the Combination Rules on Composite Accuracy

When operating under scheme “ A and B ,” we have $fp = fp_A fp_B \leq fp_A$ and similarly $fp \leq fp_B$, $tp \leq tp_A$, $tp \leq tp_B$. We see that when compared to A

or B alone, this rule generally decreases sensitivity tp but increases specificity $1 - fp$, as one might expect for a scheme that requires a consensus to return a positive result. For the scheme “ A or B ,” we have $fp = fp_A + fp_B - fp_A fp_B = fp_A + fp_B(1 - fp_A) \geq fp_A$, and similarly $fp \geq fp_B$, $tp \geq tp_A$, $tp \geq tp_B$. So the “ A or B ” rule generally increases sensitivity but decreases specificity, again as one might expect. Thus in each of these cases the operating rate (fp, tp) is not demonstrably superior to either (fp_A, tp_A) or (fp_B, tp_B) . However, an advantage is gained by an analysis of the entire range of operating rates, as we describe below.

3.4 Calculating Attainable True and False Positive Rates

To combine processes A and B , we begin by calculating for each value of the parameter pair (k_A, k_B) , and corresponding 4-tuple of false-positive and true-positive rates (fp_A, tp_A, fp_B, tp_B) , the correct ML scheme to use according to Table 2 above, and the resulting combined rates (fp, tp) for that scheme using the formulas in Table 4. In practice, we take discrete values for k_A and k_B , say by sampling them evenly. The resulting set of points (fp, tp) for two example ROCs are shown in Figure 1.

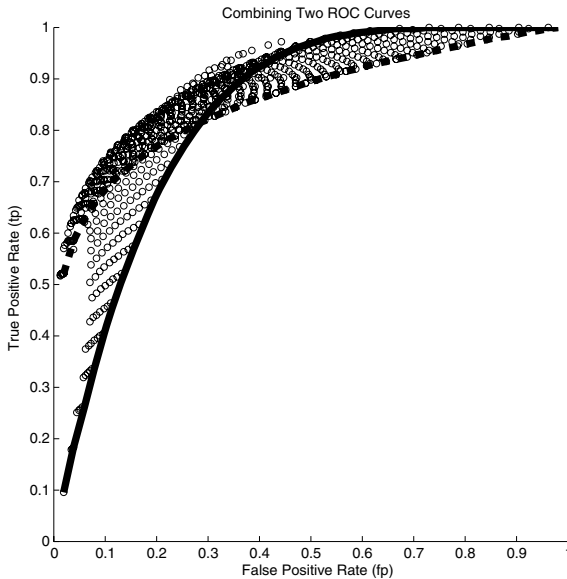


Fig. 1. Two ROCs (solid line, broken line), together with set of points (circles), the outer boundary of which represents the ROC of the combined ML process

3.5 ROC Boundary Curve

The set of points (fp, tp) represent possible operating points for our joint process. However, we do not need to consider points in the interior of the region

containing these points. For each point in the interior, there is a point on the outer boundary of the region which is superior, and thus a better operating point. For example, there is a point on the boundary which has the same false positive rate and a greater true positive rate. Thus, we discard these interior points, and consider only those points along the outer boundary. These points form a curve which is the ROC of our combined process. This combined ROC is the graph of the combined true positive rate thought of as a function of the combined false positive rate fp . We take $fp \in [0, 1]$ to be the parameter of our combined process. In practice, the outer boundary ROC can be estimated by splitting the interval $[0, 1]$ into a number of sub-intervals *i.e.* bins, and within each bin finding the pair (fp, tp) having the largest value of tp . The choice of the number of bins to use requires some care, but this is a common concern which appears whenever data histogramming is required, and standard solutions can be applied. We are currently researching a method by which an exact calculation of the joint ROC curve can be obtained. Along with each point (fp, tp) on the combined ROC curve, we keep track of a pair of parameters (k_A, k_B) , and a ML combination scheme which allows us to operate at (fp, tp) .

3.6 Calculating a Combined Statistic

In theory, the classifiers A and B can be any binary decision process governed by parameters k_A and k_B , where these parameters may be vector valued. In practice however, it is often the case that k_A and k_B are simple thresholds applied to scalar outputs s_A and s_B calculated as part of the A and B decision processes respectively. Thus A returns the estimate $T = 1$ if and only if $k_A \leq s_A$, and similarly for B . In this case, it may be desirable to have a new derived statistic s for the combined process. Let C denote our combined classifier, created as described above. For a chosen operating point (fp, tp) on the ROC curve for C , we have associated thresholds k_A and k_B and an MLE combination rule to be applied in order to derive an estimate $C \in \{0, 1\}$ of T based on the pair of statistics s_A and s_B . We define our joint statistic s as a function of s_A and s_B and the chosen operating level as follows:

Table 5. Formulas for Joint Statistic s

Scheme	Formula for s
“ A and B ”	$\min(s_A - k_A, s_B - k_B)$
“ A ”	$s_A - k_A$
“ B ”	$s_B - k_B$
“ A or B ”	$\max(s_A - k_A, s_B - k_B)$

To use s , we treat it as a statistic and return $C = 1$ if and only if $s \geq 0$. It is easy to see that the true positive and false positive rates for this process are the same as the rates associated with the point on the joint ROC at which we wish to operate. We are currently refining a method by which a single joint

statistic can be produced without the need for an *a priori* specification of an ROC operating point.

4 Illustration of the Method

We illustrate the method described above on a synthetic example. In Figure 2 we show two normal distributions for each of two classifiers *A* and *B*. One is the probability distribution for the statistic s_A or s_B given that $T = 0$, and the other is for these statistics given $T = 1$. The thresholds to use, shown as vertical lines, are determined by our algorithm after we choose an operating point (fp, tp) on the combined ROC curve, shown circled on the in Figure 3. Also displayed in

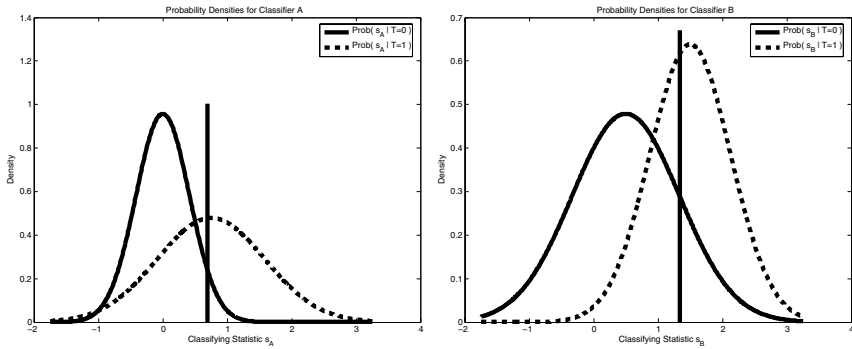


Fig. 2. Distributions associated with the statistics s_A and of a thresholding classification scheme. The thresholds to use are determined by our algorithm.

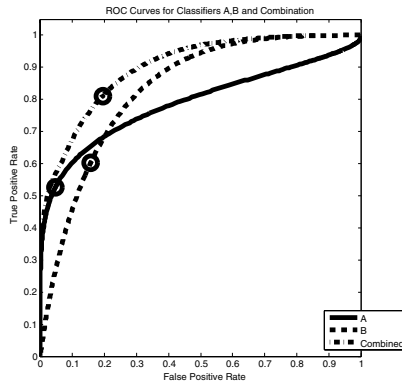


Fig. 3. ROC curves for classifiers *A* and *B*. The circled point on the combined ROC curve represents a level at which we wish to operate. Our algorithm determines the thresholds to use to attain this level, shown in Figure 2, and the corresponding operating levels for *A* and *B*, shown circled on their respective ROC curves.

Figure 3 are the ROC curves for A and B and the corresponding operating levels which result from the thresholds our algorithm chooses.

5 Conclusion and Discussion

We have developed a simple algorithm for combining multiple classifiers without the need for joint training, based on the maximum likelihood analysis of ROC curves of classifiers. Our work has been motivated by the general paucity of joint training data to use with a rapidly expanding array of new segmentation algorithms and diagnostic tests. Future work will include the testing of the method on a range of image and other clinical data, including an investigation of the validity of the conditional independence assumption across this range.

As mentioned before, ROC analysis, though standard in the medical community, has not been as widely adopted in the medical image processing field. Often, a segmentation or registration algorithm requires the specification of numerous parameters, such as kernel sizes, time steps, thresholds, weights applied in a weighted sum of functional terms, *etc.* The engineer typically varies these parameters to find the single point which gives a good result for a training data set, then applies them to a test data set. Yet finding this single point in parameter space is neither necessary nor desirable. What is more in tune with medical research outside of image processing is to report the ROC for the entire range of parameters, or the outer boundary of these possible operating points. Note that in the latter case, the outer boundary effectively reduces the degrees of freedom in the specification of parameters to one.

Medical image processing is maturing, with standardized algorithms for detection, segmentation and registration readily available to the general community in shared form through mechanisms like the ITK project. We believe more widespread use of ROC analysis will lead to greater clinical acceptance.

References

1. M.A. Black and B. A. Craig. Estimating disease prevalence in the absence of a gold standard. *Stats Med*, 21(18):2653–69, 2002.
2. I. Chan, W Wells III, R.V. Mulkern, S. Haker, J. Zhang, K.H. Zou, S.E. Maier, and C.M. Tempny. Detection of prostate cancer by integration of line-scan diffusion, t2-mapping and t2-weighted magnetic resonance imaging; a multichannel statistical classifier. *Med Phys.*, 30(9):2390–8, 2003.
3. J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas. On combining classifiers. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(3):226–239, 1998.
4. T.A. Lasko, J.G. Bhagwat, K.H. Zou, and L. Ohno-Machado. The use of receiver operating characteristic curves in biomedical informatics. *J Biomed Informatics*, In Press., 2005.
5. A. Liu, E.F. Schisterman, and Y. Zhu. Realisable classifiers: improving operating performance on variable cost problems. *Statist. Med.*, 24:37–47, 2005.
6. M.S. Pepe and M.L. Thompson. Combining diagnostic test results to increase accuracy. *Biostatistics*, 1(2):123–140, 2000.

7. M. Scott, M. Niranjana, and R. Prager. Realisable classifiers: improving operating performance on variable cost problems. *British Machine Vision Conference. BMVC.*, 1998.
8. S. K. Warfield, K. H. Zou, and W. M. Wells. Simultaneous truth and performance level estimation (staple): An algorithm for the validation of image segmentation. *IEEE Trans Med Img*, 23(7):903–921, 2004.
9. K.H. Zou, W.M. Wells III, R. Kikinis, and S.K. Warfield. Three validation metrics for automated probabilistic image segmentation of brain tumors. *Statistics in Medicine*, 23:1259–1282, 2004.

Two Methods for Validating Brain Tissue Classifiers

Marcos Martin-Fernandez^{1,2}, Sylvain Bouix^{1,3}, Lida Ungar³,
Robert W. McCarley³, and Martha E. Shenton³

¹ Laboratory of Mathematics in Imaging,
Brigham and Women's Hospital, Boston, MA, USA

² Laboratorio de Procesado de Imagen, Universidad de Valladolid, Spain

³ Department of Psychiatry, Harvard Medical School,
VA Boston Healthcare System, Brockton, MA, USA

Abstract. In this paper, we present an evaluation of seven automatic brain tissue classifiers based on level of agreements. A number of agreement measures are explained, and we show how they can be used to compare different segmentation techniques. We use the Simultaneous Truth and Performance Level Estimation (STAPLE) of Warfield et al. but also introduce a novel evaluation technique based on the Williams' index. The methods are evaluated using these two techniques on a population of forty subjects, each having an SPGR scan and a co-registered T2 weighted scan. We provide an interpretation of the results and show how similar the output of the STAPLE analysis and Williams' index are. When no ground truth is required, we recommend the use of Williams' index as it is easy and fast to compute.

1 Introduction

In today's medical imaging field when one introduces a new segmentation technique, one has to thoroughly validate it and compare it to previously published, well accepted techniques. If a true segmentation exists this task is relatively easy as one only needs to choose a metric measuring differences between the ground truth and the output of the segmentation algorithm. Common metrics are volume differences or measures of overlap [1]. Unfortunately, ground truths or even human expert segmentations are rarely available especially for brain tissue classification where labeling a single brain into gray matter (GM), white matter (WM), cortical spinal fluid (CSF) and background (BG) would take days. Nevertheless, in recent years, novel evaluation procedures have been developed to overcome this problem, and it is now possible, to a certain degree, to rate different methodologies even when a ground truth is not available [2, 3]. In this work, we introduce a novel technique to evaluate brain segmenters based on agreement level and compare it to evaluating each segmenter with STAPLE's estimated ground truth. Seven different classifiers are tested over a data set of 40 different subjects. Our findings show few differences between the results of our technique and those of STAPLE, except for the fact that our evaluation is much faster.

2 Measuring Segmentation Quality

2.1 Williams Index

Consider a set of r raters labeling a set of n voxels with labels $\{1, \dots, l\}$. Let \mathbf{D} denote the set of all labeled voxels (the label map) of all raters. D_{ij} represents the label of rater j for voxel i ; \mathbf{D}_j denotes the label map of rater j ; and $a(\mathbf{D}_j, \mathbf{D}_{j'})$ is the agreement between rater j and j' over all n voxels. Several agreement measures can be used and a few will be defined in section 2.3. Williams' index I_j for rater j is defined as [4]:

$$I_j = \frac{(r-2) \sum_{j' \neq j}^r a(\mathbf{D}_j, \mathbf{D}_{j'})}{2 \sum_{j' \neq j}^r \sum_{j'' \neq j}^{j'} a(\mathbf{D}_{j'}, \mathbf{D}_{j''})} \quad (1)$$

If the upper limit of the confidence interval of this index is greater than one, it can be concluded that rater j agrees with the other raters at least as well as they agree with each other [4]. Using the agreements defined in section 2.3 we can study the statistics of Williams' index for each algorithm, for each label and for each subject.

2.2 Multi-label STAPLE Algorithm

In this section, we describe the multi-label version of the Simultaneous Truth and Performance Level Estimation (STAPLE) Algorithm [3]. This algorithm calculates an estimated ground truth label map out from a set of r given segmentations (raters). Consider a label map with n voxels taking one of l possible labels. Let θ_j be an $l \times l$ matrix. Each element $\theta_j(s', s)$ describes the probability that rater j labels a voxel with s' when the true label is s . The perfect rater will have θ_j equal to the identity matrix. Let $\theta = [\theta_1, \dots, \theta_r]$ be the unknown set of all probability matrices characterizing all r raters. Let $\mathbf{T} = (T_1, \dots, T_n)^T$ be a vector representation of the unknown ground truth segmentation and \mathbf{D} an $n \times r$ matrix whose columns are the r known segmentations. \mathbf{D} is the incomplete data and (\mathbf{D}, \mathbf{T}) the complete data. STAPLE is an estimation process based on the EM algorithm which can estimate the ground truth \mathbf{T} and the parameter matrix θ at the same time by maximizing the complete data log likelihood $f(\mathbf{D}, \mathbf{T} | \theta)$. We refer the reader to [3] for the technical details of the optimization process. Once the ground truth is known, we can use any of the normalized metrics defined in section 2.3 for each algorithm, for each label and for each subject and study the resulting statistics.

2.3 Similarity Measurements

Consider two binary images I_1 and I_2 defined over a finite grid (lattice) L of N spatial sites x . Let $X = \{x \in L, I_1(x) = 1\}$ and $Y = \{x \in L, I_2(x) = 1\}$, and let us define four scalar measurements, $a_1 = |X \cap Y|$, $a_2 = |X| - a_1$, $a_3 = |Y| - a_1$ and $a_4 = N - |X \cup Y|$ as shown schematically in figure 1. We can then express, using these four values, the following similarity measurements, all of them taking values between 0 and 1.

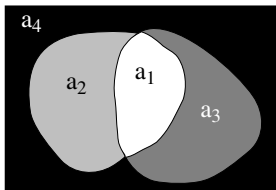


Fig. 1. Schematic diagram for sets X and Y and scalar values a_1 (white), a_2 (light gray), a_3 (dark gray) and a_4 (black)

- **Jaccard (JC)** [5]: $\frac{a_1}{a_1+a_2+a_3} = \frac{|X \cap Y|}{|X \cup Y|}$. It is zero when X and Y are disjoint and one when the sets are equal, i.e. $a_2 = a_3 = 0$.
- **Tanimoto (TN)** [6]: $\frac{a_1+a_4}{a_1+2a_2+2a_3+a_4} = \frac{|X \cap Y| + |\overline{X \cup Y}|}{|X \cup Y| + |\overline{X \cup Y}|}$ where \overline{X} is the set $L - X$. It is zero when X and Y are disjoint and $X \cup Y = L$ and it is one when the sets are equal.
- **Volume similarity (VS)**: $1 - \frac{|a_2-a_3|}{2a_1+a_2+a_3} = 1 - \frac{||X|-|Y||}{|X|+|Y|}$. It is one when the number of elements in both sets are equal, and zero when one of the sets is empty. The positions of the points is irrelevant, only their number counts.

3 Experiments

3.1 Segmentation Pipelines

Data Set: Our data set consists of forty female subjects. The acquisition protocol involves two MR pulse sequences acquired on a 1.5-T GE scanner. First, a SPoiled Gradient-Recalled (SPGR) sequence yielded a coronal MR volume of size $256 \times 256 \times 124$ and voxel dimensions $0.9375 \times 0.9375 \times 1.5\text{mm}$. Second, a double-echo spin-echo sequence gave two axial MR volumes (proton density and T2 weighted) of size $256 \times 256 \times 54$ and voxel dimensions $0.9375 \times 0.9375 \times 3\text{mm}$. For each subject, both axial volumes were co-registered and resampled to the SPGR volume coordinate space using a Mutual Information rigid registration algorithm [7]. Due to limitations on the number of inputs of some of the classification algorithms, only the resampled T2 weighted and the original SPGR were used for segmentation.

Segmentation Techniques: Seven different automatic classifiers were evaluated. The task given was to segment the brain into four classes: BG, CSF, GM and WM. The algorithms were used “as is” with no special tuning of the parameters. A description of each segmenter follows:

- **kNN:** A statistical classification, whose core is a k Nearest Neighbor classifier algorithm trained automatically by non linear atlas registration [8].
- **MINC:** A back-propagation Artificial Neural Network classifier, trained automatically by affine atlas registration [9], the pipeline also includes its own bias field correction tool [10].

Table 1. Segmentation Pipeline Features. The “O” marks a missing feature in the pipeline. In such cases, standard tools were used (see text).

	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
Filtering	O	X	X	X	X	X	O
Bias Correction	O	X	X	X	X	X	O
Brain stripping	O	O	X	O	O	X	X

- **FSL:** A classification algorithm which makes use of a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm [11].
- **SPM:** A mixture model clustering algorithm, which has been extended to include spatial priors and to correct image intensity non-uniformities [12].
- **EM:** The original implementation of the Expectation Maximization algorithm designed by Wells et al. [13].
- **EMAtlas:** An EM based segmentation incorporating a Markov Random Field Model, and spatial prior information aligned to subject’s space by non linear registration [14].
- **Watershed:** A watershed based segmentation which also incorporates spatial prior information in the form of a non linearly aligned atlas [15].

For atlas based classifiers, all techniques use differently defined spatial priors, except for EMAtlas and Watershed which share the same atlas.

Pre- and Post-Processing: Most segmentation techniques are usually a full pipeline involving (i) filtering, (ii) bias field correction, (iii) tissue classification and (iv) brain stripping. As shown in table 1, some methods did not have the all these tools embedded in their framework. We thus used the following three techniques when necessary: **filtering**, the data was smoothed using a diffusion based anisotropic filter [16]; **bias field correction**, was done using the technique of Wells et al. [13]; **brain stripping**, the brain was extracted using the Brain Extraction Tool [17].

3.2 Statistical Analyses

Let $X_{il} = \{x, \mathbf{D}_i(x) = l\}$ be the set of voxels labeled l by rater i . Let T be the estimated ground truth computed by STAPLE from all seven label maps, with T_l the set of voxels labeled l in T . For each subject, for each label, four different types of analysis were done. First, Williams’ index was computed for each label using the seven X_{il} as input and all three agreement measures (Williams 1). The mean and standard deviation of the index over all subjects for each label and segmentation algorithm are shown in table 3. Second, for each rater, the three agreement measures between X_{il} and the estimated ground truth for that label T_l were computed (Staple 1). The mean and standard deviation of the index over all subjects are shown in table 2. Third, $C_l = \cap_{1 \leq i \leq r} X_{il}$, the set of all points with the same label l in all r label maps was computed. Williams’ index was then calculated for each label l using all seven $X_{il} - C_l$ as input (Williams 2).

Table 2. Staple 1

BG	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	1.00(0.00)	0.99(0.00)	0.99(0.00)	0.99(0.00)	0.99(0.00)	0.99(0.00)	0.99(0.00)
TN	0.99(0.00)	0.99(0.00)	0.99(0.00)	0.98(0.01)	0.98(0.01)	0.98(0.01)	0.98(0.01)
VS	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)	0.99(0.00)	1.00(0.00)
GM	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	<i>0.85(0.09)</i>	0.77(0.03)	0.73(0.23)	0.82(0.03)	0.72(0.03)	0.77(0.03)	0.87(0.03)
TN	<i>0.98(0.01)</i>	0.97(0.01)	0.96(0.04)	0.97(0.01)	0.96(0.00)	0.97(0.01)	0.98(0.00)
VS	0.98(0.01)	0.96(0.02)	0.96(0.06)	0.97(0.01)	<i>0.98(0.02)</i>	<i>0.98(0.02)</i>	<i>0.98(0.02)</i>
CSF	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	0.73(0.11)	0.46(0.14)	0.63(0.24)	0.60(0.12)	0.36(0.11)	0.42(0.12)	<i>0.62(0.10)</i>
TN	0.98(0.01)	0.97(0.01)	<i>0.98(0.02)</i>	0.98(0.01)	0.97(0.01)	0.96(0.01)	0.98(0.01)
VS	0.96(0.04)	0.76(0.11)	0.89(0.12)	<i>0.91(0.07)</i>	0.52(0.12)	0.81(0.07)	0.86(0.06)
WM	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	<i>0.87(0.09)</i>	0.80(0.04)	0.80(0.13)	0.82(0.03)	0.80(0.03)	0.83(0.03)	0.90(0.02)
TN	<i>0.99(0.01)</i>	0.98(0.00)	0.97(0.02)	0.98(0.00)	0.98(0.00)	0.98(0.00)	0.99(0.00)
VS	0.98(0.01)	<i>0.97(0.02)</i>	0.92(0.10)	0.91(0.02)	0.89(0.02)	0.96(0.03)	<i>0.97(0.02)</i>

The rationale for this process is to try to get a more focused analysis on the differences between raters. The results are show in table 5. Finally, the three agreement measures between $X_{il} - C_l$ and $T_l - C_l$ was calculated (Staple 2). The results are shown in table 4. In all the tables the best score is in **bold** and the second best score in *italic*.

4 Results

Table 2 shows results for Staple 1. We can see that for BG, none of the three agreement measurements give meaningful results. For GM and WM, Watershed is best one followed very closely by kNN when looking at JC and TN. Using VS, kNN performs slightly better than Watershed, which is similar to EM and EMAtlas for GM and similar to MINC for WM. For CSF, kNN is best, and Watershed and SPM rank second.

Table 3 shows the results for Williams 1. Similarly to Staple 1, none of the measurements are meaningful for BG. For GM, using JC and TN, Watershed is best followed closely by kNN. Using VS, kNN and EMAtlas are slightly better than Watershed. For WM, using JC and TN, Watershed is also best, followed by kNN for JC and SPM and EMAtlas for TN. Using VS, kNN is best, MINC and Watershed are second. All of these results mostly agree with the ones using Staple 1. Finally, for CSF, SPM is best followed by Watershed using JC. Watershed is best using TN and VS. In this case, Staple 1 and Williams 1 only agree for TN, and there is disagreement for JC and VS.

Table 4 shows the results for Staple 2. In comparison to Staple 1 (table 2), similar rankings are obtained for GM, CSF and WM for all the agreement measures. The main difference are in evaluating BG. More significance is

Table 3. Williams 1

BG	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)
TN	1.00(0.00)	1.00(0.00)	0.99(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)
VS	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)
GM	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	<i>1.07(0.10)</i>	1.00(0.07)	0.90(0.25)	1.03(0.05)	0.94(0.07)	1.00(0.07)	1.09(0.06)
TN	1.01(0.01)	<i>1.00(0.01)</i>	0.99(0.03)	<i>1.00(0.01)</i>	0.99(0.01)	<i>1.00(0.01)</i>	1.01(0.01)
VS	1.01(0.01)	0.99(0.03)	0.97(0.06)	<i>1.01(0.02)</i>	1.00(0.02)	1.01(0.01)	<i>1.01(0.02)</i>
CSF	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	1.04(0.12)	1.08(0.14)	0.79(0.36)	1.23(0.08)	0.81(0.17)	0.99(0.13)	<i>1.18(0.09)</i>
TN	0.99(0.01)	<i>1.00(0.00)</i>	0.99(0.01)	<i>1.00(0.00)</i>	<i>1.00(0.00)</i>	<i>1.00(0.00)</i>	1.01(0.00)
VS	1.00(0.05)	1.04(0.08)	0.99(0.05)	1.08(0.06)	0.72(0.13)	<i>1.09(0.04)</i>	1.11(0.04)
WM	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	<i>1.04(0.09)</i>	0.97(0.04)	0.96(0.13)	0.99(0.04)	0.98(0.05)	1.01(0.05)	1.07(0.03)
TN	1.00(0.01)	1.00(0.01)	0.99(0.02)	<i>1.00(0.00)</i>	1.00(0.01)	<i>1.00(0.00)</i>	1.01(0.00)
VS	1.04(0.02)	<i>1.03(0.02)</i>	0.97(0.09)	0.96(0.03)	0.94(0.04)	1.02(0.03)	<i>1.03(0.02)</i>

Table 4. Staple 2

BG	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	0.37(0.16)	0.50(0.22)	0.28(0.11)	0.40(0.22)	0.40(0.23)	0.34(0.19)	<i>0.41(0.22)</i>
TN	0.99(0.00)	0.99(0.00)	0.99(0.00)	0.98(0.01)	0.98(0.01)	0.98(0.01)	0.98(0.01)
VS	<i>0.68(0.23)</i>	0.70(0.25)	0.65(0.26)	0.54(0.22)	0.54(0.24)	0.48(0.21)	0.56(0.24)
GM	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	<i>0.71(0.11)</i>	0.57(0.10)	0.55(0.24)	0.65(0.08)	0.47(0.10)	0.56(0.10)	0.74(0.07)
TN	<i>0.98(0.01)</i>	0.97(0.01)	0.96(0.04)	0.97(0.01)	0.96(0.00)	0.97(0.01)	0.98(0.00)
VS	0.96(0.03)	0.91(0.05)	0.93(0.08)	<i>0.95(0.03)</i>	0.94(0.05)	0.96(0.03)	0.95(0.04)
CSF	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	0.69(0.12)	0.36(0.17)	<i>0.59(0.23)</i>	0.53(0.14)	0.22(0.16)	0.32(0.15)	0.54(0.12)
TN	0.98(0.01)	0.97(0.01)	<i>0.98(0.02)</i>	0.98(0.01)	0.97(0.01)	0.96(0.01)	0.98(0.01)
VS	0.95(0.05)	0.68(0.16)	0.88(0.12)	<i>0.90(0.09)</i>	0.33(0.20)	0.76(0.10)	0.83(0.08)
WM	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	<i>0.63(0.11)</i>	0.46(0.08)	0.53(0.16)	0.38(0.12)	0.52(0.05)	0.48(0.11)	0.67(0.06)
TN	<i>0.99(0.01)</i>	0.98(0.00)	0.97(0.02)	0.98(0.00)	0.98(0.00)	0.98(0.00)	0.99(0.00)
VS	0.94(0.04)	<i>0.90(0.06)</i>	0.80(0.20)	0.58(0.13)	0.69(0.05)	0.84(0.12)	<i>0.90(0.06)</i>

achieved for JC and VS, but not for TN which is still not very meaningful. MINC is performing best for BG using Staple 2, followed by Watershed using JC and by kNN using VS.

Table 5 shows the results for Williams 2. Again, the results for GM, CSF and WM are similar to Williams 1 (table 3) overall. The only difference is for JC and for WM, for which Watershed is best for Williams 1, while EM is best, followed by Watershed for Williams 2. With respect to BG, TN is still not meaningful. Using JC and VS, there is significance for Williams 2. For JC, EM performs best followed by SPM and for VS, Watershed is best followed by SPM and EM.

Table 5. Williams 2

BG	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	0.36(0.23)	1.09(0.24)	0.26(0.09)	<i>1.50(0.10)</i>	1.54(0.09)	1.46(0.10)	1.42(0.12)
TN	1.00(0.00)	1.00(0.00)	0.99(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)
VS	0.64(0.23)	1.14(0.09)	0.54(0.14)	<i>1.25(0.05)</i>	<i>1.25(0.06)</i>	1.14(0.08)	1.26(0.06)
GM	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	<i>1.17(0.15)</i>	1.00(0.13)	0.89(0.38)	1.10(0.09)	0.80(0.18)	0.95(0.14)	1.21(0.10)
TN	1.01(0.01)	<i>1.00(0.01)</i>	0.99(0.03)	<i>1.00(0.01)</i>	0.99(0.01)	<i>1.00(0.01)</i>	1.01(0.01)
VS	1.03(0.02)	0.97(0.06)	0.96(0.09)	1.01(0.03)	0.99(0.05)	<i>1.02(0.03)</i>	<i>1.02(0.03)</i>
CSF	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	1.20(0.17)	1.02(0.22)	0.93(0.46)	1.42(0.09)	0.52(0.33)	0.93(0.18)	<i>1.29(0.13)</i>
TN	0.99(0.01)	<i>1.00(0.00)</i>	0.99(0.01)	<i>1.00(0.00)</i>	<i>1.00(0.00)</i>	<i>1.00(0.00)</i>	1.01(0.00)
VS	1.05(0.09)	1.05(0.13)	1.03(0.08)	1.15(0.10)	0.52(0.24)	<i>1.15(0.07)</i>	1.19(0.08)
WM	kNN	MINC	FSL	SPM	EM	EMAtlas	Watershed
JC	1.19(0.22)	0.89(0.15)	1.10(0.22)	0.60(0.20)	1.27(0.12)	0.91(0.23)	<i>1.21(0.13)</i>
TN	1.00(0.01)	1.00(0.01)	0.99(0.02)	<i>1.00(0.00)</i>	1.00(0.01)	<i>1.00(0.00)</i>	1.01(0.00)
VS	1.17(0.06)	<i>1.14(0.07)</i>	0.99(0.22)	0.71(0.18)	0.87(0.11)	1.07(0.15)	1.13(0.07)

5 Discussion

We have investigated different approaches to evaluate the quality of a segmentation only based on agreement measures. A great number of similarities have been found between Staple 1 and Williams 1, suggesting kNN is the most consistent segmentation method. In general for GM, WM and CSF, all ranking techniques give similar results. For BG, better significance is achieved for JC and VS after discarding the common agreement among algorithms and focusing only on differences (Williams 2 and Staple 2). Staple 2 presents MINC as the best method for BG and Williams 2 has EM as the better technique. In general, FSL, SPM and EMAtlas are less well ranked.

Overall, Williams' index gives similar results to STAPLE, and unless one absolutely needs the *estimated* ground truth of STAPLE for further processing, using Williams' index is sufficient. The biggest advantage is that of speed, as STAPLE in our experimental setup can take as much as 20mn to process one case, whereas Williams index takes a few seconds. We are now interested in studying the probability matrices provided by STAPLE and also finding more intuitive and compact ways to present the results of the evaluation.

Acknowledgments

We acknowledge the support of NIH (K02 MH01110, R01 MH50747, U54 EB005149 to MES, R01 MH40799 to RWM), the Dept of Vet. Affairs Merit Awards and REAP award (MES, RWM) and Fulbright (FU2003-0968 to MMF).

References

1. Zou, K.H., Wells, W.M., Kikinis, R., Warfield, S.K.: Three validation metrics for automated probabilistic image segmentation in brain tumors. *Statistics in Medicine* **23** (2004) 1280–1291
2. Bello, F., Colchester, A.C.F.: Measuring global and local spatial correspondence using information theory. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. (1998) 964–973
3. S. K. Warfield, K. H. Zou, W.M.W.: Simultaneous truth and performance level estimation (staple): An algorithm for the validation of image segmentation. *IEEE Trans. Medical Imaging* **23** (2004) 903–921
4. Williams, G.W.: Comparing the joint agreement of several raters with another rater. *Biometrics* **32** (1976) 619–627
5. Jaccard, P.: Étude comparative de la distribution florale dans une portion des alpes et de jura. *Bulletin de la Société Voudoise des Sciences Naturelles* **37** (1901) 547–579
6. Rogers, J.S., Tanimoto, T.T.: A computer program for classifying plants. *Science* **132** (1960) 1115–1118
7. Wells, W.M., Viola, P., Atsumi, H., Nakajima, S., Kikinis, R.: Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis* **1** (1996) 35–52
8. Warfield, S.: Fast knn classification for multichannel image data. *Pattern Recognition Letters* **17** (1996) 713–721
9. Zijdenbos, A.P., Forghani, R., Evans, A.C.: Automatic "pipeline" analysis of 3-d mri data for clinical trials: application to multiple sclerosis. *IEEE Transactions on Medical Imaging* **21** (2002) 1280–1291
10. Sled, J.G., Zijdenbos, A.P., Evans, A.C.: A non-parametric method for automatic correction of intensity non-uniformity in mri data. *IEEE Transactions On Medical Imaging* **17** (1998) 87–97
11. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden markov random field model and the expectation maximization algorithm. *IEEE Transactions on Medical Imaging* **20** (2001) 45–57
12. Ashburner, J., Friston, K.: Spatial normalization using basis functions. In Frackowiak, R.S.J., Friston, K.J., Frith, C., Dolan, R., Friston, K.J., Price, C.J., Zeki, S., Ashburner, J., Penny, W., eds.: *Human Brain Function*. 2nd edn. Academic Press (2003)
13. Wells, W.M., Kikinis, R., Grimson, W.E.L., Jolesz, F.: Adaptive segmentation of mri data. *IEEE Transactions on Medical Imaging* **15** (1996) 429–442
14. Pohl, K.M., Bouix, S., Kikinis, R., Grimson, W.E.L.: Anatomical guided segmentation with non-stationary tissue class distributions in an expectation-maximization framework. In: *IEEE International Symposium on Biomedical Imaging, Arlington, VA* (2004) 81–84
15. Grau, V., Mewes, A.U.J., Alcaiz, M., Kikinis, R., Warfield, S.K.: Improved watershed transform for medical image segmentation using prior information. *IEEE Transactions on Medical Imaging* **23** (2004) 447–458
16. Krissian, K.: Flux-based anisotropic diffusion applied to enhancement of 3-d angiogram. *IEEE Transactions on Medical Imaging* **21** (2002) 1440–1442
17. Smith, S.: Fast robust automated brain extraction. *Human Brain Mapping* **17** (2002) 143–155

Comparison of Vessel Segmentations Using STAPLE

Julien Jomier, Vincent LeDigarcher, and Stephen R. Aylward

Computer-Aided Diagnosis and Display Lab,
The University of North Carolina at Chapel Hill,
Department of Radiology, 27510 Chapel Hill, USA
{jomier, aylward}@unc.edu
vincent.ledigarcher@cpe.fr

Abstract. We propose a novel method for the validation of vascular segmentations. Our technique combines morphological operators and the STAPLE algorithm to obtain ground truth of centerline extractions as well as a measure of accuracy of the methods to be compared. Moreover, our method can be extended to the validation of any open-curves. We also present a comparison study of three vascular segmentation methods: ridge traversal, statistical and curves level set. They are compared with manual segmentations from five experts.

1 Introduction

Blood vessels and their branches vary considerably but are often critical in planning and performing neurosurgical and interventional procedures. In planning, the feeding and drain vessels of a lesion must be defined. During surgery the vessels serve as landmarks and guidelines to the lesion [1]. The more precise the vascular segmentation is, the more accurate the plan and navigation.

Segmentation of tubular structures and more specifically vascular segmentation has been of high interest in medical imaging, and several excellent techniques have been developed, incorporating a variety of different approaches. For instance, the curve evolution algorithm [4] produces accurate vascular segmentations by combining the modified curvature diffusion equation (MCDE) with a level-set based technique. On the other hand, Aylward et al. [2] use a ridge traversal technique with width estimation to extract vascular centerline and estimated radius at each point along blood vessels. Wink et al. [6] have developed a front propagation method for the extraction of blood vessels. Given a starting point, a front wave is moved to fit the tubular structure.

Surveys have been conducted to contrast the vascular segmentation algorithms [3], however, no direct comparison has been done on the accuracy of these techniques. In this paper, we propose a measure for vessel segmentation comparison and we apply it to the analysis of vascular segmentation algorithms from 3-dimensional images of the liver (CT and MR combined). The rest of this paper is structured as follow: first the STAPLE algorithm and its extension to open curves is described, second the segmentation methods to be compared and the results of the comparison are presented.

2 Method

In this section we present the STAPLE algorithm [5] and detail how we extend its use for open curves validation.

2.1 STAPLE Algorithm

The Simultaneous Truth and Performance Level Estimation (STAPLE) algorithm generates ground truth volumes from a set of binary expert segmentations as well as a simultaneous accuracy assessment of each expert.

STAPLE works as follow: considering $\mathbf{p} = (p_1, p_2, \dots, p_r)^T$ a column vector of R elements, with each element a sensitivity parameter characterising one of the R segmentations, and $\mathbf{q} = (q_1, q_2, \dots, q_r)^T$ a column vector representing the specificity parameter of one of R segmentations. Let \mathbf{D} be an $N \times R$ matrix describing the binary decisions made for each segmentation at each voxel of the image and \mathbf{T} be an indicator vector of N elements representing the hidden binary true segmentation. The complete data can be written as (\mathbf{D}, \mathbf{T}) and the probability mass function of these data $f(\mathbf{D}, \mathbf{T}|\mathbf{p}, \mathbf{q})$. An expectation-maximization algorithm then estimates the performance level of the experts characterized as a tuple (p, q) , where p represents the sensitivity (“true positive fraction”) and q the specificity (“true negative fraction”), which maximizes the complete data log likelihood function

$$(\hat{\mathbf{p}}, \hat{\mathbf{q}}) = \operatorname{argmax}_{\mathbf{p}, \mathbf{q}} \ln f(\mathbf{D}, \mathbf{T}|\mathbf{p}, \mathbf{q}) \quad (1)$$

The STAPLE algorithm treats segmentation as a pixelwise classification, which leads to an averaging scheme that accounts for systematic biases in the behavior of experts in order to generate a fuzzy ground truth volume and simultaneous accuracy assessment of each expert. One can notice that for STAPLE to work, the set of binary segmentations should overlap which is often not the case for open curves. In order to overcome this issue, we propose an iterative scheme to construct ground truth of open curve segmentations such as blood vessels. First, each centerline extraction is discretized to form a binary image volume - having an intensity of one on the centerline and zero outside. Second, our method iteratively creates new morphologically dilated segmentations from the initial volume. Each discretized centerline is dilated by the same kernel (spherical operator of increasing radius). The STAPLE algorithm is then used to compute the level of each expert as well as the ground truth at each stage of the dilation process. The resulting output is a set of probability maps as well as sensitivity and specificity levels of the experts for each dilation factor.

2.2 Toy Example

We illustrate the behavior of our technique using a simple example composed of three segmentations (one per expert) of an horizontal straight line. Figure 1 shows the three original segmentations as well as the resulting probability map. The sensitivity p and specificity q of the experts are shown in figure 2.

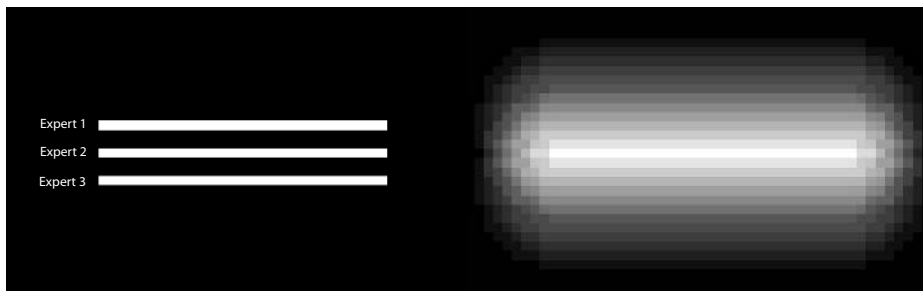


Fig. 1. Simple example composed of three segments representing segmentation of the same structure by three experts (**left**) and the resulting probability map obtained by averaging the ground truth probability volumes (**right**)

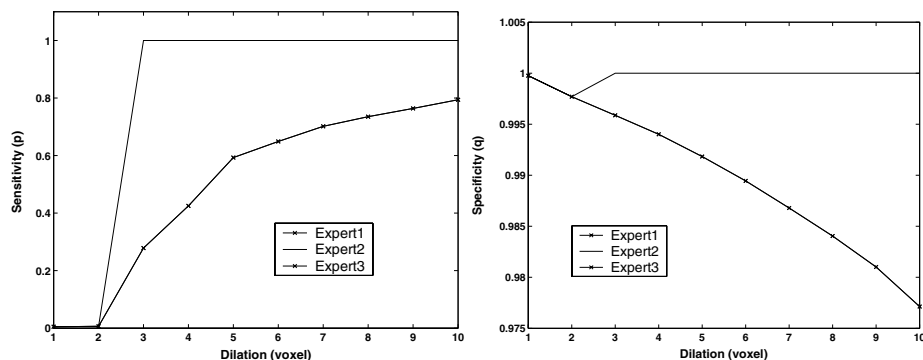


Fig. 2. Sensitivity and specificity of the experts in the simple example of three straight line segmentations. (Expert 1 and Expert 3 are overlapping).

In this example, the three initial segmentations were separated by three pixels. As we can predict the expert levels graph shows high specificity and very low sensitivity as long as the three binary segmentations do not overlap ($dilation < 3$). As soon as the dilated segmentation start to overlap, the sensitivity increases and the specificity decreases. One can notice that the sensitivity of the second expert is higher than the two others which is the intended behavior of the algorithm since the second expert has extracted the real structure. The sensitivity will increase logarithmically since in the limit ($dilation \rightarrow \infty$) the sensitivity will be one for all experts and specificity is undefined. The specificity graph shows that all experts are following the same specificity until segmentations overlap, then, once again, the second expert outperforms the other two as expected.

2.3 Open Curves Validation

In order to validate our technique, we created simulated data composed of single spiral going down the Z axis. A binary mask (high intensity inside the spiral, low

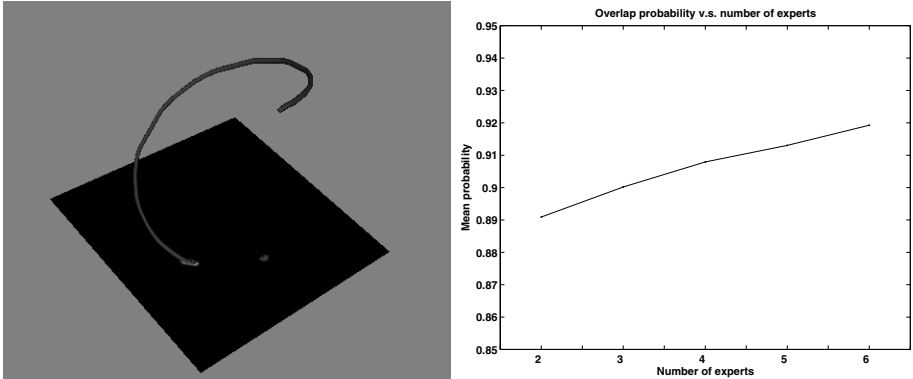


Fig. 3. Simulated open-curve: a spiral (**left**) and the mean probability along the centerline of the spiral given an increasing number of experts (**right**)

intensity outside) was computed from it. After a gaussian blurring of the binary volume ($\sigma = 2$), six experts extracted the spiral using a manual (point and click) segmentation technique. We then applied our validation method from $n = 2$ experts to $n = 6$ experts to create five normalized ($I \in [0, 1]$) probability maps of the manually segmented spiral. We then superimposed the original spiral with each probability map and computed the mean probability along the centerline. A mean probability of 1.0 would mean that the segmented spiral is the exact representation of the original. Figure 3 shows the results of the experiment.

As one can see the segmentation keeps improving with the number of experts. This is due to the STAPLE algorithm which limits the contribution of “bad” experts. A “bad” expert can be viewed as the one who produces segmentation results significantly different compared to the other experts (because that expert has fewer experience for instance). From the probability map we can then reconstruct the mean spiral.

3 Comparison of Segmentation Techniques

In this section we compared three vascular segmentation techniques on seven liver datasets (CT and MR). CT volumes are contrast enhanced $1 \times 1 \times 3mm^3$ voxels and MRA volumes are time-of-flight data with $1 \times 1 \times 1mm^3$ voxels. In order to establish truth, five experts have segmented vascular centerlines manually by point and click. We then compared the expert segmentations with three other techniques: (a) a ridge traversal and width estimation method, (b) a statistical segmentation technique using connected components and (c) a region growing via level set method using the Curves algorithm. Next we present these three techniques in details.

3.1 Ridge Traversal and Width Estimation

This vessel extraction method extracts blood vessels from 3-dimensional images using a scale space technique. The algorithm traverses a ridge in an intensity

function \mathbf{F} , using the Hessian at a point x . Let's define α and β as ascending-ordered eigenvalues of the Hessian at x , \mathbf{u} and \mathbf{v} as the corresponding eigenvectors of the Hessian, and \mathbf{P} and \mathbf{Q} as the directional derivatives: $\mathbf{P} = \mathbf{u} \cdot \mathbf{F}$ and $\mathbf{Q} = \mathbf{v} \cdot \mathbf{F}$. Therefore, if x is exactly in the middle of the ridge the following conditions must hold: $\alpha < 0$ and $\mathbf{P} = 0$. Given an initial starting point close to the ridge, the intensity ridge is computed to minimize \mathbf{P} using a direction search with respect to the Hessian. The line search is performed from x in the direction \mathbf{u} to find the local minimum of \mathbf{P} . If the resulting minimum is not within a specified tolerance a new initial point is required. This vascular segmentation method has shown sub-voxel accuracy but has never been compared with other techniques on physical datasets.

3.2 Statistical Segmentation

The second segmentation technique to be compared uses statistics on the intensity of the blood vessel as well as a region growing technique. First the user specifies a set of points inside a blood vessel, then statistics are computed to define a gaussian probability class corresponding to the blood vessel likelihood. Second, a connected components algorithm is performed based on the previously defined class. Third and last, morphological operators are used to remove unwanted structures and obtain a smooth segmentation.

Skeletization is then done using a ridge finding technique. Basically the binary image volume is blurred by an amount proportional to the expected radius value of the tubular structure and the ridge is tracked.

3.3 Segmentation Via Levelset

The third segmentation method is based on the Curves algorithm [4], an extension of geodesic active contours based on a level set implementation. The evolution of the level set is driven by the curvature equation. In the case of a 1-dimensional curve in a 3-dimensional space, the following equation should be minimized

$$\oint_1^0 g(|\nabla I(c(p))|)|C'(p)|dp \tag{2}$$

where $C(p)$ is the 1-dimensional curve, I the image, and g a strictly decreasing function such that $g(r) \rightarrow 0$ as $r \rightarrow \infty$. By computing the Euler-Lagrange equations, the curve evolution equation can be formulated as

$$C_t = kN - \frac{g'}{g} \prod \left(H \frac{\nabla I}{|\nabla I|} \right) \tag{3}$$

where H is the Hessian of the intensity function. Therefore the equation for embedding space is

$$\nu_t = \lambda (\nabla \nu(x, t), \nabla^2 \nu(x, t)) + \frac{g'}{g} \nabla \nu(x, t) \cdot H \frac{\nabla I}{|\nabla I|} \tag{4}$$

The output of the Curves algorithm is a binary image obtained from the zero level set. Thereafter, the same skeletization technique as in 3.2 is used to obtain the centerlines.

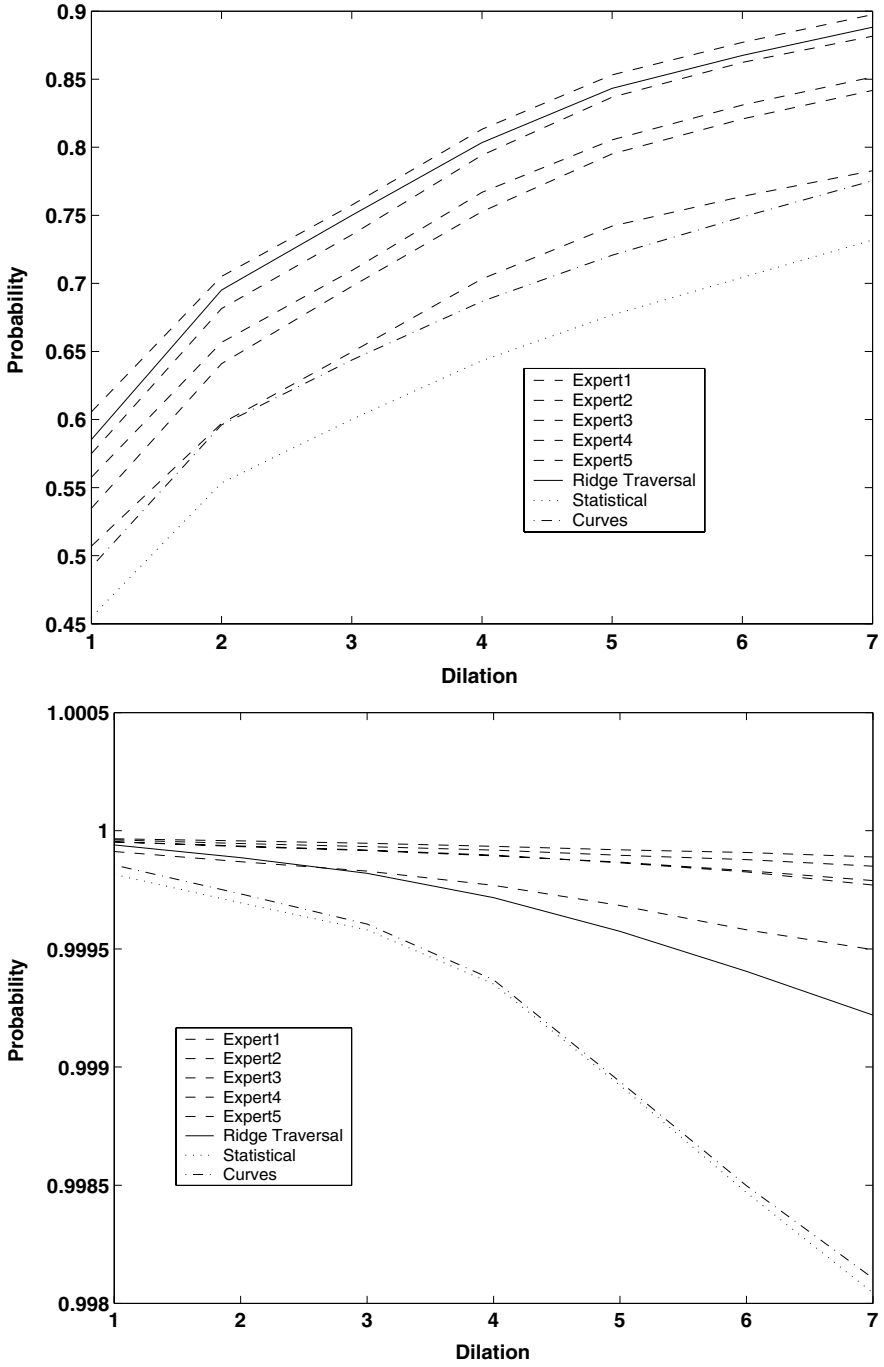


Fig. 4. Sensitivity (top) and specificity (bottom) of the 5 experts and the 3 segmentation methods to be compared

3.4 Results

Given the centerline segmentations (5 experts and 3 automated) for the seven datasets we applied our algorithm until a maximum dilation of 7 voxels. We are interested in the sensitivity and specificity of each method; these two values characterize the confidence level of a method and allow for comparisons. The mean results are presented in figure 4.

One can notice that the ridge traversal technique outperforms the statistical and curves method in both terms of sensitivity and specificity. Moreover, the sensitivity of the ridge traversal and curves technique are close to the sensitivity of the experts. However, the statistical method performs poorly compared to the experts and the other automated techniques.

From the specificity graph, one can notice that the three automated methods have less specificity than the experts. There are two reasons for this. First, the length of the segmented centerlines appear to be shorter for all the experts segmentation. This can be explained by the noise present in the image which makes small structures invisible to the human eye but the automated algorithms are able to extend further into the tubular structure. Second, Curves and statistical methods tends to “leak” outside of the structure making the skeleton less reliable.

4 Discussion and Conclusions

We have presented a novel algorithm for validation of open-curve segmentation and applied it to the evaluation of vessel segmentation methods. Our technique quantifies the accuracy of different segmentation techniques and also produces an accurate probability map of the ground truth segmentation. We have shown that ridge traversal techniques can outperform level sets and statistical method in localizing the centerline of a tubular structure.

The main drawback of our approach is that it discretizes centerlines to the pixel level. Some segmentations can have a sub-voxel accuracy, but this technique cannot be used to discern that accuracy. Moreover, we are currently extending our method to estimate the radius of segmented tubular structures assuming a constant radius at each centerline point.

This work benefited from the use of the Insight Segmentation and Registration Toolkit (ITK), an open source software developed as an initiative of the U.S. National Library of Medicine and available at www.itk.org.

This work is funded in part by the Whitaker Foundation (TF-04-0008).

References

1. Aylward, S., Jomier, J., Weeks, S., Bullitt, E.: Registration of Vascular Images International Journal of Computer Vision, November 2003, pages 15
2. Aylward, S., Bullitt, E.: Initialization, Noise, Singularities, and Scale in Height-Ridge Traversal for Tubular Object Centerline Extraction IEEE Transactions on Medical Imaging, Feb, 2002, Pages 61-75

3. Kirbas C., Quek Francis K. H.: Vessel Extraction Techniques and Algorithms: A Survey Third IEEE Symposium on BioInformatics and BioEngineering (BIBE'03)
4. Lorigo L., Faugeras O., Grimson W.E.L., Keriven R., Kikinis R., Nabavi A. and Westin C.-F.: Curves: Curve evolution for vessel segmentation. *Medical Image Analysis*, 5:195-206, 2001.
5. Warfield S., Zou K., Wells W.: Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *IEEE Trans Med Imaging*. 2004 Jul;23(7):903-21.
6. Wink O., Niessen W. J., Verdonck Bert, Viergever M. A.: Vessel Axis Determination Using Wave Front Propagation Analysis. *MICCAI 2001: Fourth International Conference on Medical Image Computing and Computer-Assisted Intervention*. 845-853

Validation Framework of the Finite Element Modeling of Liver Tissue

Hongjian Shi, Rachid Fahmi, and Aly A. Farag

Computer Vision and Image Processing Laboratory (CVIP),
University of Louisville, Louisville, KY 40292
{hshi, rachidf, farag}@cvip.louisville.edu

Abstract. In this work, we aim at validating some soft tissue deformation models using high resolution Micro Computed Tomography (Micro-CT) and medium resolution Cone-Beam CT (CBCT) images. These imaging techniques play a key role in detecting the tissue deformation details in the contact region between the tissue and the surgical tool (probe) even for small force loads, and provide good capabilities for creating accurate 3D models of tissues. Surgical simulations rely on accurate representation of the mechanical response of soft tissues subjected to surgical manipulations. Several finite element (F.E.) models have been suggested to characterize soft tissues. However, validating these models for specific tissues still remains a challenge. For our validation, ex vivo lamb liver is chosen to validate the linear elastic model (LEM), the linear viscoelastic model (LVEM), and the neo-Hooke hyperelastic model (NHM). We found that the LEM is more applicable to lamb liver than the LVEM for small force loads ($< 40g$) and that the NHM is closer to reality than the LVEM for this same range of force loads.

1 Introduction

Computer-aided surgical simulation progressed significantly in the last decade [1, 3]. Careful planning is of great importance in order to limit the damage to healthy tissue during surgery. In particular, accurate modeling of the mechanical behavior of the tissues is required. To achieve accurate prediction, biomechanical F.E. modeling of soft tissues is employed. The simplest one of such models is the LEM which was used early to estimate the force driven left ventricular deformation (e.g., [1]). Because of its simplicity, the LEM was widely used in early surgical simulation and is still used frequently [2, 4]. With the increasing speed of computer processing, the prediction accuracy attracts more attention. More complicated biomechanical models have been then employed to predict the deformations or registrations of brain, liver, skin and muscle tissues [4, 5]. Some works on the liver tissue modeling using the FEM were reviewed and reported in [3]. However, the validation of these models for the liver tissue is still a challenging step in building a real surgical simulator for clinical use.

In this paper, we propose a validation framework and use it to validate the LEM, the LVEM, and the NHM of liver tissue. The tissue's mechanical properties

are acquired by material testing using a Dynamic Materials Analyzer (DMA). A chamber consisting of a cylindrical plexiglass compartment, and a piece of liver are used for Micro-CT and CBCT scanning (details of this design can be found in [6]). These types of imaging can produce high resolution images. As a benefit, the deformation details in the contact region between the tissue and the probe can be captured and rendered even at small force loads ranging from several grams to tens of grams. Two techniques are used to validate these models. Firstly, we measure and compare the volume difference of the deformed object from simulations and the deformed object from experiments, as well as the vertical displacements of the probe from the experiments and the simulations. Secondly, the displacements of each vertex of the F.E. mesh are acquired and their root mean squares (RMS) for the three models, are computed and compared.

2 Finite Element Models of Soft Tissues

In this section, we will simply introduce the general framework of the LEM, the LVEM, and the NHM.

2.1 Linear Elastic Model

For the LEM, the material is assumed homogeneous and isotropic: the stress σ and strain ϵ are directly proportional to each other via the Hooke's law $\sigma = \mathbf{D}\epsilon$. The matrix D describes the material characteristics and depends on the Young's modulus and the Poisson's ratio (e.g., [7,8]). The relation between the stress and the displacement \mathbf{u} can be written as $\epsilon = \mathbf{B}\mathbf{u}$. Linear tetrahedral elements are used in our FE simulation. We denote by \mathbf{u}_i^e the four nodal displacements of such elements. These displacements are used to compute the displacement inside each element as a linear combination of the shape functions $N_i^e(\cdot)$. Figure 1 shows the domain subdivision into finite elements and a linear tetrahedral element. For

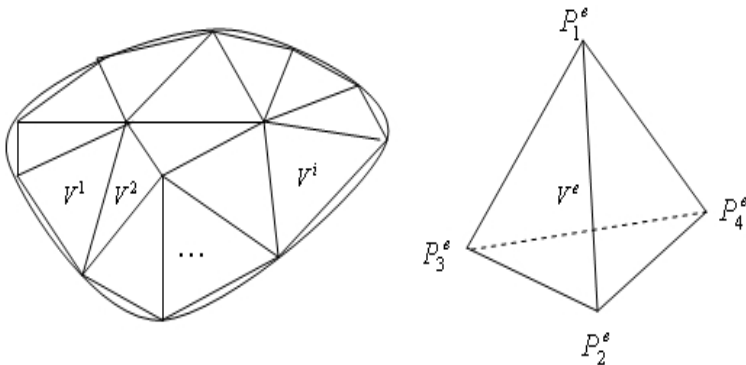


Fig. 1. Left: Subdivision of the domain into finite elements. Right: linear tetrahedral finite element. V^e : element; P_i^e 's: nodes.

each element V^e , the potential energy in terms of the total element compound displacement vector $\mathbf{u}^e = [u_1^{eT}, u_2^{eT}, u_3^{eT}, u_4^{eT}]$ is:

$$E^e(\mathbf{u}^e) = \frac{1}{2} \int_{V^e} (\mathbf{u}^e)^T \mathbf{B}^{eT} \mathbf{D} \mathbf{B}^e \mathbf{u}^e dV^e - \int_{V^e} \mathbf{f}^{eT} \mathbf{u}^e dV^e,$$

where \mathbf{f}^e is the external force acting on the element V^e , and \mathbf{B}^e is a constant matrix depending on the coefficients of the polynomial shape functions N_i^e ($i = 1, 2, 3, 4$) [9]. The global equilibrium equation, obtained by minimizing the total potential energy ($\sum_e E^e$), can be derived as a linear equation: $KU = F$. The dimension of the global stiffness matrix K is $3m \times 3m$ and the dimension of the global force vector F is $3m$, where m is the total number of nodes.

2.2 Linear Viscoelastic Model

A viscoelastic material shows a combination of viscous and elastic effects. The response of a viscoelastic material depends on both the current stress and the stress history up to the current time. Some soft tissues, such as liver, exhibit the viscoelastic property. One simple material model is the direct sum of the linear elastic effect and the linear viscous effect: $\epsilon(t) = \epsilon_e(t) + \epsilon_v(t)$ where $\epsilon_e(t)$ and $\epsilon_v(t)$ represent the strains in the elastic deformation and viscous flow respectively. Note that $\sigma(t) = E\epsilon_e(t)$, and $\sigma(t) = \eta\dot{\epsilon}_v(t)$ where $\eta > 0$ is the Newtonian viscosity and E is the Young’s modulus. One can derive the constitutive equation of this model as follows (e.g., [7, 10]):

$$\dot{\epsilon}(t) = \frac{\dot{\sigma}(t)}{E} + \frac{\sigma(t)}{\eta}.$$

In our simulation, the F.E. software Abaqus [11] will use the stress relaxation test data for the viscoelastic modeling.

2.3 Hyperelastic Model

Hyperelastic materials are described in terms of a strain energy potential. There are several forms of strain energy potentials available in Abaqus FEM software [11] to model hyperelastic materials. One of them is the neo-Hooke’s modeling. The material is assumed isotropic, and the energy potential function W is approximated as follows:

$$W = C_{10}(\bar{I}_1 - 3) + \frac{1}{D_1}(J^{el} - 1)^2,$$

where C_{10} and D_1 are material parameters; \bar{I}_1 is the first deviatoric strain invariant defined as $\bar{I}_1 = \bar{\lambda}_1^2 + \bar{\lambda}_2^2 + \bar{\lambda}_3^2$, where the deviatoric stretches $\bar{\lambda}_i = J^{-1/3}\lambda_i$; J is the total volume ratio; J^{el} is the elastic volume ratio; and λ_i are the principal stretches. The initial shear modulus and bulk modulus are given by $\mu_0 = 2C_{10}$ and $K_0 = 2/D_1$. More details about the NHM and other hyperelastic models can be found elsewhere (e.g., [9, 11]).

3 Preprocessing of Soft Tissues

To perform F.E. analysis of soft tissue deformation, preprocessing is needed: image segmentation, mesh generation, and estimation of the mechanical properties of the tissue.

Mesh generation techniques have been developed for different contexts and were aimed at different applications. In this work, we are interested in generating

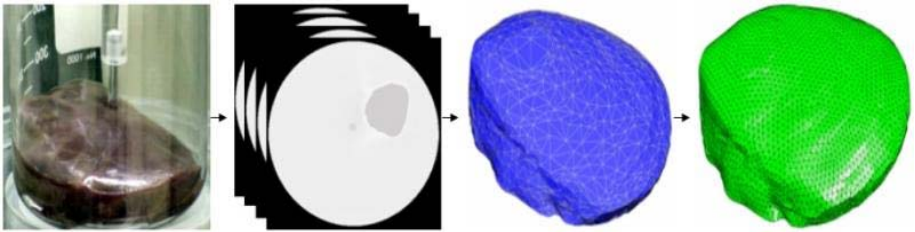


Fig. 2. From left to right, image acquisition, segmentation, stl conversion, and volume mesh generation

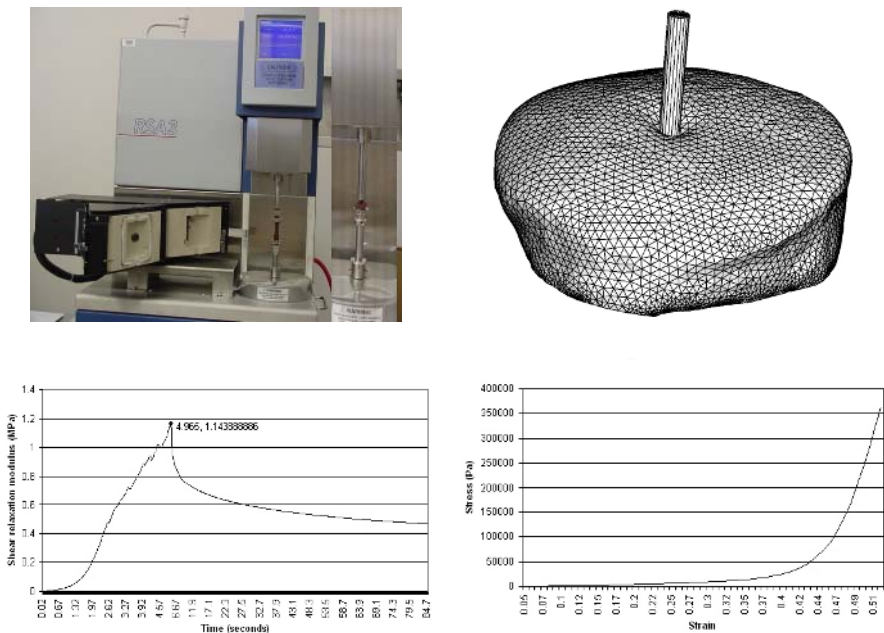


Fig. 3. Left to right: the DMA machine and deformed tissue with probe (top); the shear relaxation modulus curve; the stress vs. strain curve for strain less than 10.5% (zoomed) (bottom)

high quality tetrahedral meshes that conform to the input surface meshes. For this purpose, a level set based segmentation approach [12] is employed. Furthermore, the segmented images are converted to a stereolithographic file which is input into NetGen [13] to generate a 3D mesh of the liver. The NetGen uses the advancing front technique combined with Delaunay tessellation. Figure 2 shows our procedure to generate volume meshes from a stack of medical images.

A second step of the preprocessing phase consists in measuring the mechanical properties of the liver tissue. Generally, these properties can be measured either *in vivo* [14, 16] or *ex vivo* [17]. In this work, we cut two pieces of lamb liver, one with dimensions of $0.5 \times 2 \times 0.15$ *in* that we use for an extension test. The other piece, having a cylindrical shape with height of 0.5 *in*, is used for a compression test. These tests are performed using the DMA machine (Fig. 3:top-left). The shear relaxation test data (Fig. 3: bottom-left) are used by Abaqus software for the viscoelastic modeling. From the compression test, we got the global stress-strain curve (Fig. 3:bottom-right). The portion of the stress-strain curve for the strain less than 10.5% is used to compute the Young's modulus yielding $E = 11055 Pa$. The Poisson's ratio of 0.4 is used from literature [16,17]. For the NHM, the initial shear modulus $\mu_0 = 3948 Pa$ and bulk modulus $K_0 = 18425 Pa$ are computed from the stress-strain curve for strain less than 10.5%.

4 Proposed Finite Element Model Validation System

In this section, we present a validation standard of F.E. models to verify the LEM, and the LVEM, and the NHM.

Experimental Setup: We developed a general purpose device (“testing chamber”) that can be used to induce measurable deformations, which can be captured by two imaging sources: high resolution Micro-CT or medium resolution CBCT. The testing chamber is formed of a cylindrical plexiglass compartment. The tissue is placed on the bottom of the compartment and an aluminium probe with flat tip touches the upper surface of the tissue. Forces are applied by adding weights to a lever causing the probe to deform the tissue. The proposed soft tissue validation procedure can be summarized as follows. We first scan a piece of soft tissue using Micro-CT or CBCT machine to collect a stack of images of the undeformed tissue. Four force loads (5, 10, 20 and 40 grams) are gradually applied to the lever of the chamber. For each force load, we scan the deformed tissue when the equilibrium is reached. The collected images, both for the undeformed tissue and the deformed tissues, are then segmented and volume meshes are generated. We select a F.E. model (LEM, LVEM, or NHM) to perform F.E. analysis on the meshed undeformed object with the same force loads and same boundary conditions as those for the object in the experiment (see Fig. 3:top-right for deformed object). Finally, we superimpose the deformed objects from simulation and the corresponding ones in the experiment, and we compute their volume difference and measure the vertical displacements of the probe tip in the experiment and the simulation for comparison purpose.

5 Validation Results

To compute the deformation volume difference between the deformed object from simulation and the one from experiment, and its percentage with respect to the real deformation volume, we use a cylinder with radius of 20 *mm* and height of 20 *mm* to cover the deformed region. The cylinder is positioned vertically and centered at the probe contact and the top surface of the cylinder takes the shape of the undeformed liver surface. The deformed volume is computed by the number of voxels between the deformed surface and the top surface of the cylinder. The vertical displacements of the probe tip in real deformation and simulation are estimated from the acquired images and from the simulated object. For the LEM, the LVEM, and the NHM, all related volume differences, volume difference ratios, and the vertical displacements of the probe tip are listed in Table 1, where the abbreviations RD, LED, LVED and NHD stand for the vertical displacements of the probe tip in real deformation, and in simulation using the LEM, the LVEM and the NHM, respectively. RDV, LEVD, LVEVD, NHVD, LEVDP, LVEVDP, and NHVDP stand for the real deformation volume, the volume differences using the LEM, the LVM, and the NHM and their volume difference percentages, respectively. Figure 4 shows the superimposing of the deformed liver from simulation and the real deformed liver from experiments. From Table 1 and Fig. 4, we see that for small force loads (< 40*g*), the LEM is more applicable to the modeling of the lamb liver tissue. One can conclude that the liver deformation using the LEM is gradually going deep with the increasing force load. At the force load of 40 grams, the liver deformation using the LEM is larger than the real deformation. The liver deformations using the LVEM and the NHM are gradually approaching the real deformation with the increasing force load but the later is approaching faster. The root mean square values of the nodal displacements for each F.E. model are computed for different force loads (Table 2). These results are consistent with the results in Table 1. The strain

Table 1. Vertical displacements of the probe tip, volume differences and their percentages with respect to the real deformed volumes

	RD	LED	LVED	NHD	RDV	LEVD	LEVDP	LVEVD	LVEVDP	NHVD	NHVDP
	<i>mm</i>	<i>mm</i>	<i>mm</i>	<i>mm</i>	<i>mm</i> ³	<i>mm</i> ³	%	<i>mm</i> ³	%	<i>mm</i> ³	%
5 g	1.8	0.77	0.38	0.68	956	191	19.97	228	23.85	220	23.01
10 g	2.6	1.54	1.18	1.32	1075	240	22.33	316	29.40	302	28.01
20 g	4.2	4.39	2.36	2.57	1241	245	19.74	382	30.78	359	28.93
40 g	6.2	9.59	4.71	5.04	1527	299	19.58	506	33.14	476	31.17

Table 2. RMS values (mm) for each F.E. model for different force loads

	5 g	10 g	20 g	40 g
LEM	0.0487956	0.097591	0.195182	0.394411
LVEM	0.025235	0.050470	0.100939	0.201876
NHM	0.030717	0.061107	0.121318	0.239742

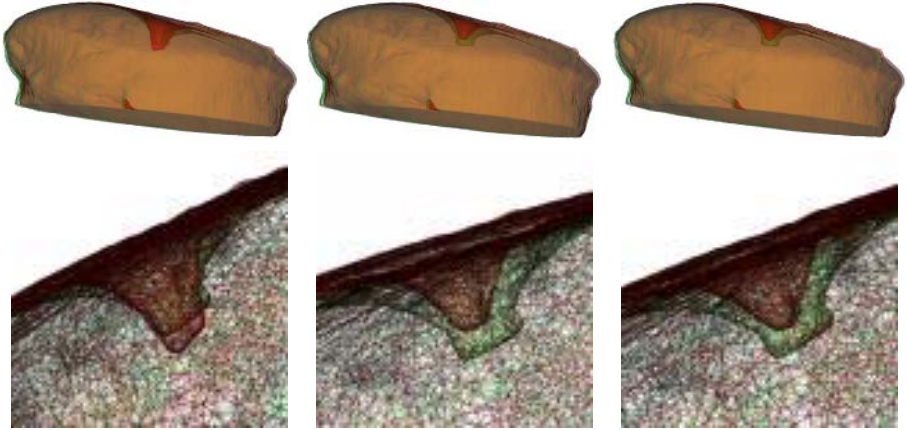


Fig. 4. Left to right: superimposing of the simulated livers and the real deformed ones under the force load 40 grams for the LEM, the LVEM, and the NHM, respectively. Wireframed zooms around the contact region are shown in the second row.

repartition plots from Abaqus for the three F.E. models and for force loads up to 40 grams, show that the maximal strains do not exceed 10%. This confirms that all simulations are carried out in the small deformation framework. The viscous component only governs the time dependent behavior of the tissue and the final displacement for any given load will only depend on the elastic property. The observed difference in displacements for the LEM and the LVEM is related to the fact that the LVEM accounted for time behavior or the reported displacements were not at the time of infinity, the elastic components of the two models might be different. Also, it should be pointed out that there is a geometric nonlinearity in the experiment and the simulation. As larger force is applied to the probe, it pushes into more tissue. Consequently, there is more surface area in contact with the probe to provide support. For the NHM, the geometric nonlinearity is considered in simulation.

6 Conclusion

We presented F.E. modeling of soft tissues from Micro-CT or CBCT images. These types of images show the volume and surface shape changes for small force loads. The F.E. method was used to analyze the lamb liver deformations under certain force loads using the LEM, the LVEM, and the NHM. The deformations were measured from the models and compared to the real deformations measured by the experiment setup. We found that the LEM is more applicable to lamb liver than the LVEM for small force loads ($< 40g$) and that the NHM is closer to reality than the LVEM for this same range of force loads. The comparability of the results presented herein to the *in vivo* tissue, which is valuable

for future surgery simulation design, is beyond the scope of this work and will be investigated in future studies.

References

1. Shi, P., Robinson, G., Constable, R.T., Sinusas, A., Duncan, J.: A model-based integrated approach to track myocardial deformation using displacement and velocity constraints. In: Fifth International Conference on Computer Vision. (1995) 687–692
2. Bro-Nielsen, M.: Finite element modeling in surgery simulation. *Proceedings of the IEEE* **83** (1998) 490–503
3. Delingette, H., Ayache, N.: Soft tissue modeling for surgery simulation. In Ayache, N., ed.: *Computational Models for the Human Body. Handbook of Numerical Analysis* (Ed : Ph. Ciarlet). Elsevier (2004) 453–550
4. Skrinjar, O.M., Spencer, D., Duncan, J.S.: Brain shift modeling for use in neurosurgery. In: MICCAI '98: Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention, London, UK, Springer-Verlag (1998) 641–649
5. Larrabee, W.: A finite element model of skin deformation: biomechanics of skin and tissue: a review. *Laryngoscope* **96** (1986) 399–405
6. Farag, A., Shi, H., Fahmi R., Voor, M.: Basic experiments and finite element analysis of soft tissues. In Payan, Y., ed.: *Biomechanics Applied to Computer Assisted Surgery. Research Signpost* (2005)
7. Fung, Y.C.: *Biomechanics: Mechanical Properties of Living Tissues*. Springer-Verlag, New York (1981)
8. Reddy, J.N.: *Energy and Variational Methods in Applied Mechanics*. John & Sons, Inc (1984)
9. Zienkiewicz, O.C., Taylor, R.L.: *The Finite Element Method*. Fifth edn. Butterworth-Heinemann (2000)
10. Creus, G.J.: Viscoelasticity - basic theory and applications to concrete structures. In Brebbia, C.A., Orszag, S.A., eds.: *Lecture Notes in Engineering*. Volume 17. Springer-Verlag (1986)
11. Abaqus: *Abaqus User's Manuals, Abaqus Version 6.4*. HKS Inc. (2003)
12. Farag, A.A., Hassan, H.S.: Adaptive segmentation of multi-modal 3d data using robust level set techniques. In: MICCAI (1). (2004) 143–150
13. Schöberl, J.: Netgen - an advancing front 2d/3d-mesh generator based on abstract rules. *Comput. Visual. Sci.* (1) 41–52
14. Carter, F.J., Frank, T.G., Davies, P.J., McLean, D., Cuschieri, A.: Measurements and modeling of the compliance of human and porcine organs. *Medical Image Analysis* **5** (2001) 231–236
15. Brouwer, I., Ustin, J., Bentley, L., Sherman, A., Dhruv, N., Tendick, F.: Measuring in vivo animal soft tissue properties for haptic modeling in surgical simulation. In: *Medicine Meets Virtual Reality 2001*. (2001) 69–74
16. Ottensmeyer, M.P., Kerdok, A.E., Howe, R.D., Dawson, S.: The effects of testing environment on the viscoelastic properties of soft tissues. In: ISMS. (2004) 9–18
17. Hu, T., Desai, J.P.: A biomechanical model of the liver for reality-based haptic feedback. In: MICCAI (1). (2003) 75–82

A Complete Augmented Reality Guidance System for Liver Punctures: First Clinical Evaluation

S.A. Nicolau², X. Pennec¹, L. Soler², and N. Ayache¹

¹ INRIA Sophia, Epidaure, 2004 Rte des Lucioles, F-06902 Sophia-Antipolis, Cedex
{Stephane.Nicolau, Xavier.Pennec, Nicholas.Ayache}@sophia.inria.fr

² IRCAD-Hopital Civil, Virtual-surg, 1 Place de l'Hopital, 67091, Strasbourg, Cedex
{stephane.nicolau, luc.soler}@ircad.u-strasbg.fr

Abstract. We provided in [14] an augmented reality guidance system for liver punctures, which has been validated on a static abdominal phantom [16]. In this paper, we report the first in vivo experiments.

We developed a strictly passive protocol to directly evaluate our system on patients. We show that the system algorithms work efficiently and we highlight the clinical constraints that we had to overcome (small operative field, weight and sterility of the tracked marker attached to the needle...). Finally, we investigate to what extent breathing motion can be neglected for free breathing patient. Results show that the guiding accuracy, close to 1 cm, is sufficient for large targets only (above 3 cm of diameter) when the breathing motion is neglected. In the near future, we aim at validating our system on smaller targets using a respiratory gating technique.

1 Introduction

The treatment of liver tumors by Radio-Frequency (RF) is a new technique which begins to be widely used [11]. The guidance procedure to reach the tumors with the electrode is generally made visually with per-operative 2D cross-sections of the patient using either Ultra-Sound (US), Computed Tomography (CT) or Magnetic Resonance Images (MRI). Practitioners empirically evaluate their targeting accuracy around 3 mm. However, the insertion needs repetitive CT/MRI images for needle adjustment (if CT/MRI guided) and sometimes several reinsertion attempts. This lengthens the intervention duration, and increases post procedure complications and radiation exposure (when CT-guided). Moreover, the targeting accuracy depends on the practitioner's experience.

Some systems have already been designed for puncture application. However they are quite expensive, or unsuitable for liver punctures. Some of them need a robotic arm [4, 7] or a head mounted display associated to a cumbersome reference structure attached to the MRI table [18]. Others [6] guide the needle placement when it is contained in a single axial CT slice. Their system is suited to the majority of cases but not to liver punctures. Indeed, most of liver punctures

cannot be realized along an axial plane because of the critical structures that have to be avoided (like vessels). Finally, the C-arm systems presented in [2, 12] do not allow to visualize what structures are on the needle path, which prevents their use for liver punctures.

To help practitioners, we developed in [14, 16] a low cost augmented reality (AR) guidance system for CT/MRI assisted liver punctures: only one PC, two cameras and some radio-opaque markers are necessary. Despite breathing motion, we choose to use a rigid model since the patient is usually under general anesthesia and intubated to limit the pain (60% of CT-guided RF ablation are done under general anesthesia at the Strasbourg Hospital). Thus it is possible to ask for a breath control that can provide a tumor repositioning error about 1 mm (see [20, 1, 19]). This system was fully validated on a *static* abdominal phantom [14, 16]: several surgeons and engineers were able to reach targets with an accuracy of 2 mm and a duration under 40 sec.

In this paper, we report the first introduction of our system in the operating room (OR) during interventions on patients. After a presentation of our system, we explain how we evaluate its accuracy on patient with a safe protocol that we developed. Then, we present encouraging results which show that a full validation is possible with few additional constraints in the OR.

2 System Components

In our setup, two jointly calibrated cameras are viewing the skin of the patient who is lying on the CT-table. A 3D model of the patient (including his skin, liver and tumors) is automatically obtained from the first CT scan [17], and is rigidly registered in the camera frame thanks to radio-opaque markers previously stuck on the patient's skin. The needle being tracked by the cameras, we display on a screen its relative position with respect to the patient model.

In this section, we summarize the basic algorithmic principles of our system. Firstly, we explain how we automatically extract and match the radio-opaque markers (to avoid time loss). Then, we describe the 3D/2D criterion we use to register the 3D patient model in the camera frame, and show how we track the needle position in real-time. Finally, we present the guiding interface designed with surgeons.

2.1 Automated Data Processing

The principle of the marker localization in the video images is based on a HSV color analysis, followed by a component size and shape thresholding. The markers in the CT-image are extracted by a top-hat characterization that emphasizes small singularities on the skin surface.

The matching between the video markers is realized with epipolar geometry, and the correspondences between video and CT markers are carried out by a prediction/verification algorithm. A validation study [14] showed that these algorithms are robust and that the overall computation time of the extraction and

matching process is below 120 sec (CT marker extraction - done once - takes 118 sec. on average, video marker tracking and registration are refreshed at 10 Hz).

2.2 Registration of the Virtual Model in the Camera's Frame

We choose a 3D/2D point registration approach to provide the rigid transformation that relates scanner frame and camera frame. The classical choice is to optimize the SPPC criterion (see [15]):

$$SPPC(T) = \sum_{k=1}^S \sum_{i=1}^N \xi_i^k \cdot \frac{\|\tilde{m}_i^{(k)} - P^{(k)}(T \star \tilde{M}_i)\|^2}{2 \cdot \sigma_{2D}^2}$$

where S (resp. N) is the number of cameras (resp. markers), $\tilde{m}_i^{(k)}$ is the observed 2D coordinates of the i^{th} markers in the k^{th} video image, \tilde{M}_i is the observed 3D coordinates of the i^{th} markers in the CT-image, $P^{(k)}$ the projective function, ξ_i^k is a binary variable equal to 1 if the i^{th} marker is visible in the k^{th} video image and 0 if not, and T the sought transformation. However, this criterion considers that noise only corrupts the 2D data and that 3D data are exact. In our context, this assumption is erroneous as the marker extraction from the CT-image is corrupted by noise as well.

A more realistic statistical hypothesis is that we are measuring noisy versions \tilde{M}_i of the unknown exact 3D points M_i (more details are given in [15]). A ML estimation of the transformation T and the *auxiliary variables* M_i leads to minimize the *Extended Projective Points Criterion* (EPPC):

$$EPPC(T, M_1, \dots, M_N) = \sum_{i=1}^N \frac{\|\tilde{M}_i - M_i\|^2}{2 \cdot \sigma_{3D}^2} + \sum_{k=1}^S \sum_{i=1}^N \xi_i^k \cdot \frac{\|\tilde{m}_i^{(k)} - m_i^{(k)}\|^2}{2 \cdot \sigma_{2D}^2}$$

2.3 Needle Tracking

To track the needle location and orientation in the camera frame, we attach an oriented square marker whose corners are automatically localized on video images in real-time using an adapted version of the ARTkit library [9]. Then, knowing the size of the square, we are able to localize it in the camera reference frame by minimizing the classical 3D/2D SPPC criterion. Calibrating the relative needle position with respect to the square marker with the pivot method [10], we are finally able to superimpose the virtual model on the real one on video images. An accuracy evaluation realized in simulated clinical conditions (cameras 1.5 m away from the needle with a test volume of $40 \times 30 \times 25$ cm³), showed that the average superimposition error of a radiofrequency needle was 0.95 mm.

2.4 A Secured and Ergonomic Guidance Interface

Our interface has been optimized with surgeons, in order to provide them a clear and intuitive tool. It is divided into three screens (see Fig. 1) described below.

The bottom left screen displays one of the two video images returned by our cameras, on which can be superimposed the virtual needle and the 3D patient model. The user can check visually the registration quality by superimposing virtual elements (like skin and radio-opaque markers). In the right screen, the 3D virtual scene, composed of the 3D patient model and the needle representation, is rendered from a viewpoint controlled by the user. Moreover, it is possible to display the CT-scan from which the reconstruction is made, and navigate through its slices. In the top left screen, we display a view that corresponds to what would see a camera positioned on the needle tip and oriented along its axis. This view facilitates the orientation of the needle toward the point to be reached before its insertion.

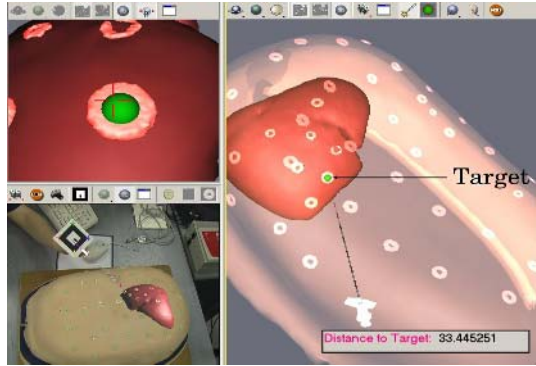


Fig. 1. Illustration of the three screen guidance interface

3 Clinical Experiments on Patients

We have previously evaluated the whole system on an abdominal phantom [16], and we have shown that, with simulated clinical conditions, our system accuracy and ergonomics are fully validated. However, in an OR, we will undergo additional constraints (lack of room, luminosity variation, validity range of rigidity assumption) that may hamper marker extraction, needle tracking and 3D model registration accuracy.

To check the efficiency of a system in real clinical conditions, the standard procedure is to make several tests on animals (like Vogt [18] or Mourgues [13]). However, it does take a long time to get animal study authorizations (at least 6 months in Strasbourg). Moreover, animals are not optimal models for evaluation: their anatomy size and shape, skin color, vessel branching structure and tissue mechanical properties are different from human ones. Thus, we propose to introduce our system in the OR directly on human patients, thanks to a strictly passive setup: the intervention is realized with the setup needed by our system, yet with the standard CT-guided protocol. In these conditions, no ground truth is available to evaluate the accuracy of our system (invasive techniques are needed, which is not clinically acceptable). Thus, we designed a protocol that allows to assess the system accuracy w. r. t. the practitioner accuracy, without any risk for the patient.

In the sequel, we firstly describe the evaluation protocol we developed. Then we present the first experiments we realized on patients. Let us highlight that these experiments are not aimed at validating the system but just at evaluating

it to determine the problems that need to be overcome to introduce it in the OR. The standard protocol was not modified, which means that patients breathed freely during the interventions. Then, it allowed us to see to what extent breathing motion could be neglected or not.

3.1 A Safe Validation Protocol with Patients

The purpose of the method is to measure the distance between the real tumor and the virtual tumor registered by our system. Since no ground truth is available (without modifying the intervention protocol with an invasive procedure) for the real tumor position, we consider the final position of the needle as a reference. Then, we measure, at the end of the needle insertion, the distance between the needle tip and the target registered by our system. This target is a point C_{scan} within the tumor that the practitioner defines (in the first acquisition CT_0) as the point he is targeting. Obviously, our reference depends on the practitioner’s accuracy. Radiologists and surgeons evaluate their targeting accuracy to 3 mm. Therefore, to show the equivalence between our system and the practitioner, we have to obtain an average distance close to 3 mm (assuming that both practitioner targeting and system registration errors are unbiased).

Figure 2 summarizes the different steps of our protocol and how we get a measure. To be in our system setup, a square marker is attached to the needle, radio-opaque markers are stuck on the patient skin, and we set our two cameras in the OR. The practitioner performs the intervention with his radiological protocol (repetitive CT acquisitions) without any advice nor instruction from our system. Note that the radiation exposure is not negligible since 40 sets of 4 slices are at least taken for each tumor. When the needle insertion is completed, the practitioner checks visually (using a control scan) that the needle tip was sufficiently close to the initial target. This is done to avoid a bias. Then, we make a stereoscopic video acquisition of the patient abdomen (with the radio-opaque

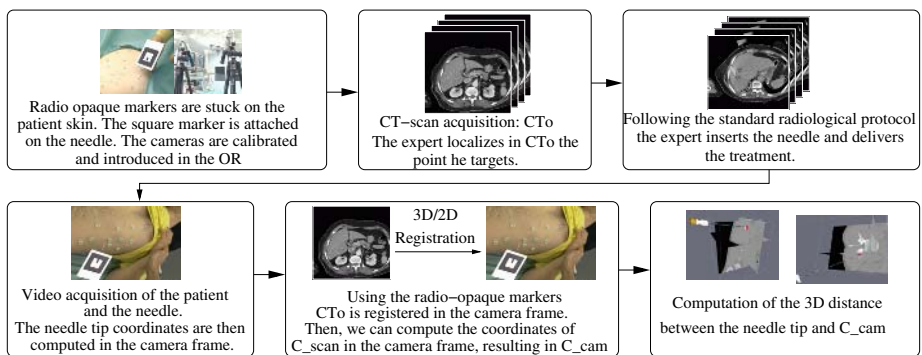


Fig. 2. Description of the validation protocol without any risk for the patient. It allows to obtain an assessment of the system guidance error with respect to the targeting error of the practitioner.

markers) and of the square marker attached to the needle. From this acquisition, we estimate the needle tip position P_{cam} and register the target C_{scan} (originally defined in CT_0) in the camera frame. The registered target C_{cam} is the point toward which we would have guided the practitioner if he had used our system. Finally, our accuracy measure is given by the distance $d(P_{cam}, C_{cam})$.

3.2 Results

We have introduced our system three times in the OR. All the data processing algorithms worked efficiently (marker extraction and matching, needle tracking, 3D/2D registration). Figure 3 shows several augmented images obtained during the interventions. Quantitative accuracy results are shown in table 1.

The measured average accuracy is 9.5 mm, which is two to three times larger than the value needed to show the equivalence between our system and a practitioner. We think that most of the error can be attributed to the free breathing of the patient during the intervention. Studies showed that the liver moves over 1 cm with respect to the skin with free breathing [5]. This means that the 3D

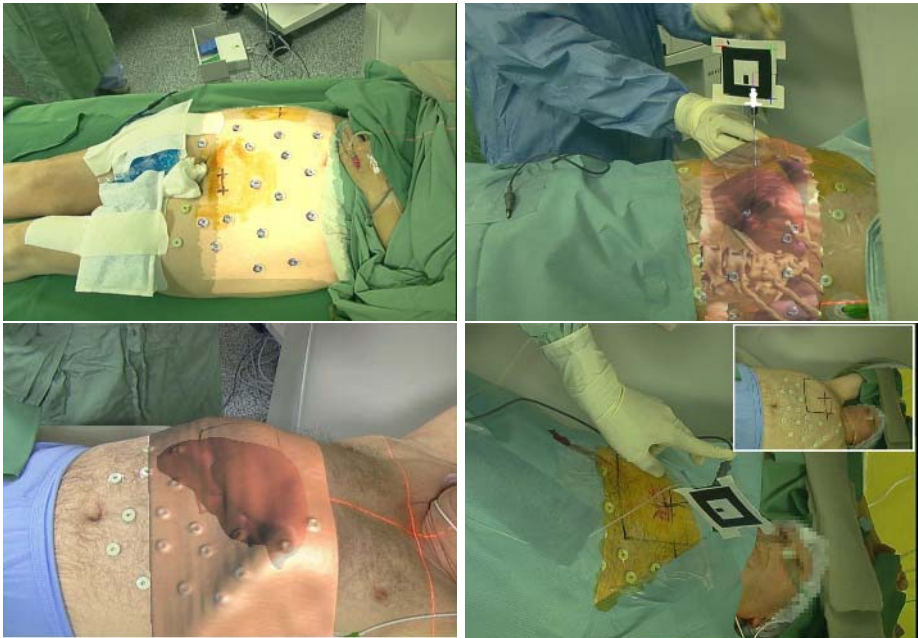


Fig. 3. Top left: visual check of the 3D model registration quality. The radio-opaque markers of the 3D reconstructed model are well superimposed on these visible in the video images. Top right: augmented view of the patient at the end of the needle insertion. Bottom left: view in transparency of the patient. The edge of the patient skin seemed perfectly superimposed with the 3D reconstruction edge. Bottom right: for this patient, only 7 markers were detectable among the 18 initially stuck.

Table 1. Results of the system evaluation on three patients

	Tumor type	General anesthesia	Gas volume monitoring	Number of markers used/stuck	Accuracy R-S (mm)
Patient 1	Liver	no	no	7 / 18	12.0
Patient 2	Bone	yes	no	16 / 19	6.7
Patient 3 : tumor 1	Liver	no	no	9 / 15	8.0
Patient 3 : tumor 2	Liver	no	no	9 / 15	11.3
Average	-	-	-	10.6 / 17.3	9.5

model reconstructed from the initial CT-scan was not close enough to reality to obtain an accurate rigid registration.

The second error cause is the low number of radio-opaque markers used for the transformation computation. Indeed, although we stuck around 17 markers on the patient skin, only 10 on average were imaged in the initial CT-scan CT_0 (one can see on Fig. 3 that the 3D model does not contain all the radio-opaque markers). Moreover, on the first patient, the operative field was not large enough (for sterility reasons) and reduced the number of visible skin markers. This issue was solved afterward with the use of sterile transparent stickers that were stuck on the patient skin (see top right Fig. 3). Another explanation of the error is the relative position of the target with respect to the radio-opaque marker used for the registration. Since many unused markers were the closest ones to the target, the theoretical inaccuracy increased.

Note that we obtained our best accuracy result (6.7 mm) for the bone radio-frequency intervention. For this patient, the target was far away from the upper abdomen in a more static zone (see top left Fig. 3). Therefore, we are convinced that we can provide better results if a respiratory gating technique is used.

4 Conclusion

In previous papers [14, 16], we developed and validated on phantom an AR system to guide liver percutaneous punctures in interventional radiology. This video based system tracks in real time the puncture needle and automatically registers (at a 10 Hz rate) the 3D patient model in the camera frame using radio-opaque markers. To show the robustness of the system in the OR (needle tracking, marker detection), we proposed to introduce it passively during interventions on patients. Since a passive approach limits the possibility of accuracy evaluation, we designed a new protocol that allows to assess the system accuracy without any risk for the patient. To our knowledge, it is the first time that an evaluation of AR-guided liver punctures is experimented on patients.

The three experiments we realized were an essential step to refine the constraints for using our system in a clinical setup. Firstly, we needed to build a square marker in sterile plastic for the needle tracking. On the one hand, it had to be sufficiently light to avoid curving the needle once inserted. On the other hand, its size had to be small to avoid touching the CT-gantry during the control

scans (currently it is too large for an MRI gantry). Secondly, we have experimented problems with the positioning and visibility of the radio-opaque markers. They have to be stuck carefully in the liver range to be imaged in the first CT-scan. Moreover, they are not visible in the video images if the practitioner does not enlarge the operative field (this could be realized with transparent sterile stickers). Finally, markers being non sterile, we could not stick them close to the entry point, inducing a loss of registration accuracy. This has to be compensated with a greater number of markers.

From a technical point of view, we showed that our system worked well in clinical conditions. However, we observed a positioning error which is larger than the practitioner's one, due to the breathing motion. We can take this motion into account either by modifying the clinical setup, i.e. by requiring a respiratory gating technique, or by modifying our system, for instance using a non-rigid model [3, 8]. Although the current accuracy of the final positioning is sufficient only for large targets (diameter above 3 cm), practitioners estimate that the system provides relevant informations for orienting the needle at the beginning of the insertion. This initial guidance could avoid several insertion attempts.

In the next step, we will perform an evaluation with intubated patients under breathing control. Practitioner movements will be synchronized on the respiratory cycle point corresponding to the first CT acquisition. These experiments should show that the system can be used in pseudo static conditions.

We are currently adapting this work for laparoscopic application. We track the endoscopic camera and surgical tools to provide surgeons with AR view in the endoscope coupled with the relative position of tools. We hope to test that new system on static organs linked to the spine in less than six months. Eventually, since it would be more comfortable for patients and less complicated for practitioners to avoid respiratory gating techniques, we intend to use a non-rigid registration coupled with real time US images.

References

1. J.M. Balter et al, K.L. Lam, C.J. McGinn, T.S. Lawrence, and R.K. Ten Haken. Improvement of CT-based treatment-planning models of abdominals targets using static exhale imaging. *Int. J. Radiation Oncology Biol. Phys.*, 41(4):939–943, 1998.
2. B. Bascle, N. Navab, M. Loser, B. Geiger, and R Taylor. Needle placement under x-ray fluoroscopy using perspective invariants. *In MMBIA '00*, pages 46–53, 2000.
3. J.M. Blackall et al. Tracking alignment of sparse ultrasound with preoperative images of the liver and an interventional plan using models of respiratory motion and deformation. *In Proceedings SPIE Medical Imaging*, 2004.
4. EM Doctor et al. Robotically assisted intraoperative ultrasound with application to ablative therapy of liver cancer. *SPIE Med. Imaging 03*, vol. 5029, pp. 281–291.
5. M. Clifford et al. Assessment of hepatic motion secondary to respiration for computer assisted interventions. *Computer Aided Surgery*, 7:291–299, 2002.
6. G. Fichtinger et al. Needle insertion in CT scanner with image overlay - cadaver studies. *In MICCAI'04*, LNCS 3217, pages 795–803. Springer-Verlag, 2004.
7. G. Fichtinger et al. System for robotically assisted prostate biopsy and therapy with intraoperative CT guidance. *Journal of Academic Radiology*, 9(1):60–74, 2002.

8. D.J. Hawkes et al. Tissue deformation and shape models in image-guided interventions: a discussion paper. *Medical Image Analysis*, 9(2):163–175, 2004.
9. Hiro. Human interface technology lab., <http://www.hitl.washington.edu/artoolkit/>.
10. S. Lavalle, P. Cinquin, and J. Troccaz. *Computer Integrated Surgery and Therapy: State of the Art*, chapter 10, pages 239–310. IS Press, Amsterdam, NL, in C. Roux and J.L. Coatrieux edition, 1997.
11. J.F. McGahan and G.D. Dodd III. Radiofrequency ablation of the liver: Current status. *American Journal of Roentgenology*, 176(1):3–16, 2001.
12. M Mitschke, A. Bani-Hashemi, and N. Navab. Interventions under video-augmented x-ray guidance: Application to needle placement. In *MICCAI'00*, pages 858–868, Pittsburgh, Pennsylvania USA, October 11-14 2000.
13. Fabien Mourgues, Thierry Vieville, Volkmar Falk, and Ève Coste-Manire. Interactive guidance by image overlay in robot assisted coronary artery bypass. In *MICCAI'03*, LNCS 2878, pp. 173–181, November 2003.
14. S. Nicolau, A. Garcia, X. Pennec, L. Soler, and N. Ayache. An augmented reality system to guide radio-frequency tumor ablation. In *the Journal of Computer Animation and Virtual World*, 16(1):1–10, 2005.
15. S. Nicolau, X. Pennec, L. Soler, and N. Ayache. An accuracy certified augmented reality system for therapy guidance. In *ECCV'04*, LNCS 3023, pages 79–91. 2004.
16. S. Nicolau, J. Schmid, X. Pennec, L. Soler, and N. Ayache. An augmented reality & virtuality interface for a puncture guidance system: Design and validation on an abdominal phantom. In *MIAR 2004*, LNCS 3150, pp. 302–310, 2004.
17. L. Soler, H. Delingette, G. Malandain, N. Ayache, and J. Marescaux. Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery. *Computer Aided Surgery*, 6(3), August 2001.
18. S. Vogt et al. Augmented reality system for MRI-guided interventions: phantom studies and first animal test. In *SPIE Med. Imaging*, vol. 5367, pp 100-109, 2004.
19. R. Wagman, E. Yorke, E. Ford, P. Giraud, G. Mageras, B. Minsky, and K.. Rosenzweig. Respiratory gating for liver tumors: use in dose escalation. *Int. J. Radiation Oncology Biol. Phys.*, 55(3):659–668, 2003.
20. J. Wong, M. Sharpe, D. Jaffray, V. Kini, J. Robertson, J. Stromberg, and A. Martinez. The use of active breathing control (abc) to reduce margin for breathing motion. *Int. J. Radiation Oncology Biol. Phys.*, 44(4):911–919, 1999.

A Novel Approach to High Resolution Fetal Brain MR Imaging

F. Rousseau¹, O. Glenn¹, B. Iordanova^{1,2}, C. Rodriguez-Carranza^{1,2},
D. Vigneron¹, J. Barkovich¹, and C. Studholme¹

¹ University of California, San Francisco
² NCIRE, San Francisco

Abstract. This paper describes a novel approach to forming high resolution MR images of the human fetal brain. It addresses the key problem of motion of the fetus by proposing a registration refined compounding of multiple sets of orthogonal fast 2D MRI slices, that are currently acquired for clinical studies, into a single high resolution MRI volume. A robust multi-resolution slice alignment is applied iteratively to the data to correct motion of the fetus that occurs between 2D acquisitions. This is combined with an intensity correction step and a super resolution reconstruction step, to form a single high isotropic resolution volume of the fetal brain. Experimental validation on synthetic image data with known motion types and underlying anatomy, together with retrospective application to sets of clinical acquisitions are included. Results indicate the method promises a unique route to acquiring high resolution MRI of the fetal brain in vivo allowing comparable quality to that of neonatal MRI. Such data is highly valuable in allowing a clinically applicable window into the process of normal and abnormal brain development.

1 Introduction

Ultrasound is the modality of choice for clinical evaluation of the developing fetus. However, clinical diagnosis with ultrasound is sometimes unclear and in these cases, alternative studies with magnetic resonance imaging (MRI) can be essential [6, 4]. One such critical application, which is the focus of this paper, is in the evaluation of isolated ventriculomegaly, which despite the absence of other clinical findings, can be associated with neurodevelopmental disabilities in childhood and infancy.

The development of ultrafast 2D acquisition sequences has led to significant improvements in the clinical utility of fetal MRI ([10]). However, the slice acquisition time is still very critical and has to be as short as possible to reduce the impact of fetal motion on the exam, since fetal MRI is performed without sedation. As a result, sets of thick 2D slices are generally acquired in clinical studies, with motion commonly occurring between slices. Overall, the resulting image data is limited in its geometric integrity between slices due to motion, and in its through plane spatial resolution. In addition, body coils used in the imaging processes do not provide a homogeneous sensitivity over the field of view and, because of motion during the acquisition, can produce different a distortion profile over time (see figure 2).

This work is motivated by the observation that current clinical imaging protocols make use of multiple orthogonal 2D multi-planar acquisitions to study the fetal brain. However, because of the motion between the slices, interpretation is limited to visual inspection by a trained radiologist, and does not allow direct quantitative measurements to be made on the 3D anatomy. The aim of this work is to develop and apply registration based reconstruction methods to the problem of correcting motion and intensity distortions between the different 2D acquisitions, and to then reconstruct a single higher resolution MRI image of the fetal brain.

As far as we know, this problem has not been discussed before in the literature. Moore *et al.* [8] built a high resolution dynamic heart model from isotropic coronal slices. The misalignment was corrected by registering a volume with sagittal and axial scout images. Image intensities were averaged together to generate a high resolution volume. Lötjönen *et al.* [7] have proposed an algorithm to correct translation-only movement artifacts simultaneously from short- and long-axis cardiac MR series. Locations of short-axis slices were optimized based on data from long-axis slices and vice versa.

In this article we report our work on this challenging problem and describe a fully automatic and accurate algorithm to correct slice misalignments, correct intensity distortions, and reconstruct a single high resolution image from sets of clinically typical low resolution fetal MRI data.

2 Method

The resolution of the low resolution source data is typically 1x1mm in plane with 3mm thick slices. Multiple sets consisting of between 30 and 40 slices each are acquired over a period of around 20 seconds for each set. The slices are commonly acquired in an interleaved pattern (usually alternating odd-even) and significant movement of a centimeter or more can occur during the acquisition of one set of slices. Sets of slices are typically acquired in approximately orthogonal axes (coronal, sagittal and axial), providing complementary resolution. It is reasonable (and is confirmed visually for most slices) to assume that motion within the time period of one slice is negligible. We are posed the task of recovering the local relative alignment of each slice in 3 dimensions, but can make use of the consistent geometry provided in the through plane direction of a slice that is provided by the other orthogonal sets of slice data, to constrain the collective alignment of the multiple low resolution images.

Notations. Let us consider n low resolution 3D images, denoted by I_{LR}^i , $i \in \{1...n\}$. A so-called low resolution 3D image is a stack of 2D thick slices. The reconstructed high resolution 3D image is denoted by I_{HR} .

2.1 Motion Correction

In order to correct fetal motion, we propose to use a registration based method. Registration methods are usually classified as being either sparse (landmark)

or dense (voxel) based. In the case of fetal brain MRI registration, because of the 3D non-consistency of the scans, finding homologous 3D landmarks is a very challenging task and have therefore approached the problem using a voxel based method. Due to the motion and contrast artifacts corrupting fetal MRI, it is highly desirable to use a similarity measure that is not based on strong assumptions between the two images. We use the normalized mutual information [11] which is: $NMI(X, Y) = \frac{h(X)+h(Y)}{h(X, Y)}$ where $h(X)$ is the marginal entropy of X and $h(X, Y)$ is the joint entropy. Motion correction is performed by aligning globally the low resolution images together and then by aligning every slice of the low resolution images to the reconstructed high resolution image.

Low Resolution Image Registration. The first step consists of approximately globally aligning the n low resolution images. One low resolution image is chosen arbitrarily as the reference image and is used to define a global coordinate system. The other low resolution images are rigidly registered to the reference image (6 degrees of freedom: 3 translations and 3 rotations). The 3D rigid transformation between an image I_{LR}^i and the reference image is denoted T_1^i . The high resolution volume coordinate system is chosen to match with the global coordinate system defined previously (see Figure 1).

Slice Registration. Once the low resolution images are roughly globally aligned, slice motion artifacts are corrected by rigidly registering a slice to the current reconstructed high resolution volume. We register the slices of low resolution image I_{LR}^i with the volume reconstructed using the $n - 1$ other low resolution images $I_{LR}^j, j \in \{1 \dots n, j \neq i\}$. A local coordinate system is defined for every slice and the transformation between a slice and the high resolution volume is composed by two rigid transformations: $\mathbf{y}_k^i = T_1^i T_{2_k}^i \mathbf{x}_k^i$ where $T_{2_k}^i$ denotes the 3D rigid transformation from the slice k of the low resolution image i to the low resolution image i coordinate system, \mathbf{x}_k^i is the slice pixel coordinates and \mathbf{y}_k^i the voxel coordinates in the compounded high resolution volume (see Figure 1).

Optimization. For both image and slice registration, a gradient ascent method is used to maximize the normalized mutual information. Concerning slice registration, we exploit the interleaved acquisition to propose a hierarchical approach

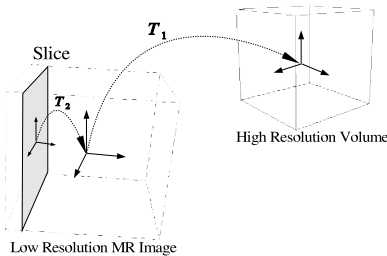


Fig. 1. Coordinate systems used during the slice and image registration

to provide a starting estimate to the optimization procedure. Slices of a low resolution image are ordered by acquisition time and are separated into two groups. The optimization consists in finding the optimal rigid transformations T_2 between the new groups of slices and the low resolution image coordinate system. The similarity measure is computed between the current reconstructed volume and the group of slices. This is performed for the n low resolution images. Once the convergence criterion is reached, each group is divided into two groups and the optimization procedure is then performed on these new groups of slices.

For the final phase of alignment, we maximize independently the normalized mutual information between each slice and the current reconstructed high resolution volume. This uses an iterative slice-by-slice scheme, looping over all slices until the sum of the slice-to-volume criteria (called the global convergence criterion) does not increase. For each 2D-3D registration step, a two-level multiresolution strategy is applied to avoid the pitfall of local optima.

2.2 Volume Reconstruction

As the volume reconstruction algorithm is required during registration process and contrast correction algorithm, the reconstruction has to be as fast as possible. To deal with the computation time constraint, a local neighborhood approach has been adopted. For each pixel \mathbf{x} in the slices of the original low resolution scans, the 3D coordinates \mathbf{y} of \mathbf{x} in the high resolution volume are calculated taking into account the results of the image and slice registration. The intensity value of \mathbf{x} is injected into the reconstructed volume using a kernel f centered at \mathbf{y} . To approximate the slice profile of the MRI data, we have used a Gaussian kernel ([9]).

2.3 Contrast Correction

We employ a contrast correction step to correct the local relative intensity distortion between the low resolution images to allow accurate compounding. Intensity changes from one 2D slice to another are both global and local. To address this issue, one low resolution image is used as a reference for tissue intensities and the other low resolution images are corrected to it. We consider the relationship between contrasts in the low resolution image to be corrected and the reference image as a spatially varying multiplicative field which is assumed to vary smoothly over the field of view. A direct approach to estimation is used where the two volumes are low pass filtered with a Gaussian kernel \mathcal{G} . The relative scaling in intensity $\beta_i(\mathbf{x})$ at location \mathbf{x} of the low resolution image $I_{LR}^i(\mathbf{x})$ is estimated for all slice pixels \mathbf{x} as follows:

$$\beta_i(\mathbf{x}) = a_i * \frac{\mathcal{G}(IR(\mathbf{x}))}{\mathcal{G}(I_{LR}^i(\mathbf{x}))} \quad \text{with} \quad a_i = \frac{\sum_{\mathbf{x}} I_{LR}^i(\mathbf{x})}{\sum_{\mathbf{x}} I_{LR}^i(\mathbf{x}) \frac{\mathcal{G}(IR(\mathbf{x}))}{\mathcal{G}(I_{LR}^i(\mathbf{x}))}} \quad (1)$$

where IR is the intensity reference image, I_{LR}^i is the low resolution image considered and a_i is used to keep the mean value of the image I_{LR}^i unchanged. IR is an arbitrary low resolution image projected in the reconstruction space. The corrected low resolution image \hat{I}_{LR}^i is then: $\hat{I}_{LR}^i(x) = \beta_i(x)I_{LR}^i(x)$.

2.4 Overview of the Algorithm

The proposed method consists in the following steps:

1. Apply global rigid registration between all the low resolution images.
2. Create initial slice registration estimate by hierarchical application of Slice Registration Algorithm
3. For each Resolution:
 - (a) Apply Slice Registration Algorithm
 - (b) Test Global Criteria Reached
4. Apply relative intensity distortion correction
5. Reconstruct the final high resolution volume

Where the slice registration algorithm is the following:

1. For each low resolution image I_{LR}^i :
 - (a) Reconstruct a high resolution volume using the low resolution images $I_{LR}^j, j \in \{1 \dots n, j \neq i\}$,
 - (b) For each slice in I_{LR}^i , perform the 2D-3D registration.

3 Material and Evaluation Procedure

3.1 Simulated Motions on Real MR Datasets

In order to perform the most realistic simulation possible, we used as a starting point a 3D MRI acquisition of a premature neonate of similar gestational age (26 weeks gestational age) to the fetal MRI studies, but with high image quality and no motion. The images consisted of a T1 weighted SPGR image with a spatial resolution of $0.7 \times 0.7 \times 2.1\text{mm}$. The data were interpolated to 1mm isotropic voxels.

Anisotropic subsampled volumes with 6 DOF motion artifact were simulated from the isotropic neonatal image. The resolution of these subsampled volumes is $1 \times 1 \times 3\text{mm}$. Motion artifacts were defined by the maximum magnitude of the head motion and by the head motion type. The first simulated motion was a sudden motion. It simulated a short fast movement of the head during a short portion of the acquisition. The second type of motion was a temporally smooth motion simulated using B-Splines. The displacements were chosen from a uniform distribution between $[-x, x]\text{mm}$ in each direction and between $[-x, x]\text{degrees}$ for each rotation, $x \in \{1, 5, 10\}$.

The accuracy was assessed by computing a registration error measured on a set of 4 points (bounding square) P_i distributed within every slices as follows: $RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n TRE_i^2}$, where TRE is the target registration error defined as follows: $TRE_i = \|P_i - \hat{T}^{-1}(T^*(P_i))\|^2$. T^* denotes the theoretical transformation and \hat{T} is the estimated geometric transformation.

3.2 Fetal Datasets

A typical clinical study of a fetus at our institute follows the following procedure: a quick localizer sequence is obtained in less than 30 seconds during maternal

free breathing in order to determine the location of the fetal head. No sedation or contrast agents were administered during the study. Conventional Single-shot Fast Spin-echo (SSFSE) T2-weighted images are then acquired during normal maternal breathing (a single slice is acquired in less than one second). Sets of slices were acquired in the axial, sagittal, and coronal planes with respect to the fetal brain, based on the initial localizer. All images were acquired in an interleaved manner to reduce cross-talk between slices. The in plane resolution was about 1mm (3mm slice thickness).

4 Results

From the simulated motion experiments on premature neonatal data, we evaluated the RMS registration error for four points at the corners of a box within the brain tissue of size $100mm \times 100mm$ for each slice. These are presented in Table 1. For all cases the final overall slice alignment error was significantly reduced by the alignment process. Overall, even with the presence of strong motion arte-

Table 1. Accuracy evaluation for the simulated data sets (RMS errors are expressed in *mm*)

Maximum Magnitude	Sudden Motion		Smooth Motion	
	RMS start.	RMS end	RMS start.	RMS end
± 1	0.94	0.63	0.72	0.66
± 5	4.79	0.78	2.55	0.73
± 10	9.71	1.22	3.88	0.66

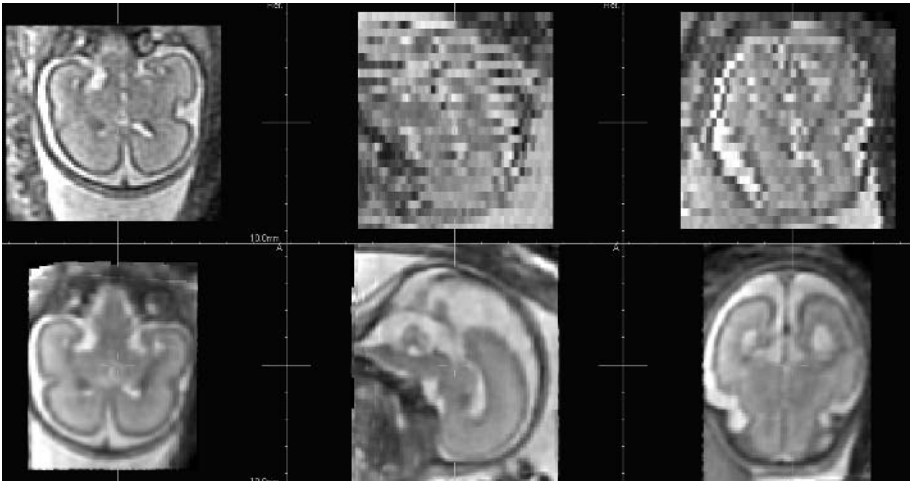


Fig. 2. Subject (a): acquired fetal MR coronal image data (top row) with a resolution of $1 \times 1 \times 3$ mm. Lower row shows the resulting reconstructed high resolution image.

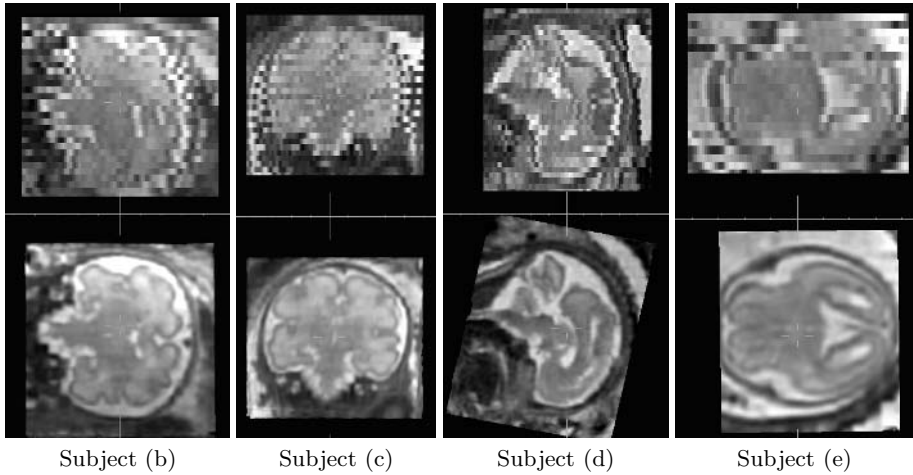


Fig. 3. Further examples of acquired fetal MR image data (top row) and reconstructed high resolution images (bottom row) from four subject studies

facts, RMS errors remained about the in plane resolution which is $1mm$. In all the cases, visual inspection by clinicians revealed very promising reconstructed image quality.

We then applied the algorithm to clinically acquired fetal MRI datasets using the same optimization parameters. Results are summarized in Figures 2 and 3. Figure 2 shows one original low resolution coronal image compared to the high resolution reconstructed image for coronal, sagittal and axial views. Figure 3 shows results from four different datasets one original low resolution image and the corresponding high resolution reconstructed image. These figures clearly show the quality improvement in terms of resolution and motion correction providing by the proposed compounding method.

5 Discussion

The ability to study the developing fetal brain in high resolution promises to provide a vital source of clinical information which could contribute directly to a number of challenging clinical questions. It permits the application of many quantitative morphometric analysis methods [12, 2] developed to study the adult brain to probe the process of normal and abnormal brain development. Critically, high resolution imaging is the key to seeing the process of cortical folding [5, 1, 3], while consistent contrast allows us to study patterns of myelination. Such data is valuable both to specific clinical questions and, more fundamentally, to an improved understanding of the process of human brain development.

In this paper we have described a new methodology to reconstruct *in vivo* high resolution 3D MRI scans of the fetal brain by proposing to compound multiple sets of orthogonal 2D MRI slices. The method makes use of a novel

combination of 2D to 3D registration, relative inhomogeneity correction and high resolution reconstruction from sparse data. This is achieved by first globally registering the low resolution images, and then applying an iterative slice alignment scheme which seeks to refine the 3D positioning of each slice with respect to the current combined high resolution image. As a final step, a relative intensity correction is applied between the low resolution images to remove the differences in relative signal strength across the different acquisitions in each region of the fetal brain.

Visual inspection of the obtained results on clinical data are very encouraging. The developed algorithm automatically reconstructs a 3D high resolution geometrically consistent image and it has proved to be robust to large artifacts. This represents an important step towards 3D quantitative analysis of the fetal brain. Further work is needed to better understand the capabilities and limitations of the approach on a range of both fetal anatomies and fetal motion profiles.

This work was supported by a Whitaker foundation award RG-01-0115, a NIH Grant R01-MH65392 and a NIH Biomedical Research Partnership grant, R01-EB0822.

References

1. Batchelor, P.G. *et al.* Measures of folding applied to the development of the human fetal brain *IEEE Trans. Medical Imaging*, 21(8), 953-965, 2002.
2. J. P. Boardman, J.P. *et al.* An evaluation of deformation-based morphometry in the developing human brain and detection of volumetric changes associated with preterm birth *MICCAI*, 2003
3. Cachia, A. *et al.* A Primal Sketch of the Cortex Mean Curvature: a Morphogenesis Based Approach to Study the Variability of the Folding Patterns *IEEE Trans. Medical Imaging*, 22(6), 754-765, 2003.
4. Coakley, F.V. *et al.* Fetal MR imaging: A developing modality for the developing patient *American Journal of Roentgenology*, 182: 243-252, 2004.
5. Inder, T.E. *et al.* Periventricular white matter injury in the premature infant is followed by reduced cerebral cortical gray matter volume at term *Ann Neurol.*, 46(5), 755-60, 1999.
6. Levine, D. Fetal magnetic resonance imaging *Journal of Maternal-Fetal and Neonatal Medicine*, 15:85-94, 2004.
7. Lötjönen, J. *et al.* Correction of Movement Artifacts from 4D Cardiac Short- and Long-Axis MR Data *MICCAI*, 405-412, 2004.
8. Moore, J. and Drangova, M. and Wierzbicki, M. and Barron, J. and Peters, T. A High Resolution Dynamic Heart Model Based on Averaged MRI Data *MICCAI*, 549-555, 2003.
9. Ohbuchi, R. and Chen, D. and Fuchs, H. Incremental volume reconstruction and rendering for 3D ultrasound imaging *SPIE Visualization in Biomedical Computing*, 1808:312-323, 1992.
10. Prayer, D. and Brugger, P. C. and Prayer, L. Fetal MRI: techniques and protocols *Pediatr Radiol*, 34:685-693, 2004.
11. Studholme, C. and Hill, D.L.G. and Hawkes, D.J. An overlap invariant entropy measure of 3D medical image alignment *Pattern Recognition*, 32(1):71-86, 1999.
12. Studholme, C. *et al.* A deformation morphometry study of the influence on the pattern of brain tissue development in premature neonates HBM, 2005.

Respiratory Signal Extraction for 4D CT Imaging of the Thorax from Cone-Beam CT Projections

Simon Rit¹, David Sarrut^{1,2}, and Chantal Ginestet²

¹ LIRIS, Université Lumière Lyon 2, 5 Av. P. Mendès-France, 69676 Bron, France

² Radiotherapy Dpt, Centre Léon Bérard, 28 rue Laënnec, 69353 Lyon, France

Abstract. Current methods of four-dimensional (4D) CT imaging of the thorax synchronise the acquisition with a respiratory signal to retrospectively sort acquired data. The quality of the 4D images relies on an accurate description of the position of the thorax in the respiratory cycle by the respiratory signal. Most of the methods used an external device for acquiring the respiratory signal. We propose to extract it directly from thorax cone-beam (CB) CT projections. This study implied two main steps: the simulation of a set of CBCT projections, and the extraction, selection and integration of motion information from the simulation output to obtain the respiratory signal. A real respiratory signal was used for simulating the CB acquisition of a breathing patient. We extracted from CB images a respiratory signal with 96.4% linear correlation with the reference signal, but we showed that other measures of the quality of the extracted respiratory signal were required.

1 Introduction

Four-dimensional (4D) CT imaging, defined by Keall [1] as the *"acquisition of a sequence of CT image sets over consecutive segments of a breathing cycle"*, constitutes the first step of 4D radiotherapy, which is *"the explicit inclusion of the temporal changes in anatomy during the imaging, planning and delivery of radiotherapy"*. The 4D CT imaging of a free-breathing thorax must face one major technical problem: the time required ($\simeq 0.5$ second per scan) for acquiring several complete 3D CT images along one respiratory cycle ($\simeq 4$ second). Mori et al [2] have proposed a prototype 256-slice CT-scanner dedicated to real-time 4D imaging. Other methods currently in use address the slowness of actual CT acquisitions relatively to the respiratory cycle on both scanner geometries: spiral/helical and cone-beam (CB).

The respiratory correlated method is based on the spatial periodicity of respiration: suppose a respiratory signal $f(t)$, representing the periodic spatial variations of the position of the thorax in the respiratory cycle during acquisition; it is possible to sort a posteriori acquired data into n bins and reconstruct n 3D images [3][4][5]. Different respiratory synchronization techniques have been used: Vedam et al [3], Underberg et al [4] and Pan et al [5] synchronized CT acquisition with the position of the thorax skin surface in the antero-posterior direction; Low et al used a spirometer [6]; Damen et al [7] a thermometer under the patient's nose.

Grangeat et al [8] also proposed a dynamic reconstruction algorithm with CB projections obtained from a CBCT system with a fast rotative gantry, without such synchronization signal. However, the gantry rotation of modern CBCT systems coupled with a linear accelerator is too slow to apply Grangeat's dynamic reconstruction.

With the respiratory correlated method, the critical issue stands with the quality of the respiratory signal, whatever the scanner geometry. Contradictory studies have explored the accuracy of these signals. Depending on the studies, the method using the antero-posterior position of the skin surface does not seem totally accurate [9][10] whereas the air volume obtained with a spirometer seems to have better results [11]. This is why some have proposed methods extracting the signal directly from image data, without any marker, both on spiral CT scan slices [5] and on CBCT projections [12].

We propose to analyze the motion of thorax CBCT projections obtained using a CBCT scan system with a slow rotating gantry to accurate respiratory signal extraction from image data. This study implied two main steps: the simulation of a set of CBCT acquisitions from a 4D model and a respiratory signal, and the extraction, selection and integration of motion information from the simulation output to obtain the respiratory signal for the time of acquisition. The simulation uses a reference respiratory signal, used as a gold standard for evaluating the results of our method.

2 Cone-Beam CT Acquisition Simulation

Our simulation follows a temporal respiratory signal; at each instant t , the respiratory signal (section 2.1) gives the position $f(t)$ in the respiratory cycle and the 4D model (section 2.2) gives the corresponding 3D volume. Digitally Reconstructed Radiographs (DRR) of the volume, i.e. CBCT projections, are then computed using a home made *shearwarp* algorithm [13].

2.1 Respiratory Signal

For each given time t , the respiratory signal indicates the position of the thorax in the respiratory cycle, denoted by $f : \mathbb{R} \rightarrow \mathbb{R}^n$. We define that f must respect the following properties: when $f(t_1) = f(t_2)$, the thorax has the same spatial configuration at times t_1 and t_2 ; f is continuous, meaning that the spatial configuration at time $t + \varepsilon$ is almost equivalent to the one at time t when ε is very small. Previous studies [14] have suggested that the signal can be characterized in first approximation by a 1D function ($n = 1$). It is pseudo-periodic with a pseudo-period around 4 seconds and its extrema often correspond to maximal expiration and inspiration.

2.2 4D Model

From $f(t)$, we must determine the corresponding spatial volume (spatial position) of the thorax, i.e. the corresponding 3D image. Our team had produced

such a 4D model not elaborated from a complete 4D acquisition but from two 3D breathhold acquisitions, one at the end of normal expiration (I1) and one at the end of normal inspiration (I2), acquired using spiral CT imaging and the Active Breathing Coordinator (ABC, Elekta Oncology Systems) [15]. The non-rigid registration of I2 on I1 produced a dense vector-field representing the displacement of each point of I2 toward I1. From these displacements, and supposing in first approximation that each point of the thorax moves on a line, we could interpolate intermediate positions between the two extrema and produce a 4D image. We computed the air volume of the lungs by thresholding and morphological operations, indexing thus the different 3D images in the respiratory cycle and obtaining our 4D model.

3 Respiratory Signal Extraction

Visual observation of the motion of a set of CBCT projections, acquired from a free-breathing patient, intuitively led us to believe that the respiratory signal could be extracted a posteriori from this set of images. Following the same path, Zijp et al [12] have focused on the diaphragmatic cupola and projected their CB projections in the cranio-caudal direction, which produced a set of 1D signals from which they extracted the respiratory signal. The result is not compared to a reference for validation of its accuracy.

We propose here a study of the motion in a sequence of 2D CBCT projections for respiratory signal extraction. The method comports three sequential parts: selection of points of interest, motion extraction and trajectory processing.

3.1 Points of Interest

Some specific parts of the thorax are generally observed for motion extraction [14][12][10], like diaphragmatic cupolas or lung walls. Instead, we chose to consider a uniform set of points constituting a sub-sampling of the pixels of CB projections for a complete study of motion in sequential CBCT projections, with no limitation to any anatomic part. We typically used between 100000 and 200000 points of interest in total.

3.2 Motion Extraction

Our aim was to follow each point from CB projection to CB projection, i.e. the motion of points in the 2D projective space over the time of acquisition. We used a Block Matching Algorithm (BMA) [8].

CBCT imaging does not provide the projection of 3D points in a 2D space, but the integration of tissue densities on lines, which implies that the application of the BMA in the 2D projective space does not follow the projection of a 3D point, but the projection of a high tissue contrast 3D area, like chest and lung walls. It is then almost inevitable to lose the trajectory of the projected area after a significant rotation.

We took this observation into account when using the BMA. Firstly, we computed the most probable trajectory of each point of interest by using the BMA from CB projection to CB projection, using the maximum linear correlation coefficient CC as similarity measure. Secondly, we detected when the BMA is not following the original high contrast area projection: we compared each newly detected block to the original block and stopped when the CC fell under a given user-defined threshold.

3.3 Trajectory Processing

Motion analysis of the set of points of interest produced a set of trajectories $E\{T\}$ where $T : \mathbb{N} \rightarrow \mathbb{N}^2$ defines the spatial position (x, y) in the 2D projective space of a pixel, for CB projections P_i with $0 \leq i < N$, and N being the total number of CB images. These pieces of signal were processed to reconstruct the respiratory signal over the complete time of acquisition, respecting the characteristics of the respiratory signal described in 2.1 above.

Trajectory projection. We obtained a 1D respiratory signal by transforming the function $T : \mathbb{N} \rightarrow \mathbb{N}^2$ into a function $f : \mathbb{N} \rightarrow \mathbb{R}$, using a method that preserved maximum motion information, i.e. variations of the position of the point of interest over time. We firstly calculated the unit vector \vec{u} corresponding to the best approximation line of all positions $M_i(x, y)$ with a linear regression and the mean 2D spatial position of points in time $C(\bar{x}, \bar{y})$. Then, for each projection i , the 1D projection value of the 2D spatial position $M_i(x, y)$ was given by $f_T(i) = \| \overline{CM}_i \cdot \vec{u} \|$ which is the norm of the projection of \overline{CM}_i on the best-fit line (figure 1).

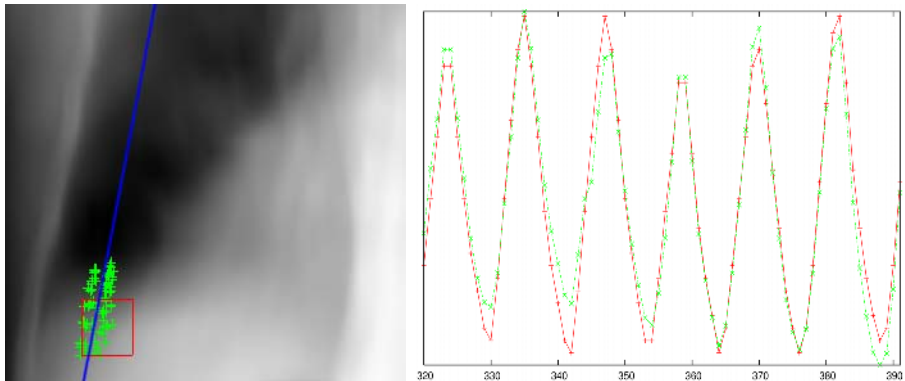


Fig. 1. On the left: the bottom left profile of a simulated CB projection of the lungs, with the selected trajectory (green) of an original block (red) and the best-fit line (blue). On the right: the computed 1D piece of respiratory signal (green) compared to the reference (red). The X axis represents the number of CB projections (chronologically numbered) and the Y axis the respiratory signal value $f(t)$.

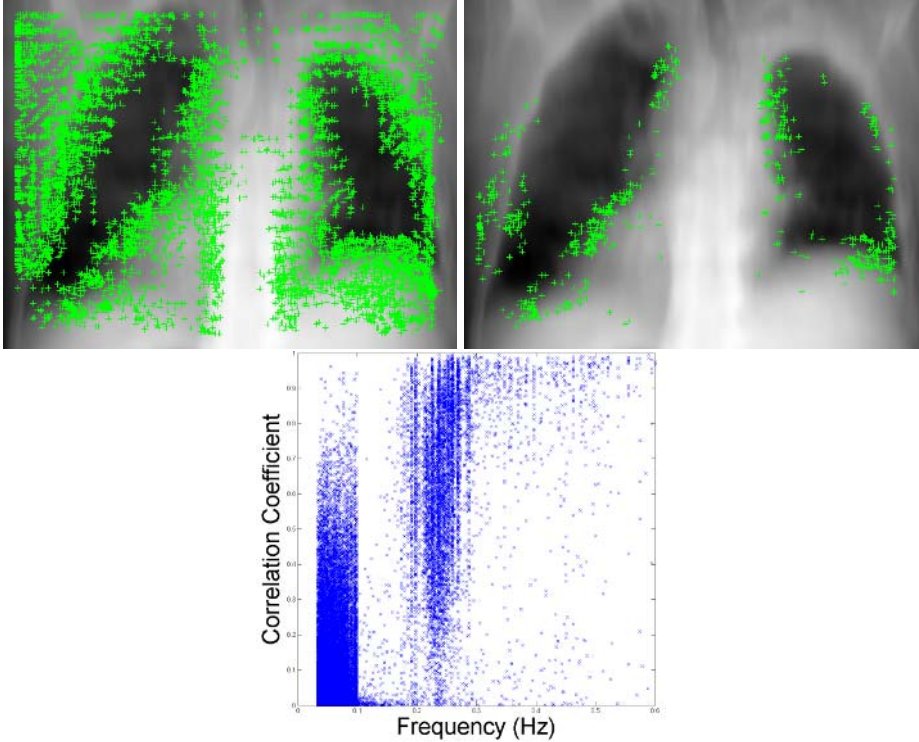


Fig. 2. On the left: a CBCT projection with points extracted by the BMA; on the right: the selected final points. In between: the point (u_{max}, CC) drawn for each piece of signal demonstrates the efficacy of frequency selection. Most of the signal pieces highly correlated to the reference have a period between 3 and 5 sec (0.33 and 0.2 Hz).

Filtering. These 1D pieces of signal generally contain information about the respiratory signal, but this information is correlated with the motion due to cone-beam rotation. As this motion corresponds to an ellipsoidal motion with low frequency, we simply processed the data through a Fourier high pass filter.

Selection. Integration of the different pieces of signal provided good results, but an individual visual observation of the different trajectories showed that most were not relevant to the respiratory signal and a selection was required. We first eliminated short signals (< 7 seconds) or signals with weak mean absolute amplitude (< 2 pixels) because their spatial and/or temporal resolutions were not satisfactory.

To find an interesting selection parameter p , we referred to an important characteristic of the respiratory signal, its temporal pseudo-periodicity. We measure the higher amplitude value of the signal in the Fourier domain, i.e. $p = u_{max}$ with $F(u_{max}) = \max(|F(u)|)$, $u \rightarrow |F(u)|$ being the Fourier amplitude spectrum. Selected signals have $p \simeq p_{ref}$, where p_{ref} corresponds to the pseudo-period of respiration.

This selection criterion was validated using the gold standard as reference by calculation of CC for each piece of signal. After drawing $CC(p)$ (figure 2), we observed that p for all signal pieces with high CC approximated a particular value $p_{ref} \simeq 4$ sec., which confirms our supposition.

Integration. Selected signals represent pieces of the respiratory signal over overlapping intervals of the time of acquisition. We integrated all the signals to determine the whole respiratory signal over the time of acquisition. We considered that extreme positions were matched for all thorax points, thus making integration simpler. Firstly, we linearly normalized all signal pieces between 0 and 1 and eventually put them in phase when they were in phase opposition. For each time t , the respiratory signal value was equal to the mean of the values of all different pieces at this time.

4 Experiments and Results

Experiments. We simulated 720 CB projections, one every 0.5° and every 0.36 second (Elekta Synergy parameter [12]), with a resolution of 600×460 . Simulation was based on a real respiratory signal (irregular in phase and amplitude), acquired with an ABC system and two CT 3D breath-hold images of $512 \times 512 \times 65$ voxels. Motion extraction used blocks of 40×40 pixels and stopped when CC fell under 93%. We selected trajectories with a maximum period in the Fourier domain between 3 and 5 seconds (0.33 and 0.2 Hz).

Measurement. The measure commonly used for evaluating respiratory signals with respect to a reference is the CC [10][11]. Here we also calculated another measure in direct relation with our use of the respiratory signal. We sorted both the respiratory signal extracted from CB images (result bins) and the gold standard (reference bins), then measured the percentage of misplaced samples in the result bins compared to the reference ones. We also calculated $\bar{\sigma}$, the average of σ , the respiratory signal standard deviation from reference values in each bin, for both reference and result bins, and their ratio $\bar{\sigma}_{ref}/\bar{\sigma}_{res}$. The smaller $\bar{\sigma}$ is, the more data in each bin were in phase with the respiratory cycle. $\bar{\sigma}_{ref}$ automatically decreases with the number of bins, but $\bar{\sigma}_{res}$ stops decreasing when the number of bins is too big because of the inaccuracy of the respiratory signal result. $\bar{\sigma}_{ref}/\bar{\sigma}_{res}$ is equal to 1 when both sortings are equally good and decreases when the quality of the result sorting decreases.

Binning. There are different methods for sorting CB images with the respiratory signal, depending whether one takes hysteresis into account and whether bins have the same size. If bins have the same size, the 3D images have homogeneous resolutions but variation of the respiratory signal in each bin is variable; else, extrema have a higher resolution because there are more data used for their reconstruction. We chose to take hysteresis into account and to divide the signal into equal bins.

We extracted a respiratory signal with 96.4% correlation with the reference. After sorting, we obtain:

Number of bins	Misplaced samples	$\bar{\sigma}_{ref}$	$\bar{\sigma}_{res}$	$\bar{\sigma}_{ref}/\bar{\sigma}_{res}$
1	0	0.249	0.249	1
2	23 (3%)	0.123	0.127	0.97
4	57 (8%)	0.093	0.101	0.92
8	122 (17%)	0.055	0.068	0.81
12	184 (25%)	0.038	0.055	0.69

5 Discussion and Conclusion

An interesting result of this method is the location of selected trajectories on CB images, i.e. points of CBCT images having a motion directly related to respiration. Most of these points are located around the diaphragmatic cupola, which corroborates Zijp et al [12] hypothesis, but others are also located on lung walls in the inferior lobe (figure 2). Their motion direction depends on their location: mostly cranio-caudal for points around the cupola, and perpendicular to the lung walls otherwise. The location of the points of interest has now been determined; future work will improve the method with *a priori* detection of the points of interest.

The computed respiratory signal is visually very close to the reference, as confirmed by the high *CC*. But the result bins has a proportion of misplaced points compared to the reference bins which increases with the number of bins. The impact of these misplacements is measured by $\bar{\sigma}_{ref}/\bar{\sigma}_{res}$. It is unnecessary to increase the number of bins if $\bar{\sigma}_{ref}/\bar{\sigma}_{res}$ is too low, and/or if it does not decrease sufficiently $\bar{\sigma}$. This measurement points out the importance of an accurate respiratory signal for 4D imaging, and the not so good capacity for the *CC* to measure this accuracy. However, the number of bins is limited because the quantity of data in each bin is a primordial criterion for the quality of the reconstruction.

4D CT images will be reconstructed with a number of bins deduced from these observations. They will be used for modeling the thorax motion, and will offer the possibility to take motion into account for treatment planning. Future work will also include validation on real CBCT data.

Acknowledgement. This work was supported in part by Elekta Oncology Systems.

References

1. Keall, P.: 4-dimensional computed tomography imaging and treatment planning. *Radiation Oncology* **14** (2004) 80–90
2. Mori, S., Endo, M., Tsunoo, T., Kandatsu, S., Tanada, S., Aradate, H., Saito, Y., Miyazaki, H., Satoh, K., Matsushita, S., Kusakabe, M.: Physical performance evaluation of a 256-slice CT-scanner for four-dimensional imaging. *Medical Physics* **31** (2004) 1348–1356

3. Vedam, S., Keall, P., Kini, V., Mostafavi, H., Shukla, H., Mohan, R.: Acquiring a four-dimensional computed tomography dataset using an external respiratory signal. *Physics in Medicine and Biology* **48** (2003) 45–62
4. Underberg, R.W., Lagerwaard, F.J., Cuijpers, J.P., Slotman, B.J., van Sörnsen de Koste, J.T., Senan, S.: Four-dimensional CT scans for treatment planning in stereotactic radiotherapy for stage I lung cancer. *International Journal of Radiation Oncology Biology Physics* **60** (2004) 1283–1290
5. Pan, T., Lee, T.Y., Rietzel, E., Chen, G.T.Y.: 4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT. *Medical Physics* **31** (2004) 333–341
6. Low, D.A., Nystrom, M., Kalinin, E., Parikh, P., Dempsey, J.F., Bradley, J.D., Mutic, S., Wahab, S.H., Islam, T., Christensen, G., Politte, D.G., Whiting, B.R.: A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing. *Medical Physics* **30** (2003) 1254–1263
7. Damen, E., Wolthaus, J., Herk, M.V., Sonke, J.J., Remeijer, P., Zijp, L.: Image-guided radiotherapy for lung cancer: Respiration correlated cone-beam CT to verify tumor position and motion characteristics during treatment delivery (abstract). In: *ASTRO*. (2004)
8. Grangeat, P., Koenig, A., Rodet, T., Bonnet, S.: Theoretical framework for a dynamic cone-beam reconstruction algorithm based on a dynamic particle model. *Physics in Medicine and Biology* **47** (2002) 2611–2625
9. Ford, E., Mageras, G., Yorke, E., Ling, C.: Respiration-correlated spiral CT: A method of measuring respiratory-induced anatomic motion for radiation treatment planning. *Medical Physics* **30** (2003) 88–97
10. Koch, N., Liu, H.H., Starkschall, G., Jacobson, M., Forster, K., Liao, Z., Komaki, R., Stevens, C.W.: Evaluation of internal lung motion for respiratory-gated radiotherapy using MRI: Part I-correlating internal lung motion with skin fiducial motion. *Int Journal of Radiation Oncology Biology Physics* **60** (2004) 1459
11. Lu, W., Parikh, P.J., Naqa, I.M.E., Nystrom, M.M., Hubenschmidt, J.P., Wahab, S.H., Mutic, S., Singh, A.K., Christensen, G.E., Bradley, J.D., Low, D.A.: Quantitation of the reconstruction quality of a four-dimensional computed tomography process for lung cancer patients. *Medical Physics* **32** (2005) 835–1228
12. Zijp, L., Sonke, J.J., Herk, M.V.: Extraction of the respiratory signal from sequential thorax cone-beam X-ray images. *International Conference on the Use of Computers in Radiation Therapy* (2004)
13. Lacroute, P.G.: Fast volume rendering using a shear-warp factorization of the viewing transformation. PhD thesis, Stanford University (1995)
14. Kondo, T., Kobayashi, I., Taguchi, Y., Ohta, Y., Yanagimachi, N.: A dynamic analysis of chest wall motions with MRI in healthy young subjects. *Respirology* **5** (2000) 19
15. Boldea, V., Sarrut, D., Clippe, S.: Lung deformation estimation with non-rigid registration for radiotherapy treatment. In: *Medical Image Computing and Computer-Assisted Intervention MICCAI'2003*. Volume 2878. (2003) 770–7

Registering Liver Pathological Images with Prior In Vivo CT/MRI Data

Huadong Wu, Alyssa M. Krasinskas, Mitchell E. Tublin, and Brian E. Chapman

Abstract. Liver transplantation affords a unique opportunity to assess and improve radiological imaging of the liver, as the full explanted liver is available for review and comparison. Quantitative comparison between the explanted liver and in vivo images acquired prior to transplantation requires accurate registration of images of the explanted liver to the radiological images. However, this registration problem is challenging because the orientation change and the deformation magnitude between the two image sets exceed the level assumed for most registration algorithms. This paper suggests a two-step registration process to overcome the difficulty: to first align the orientation of 3D liver models built from two sets of image data using maximum volume overlap as their similarity measurement, and second to deform one model to match the other. The key contribution of this paper is that it utilizes the global volumetric information and the asymmetry property of the liver model to determinately provide a simple and reliable initial point for further deformable model based registration. Our experimental data demonstrate the effectiveness of this approach.

1 Introduction

1.1 Motivation

Imaging of the liver is an important problem in radiology, particularly regarding cancer. Not only is the liver an important target for metastatic disease, but the incidence of primary liver tumors, such as hepatocellular carcinoma (HCC), has reach epidemic levels in some parts of the world. The health impact of HCC world-wide is tremendous [8]. However, the early clinical and radiological manifestations of HCCs are subtle and optimizing imaging strategies for detecting them is difficult [9]. We believe that explanted livers provide a unique basis for optimizing MR or CT strategies for imaging the liver, particularly regarding the early detection of HCC.

Correlating in vivo liver images with photographs of the sectioned explanted liver is difficult since the liver geometry changes dramatically when the flexible liver is no longer supported by its surrounding anatomic structures. Additional deformations occur due to loss of tissue distension from blood pressure and tissue relaxation when the liver capsule is cut. However, the liver boundary shape (3D surface) is a relatively rich feature for matching the pathology images to the in vivo images. This paper suggests the overall geometric shape retains sufficient information to do a preliminary alignment of the ex vivo and in vivo images. Application of a deformable registration can then be applied to provide final alignment of the images.

1.2 Registration Challenge

Modeling deformable anatomical structures with significant variability is currently being investigated. Deformable geometric models allow structures to change shape (from static point of view) or to evolve (from dynamic point of view) subject to constraints derived from physical principles. These models provide a reasonable basis for image registration [2], [4]. The most widely used physical assumption for static models is an elastic body, whose substrate responding naturally to applied forces and internal structural constraints. Dynamic models usually assume the viscous fluid property that the applied forces are proportional to the time rate of change in displacement [1], [3].

Almost all registration methods today assume that the initial estimation of location and pose is close to the ground truth so that the optimal point can be assumed to be within a local optimum. Obviously our pathological liver registration situation violates this basic assumption. Therefore, for our liver registration problem we need to align the images so as to find an acceptable initial start in registration.

The generally used approach to initialize the image registration process is to take course-to-fine multiple resolution registration scheme, using the result from the course-resolution registration as the initial point for the finer resolution registration. Ourselin et al. extended this approach (named “Block Matching” method) for rigid body registration, in which they use the majority of sub sets with maximized similarity to estimate a start point [5]. Lau et al. also used a multiscale approach and addressed the initial estimation problem [6]. They build a deformation vector field by matching block sub-regions between the images using a voxel-similarity measure. A median filter is then used to ensure a locally smooth transformation [11].

It is not difficult to conclude that these approaches are not suitable for our liver registration tasks, where orientation is totally lost, large deformation is assumed, and the voxel information from pathological images can be at best sparsely available due to the thick sectioning of the explanted liver¹.

Intuitively, the best solution to this challenge would require globally abstracting the geometry characteristics of the entity of concern. Using the medial axis description of geometric body, a research topic getting more popular recently, might be a good candidate to realize such abstraction [2]. However, the medial axis extraction process is intrinsically sensitive to surface variation of the object [7], and in our case the constructed 3D models from pathological images are not sufficiently stable in the sense that the slices are arbitrarily laid out² and contour line transitions between adjacent images are not smooth.

This paper will show that, because of the asymmetry characteristics of the liver geometric shape, the geometric principle axis description can be used as a simple but effective clue for the orientation alignment to be used as the initial step in the registration process. Following the gross alignment of the respective volumes using the geometric principle axes, local deformations can easily be applied.

¹ Even if the liver is well cut into slices during the pathology examination, the slices are usually 10~20 mm thick. Therefore, if a 3D volume model is built out of the slices, no real voxel information between slices is available, thus limiting the usefulness of truly 3D to 3D deformable registration techniques which are currently available (e.g. www.itk.org).

² The slices are laid out for photographing in a relaxed shape that resembles their in vivo nature status — but only according to the pathology technician’s own imagination.

2 Image Registration Via Volume Match

2.1 Registration Process Overview

First, a 3D model is built by segmenting the liver from the CT/MRI data on a slice by slice manner; this model serves as the reference model. The model is interpolated in the slice direction in order to obtain isotropic voxels.

Second, a 3D model is built by stacking the photographs of the sectioned explanted liver, where adjacent slices are aligned via translation, rotation and scaling. Alignment is driven by voxel similarity. This 3D model is subsequently scaled to match the volume of the 3D model built from CT/MRI data.

For these two 3D models, the centroid and voxel distribution are calculated respectively. The voxel distribution is described for each volume using a covariance matrix. Decomposing the two covariance matrices respectively, we can find the geometric principle component axes (eigenvalues) of each object.

Then, by rotating the CT/MRI 3D model around its centroid, we can align the geometric principle axes with respect to the corresponding geometric principle axes of the 3D model built from pathological images.

This global orientation alignment provides an initial point for further model deforming. Currently we correlate pathological slice images with the CT/MRI 3D model's corresponding intersections and do 2D image warping to increase the registration accuracy.

2.2 Building 3D Models from CT/MRI Images

The liver CT images are first hand-segmented from the reminder of the images in each cross section slice image, as shown in the left image of Fig. 1. The extracted liver model is represented as a 3-dimensional matrix where the value of pixels outside the liver is set to zero. To describe the overall geometric properties of the liver model (Ω), we set the liver voxel V_i to a unit value and calculate the centroid point coordinates (the mass center as if the liver has unit density)

$$[m_x = \frac{\sum_{i \in \Omega} x_i V_i}{N}, m_y = \frac{\sum_{i \in \Omega} y_i V_i}{N}, m_z = \frac{\sum_{i \in \Omega} z_i V_i}{N}], V_i = 1 \tag{1}$$

and the voxel distribution covariance matrix:

$$\text{Cov} = \frac{1}{\sum_{i \in \Omega} V_i} \begin{bmatrix} \sum_{i \in \Omega} (x_i - m_x)^2 & \sum_{i \in \Omega} (x_i - m_x)(y_i - m_y) & \sum_{i \in \Omega} (x_i - m_x)(z_i - m_z) \\ \sum_{i \in \Omega} (y_i - m_y)(x_i - m_x) & \sum_{i \in \Omega} (y_i - m_y)^2 & \sum_{i \in \Omega} (y_i - m_y)(z_i - m_z) \\ \sum_{i \in \Omega} (z_i - m_z)(x_i - m_x) & \sum_{i \in \Omega} (z_i - m_z)(y_i - m_y) & \sum_{i \in \Omega} (z_i - m_z)^2 \end{bmatrix} \tag{2}$$

Since the covariance matrix is real, symmetric, and positive semi-definite, it can be eigen-decomposed into principle axes:

$$\text{Cov} = \mathbf{E} \mathbf{D} \mathbf{E}^T, \quad \mathbf{E} = [E_1, E_2, E_3] \tag{3}$$

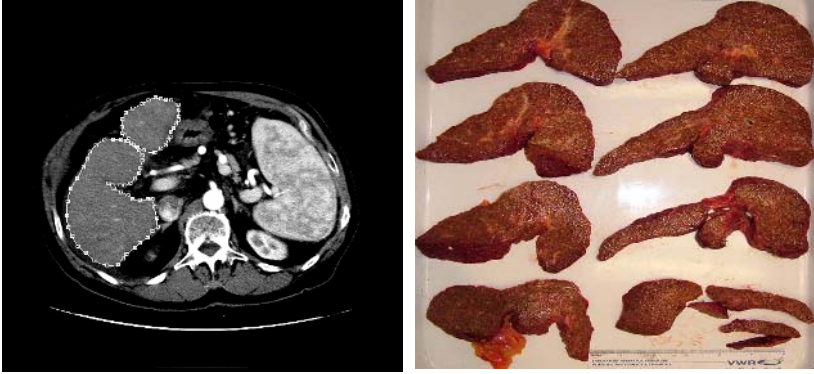


Fig. 1. Left: CT image data with the liver manually segmented; Right: pathological sliced liver images

The eigen vectors, E_1 , E_2 and E_3 , describe the 3 orthogonal directions, i.e., the orientation of the geometric principle axes, in which the liver volume spans. The eigen values, or the element e_1 , e_2 and e_3 of the diagonal matrix \mathbf{D} , correspondingly describe the voxel distribution in these principle axis directions: the larger the value the more the liver is spread-out in this direction.

2.3 Building 3D Models from Pathological Liver Slice Images

We create a 3D model by virtually stacking slices to form a 3D liver model. The physically analogy of this process is shown in Fig. 2. The images were first preprocessed to segment the liver region, with the background pixel values being set to zero



Fig. 2. Pathologists would stack slices in trying to register pathological findings

and the voxel dimensions scaled to match the in vivo voxel size. Choosing a slice image that is close to the center of the liver as base, the stack grows out in both directions perpendicular to the base slice image, with local registration between adjacent slice images. Slice-to-slice alignment is achieved via translation, rotation around the center of the ROI (Region of Interest, i.e., the liver pixels inside the image), and isotropic scaling transformation $T(\cdot)$. This is represented mathematically as:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \cdot \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \cdot \begin{bmatrix} x - m_x \\ y - m_y \end{bmatrix} + \begin{bmatrix} T_x + m_x \\ T_y + m_y \end{bmatrix} \quad (4)$$

The pixel-wise mean square difference between the image \mathbf{A} in the base stack and the transformed image (the image to be added to the stack) $T(\mathbf{B})$ is used as the metric to optimize the transformation parameters.

$$MS(\mathbf{A}, T(\mathbf{B})) = \frac{1}{N} \sum_i^N (A_i - B_i)^2 \quad (5)$$

where A_i and B_i are pixel values in image \mathbf{A} and \mathbf{B} , and N is the number of total pixels considered.

To achieve global optimization, the process is implemented in two steps. In the first step, the image pixel values are reduced to binary data with pixels inside or on the liver ROI set to 1 and the background set to 0. Starting at different initial rotation angles with increment of 10° , a coarse optimization is conducted. Then the best resulting transformation parameters in the first step are used as the starting position in the second step, where a finer optimization step-length is used and the original gray-scale image pixel value is used.

With the calculated transformation parameters and by setting the scaling factor back to $\lambda = 1$, new slice images are added to the liver model stack. One image at a time, and taking the slice thickness into account, the 3D liver model is built from the 2D pathological images. The volume of the resulting 3D liver model typically differs from the model built from the in vivo CT or MRI images. Since the precise nature of the volumetric changes is unknown, a uniform scaling scheme is used to approximately undo the effect.

We use the same measurements described in Equation 1, 2, and 3 in previous subsection to present the characteristics of the model geometry. The eigen values from the largest to the least are denoted as e_1 , e_2 , and e_3 , and their corresponding eigenvectors (geometric principle axes) are denoted as E'_1 , E'_2 , and E'_3 .

2.4 Aligning the 3D Models

Even though the 3D model built from in vivo CT scan images may seem substantially different from the 3D model built from the pathological images of the ex-transplanted liver slices, because of the geometric asymmetry, the two models can still be approximately aligned using their geometry principle axis properties. The alignment proceeds as follows. First we rotate the 3D liver model from CT images along the direction vector:

$$E_R = E_1 \times E'_1 \quad (6)$$

with an angle:

$$\theta = \arccos\left(\frac{E_1 \cdot E'_1}{\|E_1\| \cdot \|E'_1\|}\right) \quad (7)$$

This aligns the largest geometric principle axis E_1 of the 3D liver model from CT images to the direction of the largest principle axis E'_1 of the 3D liver model from pathological images. Second, the next-to-last geometric principle axis E_2 of the model is rotated to a new orientation as $E_{R\theta}(E_2)$. Third, we rotate the 3D liver model from CT images along the direction of E'_1 with an angle:

$$\beta = \arccos\left(\frac{E_{R\theta}(E_2) \cdot E'_2}{\|E_{R\theta}(E_2)\| \cdot \|E'_2\|}\right) \quad (8)$$

This rotation aligns the next largest geometry principle axes of the two models. If the two models have a similar geometry shape, their orientation should have now properly aligned in the sense that their second order voxel distribution moment has a maximum overlap. Fourth, since the difference of eigen values corresponding to second and third geometry principle axes are typically less conspicuous, to make sure that the two models are best aligned in terms of volume overlap, fine search is conducted via changing the rotation angle β .

2.5 2D Image Warping Deformation

After application of our principle axes alignment and volume scaling, the images are sufficiently aligned to apply a deformable registration technique. For these initial experiments we have implemented a 2D slice-by-slice deformation using the IDL function `WARP_TRI(.)` [10]. Thus our final registration accuracy largely depends on properly choosing the control points in the image warping process.

3 Results

Examples of the resulting models after rigid body alignment is completed are shown in Fig. 3. In this example, the geometric variation eigenvalues of the 3D model built from

the in vivo CT images are: {3791.8, 1265.4, 538.6}, and the geometric variation eigenvalues of the volume-normalized 3D model built from the pathological images are {5236.0, 678.5, 523.4}. It seems that the geometric asymmetry property is largely preserved but the spread-out effect has increased the largest principle vector and decreased the next-largest principle vector.

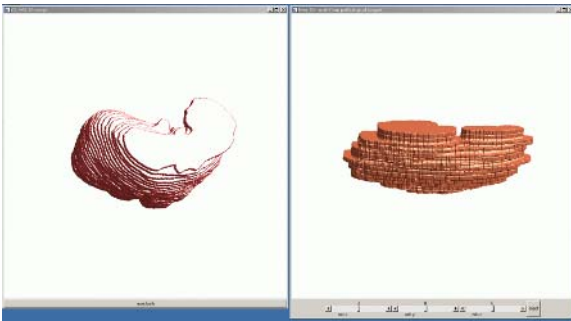


Fig. 3. Two models' orientation aligned

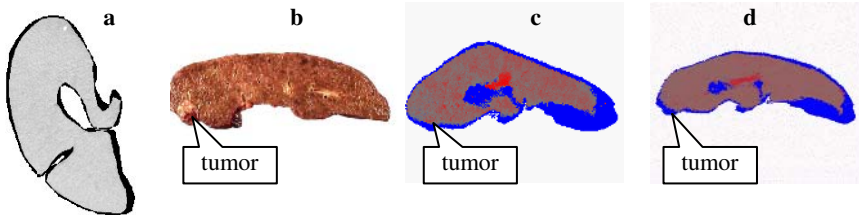


Fig. 4. Pathological liver slice image registration, **b**: pathological image with a tumor at lower-left; **a**: human technician estimated corresponding CT scan slice (a failure); **c**: global geometry principle axis alignment scheme successfully finds the plane in CT data that contains the tumor; **d**: registration is improved by simple 2D image warping deformation

Fig. 4 shows corresponding slices extracted from the rigidly aligned models. It is clear that although the two slices still differ by an in-plane bending, our registration scheme successfully finds the correct region. In this given example, the registration error in the direction parallel to the slice is in the range of 5–10 mm. Application of the 2D warping provides a nice alignment of the tumor in the two images. We estimated the accuracy of the registration by measuring the displacement between tumor nodules in the radiological and pathological images. Using 10 tumors as landmarks, we estimated the residual displacement to be on the order of 10 to 30 mm.

4 Conclusions and Discussion

Because of large deformation due to tissue flexibility as well as the lack of consistent fiducial marks between modalities, correlating pathological findings in explanted liver slices to its prior *in vivo* radiological imaging manifestation poses a big challenge. After a translation to make the centroids coincide, the key to solving this difficult problem is to properly align the geometric orientation of the 3D model — so that overall correspondence between models can be established and deformable model based registration process can have a proper initial starting point.

The underlying principle for this registration initialization scheme is based on the assumption that although the liver has undergone dramatic deformation after being explanted and sliced for pathological examination, its overall asymmetric shape property is still preserved in terms of the geometric principle axes' orientation. In the cases where this condition fails to be held due to badly laid out slices (*i.e.*, the overall shape is too far away from its natural status) or the slices were not evenly cut, our solution is to manually estimate the approximate orientation and then do a local optimum orientation search based on maximum volume overlap.

When available, a radiological image of the explanted specimen can be extremely useful, because it preserves an intermediate state of deformation (prior to sectioning). Such an *ex vivo* image can be used as an intermediate target to which the sectioned photographs and *in vivo* images can be independently matched.

The aligned 3D model pair has global maximum volume overlap. However, the shape deformations that exist between the objects are still large and we believe that a large scale bending needs to be done prior to the application of a typical deformable registration model. We hypothesize that the bending can be estimated by examining the medial axes extracted from smoothed versions of our 3D models. Our results, however, are already good enough for practical visual correlation of the pathological and radiological images.

Acknowledgement

This work is supported in part by NIH grant CA095759.

References

1. D.L.G. Hill, P.G. Batchelor, M. Holden and D.J. Hawkes "Medical Image Registration," *Physics in Medicine and Biology* vol. 46, no. 3, pp. 1-45, 2001.
2. S.M. Pizer, T. Fletcher, S. Joshi, A. Thall, J.Z. Chen, Y. Fridman and e. al. "Deformable M-Reps for 3D Medical Image Segmentation," *International Journal of Computer Vision* vol. 55, no. 2/3, pp. 85-106, 2003.
3. T. McInerney and D. Terzopoulos "Deformable Models in Medical Image Analysis: A Survey," *Medical Image Analysis* vol. 1, no. 2, pp. 91-108, 1996.
4. J. Montagnat, H. Delingette and N. Ayache "A Review of Deformable Surfaces: Topology, Geometry and Deformation," *Image and Vision Computing* vol. 19, no. 14, pp. 1023-1040, 2001.
5. S. Ourselin, A. Roche, S. Prima and N. Ayache "Block Matching: a General Framework to Improve Robustness of Rigid Registration of Medical Images," *Third International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'00)* vol. 1935, pp. 557-566, 2000.
6. Y. Lau, M. Braun and B. Hutton "Non-rigid Image Registration Using a Median-filtered Coarse-to-fine Displacement Field and a Symmetric Correlation Ratio," *Physics in Medicine and Biology* vol. 46, no. 4, pp. 1297-1319, 2001.
7. S.M. Pizer, K. Siddiqi, J.N. Damon and S. Zucker "Multiscale Medial Loci and Their Properties," *International Journal of Computer Vision* vol. 55, no. 2-3, pp. 155-179, 2003
8. Hashem B. El-Sera and Andrew C. Mason "Rising Incidence of Hepatocellular Carcinoma in the United States," the *New England Journal of Medicine* Volume 340, no. 10, pp. 745-750, March 1999
9. Shahid Hussain, P. Zondervan et al. "Benign versus Malignant Hepatic Nodules: MR Imaging Findings with Pathological Correlation," *Radiology Graphics*, vol. 22, no. 5, pp. 1023-1039, 2000
10. Robert J. Renka, "Interpolation of Data on the Surface of a Sphere and Triangulation and Interpolation at Arbitrary Distributed Points in the Plane," *ACM Transactions on Mathematical Software*, Vol.10, no. 4, pp.437-439, Dec. 1984

Support Vector Clustering for Brain Activation Detection

Defeng Wang¹, Lin Shi², Daniel S. Yeung¹, Pheng-Ann Heng²,
Tien-Tsin Wong², and Eric C.C. Tsang¹

¹ Department of computing, The Hong Kong Polytechnic University,
Hung Hom, Kowloon, Hong Kong, China

{csdfwang, csdaniel, csetasang}@comp.polyu.edu.hk

² Department of Computer Science and Engineering,
The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, China

{lshi, pheng, ttwong}@cse.cuhk.edu.hk

Abstract. In this paper, we propose a new approach to detect activated time series in functional MRI using *support vector clustering* (SVC). We extract Fourier coefficients as the features of fMRI time series and cluster these features by SVC. In SVC, these features are mapped from their original *feature space* to a very high dimensional *kernel space*. By finding a compact sphere that encloses the mapped features in the kernel space, one achieves a set of cluster boundaries in the feature space. The SVC is an effective and robust fMRI activation detection method because of its advantages in (1) better discovery of real data structure since there is no cluster shape restriction, (2) high quality detection results without explicitly specifying the number of clusters, (3) the stronger robustness due to the mechanism in outlier elimination. Experimental results on simulated and real fMRI data demonstrate the effectiveness of SVC.

1 Introduction

Functional magnetic resonance imaging (fMRI) is a non-invasive tool to observe the brain neural activities when the subject is undertaking cognitive or motor tasks. Most activation detection techniques can be categorized as *model-driven* and *data-driven* approaches. Apart from principle component analysis (PCA) and independent component analysis (ICA), clustering is a family of effective data-driven approaches to identify unknown responses in fMRI data. The aim of clustering approach is to separate the time series into clusters — each contains voxels with similar activation patterns. Existing clustering approaches to fMRI data analysis include but not limited to crisp C-means [1], Kohonen clustering neural network [2], fuzzy C-means (FCM) [3], hierarchical clustering [1], etc.

However, most existing methods are suffering from the problem of choosing the number of clusters. The detected set of voxels varies significantly when the number of clusters is set to different values. For example, in crisp C-means and FCM, the user has to define a suitable value for the number of clusters C ; in hierarchical clustering, there is similar dilemma in determining a proper cut for the

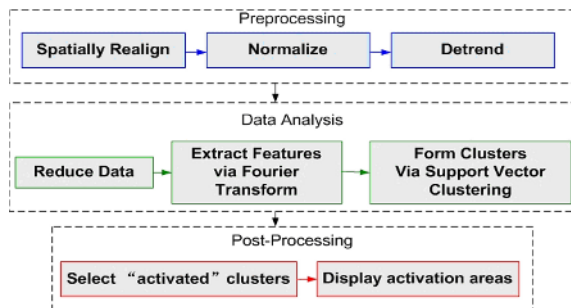


Fig. 1. fMRI Data Processing Scheme

dendrogram. Moreover, the existing fMRI clustering methods implicitly assume the potential clusters are scattered in certain shapes, such as hyper-spherical (when using Euclidean distance) and hyper-ellipsoidal (when using Mahalanobis distance). This oversimplified assumption leads to partitions against the natural data structure and thus achieves results with high false alarm rate.

Support vector machine (SVM) [4] is a preferred classifier in many theoretical and empirical areas, due to its salient properties such as margin maximization and kernel substitution for classifying data in a high dimensional kernel space. Inspired by SVM, the support vector clustering (SVC) algorithm is proposed [5].

In this paper, we explore the application of SVC to solve the fMRI activation detection problem. The whole scheme is illustrated in Fig. 1. We first extract features from fMRI time series using the Fourier transform, and then apply SVC to cluster these Fourier coefficients. They are mapped to a high dimensional kernel space using a kernel function, e.g. the Gaussian kernel. In the kernel space, we calculate the smallest sphere enclosing these mapped features, which appears in the feature space as a set of contours enclosing the original features. These contours are interpreted as cluster boundaries. The points within the same continuous region encompassed by the contour are considered as a cluster. SVC has been shown to be able to generate clusters with arbitrary shapes and eliminate the outlier by using a regularization constant that controls the compactness of the sphere in the kernel space – points out of that sphere is considered as outliers.

2 Materials and Methods

2.1 Dataset

Simulated Dataset Generation. According to a new BOLD response modeling technique [6], the BOLD response $s(t)$ is modeled as the combination of two gamma functions $g_a(t)$ and $g_b(t)$ convoluted with stimulus $c(t)$:

$$g_a(t) = (1 - e^{-1/d_a})^2 (t + 1)e^{-t/d_a}, \quad g_b(t) = (1 - e^{-1/d_b})^2 e^{-t/d_b},$$

$$s(t) = f_a(g_a * c)(t - d_0) + f_b(g_b * c)(t - d_0) + f_c(g_a * c)(t - d_0)(g_b * c)(t - d_0)$$

Table 1. Parameters for the simulated data

Region	f_a	f_b	f_c	d_a	d_b	d_0
A	0.7	0.04	0.6	2	12	3
B (left and right)	0.5	0.3	0.5	3	6	2
C	0.35	0.1	1	5	5	4

To simulate the real-world fMRI data and to get concrete performance measurements for comparison, we construct four artificial activation regions, i.e. A, B(left and right), and C regions in Fig. 2(b), with three sets of values for the parameters (see Table 1) in the above-mentioned BOLD response model. The baseline image is constructed by averaging time courses on the 29th slice in the processed real auditory fMRI dataset (cf. Fig. 2(a)) from Wellcome Department of Cognitive Neurology at University College London. The size of simulated data is $79 \times 95 \times 1$, and the length for each time series is 200. Then we add onto all the voxels the additive white Gaussian noise with intensities proportional to the baseline voxel intensities. In the experiment, we simulate various contrast-to-noise ratio (CNR).

In Vivo **fMRI Dataset.** We also use the auditory fMRI data acquired by Wellcome Department of Cognitive Neurology at University College London to validate the effectiveness of our method. We use it with the permission from the Functional Imaging Laboratory. The experiment was conducted to measure the

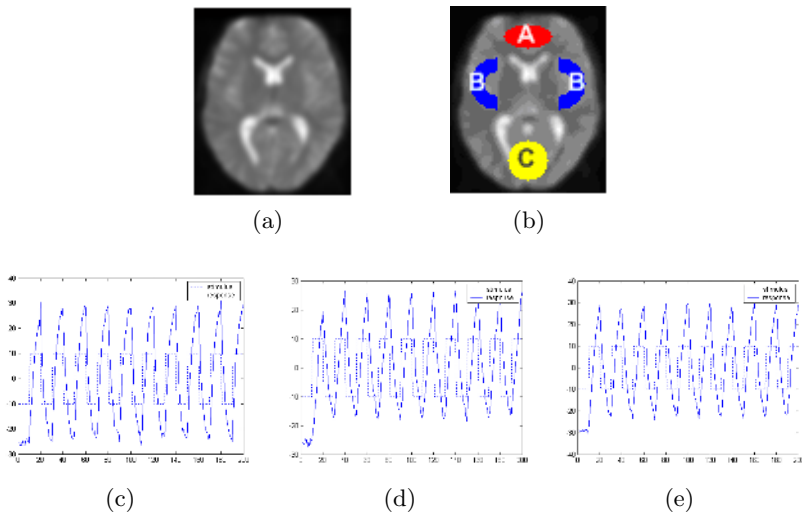


Fig. 2. Simulated Dataset (a) baseline slice generated from the 29th slice of real data; (b) spatial layout of simulated BOLD activation; (c) average activation time series in A; (d) average activation time series in B; (e) average activation time series in C

activation in the subjects brain when given an auditory bi-syllabic stimulation. This whole brain BOLD/EPI data was acquired with a modified 2-T SIEMENS scanner and the TR is set to 7 sec. The spatial resolution of the volume data is $64 \times 64 \times 64$, and the size of each voxel is $3mm \times 3mm \times 3mm$. In this experiment, the subject is given a boxcar stimulation: first begins with rest, and then switches between rest (6 scans, last for 42 sec) and stimulation (6 scans, last for 42 sec). By discarding the first few scans, the acquired 96 scans are reduced to 84 scans. The structural dataset that has been aligned with the function data is also provided for reference.

2.2 Methods

Preprocessing. To remove the head motion effect in raw fMRI data, we first spatially realign the slices with the first slice. Then the dataset is normalized by subtracting the mean amplitude from each time series and adjust their variances to one in order to emphasize their variation patterns. The last preprocessing step is first-order polynomial detrending [7] that removes the undesired drift.

Data Analysis. (1)*Data Reduction:* To accelerate feature extraction and data clustering, also to avoid ill-balance clustering result, we use t-test to remove the data points that will definitely not be considered as the activated voxels. All the voxels that are likely to be activated will be acquired by setting a low threshold in the t-test. (2)*Feature Extraction:* BOLD response to the periodic stimulus can be well characterized by the Fourier coefficients [8]. As the paradigm in our experiment is typical box-car, which is periodic, we use the Fourier transform to acquire the features for clustering. Another merit of using Fourier transform to extract features is to avoid explicitly modeling the delay from the stimulation, as the delay changes in various brain locations and for different subjects. Suppose the fMRI dataset is \mathbf{Y} , a $T \times N$ matrix, where N is the number of time series, each with a length T . Columns in \mathbf{Y} are time series and rows in \mathbf{Y} correspond to scans. Since we have performed the normalization and detrending in the preprocessing stage, the columns of \mathbf{Y} are zero mean, zero drift time series. The harmonic components are calculated as

$$\alpha_\ell(t) = \sin\left(\frac{\ell}{2}\omega t\right) \cdot (\ell \bmod 2 + 1) + \cos\left(\frac{\ell+1}{2}\omega t\right) \cdot (\ell \bmod 2), \ell = 1 \dots L, (1)$$

where ω is the fundamental frequency of the experimental paradigm, $L \leq (2\pi)/(\Delta t \cdot \omega)$ is the number of harmonic components, and Δt is the temporal sampling period. We then form a $T \times L$ design matrix $\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_L]$. So far, the data can be expressed as $\mathbf{Y} = \mathbf{A} \cdot \mathbf{X} + \boldsymbol{\varepsilon}$ [9], where \mathbf{X} is an $L \times N$ harmonic image matrix containing the linear coefficients, and $\boldsymbol{\varepsilon}$ is the noise matrix. An estimate of \mathbf{X} can be computed using a least squares fit, i.e. $\hat{\mathbf{X}} = (\mathbf{A}^t \mathbf{A})^{-1} \mathbf{A}^t \mathbf{Y}$. We treat $\hat{\mathbf{X}}$ as the input data in the next SVC step.

Support Vector Clustering. There are two major steps in support vector clustering: a training step to train a kernel radius function and a cluster labeling step to assign a cluster index to each data point.

(1) *Training:* Assume $\mathbf{X} = \{\mathbf{x}_i\}, i = 1 \dots N$ is the input data. Using a non-linear transformation, \mathbf{X} can be transformed to a high-dimensional kernel space, where we can find the smallest hyper-sphere that encloses the maps of $\{\mathbf{x}_i\}$, i.e. $\{\phi(\mathbf{x}_i)\}$ with the radius R :

$$\|\phi(\mathbf{x}_j) - \mathbf{a}\|^2 \leq R^2 + \xi_j, \quad (2)$$

where \mathbf{a} is the center and $\xi \geq 0$ are the slack variables that enables soft boundaries calculation. Using the Lagrangian with a regularization constant C in its penalty term, $L = R^2 - \sum_j (R^2 + \xi_j - \|\phi(\mathbf{x}_j) - \mathbf{a}\|^2) \beta_j - \sum_j \xi_j \mu_j + C \sum_j \xi_j$, problem (2) can be solved by dealing with its dual problem:

$$\max W = \sum_j \phi(\mathbf{x}_j)^2 \beta_j - \sum_{i,j} \beta_i \beta_j \phi(\mathbf{x}_i) \phi(\mathbf{x}_j) \quad (3)$$

$$s.t. \quad 0 \leq \beta_j \leq C, \quad \sum_j \beta_j = 1, \quad j = 1 \dots N. \quad (4)$$

Support vectors are the points on the boundary of the sphere with $0 < \beta_j < C$. The trained kernel radius function is defined as $f(\mathbf{x}) := R^2(\mathbf{x}) = \|\phi(\mathbf{x}) - \mathbf{a}\|^2 = K(\mathbf{x}, \mathbf{x}) - 2 \sum_j \beta_j K(\mathbf{x}_j, \mathbf{x}) + \sum_{i,j} \beta_i \beta_j K(\mathbf{x}_i, \mathbf{x}_j)$, where the kernel function $K(\mathbf{x}_i, \mathbf{x}_j)$ substitutes the inner products $\phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j)$. In this work, we use the most commonly used kernel function, Gaussian kernel, which is defined as $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/2\delta^2)$.

(2) *Cluster labeling:* As the trained kernel radius function is topologically invariant, the level set of $f(\cdot)$ can be decomposed into several disjoint sets [10], $L_f(\hat{R}^2) := \{\mathbf{x} : f(\mathbf{x}) \leq \hat{R}^2\} = C_1 \cup \dots \cup C_p$, where $\hat{R} = R(\mathbf{x}_i), i = 1 \dots N$ is the radius in kernel space, p is the number of clusters determined by $f(\cdot)$, and $C_i, i = 1 \dots p$ are different clusters.

3 Experiment and Results

3.1 On Simulated Dataset

For the simulated dataset, the data processing scheme (cf. Fig. 1) is performed. After preprocessing, the t-test was used to reduce the dataset, which generated the result as shown in Fig. 3(a). As there are three different activation patterns in the simulated dataset, our target is to find four clusters: three contain three types of activated voxels, and the fourth is for the non-activated voxels. The performance of SVC is compared with those of crisp C-means, ICA [11], and FCM. For crisp C-means and FCM, we set the number of clusters to be four. ICA was constrained to generate three independent components. For the SVC, we set the kernel width σ to 0.2 and the regularization parameter C to 10^3 . Four clusters are detected by SVC. In FCM, the cluster label for each voxel is determined by

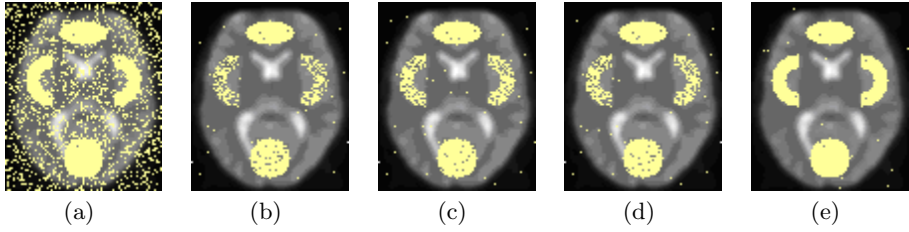


Fig. 3. Detection results from (a) after t-test (b) C-means (c) ICA (d) FCM (e) SVC

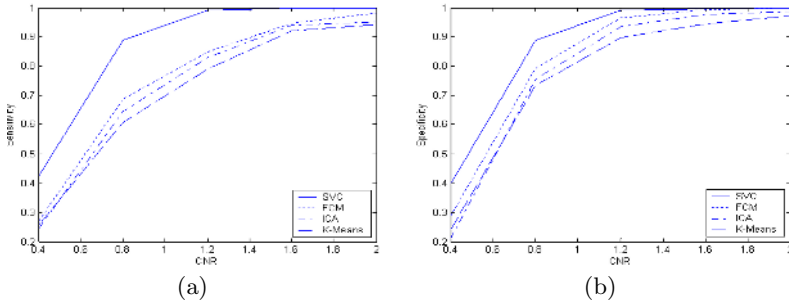


Fig. 4. Performance of C-Means, ICA, FCM, and SVC (a) Sensitivity (b) Specificity

finding the largest membership value; and in ICA, each voxel is assigned to the largest component. For the detected clusters, we find the "activated" clusters whose centroids are the most correlated ones to the stimulus paradigm.

Fig. 3 shows the detection results from C-means, ICA, FCM, and SVC, when CNR is set to 1.2. When CNR increases from 0.4 to 2, the sensitivity (the percentage of voxels correctly detected as activated) and specificity (the percentage of voxels correctly detected as non-activated) of the four methods are plotted in Fig. 4. One can find that the proposed SVC achieves both higher sensitivity and specificity than other methods under different CNR's.

3.2 On *In Vivo* Dataset

For the *in vivo* dataset, as we do not know the desired number of clusters, in order to achieve a set of finer clusters, we set the width parameter σ to 0.15 in SVC; while for FCM and C-means, the number of clusters is set to 30. Other settings are similar as discussed in 3.1. Fig. 5 shows that activated voxels on the 29th slice detected by C-means, FCM, ICA, and SVC cover the Brodmanns area (BA) 42 (primary auditory cortex) and BA 22 (auditory associated area) when they are superimposed on the accompanied structural data. But SVC detected more continuous activated regions, which are demonstrated to have stronger relationship with the stimulus (see Fig. 6(a)); however the voxels detected by C-means, but not by SVC, are shown to be almost irrelevant (see Fig. 6(b)).

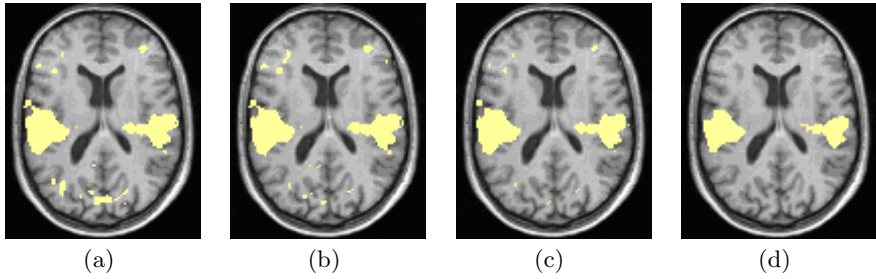


Fig. 5. The activated clusters detected by (a) C-means, (b) ICA, (c) FCM, (d) SVC

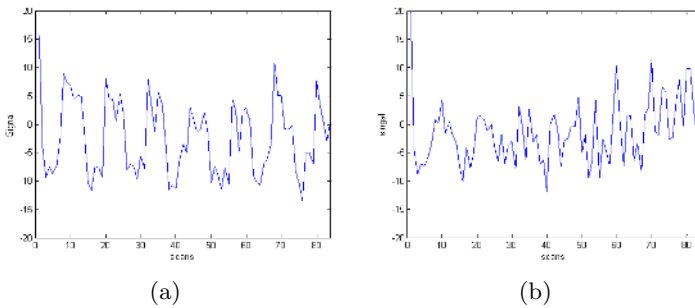


Fig. 6. The *in vivo* activation detection results (a) the averaged signal detected by SVC (b) The averaged signal detected by C-means but not by SVC

4 Discussions and Conclusion

The main contribution of this paper is finding an effective and reliable activation detection method, support vector clustering (SVC), which is free of the cluster shape assumption existing in most fMRI clustering approaches, and is able to remove the outlier points. The experimental results show the effectiveness of SVC in comparison with commonly used crisp C-means, FCM and ICA.

In the SVC algorithm that uses Gaussian kernel, the width parameter σ determines how fine the samples are clustered. Users can have a control over the clustering result by tuning this parameter: increasing σ will lead to a coarse clustering result, and vice versa. Empirical value for σ is in the range from 0.1 to 3. The regularization constant C affects the amount of outlier points, and hence influences the compactness of the generated clusters. Further research may be on how to automatically determine the values for the two parameters in fMRI activation detection by incorporating the field expert knowledge.

Acknowledgment

The work described in this paper was supported by Grant B-Q519 from the Hong Kong Research Grant Council and a grant from the Research Grants

Council of the Hong Kong Special Administrative Region, China (Project No. CUHK4223/04E).

References

1. Goutte, C., Toft, P., Rostrup, E., Nielsen, F.A., Hansen, L.K.: On clustering fmri time series. *NeuroImage* **9** (1999) 298–310
2. Chuang, K., Chiu, M., Lin, C.C., Chen, J.: Model-free functional mri analysis using kohonen clustering neural network and fuzzy c-means. *IEEE Trans. Medical Imaging* **18** (1999) 1117–1128
3. Baumgartner, R., Scarth, G., Teichtmeister, C., Somorjai, R., Moser, E.: Fuzzy clustering of gradient-echo functional mri in the human visual cortex. *Magnetic Resonance Imaging* **7** (1997) 1094–1101
4. Vapnik, V.: *The Nature of Statistical Learning Theory*. Springer-Verlag, New York (1999)
5. Ben-Hur, A., Horn, D., Siegelmann, H., Vapnik, V.: Support vector clustering. *Machine Learning Research* **2** (2001) 125–137
6. Purdon, P., Solo, V., Brown, E.M., Weisskoff, R.: Functional mri signal modeling with spatial and temporal correlations. *NeuroImage* **14** (2001) 912C923
7. Bandettini, P., Jesmanowicz, A., Wong, E., Hyde, J.: Processing strategies for time-course data sets in functional mri of the human brain. *Magnetic Resonance in Medicine* **30** (1993) 161–173
8. Lange, N., Zeger, S.: Non-linear fourier time series analysis for human brain mapping by functional magnetic resonance imaging. *Applied Statistics* **46** (1997) 1–29
9. Friston, K., Jezzard, P., Turner, R.: Analysis of functional mri time-series. *Human Brain Mapping* **2** (1994) 69–78
10. Lee, J., Lee, D.: An improved cluster labeling method for support vector clustering. *IEEE Trans. Pattern Analysis Machine Intelligence* **27** (2005) 461–464
11. Makeig, S.: *Eeg/ica toolbox for matlab* [online]. Available: <http://www.sccn.ucsd.edu/scott/ica.html> (2005, April)

Inter-frame Motion Correction for MR Thermometry

S. Suprijanto¹, M.W. Vogel³, F.M. Vos¹,
H.A. Vrooman^{2,3}, and A.M. Vossepoel^{1,2,3}

¹ Quantitative Imaging, Delft University of Technology,
2628 CJ Delft, The Netherlands

{supri, frans, albert}@qi.tnw.tudelft.nl

² Biomedical Imaging Group, Erasmus University Medical Center

³ Department of Radiology, Erasmus University Medical Center,
P.O. Box. 1738, 3000 DR Rotterdam, The Netherlands

{m.vogel, h.vrooman}@erasmusmc.nl

Abstract. Noninvasive temperature measurement is feasible with MRI to monitor changes in thermal therapy. Phase shift based MR thermometry gives an estimate of the relative temperature variation between thermal and baseline images. This technique is limited, however, when applied on targets under inter-frame motion. Simple image registration and subtraction are not adequate to recover the temperature properly since the phase shift due to temperature changes is corrupted by an unwanted phase shift. In this work, the unwanted phase shift is predicted from the raw registered phase shift map itself. To estimate the unwanted phase shift, a thin plate smoothing spline is fitted to the values *outside* the heated region. The spline value *in* the heated area serves as an estimate for the offset. The estimation result is applied to correct errors in the temperature maps of an ex-vivo experiment.

1 Introduction

Local thermal ablation is becoming an established surgical technique for destroying tumors and metastases. A particularly attractive method to monitor thermoablative treatment is by real-time MR thermometry based on the proton resonance frequency (PRF) shift. The accuracy of PRF shift has been demonstrated in immobile targets [1]. However, many targets for thermo-ablative interventions are in the abdomen, where motion is ubiquitous. Since a subtraction technique is used, errors in the temperature map may occur due to misregistration between thermal and baseline images. Additionally, magnetic flux is dependent on position, and any change in position will thus lead to a specific change in magnetic field that confounds the phase shift. Therefore, although misregistration can be solved using an image transformation, an unwanted phase shift map is not fully compensated by subtraction techniques.

Relatively few publications deal with the issue of inter-frame motion correction. A correction technique was proposed in [2, 3] to avoid errors in the

temperature maps due to accidental object motion. This method boils down to excluding the $(n)^{th}$ temperature map to avoid error progression if motion occurs between the $(n - 1)^{th}$ and the $(n)^{th}$ frame. To conserve temperature information when a periodic motion occurs, multiple baseline images acquired in several positions in the expected range of motion were used to generate temperature maps [4]. Those correction techniques [2, 3, 4] have been validated under the assumption that the largest amount of motion is along the static magnetic field (superior/inferior) direction. Moreover, a temperature map is calculated only if thermal and baseline images are scanned in the same location in order to avoid unwanted phase shift problems. Therefore, the fraction of reliable temperature maps in a series may be limited. In contrast with those correction schemes, a referenceless method is proposed in [5]. In this method, extrapolation of a surface polynomial fit to values outside the heated region serves as the baseline. The referenceless method may work well for slightly shifted images in which the heated region is always within some inner ROI frame. What is more, it is assumed that a local phase change due to the interface between the thermal applicator and tissues is insignificant. This condition may only be achieved if the selected scan-plane orientation is perpendicular to the long axis of the thermal applicator. To obtain an accurate prediction of thermal damage, however, a scan plane orientation parallel to the long axis of thermal applicator is suggested[9]. Thus, the local phase change cannot be neglected. In such a case, the referenceless method may fail to determine temperature maps accurately.

In this work, we propose a dedicated image registration scheme for MR thermometry. This scheme can be viewed as an extension of the multi-baseline approaches. An unwanted phase shift may arise if an image transformation is required to align thermal and baseline images. In order to solve for this, we explored a strategy to estimate the distribution of the phase shift offset over the image. Initially, a thermal image and a baseline image are registered to yield a *raw registered phase shift* map. Subsequently, we fit a thin plate smoothing spline to the values of outside the heated region. The offset is corrected for by subtracting the extrapolated spline values from the values inside the heated region. For validation, the method is tested to monitor temperature evolution during interstitial laser therapy in an ex-vivo sample of the liver under 2D-inter-frame motion.

2 Methods

2.1 Temperature Calculation

We define the raw registered phase shift as

$$\Delta\Psi = \arctan[I(\vec{r}, T).I\{I^*(\vec{r}_{bl}, T_o)\}] + \Delta\phi(\Gamma) \quad (1)$$

where $I(\vec{r}, T)$, $I(\vec{r}_{bl}, T_o)$ are thermal and baseline images, respectively. $\Delta\phi(\Gamma)$ is the offset phase shift and Γ is an image transformation that is required to align image coordinates between $I(\vec{r}, T)$ and $I(\vec{r}_{bl}, T_o)$ and $*$ is the complex conjugate operator.

A proper relative temperature map is determined by

$$\Delta T = \frac{\overline{\Delta\Psi}}{\gamma\alpha B_o T_E} \quad \overline{\Delta\Psi} = \Delta\Psi - \Delta\phi(\Gamma) \quad (2)$$

where $\overline{\Delta\Psi}$ is the corrected registered phase shifts, $\gamma(=42.58 \text{ MHz}\cdot\text{T}^{-1})$ is the gyromagnetic ratio, $\alpha(=0.01 \text{ ppm}/^\circ\text{C})$ is the temperature coefficient for aqueous tissue, T_E is the echo time, B_o is the main magnetic field [1].

To calculate $\Delta\Psi$, one must select the appropriate $I(\vec{r}_{bl}, T_o)$. However, the number of baseline images may be limited, and we cannot ensure that each $I(\vec{r}, T)$ corresponds exactly to the selected $I(\vec{r}_{bl}, T_o)$. If a significant transformation (Γ) is required to align both images, the term $\Delta\phi(\Gamma)$ may contribute to a spatial offset in the temperature maps. Clearly, this must be eliminated to avoid a systematic error. Therefore, a strategy to estimate $\Delta\phi(\Gamma)$ is proposed, and the result is used to obtain $\overline{\Delta\Psi}$.

2.2 Image Registration

Image registration is needed to determine $\Delta\Psi$. We have chosen a modified optical flow strategy to do so ([6]), because its formulation is elegant, and the resulting computations are relatively fast and non-feature based. Standard optical flow methods assume a constant brightness level over the image. The strategy described in [6] extends the method to cope with intensity variations, eventually yielding the six parameters of an affine transformation model.

Aligning temperature maps to a fixed target map is useful for a clinical application. By doing so, the temperature is monitored with respect to a steady location even though the heated region may move between frames. To do so, we choose one of the baseline images as a fixed target. The transformation parameters among the set of baseline images enables aligning the raw phase shift map to the fixed target.

2.3 Estimation of the Phase Shift Offset

$\Delta\phi(\Gamma)$ may arise due to spatial magnetic field deviations that are not fully compensated by registered image subtraction. It can be expected that the distribution of $\Delta\phi(\Gamma)$ is a continuous function and that a pattern of $\phi(\Gamma)$ may vary relative to the center of imaging volume.

In the heated region, the registered raw phase shift may be corrupted by $\Delta\phi(\Gamma)$, resulting in errors in the temperature map. Fortunately, in thermal therapy, usually only a small fraction of the image region is affected by temperature change. Under this assumption, the raw phase shift values outside the heated region may be used to estimate the phase shift offset in the heated region. In the current implementation, first, a tissue region that covers both the non-heated region and the heated-region is defined. This region is denoted by \mathfrak{R}_{high} . Subsequently, the predicted area of the heated region (\mathfrak{R}_{hot}) is selected (see e.g.[9]). Finally, a ROI frame *around* the area to be heated (\mathfrak{R}_{ROI}) is defined by subtracting \mathfrak{R}_{hot} from \mathfrak{R}_{high} . Practically, the highest signal values in the intensity

images can be used to guide the definition of \mathfrak{R}_{high} . This is to increase the reliability of the offset calculation (not letting it to be disturbed by noise values). A threshold may be employed to do so.

Ideally \mathfrak{R}_{hot} is a heated region for which the temperature at the boundary is equal to the background temperature. It is important to determine \mathfrak{R}_{ROI} outside \mathfrak{R}_{hot} , since a temperature change within \mathfrak{R}_{ROI} will confound the extrapolation. To anticipate the maximum area of \mathfrak{R}_{hot} , the appropriate \mathfrak{R}_{hot} may be calculated via a characteristic thermal response of laser irradiated tissue under ideal conditions (e.g. neglecting a perfusion factor). To estimate the offset phase shift in \mathfrak{R}_{hot} , the fitted phase shift is extrapolated to the center of \mathfrak{R}_{ROI} . We assume that a *thin-plate smoothing spline* is appropriate to estimate $\Delta\phi(T)$ in \mathfrak{R}_{hot} .

Thin-Plate Smoothing Spline. Thin-plate splines are members of a family of surface splines that result from the solution of a variational problem[7]. Given the set of raw phase shifts in \mathfrak{R}_{ROI} $(x_i, y_i, \Delta\Psi_i)$, $i = 1, \dots, m$, we now consider the approximating thin-smoothing spline f , that takes the form

$$f(x, y) = \sum_{i=1}^m a_i K_i(x, y) + a_{m+1} + a_{m+2}x + a_{m+3}y, \quad (3)$$

where $K_i(x, y) = \frac{1}{8\pi} r_i^2(x, y) \log(r_i(x, y))$ with $r_i(x, y) = \sqrt{(x - x_i)^2 + (y - y_i)^2}$ ($K_i = 0$ when $r_i = 0$).

The thin-plate smoothing spline f is the unique minimizer of the weighted sum, given by $S(\lambda) = \frac{1}{m} \sum_{i=1}^m (\Delta\Psi_i - f)^2 + \lambda J_2(f)$. The first term measures the goodness of fit and the second term measures the smoothness associated with f . The smoothness of a fit is measured by the integrated squared second derivative $J_2(f) = \int \int ((\frac{\partial^2 f}{\partial x^2})^2 + 2(\frac{\partial^2 f}{\partial x \partial y})^2 + (\frac{\partial^2 f}{\partial y^2})^2) dx dy$ where f belongs to the set of all continuously differentiable functions with square integrable second derivatives and λ is a positive constant. To determine an optimal smoothing parameter λ , the method of generalized cross validation was applied. The spline parameters a_i , $i = 1, \dots, m + 3$ are determined as described in [7].

3 Materials

To generate inter-frame motion of a sample porcine liver, a custom made, automated and air-pressure driven table was used in the experiments. The table moved along (S/I) and perpendicular to (R/L) the magnet's main field direction (see Fig-1). The maximum range of the motion device is about 80 mm in both directions. Over a range of 80 mm, 25 evenly spaced positions were defined. A fiber optics laser catheter was inserted into a large sample of porcine liver. We triggered the MR imaging when the sample was held stationary. MR imaging was performed on a 1.5T GE scanner. For thermal imaging, we used an echo shifted gradient echo sequence ($\theta/T_E/T_R = 15^\circ/18ms/12ms$) with 1.25 mm in-plane resolution, matrix 256x256, a slice thickness of 5 mm and a 3'' surface coil.

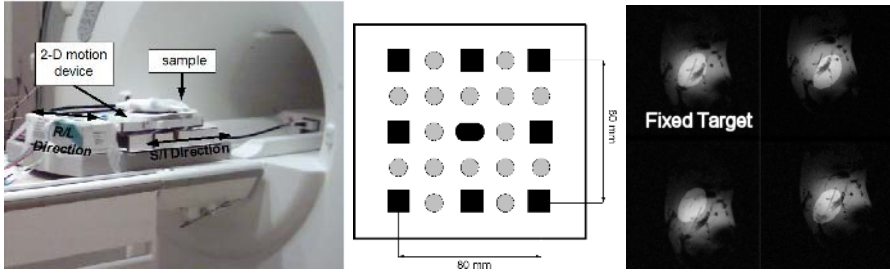


Fig. 1. (left) The custom made air-pressure driven table. (center)The approx. 25 different points (with a variation of ± 5 mm). (right) Examples of ex-vivo liver images placed on the 2D motion device.

Non Heating Experiment. Before heating started, the set of baseline images was selected by controlling the motion device to hold on the approximate positions that are marked with black squares and the ellipse in Fig-1. As the fixed target map, the baseline image scanned in the approximate central position (marked by the ellipse in Fig-1) was selected. After that, we acquired a series of images in various locations in about 6 minutes.

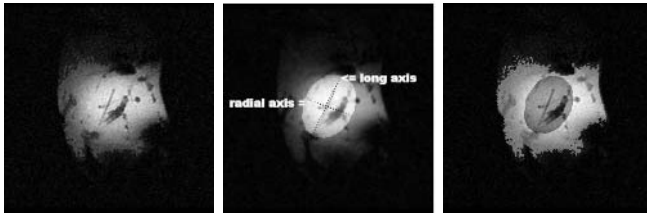


Fig. 2. Intensity image of the interstitial laser applicator inserted into a sample of liver (left). The predicted area of \mathcal{R}_{hot} (middle). The tissue region defined as \mathcal{R}_{ROI} (right).

To validate the performance of the thin plate smoothing splines, we attempted to extrapolate $\Delta\phi(\vec{T})$ in the region predicted as \mathcal{R}_{hot} using the value of $\Delta\Psi$ on the location outside of the heated region \mathcal{R}_{ROI} . To define \mathcal{R}_{ROI} , we, first defined a region of high SNR phase shift \mathcal{R}_{high} . It was guided using the signal in the intensity images. Next, the shape of \mathcal{R}_{hot} was defined under the assumption that the laser-induced thermal areas is roughly ellipsoidal[9]. We define the length of long axis the \mathcal{R}_{hot} as about 125 mm, 45 mm longer than the length of laser tip appearing in the intensity image. Next, the length of radial axis is about 90 mm (See Fig-2-middle). Finally, the \mathcal{R}_{ROI} is defined as the subtraction of \mathcal{R}_{high} and \mathcal{R}_{hot} (See Fig-2-right).

Heating Experiment. Using a similar setup as in the non-heating experiments, we started the laser ablation (wavelength of 1024 nm, 20 W) to the liver sample,

inducing focal heating at the diffuser tip. Images were acquired for about 14 minutes under heating and inter-frame motion.

4 Results

4.1 Non-heating Experiment

Examples of the raw ΔT_{back} (without correction) in \mathfrak{R}_{hot} for different positions are shown in Fig-3. Notice that the outcome basically shows the offset maps, since the temperature is constant. Only with translation parameters $[0,0]$ mm, the value is closest to the ideal offset of $\Delta T_{back}(0^\circ C)$. A planar slope and a non-linear surface of offset-raw ΔT_{back} can be seen in Fig-3 for other translation parameters.

From 8 different selected translation parameters, we tabulate the values of min., max. and mean square errors (MSE)(with respect to $0^\circ C$) of the raw ΔT_{back} in \mathfrak{R}_{hot} . The results are shown in Table-1. Next, the estimation of the offset value in \mathfrak{R}_{hot} obtained by extrapolation using the thin plate smoothing spline is validated. The estimated offset was subtracted from the raw ΔT_{back} . Again, we calculated value of min., max. and MSE of the corrected ΔT_{back} . The results are also tabulated in the Table-1.

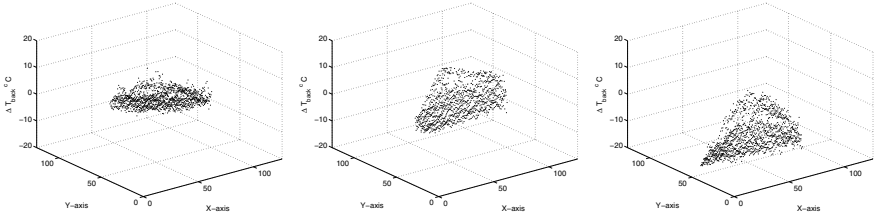


Fig. 3. Examples of the offset in raw (uncorrected) ΔT_{back} in \mathfrak{R}_{hot} for different translation parameters (in mm). The translation parameters (in mm) from left to right are $[0,0]$, $[10,0]$ and $[10,-12.5]$, respectively.

Table 1. Comparison of raw- and corrected background temperature (ΔT_{back})

Trans.(mm)	raw ΔT_{back}			corrected ΔT_{back}		
	Min	Max	MSE	Min	Max	MSE
$[0, 0]$	$-2.07^\circ C$	$1.97^\circ C$	$0.92^\circ C$	$-0.73^\circ C$	$0.75^\circ C$	$0.35^\circ C$
$[10, 0]$	$-12.31^\circ C$	$11.15^\circ C$	$4.71^\circ C$	$-1.09^\circ C$	$0.56^\circ C$	$0.55^\circ C$
$[-10, 0]$	$-12.12^\circ C$	$11.32^\circ C$	$4.90^\circ C$	$-0.61^\circ C$	$1.42^\circ C$	$0.82^\circ C$
$[0, 12.5]$	$5.21^\circ C$	$25.13^\circ C$	$15.13^\circ C$	$-0.72^\circ C$	$1.18^\circ C$	$0.92^\circ C$
$[0, -12.5]$	$-22.74^\circ C$	$-2.71^\circ C$	$14.92^\circ C$	$-0.81^\circ C$	$1.55^\circ C$	$0.81^\circ C$
$[-10, 12.5]$	$7.93^\circ C$	$28.40^\circ C$	$15.21^\circ C$	$-0.31^\circ C$	$0.53^\circ C$	$0.94^\circ C$
$[10, -12.5]$	$-23.62^\circ C$	$-3.10^\circ C$	$13.10^\circ C$	$-0.32^\circ C$	$0.85^\circ C$	$0.56^\circ C$
$[-10, -12.5]$	$-21.61^\circ C$	$-4.04^\circ C$	$14.11^\circ C$	$-0.62^\circ C$	$0.95^\circ C$	$0.74^\circ C$

4.2 Heating Experiment

After heating started, a series of images was acquired from an object under frame-to-frame motion as described above. Furthermore, the raw- and corrected- ΔT overlays in the fixed target maps are shown in Fig-4, and the temperature profile in the defined \mathfrak{R}_{high} along an arbitrarily chosen horizontal line during heating at $t_1=5.4, t_2=5.6$ and $t_3=5.8$ minutes are shown in Fig-5. We can see that without correction, the raw ΔT profiles become unrealistic, since a longer duration of laser irradiation (t_3) yields a lower ΔT profile than the shorter duration (t_1 and t_2). We also tested the performance of our correction scheme to calculate ΔT over time using the whole series of images during inter-frame motion and heating. We chose averages of 3×3 voxels in the point indicated with an arrow

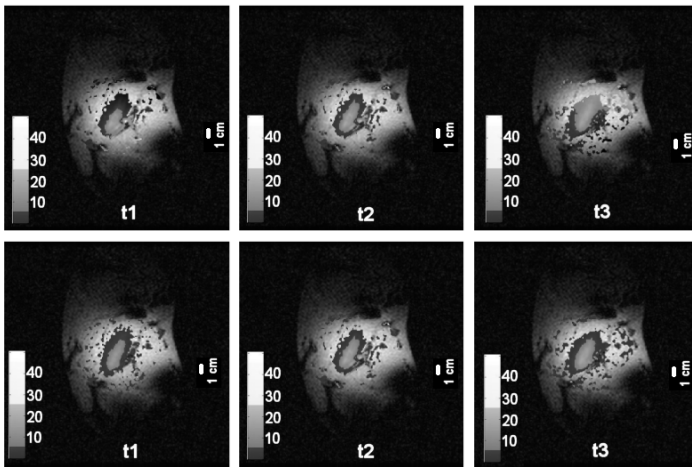


Fig. 4. Comparison of raw ΔT (top row) and corrected ΔT (bottom row) during heating at $t_1=5.4, t_2=5.6$ and $t_3=5.8$ minutes. One can see in the bottom images that the hot area in the center is enclosed by a colder "ring". This distinction cannot be made in the top images (specifically at t_1 and t_3).

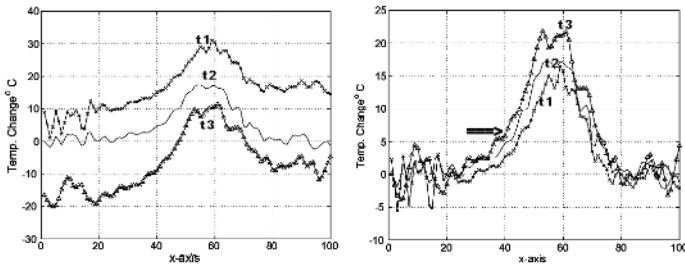


Fig. 5. Profiles in \mathfrak{R}_{high} along a horizontal axis of raw ΔT (left) and corrected ΔT (right) during heating at $t_1=5.4, t_2=5.6$ and $t_3=5.8$ minutes

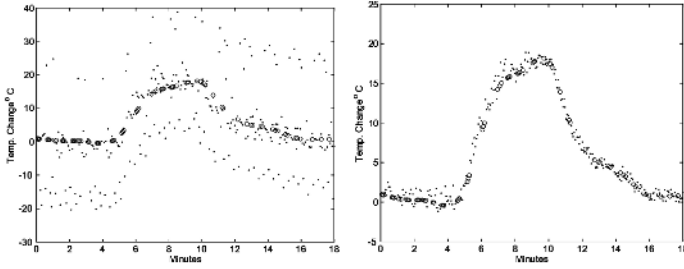


Fig. 6. The raw(left) and corrected(right) ΔT evolution. Each \cdot corresponds to an acquired image in which the temperature is determined in a single point. It may seem that two points coincide, but there is always an (albeit) small time interval). When the absolute translation along both axes was < 1 mm, ΔT is marked with 'o', indicating the reference standard.

in Fig-5. The results are shown in the Fig-6. For *reference*, the 'true' ΔT was determined when absolute translation in both axis was less than 1 mm (then the images are aligned and there is no offset, see Fig-3). An exponential function was fitted to the reference result (MSE= $0.13^{\circ}C$, $R^2=0.98$). We found that the mean difference and corresponding standard deviation between uncorrected temperature elevation and reference was $-2.13^{\circ}C \pm 10.71^{\circ}C$. The corrected temperature elevation yielded a mean difference and standard deviation of $0.88^{\circ}C \pm 1.05^{\circ}C$, indicating a significantly improved precision.

5 Conclusions

The results showed that inter-frame motion correction is feasible. Extrapolation of a thin plate spline fit to the values outside the heated region served as an estimation of the offset ΔT_{back} in the heated region. Experimental results showed that $\Delta\phi(I)$ is a continuous function and that $\phi(I)$ varies relative to the center of the imaging volume.

We registered the raw phase shift maps to a fixed target in order to always have a temperature map collocated with the same reference image (although the actual heated area actually moves between frames).

To accurately define \mathfrak{R}_{hot} , model calculations of laser-induced tissue coagulation using Monte Carlo Simulation (MCS) [9] can be applied.

We demonstrated that a translation in S/I and R/L direction >10 mm can induce an maximum offset in temperature change of $>10^{\circ}C$. As long as the heating region can be maintained in the center of \mathfrak{R}_{ROI} , the offset can be eliminated by our correction scheme.

Considering the human liver as a target of thermal therapy, the average inter-frame liver motion is 13 mm (range of 5-17 mm) under normal respiration [8]. In many cases these values refer to the S/I direction. However, depending on the tumor location in the liver, the liver motion in other directions often can not be

neglected. Liver motion of less than 2 mm in other directions have been reported under normal respiration conditions [8].

The challenge for future research is to apply the method to in-vivo acquired liver images. The proposed scheme has the potential to work together with physical techniques such as MR active tracking or navigator echoes. Using such techniques, an appropriate baseline selection is relatively straightforward since the location is implicitly measured. The translation values provided by physical techniques may also be used as an initial value for the image registration algorithms. The undesired offset due to the image transformation can be eliminated by estimation of the background offset as described here.

References

1. Wlodarczyk W, Hentschel M, Wust P, Noeske R, Hosten N, Rinneberg H, Felix R. *Comparison of four magnetic resonance methods for mapping small temperature changes*. Phys Med Biol. 1999 Feb;44(2):607-624.
2. de Zwart JA, Vimeux FC, Palussiere J, Salomir R, Quesson B, Delalande C, Moonen CT. *On-line correction and visualization of motion during MRI-controlled hyperthermia*. Magn Reson Med. 2001 Jan;45(1):128-137.
3. de Senneville BD, Desbarats P, Salomir R, Quesson B, Moonen CTW. *Correction of Accidental Patient Motion for Online MR Thermometry* MICCAI 2004, LNCS Vol. 3217:637-644.
4. Vigen KK, Daniel BL, Pauly JM, Butts K. *Triggered, navigated, multi-baseline method for proton resonance frequency temperature mapping with respiratory motion*. Magn Reson Med. 2003 Nov;50(5):1003-1010.
5. Rieke V, Vigen KK, Sommer G, Daniel BL, Pauly JM, Butts K. *Referenceless PRF shift thermometry*. Magn Reson Med. 2004 Jun;51(6):1223-1231.
6. Periaswamy S, Farid H. *Elastic registration in the presence of intensity variations*. IEEE Trans Med Imaging. 2003 Jul;22(7):865-874.
7. P. Dierckx. *Curve and Surface Fitting with Splines*, Oxford Science Publications, 1993.
8. Langen KM, Jones DT. *Organ motion and its management*. Int J Radiat Oncol Biol Phys. 2001 May 1;50(1):265-278.
9. Puccini S, Bar NK, Bublat M, Kahn T, Busse H. *Simulations of thermal tissue coagulation and their value for the planning and monitoring of laser-induced interstitial thermotherapy (LITT)*. Magn Reson Med. 2003 Feb;49(2):351-62.

Adaptive Multiscale Ultrasound Compounding Using Phase Information

Vicente Grau and J. Alison Noble

Wolfson Medical Vision Laboratory, Department of Engineering Science,
University of Oxford, Parks Road, Oxford OX1 3PJ,
{vicente, noble}@robots.ox.ac.uk

Abstract. The recent availability of real-time three-dimensional echocardiography offers a convenient, low-cost alternative for detection and diagnosis of heart pathologies. However, a complete description of the heart can be obtained only by combining the information provided by different acoustic windows. We present a new method for compounding 3D ultrasound scans acquired from different views. The method uses multiscale information about local structure definition and orientation to weight the contributions of the images. We propose to use image phase to obtain these image characteristics while keeping invariance to image contrast. The monogenic signal provides a convenient, integrated approach for this purpose. We have evaluated our algorithm on synthetic images and heart scans from volunteers, showing it provides a significant improvement in image quality when compared to traditional compounding methods.

1 Introduction

Real-Time 3D Echocardiography. Several imaging modalities (X-ray, angiography, MRI, 2D echocardiography) have been proven to be useful in cardiology. However, echocardiography presents unique characteristics that make it the most commonly applied imaging diagnostic technique in clinical practice. Traditionally, two-dimensional echocardiography has been used widely as a relatively cheap, portable and real-time interactive assessment of heart function. Using a 2D imaging modality introduces disadvantages such as the difficulty to characterize three-dimensional structures or the dependence on the probe position, which makes it difficult to compare images if acquired by different clinicians or at different times.

Recently developed technology has allowed, for the first time, to obtain three-dimensional images of the heart in real time. This new imaging modality opens a wide range of possibilities for echocardiography in clinical routine. However, at its present stage, due to the limited field of view of the transducer, it is not possible to scan the whole heart in a single acquisition (and, in some cases, not even the left ventricle). Furthermore, some cardiac structures can be appreciated only from particular acoustic windows. For this reason, a complete diagnostic study in real time 3D ultrasound (RT3DUS) will consist of several acquisitions obtained from different positions. The development of tools to combine these acquisitions and present a single, optimal dataset to the clinician could greatly improve the clinical uses of the technology.

Compounding Ultrasound Scans. Compounding ultrasound images acquired at different times or from different locations is a common way to improve image quality, mainly through speckle reduction. Though using directly RF data for compounding is also possible, these data are not usually available in the clinical environment, so our research focuses on compounding intensity images. Usually, simple techniques, such as the mean or the maximum intensity, are used to combine the images. Little work has been done to explore alternative, more sophisticated compounding techniques. In [1], the authors propose to exploit the dependence of echo intensity on the angle of incidence of the ultrasound beam on a reflecting surface. Resulting images were significantly enhanced when compared to using the mean value to combine images. However, that method is based on an initial fitting of a surface to the data. This makes it difficult to apply to complex images, such as the cardiac ones, in which several structures are present and segmentation is a challenge.

Phase-Based Image Analysis Using the Monogenic Signal. Phase-based analysis has been proposed as an alternative to intensity for many image processing tasks [4]. Phase provides invariance to changes in brightness and contrast within the image: this property makes it particularly fit for ultrasound images, in which beam attenuation is present and echo intensity depends on the angle of incidence of the ultrasound beam. Local phase is usually calculated by combining the output of a set of filters with different angles. In [2], the authors introduce a new approach to calculate phase in n -dimensional signals: the monogenic signal, an isotropic extension of the analytic signal which preserves its basic properties. Analogous to the analytic signal for 1-D, to build the monogenic signal the Riesz transform, a generalization of the Hilbert transform for higher dimensional signals, is used. In the frequency domain, the Riesz transform of an N -dimensional image can be calculated by multiplying the original image with the filters H_i :

$$H_i(\mathbf{u}) = \frac{u_i}{|\mathbf{u}|}, \quad (1)$$

where $\mathbf{u}=[u_1 \dots u_N]^T$, with u_i representing the i th coordinate unit vector; there are thus as many filters as image dimensions. The monogenic signal is the combination of the original signal and the Riesz transform, so it assigns an $N+1$ -dimensional vector to each spatial location. The angles defined by this vector represent the local phase and structure orientation, respectively. A complete explanation of the mathematical foundations and the equations needed to calculate phase and orientation from the monogenic signal can be found in [2].

In practice, it is interesting to localize features both in space and in frequency. To achieve this purpose, band-pass filters $G_s(u)$ are combined with the $H_i(u)$ filter in (1). A comprehensive analysis of different alternatives for band-pass filters was done in [3].

In this paper, we propose a novel compounding algorithm for ultrasound images in general, and for multi-view, RT3DUS echocardiography images in particular. The aim of the compounding algorithm is to maximize the information content of the combined images, for which we preserve those images which higher feature saliency; furthermore, we assign weights depending on the incidence angle of the ultrasound beam. We explore the use of phase to obtain information about both feature saliency and orientation: the monogenic signal constitutes a convenient

framework for this purpose. Preliminary results on both phantom and real-time 3D cardiac images are presented, showing significant improvement over common compounding techniques.

2 Compounding Using Feature Saliency and Orientation

Image Registration. Registering echocardiographic images obtained from different windows is a non-trivial task. Both temporal and spatial registration are required. In this paper, however, we focus on the combination of images, so we assume they have been previously registered with enough accuracy. For the algorithm evaluation presented in Section 3, we performed an initial rigid registration by aligning manually selected landmarks (the valve rings) followed by a manual fine correction of the result. For the temporal registration, we identified corresponding frames in the sequences, and performed a linear interpolation between them.

Compounding Algorithm. In order to maximise the information content of the compound image, we need to keep the features which appear only in a subset of the original N images, without degrading these features by averaging them with the rest. Furthermore, it is well known that the backscattered energy, and thus the ultrasound image appearance, depends on the incidence angle (the angle formed by the ultrasound beam and the normal to the surface at the incidence point). Averaging the intensities obtained from two different views is not an optimal solution, as it would degrade the strong echoes generated by small incidence angles by introducing weaker echoes from more oblique incidences. Thus, where several images contain significant features, we should combine them giving preference to the ones in which the normal to the detected feature is better aligned with the ultrasound beam. For this purpose, we define two measures, $S_i(r)$ and $M_i(r)$, $0 \leq S_i(r) \leq 1$, $0 \leq M_i(r) \leq 1$, which quantify how important the information contained by image i at position r is, and how well aligned features at this location are with the ultrasound beam, respectively. The way we calculate S_i and M_i in this study is described below. We can then define the following rules to identify the characteristics of an image point, and thus select the best strategy for combining the values:

- If S_i values are low in all images but one, the value from this image is adopted.
- If S_i values are high in more than one image, the values in the high S_i images are combined according to their respective alignment values M_i (those images in which the point is better aligned will contribute a higher amount to the combination).
- If S_i values in all images are low, the point is treated as pure speckle, so the average value is taken.

Here, the S_i value is treated as the primary condition, and only in the case that it is not sufficient to determine the intensity value by itself, the alignment values M_i are considered. We can transform the *if-then* conditions described above into equations to calculate the weight that each image should have in the final compound image, if we treat S_i and M_i as probabilities of having a significant feature and having a well-

aligned feature, respectively. This can be done in a multiscale way by calculating multiscale decompositions of the images, applying the compounding process individually to each scale, and finally recombining the compound multiscale images. In the following, for simplicity, we have omitted scale subindices. As an example, we show the weighting factor for image 1, λ_1 , in the case that three images are combined:

$$\begin{aligned} \lambda_1 = & \left[S_1 \bar{S}_2 \bar{S}_3 + S_1 S_2 \bar{S}_3 M_1 \bar{M}_2 + S_1 \bar{S}_2 S_3 M_1 \bar{M}_3 + S_1 S_2 S_3 M_1 \bar{M}_2 \bar{M}_3 \right] + \dots \\ & \dots + \frac{1}{2} \left[S_1 S_2 \bar{S}_3 M_1 M_2 + S_1 \bar{S}_2 S_3 M_1 M_3 + S_1 S_2 S_3 M_1 (M_2 \bar{M}_3 + \bar{M}_2 M_3) \right] + \dots \quad (2) \\ & \dots + \frac{1}{3} \left[S_1 S_2 S_3 (M_1 M_2 M_3 + \bar{M}_1 \bar{M}_2 \bar{M}_3) \right] + \frac{\alpha}{3} \left[\bar{S}_1 \bar{S}_2 \bar{S}_3 \right] \end{aligned}$$

with $\bar{S}_i = (1 - S_i)$. Note that S_i, M_i are not Boolean variables, but continuous probability values.

We can intuitively interpret the four main terms in (2). The first three terms in square brackets correspond, respectively, to the probabilities of having just one, two or three of the images (always including image 1) containing significant, well-aligned features. Finally, the last term represents the probability of having no structural information (i.e., pure speckle) at this point. We have introduced here a coefficient $0 \leq \alpha \leq 1$ which can be used for noise reduction. Note that $\alpha=1$ corresponds to an averaging of all images (the same effect produced by average compounding, but in this case applied only to regions with no feature information), while $\alpha=0$ produces a total elimination of detected speckle in the compound image. It is easy to generalize the above equation for any number of images. In our algorithm, a coefficient λ_{is} would be calculated for each image i and for each scale s , using alignment coefficients M_{is} . Finally, the compound image is calculated by combining all scales.

Using phase to calculate the feature significance and alignment parameters. The evolution of a phase measure across spatial scales can be used as a clue to differentiate image features from noise. One of the possibilities that have been proposed for this purpose is phase congruency [4].

$$PC(\bar{r}) = \max_{\bar{\varphi}(\bar{r}) \in [0, 2\pi]} \frac{\sum_s A_s(\bar{r}) \cos(\varphi_s(\bar{r}) - \bar{\varphi}(\bar{r}))}{\sum_s A_s(\bar{r})} \quad (3)$$

where A_s represents the amplitude of the s -th Fourier component of the signal, and φ_s is the local phase of the Fourier component at position \bar{r} . This parameter quantifies phase change over different scales. In [5], an alternative way to calculate the phase congruency was given. Phase congruency quantifies how phase varies across scales: a high value corresponds to a consistent phase value and is thus an indicator of a strong image feature. The utility of phase congruency to detect edges and lines in ultrasound images has been demonstrated before [6]. If the adequate scales are selected, phase congruency can be used to detect larger-scale structures, such as the ventricular walls in echocardiographic images. We can thus use $S_i(\bar{r}) = g(PC_i(\bar{r}))$ in Equation 2, where $g(\cdot)$ can be any monotonic function with $g(0)=0$ and $g(1)=1$.

As described in Section 1, the monogenic signal can be also used to calculate the orientation of image features. The incidence angle at scale s is defined as the angle between the ultrasound beam and the local image orientation at scale s , θ_s . We define the parameter $M_{is} = \cos(\theta_{is} - \varphi)$, where φ is the angle of the ultrasound beam.

To construct the multiscale image description required, many alternatives for band-pass filter selection exist [3]. We have chosen the Difference of Gaussians (DoG), which allows us to easily recombine the band-pass images. The filter bank is formed by S filters: $S-2$ band-pass filters, plus a low-pass filter and a high-pass filter which are used to generate a complete representation of the original image. All band-pass filters in the bank are constructed with an equal bandwidth value. In the experiments presented, we have found $S=5$ to be an adequate value.

3 Results and Discussion

Tests Using a Synthetic Phantom. Simulated images were generated using the Field II program [7]. An elliptical ring phantom was generated and scanned using a simulated 5 MHz sector probe. In Fig. 1, the results of compounding two images, acquired with a difference of 80° , are shown. Improvement of contrast and better edge definition, when compared to intensity averaging, can be clearly appreciated. We calculated contrast to noise ratio (CNR) as the difference between the mean values of the ring and the background, divided by the standard deviation of background intensity. CNR obtained with our method is 37% higher than with intensity averaging. An important indicator of the quality of our method is its effect on the magnitude and direction of the intensity gradient at the object contours: this is a crucial parameter for edge-based segmentation methods. We have calculated the intensity magnitude gradient in the direction normal to the ring contour. The increase in this parameter in our method, compared to intensity averaging, is shown in Fig 1(h): where the differences in alignment are high, increases of more than 30% are obtained.

Tests on Three-Dimensional Real-Time Echocardiographic Images. Finally, we applied our algorithm to RT3DUS images of the heart. Fourteen datasets were obtained by scanning two volunteers from both the apical and the parasternal windows. Images were registered and combined as explained in Section 2. Results from one of the volunteers are shown in Fig 2. When compared to intensity averaging (Fig 2(b)), our method with $\alpha=1$ (Fig 2(e)) shows a superior definition of significant heart structures. We have used smaller values of α to test the speckle reduction behaviour of the algorithm in a qualitative way: in Fig 2(c), where $\alpha=0.9$, speckle is reduced without a decrease in contrast in the important features. Finally, with $\alpha=0.6$ (Fig 2(f)), the most salient features are predominantly kept. The reader can better appreciate the improvement in the time sequences available online at <http://www.robots.ox.ac.uk/~mvl/htdocs/research/gallery.html>.

Discussion. We have presented a method for compounding ultrasound images acquired from different views. The novel aspect of our work is that, unlike the traditional methods used for compounding (average, maximum,...) we introduce information about the structural content of the image and its orientation. We also propose a particular way to calculate this information, based on the image phase, and implement it in an

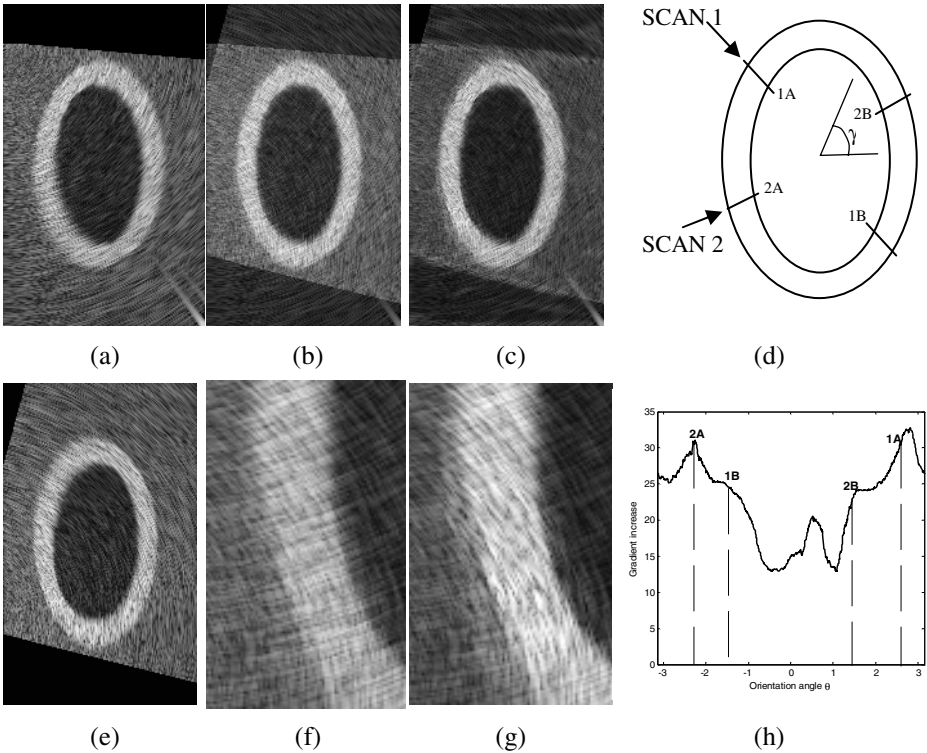


Fig. 1. Two views of a synthetic phantom (a, e). Comparison of the compounding results using intensity averaging (b) and our algorithm (c). Magnified views of a smaller region are presented in (f), for the intensity averaging, and in (g) for our algorithm. Note the contrast increase and noise reduction. (d): schematic depiction of the location and orientation of the probe, relative to the scanned ring, in both scans. Points where the ultrasound beam is approximately normal to the ring are marked (1A, 1B, 2A, 2B). (h): Increase in gradient magnitude at the ring contour with our method, when compared to intensity averaging. The horizontal coordinate represents the orientation angle γ . As expected, the maximum increase is obtained where the beam is approximately perpendicular to the ring.

integrated way by using the monogenic signal. The use of phase provides the benefit of contrast invariance, which can be key in ultrasound image processing.

An additional advantage of the presented technique is that the framework we propose is independent on the actual selection of the functions that quantify structural information and orientation. For other applications, it would be possible to introduce alternatives, while keeping the main ideas from our work. Particularly, in some cases the successful use of phase congruency for the estimation of feature significance can be quite dependent on the selected scales and the choice of function $g(S_i)$. Other techniques might prove to be more robust than phase congruency for this purpose.

The introduction of a constant α in (2) allows to control the amount of speckle reduction depending on the application. For visual diagnosis, it can be dangerous to remove information from the original image, as important information could be there;

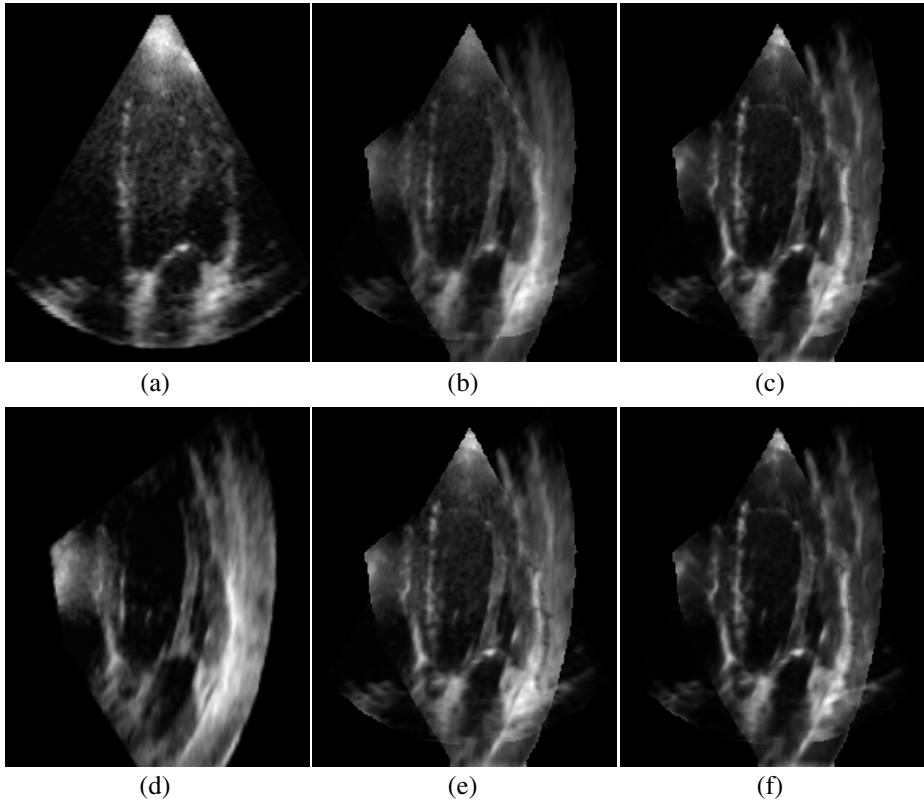


Fig. 2. Application of our method to RT3DUS images. (a): original apical image. (d): Original parasternal image. (b): Intensity averaging. (e): Our method with $\alpha=1$. Note the enhancement of the vertical structure on the right and left sides of the image. (c): Our method with $\alpha=0.9$: apart from the structure enhancement, a significant noise reduction is obtained. (f): Our method with $\alpha=0.6$: a strong noise reduction is applied without loss of structure definition. Differences are better appreciated in the animations provided online as additional material.

while for automatic segmentation algorithms, a drastic reduction of the speckle content can be advisable. Another possibility would be to keep the speckle removed from one image (the last term in (2)) and display it separately: significant information about aspects such as motion could be obtained from it. Planned clinical studies will provide us with more insight on this.

The band-pass filters used in the method have to be selected, depending on the size of the structures we want to detect. For applications in which the basic structures are always similar (e.g., echocardiography), filter frequencies can be kept unmodified for all patients.

Though registration has not been addressed in the present work, it is a fundamental task and the results will depend dramatically on the registration accuracy. Initially, as the images are acquired from the same patient with a few minutes between them, we expect rigid registration to be accurate enough for our purposes; however, patient

movement or breathing may prove to produce heart deformation that would make an elastic registration step necessary.

Preliminary results on simulated and real ultrasound images show a significant improvement, both in visual quality and quantitative measurements. Further validation, including segmentation algorithms applied on the original and combined images, is currently under way in our group.

Acknowledgments

The authors wish to thank Drs. Harald Becher and Carla Lofiego, from the John Radcliffe Hospital in Oxford, for providing the datasets used in this study. Funds were provided by the EPSRC, grant GR/S72801.

References

1. Leotta, D.F., Martin, R.W., "Three-dimensional spatial compounding of ultrasound scans with weighting by incidence angle", *Ultrasonic Imaging* 22, 1-19 (2000)
2. Felsberg, Sommer, "The monogenic signal", *IEEE Transactions on Signal Processing*, 49(12):3136-3144, (2001)
3. Boukerroui, D., Noble, J.A., Brady, M., "On the choice of band-pass quadrature filters", *Journal of Mathematical Imaging and Vision*, 21:53-80 (2004)
4. Morrone, M.C., Owens, R.A., "Feature detection from local energy", *Pattern Recognition Letters*, 6, 303-313 (1987)
5. Kovessi, P., "Phase congruency: a low level image invariant", *Psych Res* 64: 136-148 (2000)
6. Mulet-Parada, M., Noble, J.A., "2D+T Acoustic Boundary Detection in Echocardiography", *Medical Image Analysis*, 4(1):21-30 (2000)
7. Jensen, J.A. "Field: A Program for Simulating Ultrasound Systems", *Medical & Biological Engineering & Computing* 34: 351-353 (1996)

3D Freehand Ultrasound Reconstruction Based on Probe Trajectory

Pierrick Coupé, Pierre Hellier, Noura Azzabou, and Christian Barillot

Project VisAGeS, IRISA - INRIA - INSERM,
IRISA campus Beaulieu 35042 Rennes Cedex, France
{pcoupe, phellier, cbarillo}@irisa.fr, azzabou@certis.enpc.fr
<http://www.irisa.fr/visages>

Abstract. 3D freehand ultrasound imaging is a very attractive technique in medical examinations and intra-operative stage for its cost and field of view capacities. This technique produces a set of non parallel B-scans which are irregularly distributed in the space. Reconstruction amounts to computing a regular lattice volume and is needed to apply conventional computer vision algorithms like registration. In this paper, a new 3D reconstruction method is presented, taking explicitly into account the probe trajectory. Experiments were conducted on different data sets with various probe motion types and indicate that this technique outperforms classical methods, especially on low acquisition frame rate.

1 Introduction

Due to its low cost, real time image formation capability and non invasive nature, 2D ultrasound is a popular medical imaging modality. Nevertheless, the lack of 3D information prevents reproductivity, longitudinal follow-up and precise quantitative measurements. 3D ultrasound imaging addresses these disadvantages in order to obtain an objective representation of the scanned volume. Among the various options to acquire 3D ultrasound, this work focuses on 3D freehand ultrasound. This technique consists in tracking in the space a standard 2D probe by using 3D localizer (magnetic or optic). The tracking system continuously measures the 3D position and orientation of the probe. Contrary to mechanical built-in probes, freehand imaging is cheap, can address a large variety of clinical applications and allows large organs examination. However, the reconstruction step is still an acute problem with regards to computation time and reconstruction quality.

The sparsity of data is the main difficulty to transform a non uniformly distributed set of B-scans into a regular 3D volume. A correct reconstruction should not introduce geometrical artifacts, degrade nor distort the images. The most common techniques to resolve this problem are pixel nearest-neighbor (PNN) [3], voxel nearest-neighbor (VNN) [6] and distance-weighted (DW) interpolation [1, 8]. These approaches are designed to reduce computation time, but lead to a moderate reconstruction quality. More elaborated and recent methods

build on non rigid registration [2], radial basis functions (RBF) interpolation [4] or Rayleigh model for intensity distribution [5]. Quality improvement is obtained at the expense of computational burden.

Another strategy in 3D freehand imaging consists in analyzing the 3D volume without reconstruction. The StradX system [3] is built on this approach. The sequence of B-scans can be arbitrarily resliced and distance/volume measurements are performed without reconstruction. This strategy is very powerful for manual analysis of 3D datasets. However, we do think that 3D isotropic reconstruction is still necessary for specific clinical context, especially when automatic segmentation or registration procedures are required.

This paper is organized as follows. Section 2 describes the proposed reconstruction method based on utilization of the probe trajectory (PT) information. Section 3 describes briefly the evaluation framework and compare the proposed method with pixel nearest-neighbor (PNN) and distance-weighted (DW) interpolation techniques.

2 Method

This work builds on the distance weighted (DW) interpolation and proposes to incorporate probe trajectory information. The DW interpolation is first presented in section (2.1). Then the probe trajectory information is incorporated in section (2.2).

2.1 Distance Weighted (DW) Interpolation

At each point X of the reconstructed volume, the linear interpolation amounts to computing:

$$f_n(X) = \frac{1}{G} \sum_{i \in K_n(X)} g_i \tilde{f}(X_i)$$

where K_n is the interpolation kernel. In other words, K_n is the set of the different indexes of the B-scans that are involved in the interpolation. n is the interpolation order. For a given interpolation degree, the n closest B-scans before X and the n closest B-scans after X are considered. For the DW interpolation, X_i is the orthogonal projected of X on the i^{th} B-scan. $\tilde{f}(X_i)$ is the intensity on the position X_i and is obtained by a bilinear interpolation. Finally, G is the normalization constant with $G = \sum g_i$, and g_i is the distance between X and X_i (see Fig. 1). It might happen that a part of the reconstructed volume is visible at different time stamp (or view points) of the B-scans sequence. This is also known as spatial compounding. These different time stamps are computed so as to track this information and fully exploit the speckle decorrelation.

2.2 Probe Trajectory

Contrary to the classical DW interpolation approach where orthogonal projections are performed, the set of points X_i is defined as the intersection between

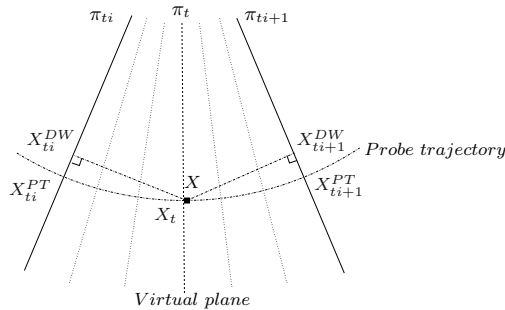


Fig. 1. Illustration of DW and PT principles. The two orthogonal projections for DW interpolation method and the construction of a “virtual” plane π_t containing X for PT method.

the closest B-scans and the probe trajectory at point X . An option would be to compute numerically this trajectory using a Runge-Kutta method. However, this would be computationally prohibitive. This problem is solved by computing the “virtual” plane that passes through X in the sense of the probe trajectory (see Fig 1). For this purpose, we need first to compute the plane position in the 3D space. This can be achieved by determining the probe position relative to this “virtual” plane and then perform a rigid transformation to map from receiver to transmitter coordinates. This transformation is based on six parameters position (3 translations and 3 rotations) that allow a perfect localization of the plane in the 3D space. The tracking of the probe motion over time provides six different signals that corresponds of the variation of the position parameter with time. Thus, computing the position of the “virtual” plane amounts to computing its acquisition time. Let us denote π_t this plane, two steps are necessary to determine the coordinates of X in π_t . Firstly, the time stamp of π_t must be evaluated. Secondly, this time t is used to estimate the probe position at this moment.

Under the assumption that the probe speed is constant between two consecutive B-scans, the latency ratio is supposed to be equal to the distance ratio:

$$t = \frac{d_{ti+1}}{d_{ti} + d_{ti+1}}(t_i) + \frac{d_{ti}}{d_{ti} + d_{ti+1}}(t_i + 1)$$

where d_{ti} is the distance (in the sense of orthogonal projection) between the current voxel and the B-scan of time stamp t_i ($d_{ti} = \|X - X_{ti}^{DW}\|$).

Once the time stamp of the “virtual” plane is computed, the probe position can be interpolated. The second step is a cubic interpolation of position parameters at time stamp t . The Key function is used to carry out a direct cubic interpolation and is defined as:

$$\varphi(t) = \begin{cases} (a + 2)|t|^3 - (a + 3)t^2 + 1 & \text{if } 0 \leq |t| < 1, \\ a|t|^3 - 5at^2 + 8a|t| - 4a & \text{if } 1 \leq |t| < 2, \\ 0 & \text{if } 2 \leq |t|. \end{cases}$$

With $a = -\frac{1}{2}$, φ is a C^1 function and a third order interpolation is obtained [7]. In practice, four B-scans are used for cubic interpolation. This seems an optimal trade-off between computational time and reconstruction quality. For example, the interpolation of rotation parameter along x axis R_x reads as:

$$R_x(t) = \sum_{k=t_i-1}^{t_i+2} R_x(k)\varphi(t-k)$$

The “virtual” plane certainly not contains X numerically, despite the distance is infinitesimal. Therefore, X_t is used as the projection of X on the “virtual” plane (see Fig 1). Then, $X_{t_i}^{PT}$ and $X_{t_i+1}^{PT}$ are obtained directly, since they have the same 2D coordinates (defined in each B-scans) as X_t .

3 Results

3.1 Material

An Sonosite system with a cranial 7–4MHz probe was used to acquire the ultrasound images. The positions of the B-scans was given by a magnetic miniBIRD system (Ascension Technology Corporation) mounted on the US probe. The StradX software [3] was used to acquire images and 3D position data. A CIRS, Inc.¹ 3D ultrasound calibration phantom was used. The phantom contains two calibrated volumetric ellipsoids test objects. At the acquisition depth, only one of the ellipsoids is in the field of view. The two sequences used for the experiments are composed of 510×441 B-scans (204 B-scans for fan motion and 222 B-scans for translation motion, see Fig. 2).

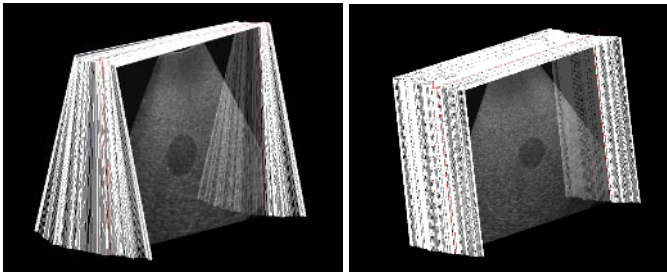


Fig. 2. B-scans sequences used during evaluation. Left: fan sequence. Right translation sequence.

3.2 Evaluation Framework

In order to perform an objective evaluation, the performance of the proposed method was compared with two other interpolation approaches: the voxel nearest neighbor (VNN) technique used in StackX [3] and the distance-weighted

¹ <http://www.cirsinc.com>

(DW) interpolation method presented in [8]. For the VNN method, each voxel is projected on the nearest B-scan and its luminance interpolated bilinearly. In the DW interpolation technique, each voxel is projected orthogonally on the $2n$ nearest B-scans and its luminance interpolated (see section 2.1 and Fig. 1). To assess the reconstruction quality, evaluation data can be created from any image sequence: given a sequence of 3D freehand US, each B-scan is removed from the sequence and then reconstructed using the other B-scans of the sequence. This technique, “leaves one out”, is performed in turn for each B-scan. Within this evaluation framework, the influence of two parameters is studied:

- “External” parameter associated with the acquisition setup: the acquisition frame rate. The sequence is sub-sampled thanks to SelectSX², which simulates a lower frame acquisition rate. The sub-sampling factor varies from 2 to 6, this means that we keep from one B-scan out of two (denoted by 1/2) to one B-scan out of six (denoted by 1/6).
- “Internal” parameter associated with the reconstruction method: the interpolation degree. The removed B-scans are reconstructed with different methods and different interpolation degree (from 1 to 2 for DW interpolation and PT method).

The Mean Square Error (MSE) is used as quality criterion:

$$MSE(t) = \frac{1}{P} \sum_{j=1}^P (\tilde{I}_t(x_j) - I_t(x_j))^2$$

where I_t is the original t image (removed from the sequence), \tilde{I}_t the reconstructed image and P is the number of pixel in this B-scan. From MSE estimation for all B-scans of the sequence, we compute the mean μ and the standard deviation σ of reconstruction error.

$$\mu = \frac{1}{N-2} \sum_{t=2}^{N-1} MSE(t)$$

$$\sigma^2 = \frac{1}{N-2} \sum_{t=2}^{N-1} (MSE(t) - \mu)^2$$

N is the total number of B-scan in the sequence. The first and last B-scans are rejected to avoid artifacts.

3.3 Experiments

Results are presented in Figure 3 and Table 1. Figure 3 shows the sub-sampling influence on reconstruction error relatively to motion nature (i.e. translation or fan), for interpolation degree equals to 1. In all cases, the Probe Trajectory

² <http://mi.eng.cam.ac.uk/~rwp/stradx/utilities.html>

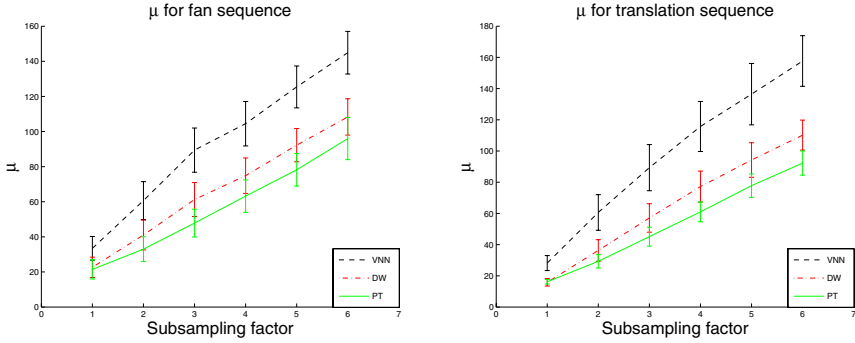


Fig. 3. Variation of error reconstruction relatively to sub-sampling factor with interpolation degree = 1. Left fan motion. Right translation motion. Three methods are evaluated : VNN, DW and PT. The PT method outperforms others methods especially on sparse data.

Table 1. Error measures composed of mean μ and standard deviation σ for the different methods. L_{time} is the time spent for labeling, while I_{time} is the time spent for interpolation. Error measures indicate that the PT method obtains better results than the VNN and DW methods. The improvement in terms of reconstruction quality is obtained at the expense of a slight computational increase.

Sequence	Order	VNN			DW			PT		
		μ	σ	I_{time}	μ	σ	I_{time}	μ	σ	I_{time}
fan $L_{time} = 414s$	$n = 1$	33.4	13.5	20s	22.6	11.6	44s	21.5	11.0	114s
	$n = 2$				26.2	12.5	46s	22.9	10.6	122s
fan 1/2 $L_{time} = 215s$	$n = 1$	60.7	21.6	21s	41.1	16.9	31s	33.0	14.2	114s
	$n = 2$				46.6	17.6	44s	38.2	14.1	124s
fan 1/3 $L_{time} = 145s$	$n = 1$	89.4	25.2	20s	61.3	19.4	30s	47.8	15.9	111s
	$n = 2$				67.6	18.6	43s	56.2	14.2	118s
translation $L_{time} = 42s$	$n = 1$	28.2	9.6	20s	15.9	4.8	37s	16.2	2.9	138s
	$n = 2$				19.1	6.4	49s	18.3	3.5	149s
translation 1/2 $L_{time} = 27s$	$n = 1$	60.6	22.9	21s	36.3	13.9	32s	29.3	8.6	138s
	$n = 2$				43.0	14.5	47s	35.8	8.7	147s
translation 1/3 $L_{time} = 20s$	$n = 1$	89.3	29.6	20s	57.1	18.3	31s	45.1	12.1	138s
	$n = 2$				63.7	17.5	46s	52.9	11.1	146s

(PT) method outperforms VNN and DW methods especially on sub-sampled sequences. On sparse data sets, the incorporation of probe trajectory leads to a more robust reconstruction. In Table 1, error measures and computation time are presented. Reconstructions were performed on P4 3.2 Ghz with 2Go RAM. In order to only compare interpolation times, the total computational time was split in two : L_{time} corresponds to labeling step and second denoted by I_{time} is interpolation time. The labeling step consist in K_n construction for each

voxel X . Our implementation of methods can be largely optimized, as compared to StackSX (22s for all reconstruction process within VNN method). Although the labeling step need significant improvement, this study aims to compare computation time between identical implementations. Increase in quality for PT method is obtained at the expense of slight computation time increase. Nonetheless, this side effect is reasonable with regards to quality reconstruction. Contrary to more elaborated techniques like non-rigid registration or RBF, which are considerably computationally expensive, the PT approach offers an attractive compromise between time computation and quality reconstruction. Figure 4 shows the differences between original and reconstructed B-scan with the VNN, DW and PT methods. Visually, the PT reconstruction appears closer to the original B-scan, which is in agreement with the numerical results of Table 1. Reconstruction results on real dataset are presented in Figure 5. The VNN method leads to a lot of discontinuities. The differences between the DW and

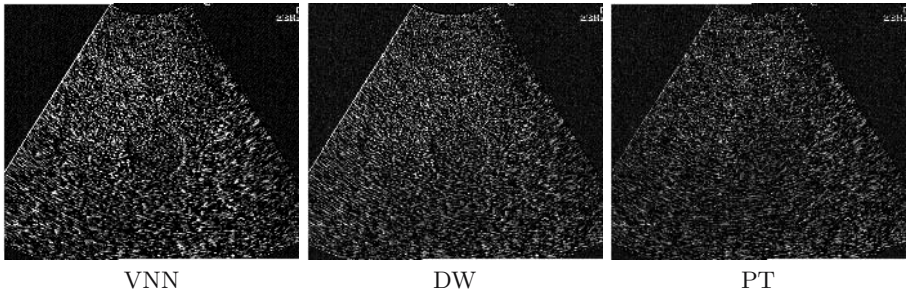


Fig. 4. Differences between original and reconstructed B-scan for not subsampled fan sequence. From left to right VNN, DW and PT methods. This shows that error between reconstructed B-scan and “ground truth” is visually better with the PT method.

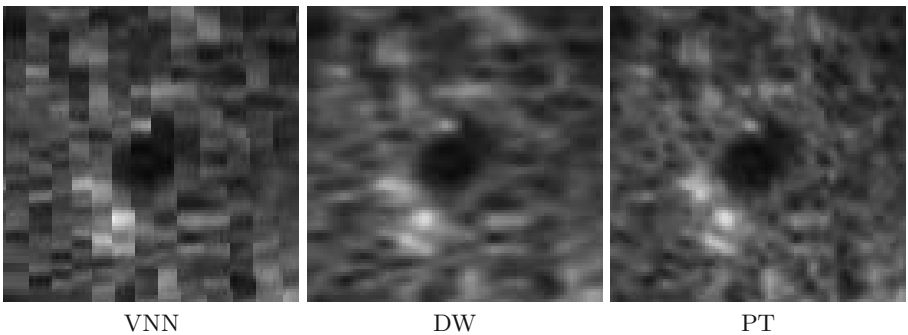


Fig. 5. Zoom on hepatic vessels extracted from liver reconstruction with degree= 1 and subsampling factor= 4. From left to right the VNN, DW and PT methods. The images are extracted from 3D volume along the temporal axis (z) in order to under-light inherent artifacts of the VNN (i.e. discontinuities) and the DW (i.e. blur) methods. The PT method is more efficient at preserving the native texture pattern of US image than the DW method.

PT methods are less visible, as assessed by numerical results. Nevertheless, in general the DW method smooths out the edges and spoils the native texture pattern of US image more than the PT method (see Fig. 5).

4 Conclusion

This paper presented a 3D freehand US reconstruction method using probe trajectory information. This method performs better than traditional reconstruction approaches with a reasonable increase in computation time. Results on sub-sampled acquisitions show that the probe trajectory brings a relevant information in the reconstruction process. The main limitation of PT method is the assumption of constant probe speed between two slices. Nevertheless, for frame rate around 5-15Hz and moderate care during acquisition (i.e. a relatively continuous motion user) this limitation can be easily overcome. Further work should pursue the comparison between the PT reconstruction approach and registration-based approaches [2]. Finally, our implementation should be largely optimized using graphic library implementation (ex : *OpenGL*) or even GPU one.

References

1. C. D. Barry, C. P. Allott, N. W. John, P. M. Mellor, P. A. Arundel, D. S. Thomson, and J. C. Waterton. Three dimensional freehand ultrasound: image reconstruction and volume analysis. *Ultrasound in medicine and biology*, 23(8):1209–1224, 1997.
2. G. P. Penney, J. A. Schnabel, D. Rueckert, M. A. Viergever, and W. J. Niessen. Registration-based interpolation. *IEEE Trans Med Imaging*, 23(7):922–926, 2004.
3. R. W. Prager, A. H. Gee, and L. Berman. Stradx : real-time acquisition and visualization of freehand three-dimensional ultrasound. *Medical Image Analysis*, 3(2):129–140, 1999.
4. R. Rohling, A. H. Gee, L. H. Berman, and G. M. Treece. Radial basis function interpolation for freehand 3D ultrasound. In Attila Kuba, Martin Sámal, and Andrew Todd-Pokropek, editors, *Information Processing in Medical Imaging*, volume 1613 of *LNCS*, pages 478–483. Springer, 1999.
5. J. M. Sanchez and J. S. Marques. A rayleigh reconstruction/interpolation algorithm for 3D ultrasound. *Pattern recognition letters*, 21:917–926, 2000.
6. S. Sherebrin, A. Fenster, R. N. Rankin, and D. Spence. Freehand three-dimensional ultrasound: implementation and applications. In *Proc. SPIE Vol. 2708, p. 296-303, Medical Imaging 1996: Physics of Medical Imaging, Richard L. Van Metter; Jacob Beutel; Eds.*, pages 296–303, April 1996.
7. P. Thévenaz, T. Blu, and M. Unser. Interpolation revisited. *IEEE Transactions on Medical Imaging*, 19(7):739–758, 2000.
8. J. W. Trobaugh, D. J. Trobaugh, and W. D. Richard. Three dimensional imaging with stereotactic ultrasonography. *Computerized Medical Imaging and Graphics*, 18(5):315–323, 1994.

Self-calibrating Ultrasound-to-CT Bone Registration

Dean C. Barratt¹, Graeme Penney¹, Carolyn S.K. Chan¹, Mike Slomczykowski²,
Timothy J. Carter¹, Philip J. Edwards³, and David J. Hawkes¹

¹ Centre for Medical Image Computing, University College London, UK
{d.barratt, g.penney, carolyn.chan, t.carter,
d.hawkes}@cs.ucl.ac.uk

² Depuy International – a Johnson & Johnson Company, iOrthopaedics, Leeds, UK

³ Department of Surgical Oncology and Technology, Imperial College London, UK

Abstract. We describe a new self-calibrating approach to rigid registration of 3D ultrasound images in which *in vivo* data acquired for registration are used to simultaneously perform a patient-specific update of the calibration parameters of the 3D ultrasound system. Using a self-calibrating implementation of a point-based registration algorithm, and points obtained from ultrasound images of the femurs and pelvis of human cadavers, we show that the accuracy of registration to a CT scan is significantly improved compared with a standard algorithm. This new approach provides an effective means of compensating for errors introduced by the propagation of ultrasound through soft tissue, which currently limit the accuracy of conventional methods where the calibration parameters are fixed to values determined preoperatively using a phantom.

1 Introduction

Freehand 3D ultrasound (US) has been proposed as an accurate and non-invasive method for intraoperative registration of bony anatomy during minimally-invasive orthopaedic, neuro- and skull-based surgery [1-3]. In this technique, an US probe is tracked by a 3D localizer and may be intuitively thought of as a non-invasive device for locating bone surfaces. This approach is an attractive alternative to standard clinical methods, which are highly invasive since they require the bone surface to be surgically exposed to allow digitization by contact with a tracked sterile pointer [4,5]. Using such methods, digitized bone surface points are matched to the corresponding surface extracted from a preoperative CT-scan or, more recently, a generic bone surface model, which is simultaneously morphed to adopt the shape of an individual patient's bone [6]. Although such invasive techniques provide an accuracy that is widely accepted to be sufficient for surgical navigation, they potentially contribute to the risk of complications associated with soft tissue trauma, such as blood loss and infection.

Fig. 1(a) illustrates a B-mode US probe with rigid attachment, which is tracked by an optical localizer to enable (freehand) 3D imaging. In order to achieve accurate bone registration using such a system, careful calibration is required to maximize the localization accuracy. Specifically, calibration is required to determine the rigid-body transformation (denoted by $\mathbf{T}_{\text{PROBE} \leftarrow \text{US}}$ in Fig. 1(a)) which relates the 3D co-ordinate

systems of the US image and the tracked object attached to the US probe. Numerous methods have been described for the calibration of freehand 3D US systems using a variety of special-purpose phantoms [7], but the accuracy of these methods is fundamentally limited by differences between the propagation of US in soft tissue compared and that in a phantom. These differences effectively introduce an unknown error into the calibration parameters when the system is used *in vivo*, limiting the localization accuracy, and, therefore, registration accuracy achievable using the 3D US technique.

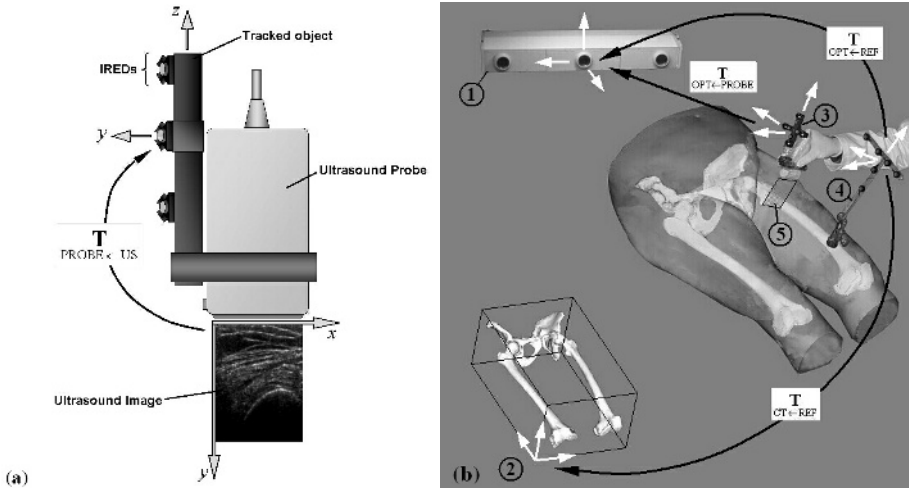


Fig. 1. (a) Left, a tracked B-mode US probe with attachment for freehand 3D US imaging. (b) Right, a schematic of the setup used for US-based registration during orthopedic surgery: (1) Optical localizer; (2) Segmented CT scan; (3) Tracked US probe; (4) Bone-implanted dynamic reference object (see text); (5) US image. The rigid-body transformations between 3D the coordinate systems of the various components are indicated by the black arrows.

A primary source of error arises from assuming a constant value for the speed of sound, typically 1540m/s. In reality, the average speed of sound in human tissue can vary from approx. 1400m/s (in fat) to 1650m/s (in muscle), with significant variations occurring between both anatomical regions and patients. Using a simplified model of US propagation, such variations can give rise to systematic error in locating the depth of a bone surface of up to about 5%, which may become particularly problematic in the obese patient. Non-linear propagation of the US beam may also become significant in some cases, potentially compromising the localization accuracy of a 3D US system further.

In the remainder of this paper we describe a self-calibrating method for US-to-CT bone registration, implemented within a point-to-surface registration scheme, which addresses these issues by updating the calibration parameters of the 3D US system using *in vivo* data acquired for the purposes of intraoperative bone registration.

2 Self-calibrating Point-Based Registration Algorithm

Fig. 1(b) shows an example set-up for hip replacement surgery in which a handheld US probe (3), tracked by an optical localizer (1), is used to obtain images of the femur and pelvis. Bone surface points derived from these images are used to calculate the rigid-body registration transformation between a tracked physical reference object (4) and a preoperative CT scan (2) by matching the US-derived points to the bone surface extracted from a CT scan. In practice, the reference object is fixed to a rigid surgical attachment implanted into a bone, as illustrated in Fig. 1(b), and is referred to as a dynamic reference object (DRO). DROs allow the patient to be repositioned during surgery without compromising the registration.

The self-calibrating, point-based registration algorithm outlined here minimizes a least-squares criterion, but differs from the popular Iterative Closest Point algorithm, in that a general non-linear optimization scheme is used to perform a least-squares minimization. If S denotes the (CT-derived) bone surface, and $(\mathbf{p}_{n,i})_{US}$ is the (homogeneous) 3D position vector containing the co-ordinates of the n^{th} bone surface point derived from the i^{th} US image, the task of registering all points to the surface S can be posed mathematically as finding the parameter vector, $\boldsymbol{\Psi}$, which minimizes the cost function, C , of the form

$$C(\boldsymbol{\Psi}) = \sum_{i=1}^I \sum_{n=1}^N d^2(\mathbf{T}_i(\boldsymbol{\Psi}) \cdot (\mathbf{p}_{n,i})_{US}, S), \quad (1)$$

where $d(\mathbf{p}, S)$ represents the Euclidean distance between S and the US point, \mathbf{p} , and the transformation matrix, $\mathbf{T}_i(\boldsymbol{\Psi})$, is defined as

$$\mathbf{T}_i(\boldsymbol{\Psi}) = \mathbf{T}_{CT \leftarrow REF}(\boldsymbol{\Psi}_{REG}) \cdot \mathbf{T}_{(OPT \leftarrow REF, i)}^{-1} \cdot \mathbf{T}_{(OPT \leftarrow PROBE, i)} \cdot \mathbf{T}_{PROBE \leftarrow US}(\boldsymbol{\Psi}_{CAL}) \cdot \mathbf{T}_{SCALE}(s_x, s_y), \quad (2)$$

where the notation $\mathbf{T}_{B \leftarrow A}$ denotes a rigid-body transformation from 3D co-ordinate system A to B. In Eqn. (2) ‘US’, ‘PROBE’, ‘OPT’, and ‘REF’ denote the 3D co-ordinate systems of the US image, the tracked object attached to the US probe, the (optical) 3D localizer, and the DRO, respectively (see Fig. 1(b)); $\boldsymbol{\Psi} = [\boldsymbol{\Psi}_{REG}, \boldsymbol{\Psi}_{CAL}, s_x, s_y]$ is a vector containing 6 rotations, 6 translations, and 2 scaling parameters; and \mathbf{T}_{SCALE} is a diagonal scaling matrix included to convert the point co-ordinates from pixels to millimetres.

Conventionally, the parameters in $\boldsymbol{\Psi}_{CAL}$ which define the transformation, $\mathbf{T}_{PROBE \leftarrow US}$, and the US image scaling parameters, s_x and s_y , are determined preoperatively from one or more calibration experiments, and remain fixed during registration. However, if we consider them to be free variables, then, given a reasonable starting estimate for $\boldsymbol{\Psi}$, a simultaneous registration and calibration can be carried out by optimizing all 14 parameters of $\boldsymbol{\Psi}$ to find the global minimum of C .

In our implementation, C was minimized using the iterative region trust algorithm provided in the Matlab Optimization Toolbox v3.0 (The Mathworks Inc, Natick, MA, USA), and, on each iteration, point-to-surface distances were computed efficiently using a 3D Euclidean distance transform of the binary image of the bone surface generated from a pre-segmented CT scan [8].

In order to avoid local minima, a good initial estimate of the registration parameters (Ψ_{REG}) was obtained by first performing a standard rigid registration with all of the calibration parameters fixed to values obtained from a prior phantom-based calibration (see Sect. 3.2 below). The performance of the self-calibrating registration algorithm was evaluated using data acquired from the pelvis and femurs of 3 complete female cadavers with intact tissue in experiments which simulated the conditions of hip replacement surgery. Importantly for this study, the use of cadavers allowed an accurate Gold Standard registration, based on bone-implanted fiducials, to be established.

3 Evaluation of Registration Accuracy Using Cadaveric Data

3.1 Cadaver Preparation and CT Scanning

With the cadavers lying in the supine position, titanium bone screws with a 2mm diameter threaded tip were implanted into both femurs and each side of the pelvis. Four bone screws were inserted into each femur and 5 into each hemi-pelvis. A custom-made fiducial marker filled with a radio-opaque contrast agent was then attached to the end of each screw, after which the cadavers were transported to the radiology department where a single high-resolution CT scan of each cadaver was performed using a Siemens SOMATOM Plus 5 spiral CT scanner. CT images were reconstructed with voxel dimensions between 0.7 and 0.8mm in the transverse plane and a slice thickness of 2mm. Following CT scanning, the cadavers were returned to the anatomy laboratory and the contrast-filled fiducial markers replaced with markers containing a 3mm diameter divot, machined so that the centre of a ball-tip digitizer placed into the divot coincided with the centroid of the contrast-filled chamber of the imaged markers. The digitizer, US probe, and a DRO implanted into the femur were all tracked using an Optotrak 3020 optical localizer (Northern Digital Inc., Ontario, Canada).

The Gold Standard image-to-physical registration transformation for each bone was found by registering the physical positions of the centroid of each contrast-filled fiducial marker, identified using the ball-tip digitizer, to the corresponding manually-identified positions in the CT scan. The surface of each bone was segmented in the CT images using a combination of manual and semi-automatic tools provided by the software package Analyze v6.0 (Mayo Foundation, Rochester, MN, USA). From the segmented CT images, a binary image which just contained the bone surface voxels was generated, and this was used to compute a 3D Euclidean distance image for the point-based registration algorithm.

3.2 Ultrasound Imaging

The freehand 3D US system used in this study was developed by our research group and is based on a standard clinical US scanner (Philips-ATL HDI-5000, Philips Medical Systems, Bothell, WA, USA). A high-frequency linear-array probe (L12-5 scan-probe; 5-12MHz broadband transducer) was chosen, which was tracked in 3D by attaching a custom-made cross-shaped piece with 20 IREDs attached as shown in Fig. 1(a). Acquisition software, also developed in our laboratory, was used to synchronize the recording of the 3D position and orientation of the US probe (relative to the DRO) with US image capture using an analogue-to-digital converter.

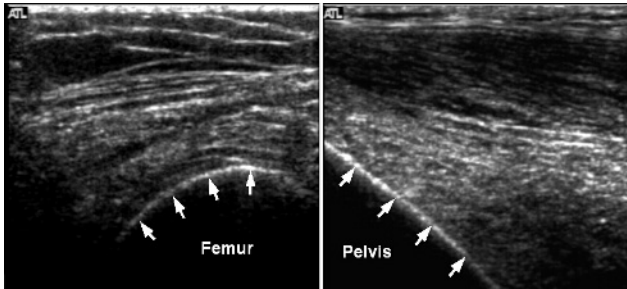


Fig. 2. Example US images of the cadaveric femur (left) and pelvis (right) obtained in this study. The arrows indicate the bone surface.

To provide a reasonable starting estimate of the calibration parameters of the 3D US system, a pinhead immersed in a water-glycerol solution was scanned and the calibration parameters calculated using a previously published method [9]. The scaling of the captured US images was determined from electronic calliper measurements provided by the US scanner. During US scanning the probe was covered with a plastic sheath containing a small quantity of gel to ensure good acoustic contact. For each cadaver, B-mode US images of the surface of the femur and pelvis were acquired using realtime 2D compounding (“SonoCT™”) with a single focal zone and a maximum penetration depth of 6cm. Example US images are shown in Fig. 2.

The bone surface was segmented manually from US images by defining a number of points on the maximum intensity of the ridge corresponding to the bone surface in each US image. Approximately 10 points were defined for each image in which the bone surface was clearly visible. Images with poor tracking data were excluded automatically by segmenting only those images where ≥ 10 out of 20 IREDs on the US probe attachment and ≥ 5 out of 6 IREDs on the DRO were visible to the Optotrak.

3.3 Registration Experiments

To test the accuracy and robustness of the US-based registration, 100 registrations were performed for each of the 9 bones (3 whole pelvises and 6 femurs) using a random starting position for each trial. Starting positions were generated by simulating a simple intraoperative, point-based registration procedure in which the surgeon identifies a series of skin points adjacent to well-defined bony landmarks. To do this, points on the skin surface close to the landmarks were first identified by hand in each CT image. Three points were identified for each femur, near to the medial and lateral epicondyles, and to the greater trochanter. For each pelvis, 4 points were selected in the proximity of the left and right superior pubic ramus and left and right anterior superior iliac spines. To simulate errors associated with intraoperative identification of anatomical landmarks, a set of random points on the skin surface were computed, each within a sphere of radius 20mm centred at the corresponding manually-selected point. These points were then transformed to the co-ordinate system of the DRO using the Gold Standard registration transformation. The maximum error of 20mm was chosen to correspond to the upper limit on the error that might reasonably occur when

landmarks are digitized manually [10]. Greater accuracy could be obtained by using self-adhesive skin markers, for example.

In the case of the femur, the estimated centre of rotation of the femoral head provided an additional anatomical point for calculating registration starting positions. The location of this point for each hip joint was determined experimentally by slowly pivoting the leg, while continuously tracking the DRO rigidly attached to the femur. A random offset of up to 10mm was added to this point to simulate the variability that might reasonably be expected using the pivot method intraoperatively.

Random starting positions for the registration trials were found by registering the original set of manually-identified points in the CT image to the corresponding set of simulated anatomical points. As indicated in Sect. 2, each registration trial required an initial 2-step rigid registration with the calibration parameters fixed – once using all US points and then repeated with outliers removed – to obtain a good estimate of the registration parameters. Then, each time using these estimates as starting values, a self-calibrating registration was performed twice, once optimizing 7 parameters, including axial scaling (s_y) as an additional free parameter, and then again with a full 14 parameter optimization. For each registration trial, starting estimates of the calibration parameters were calculated using a randomly selected subset of 30 of the 71 images of the calibration phantom to introduce variability into the starting calibration parameter estimates. In the case of the US scaling parameters, starting estimates were generated by adding a uniform random noise of up to +/-5% to the values determined using the standard calibration method.

Registration accuracy was evaluated by calculating the target registration error (TRE) for each bone surface voxel extracted from the CT scan using the formula

$$\text{TRE}(\mathbf{s}_{\text{CT}}) = \left\| \mathbf{s}_{\text{CT}} - \left(\mathbf{T}_{\text{CT} \leftarrow \text{REF}} \right)_{\text{US}} \cdot \left(\mathbf{T}^{-1} \right)_{\text{GOLD}} \cdot \mathbf{s}_{\text{CT}} \right\|, \quad (3)$$

where \mathbf{s}_{CT} is the (homogeneous) 3D position vector of the centre of a bone surface voxel in CT co-ordinates, and the subscripts ‘US’ and ‘GOLD’ denote transformations computed using US-based and Gold Standard registrations, respectively. (Note: the magnitude operator in Eqn. (3) ignores the redundant 4th element of the vector operand). The root-mean-square (RMS) TRE was computed for all surface voxels of each bone. As a further clinically important measure of registration accuracy, the RMS TRE was calculated using a subset of voxels within the surgical region of interest (SROI). For the femur, the SROI was defined as the region proximal to the lesser trochanter, which includes the femoral head and neck, whereas for the pelvis the SROI was defined as the spherical region enclosing the acetabulum with radius 50mm and centred at the approximate centre of the femoral head.

4 Results

The predicted TREs for the Gold Standard registrations, calculated over the whole bone surface using the method of Fitzpatrick et al. [11], were submillimetre in all but one case (range: 0.13 to 1.32mm RMS), indicating excellent Gold Standard

registrations. The number of US images acquired was between 257 and 576, with between 124 and 510 images segmented, giving a total of between 960 and 6112 surface points for each bone (~10 points per US image).

Table 1. Average RMS TRE (absolute maximum TRE in brackets) in millimetres for US-to-CT bone registration from 100 trials of 9 datasets. Key: The “Initial TRE” results from the registration based on simulated anatomical landmarks, identified intraoperatively; “TRE1” is the TRE for a standard rigid registration using all available US points; “TRE 2” is the TRE after removing outliers and repeating the first registration using the remaining 90% of the original points; “TRE 3” is the TRE for a 7-parameter self-calibrating registration; and “TRE 4” is the TRE for a full 14-parameter self-calibrating registration.

Region	Initial TRE	TRE 1	TRE 2	TRE 3	TRE 4
Whole Bone	10.25 (64.88)	2.57 (9.30)	2.20 (7.37)	1.87 (7.15)	1.59 (7.15)
Surgical ROI	9.07 (39.22)	2.66 (6.89)	2.36 (7.16)	1.97 (5.13)	1.74 (5.44)

The average RMS TREs for each US-based registration are summarized in Table 1. It can be seen from the results presented that, overall, the self-calibrating registration resulted in a marked improvement in registration accuracy (using the Kolmogorov-Smirnov test, the differences between TREs 1 to 4 were found to be statistically significant at the 1% level). In particular, performing a full 14-parameter self-calibrating registration (TRE 4) resulted in a reduction in TRE for the whole bone of approximately 28% compared with a standard point-based registration with automatic outlier removal (TRE 2). With a failed registration defined as one for which the RMS TRE over the whole bone surface was >5mm, the number of failed registrations using a full self-calibration (TRE 4) was 46 from 900 trials (compared with 44 for the standard registration with outliers removed (TRE 2)). However, of these, 39 occurred for one pelvis, resulting in a very low failure rate (<1%) if the results from this dataset are not considered. We believe that the reason for the relatively high number of failures for this dataset is likely to be due to the much higher initial TREs (Mean = 20.07mm RMS) compared with the other datasets.

5 Discussion and Conclusions

In this paper we have demonstrated that optimizing the 3D US calibration parameters improves the accuracy of US-to-CT bone registration compared with a standard algorithm where calibration parameters are fixed. For the purposes of this study, manually segmented US/CT data and a point-based registration scheme were adopted, but future work will focus on developing a self-calibrating image-intensity-based registration algorithm to allow fully automatic registration, and devising an optimized clinical protocol for intraoperative US imaging of bone.

Including calibration parameters into the registration optimization provides a useful method for compensating for unpredictable, patient-specific localisation errors introduced when imaging through soft tissue. Provided fairly accurate starting

estimates for the calibration parameters are available, updating the calibration parameters during registration also has the practical advantage that US scanner settings could be adjusted for individual patients within the operating theatre without compromising calibration or registration accuracy.

References

1. Lavallée, S., Troccaz, J., et al.: Computer-assisted spinal surgery using anatomy-based registration. In Computer Taylor, R.H., Lavallée, S., Burdea, G.C., Mösges, R. (eds): *Integrated Surgery: Technology and Clinical Applications*. MIT Press, Cambridge MA (1996) 425-449
2. Amin, D., Kanade, T., DiGioia, A., Jaramaz, B.: Ultrasound registration of the bone surface for surgical navigation. *Comp. Aid. Surg.* **8** (2003) 1-16
3. Amstutz, C., Caversaccio, M., et al.: A-mode ultrasound-based registration in computer-aided surgery of the skull. *Arch. Otolaryngol. Head Neck Surg.* **129** (2003)1310-1316
4. DiGioia, A.M., Jaramaz, B., et al.: The Otto Aufranc Award. Image guided navigation system to measure intraoperatively acetabular implant alignment. *Clin. Orthop.* **355** (1998) 8-22
5. Ma, B. and Ellis, R.B.: Robust registration for computer-integrated orthopedic surgery: laboratory validation and clinical experience. *Med. Image Anal.* **7** (2003) 237-250
6. Stindel, E., Briard, J., et al.: Bone morphing: 3D morphological data for total knee arthroplasty. *Comp. Aid. Surg.* **7** (2002) 156-168
7. Mercier, L., Lango, T., Lindseth, F., Collins, L.D.: A review of calibration techniques for freehand 3-D ultrasound systems. *Ultrasound Med. Biol.* **31** (2005) 143-165
8. Fitzgibbon, A.: Robust registration of 2D and 3D point sets. *Proc. B.M.V.C.* (2001) 411-420
9. Barratt, D., Davies, A., et al.: Accuracy of an electromagnetic three-dimensional ultrasound system for carotid artery imaging. *Ultrasound Med. Biol.* **27** (2001) 1421-1425
10. della Croce, U., Cappelzozzo, A., Kerrigan, D.: Pelvis and lower limb anatomical calibration precision and its propagation to bone geometry and joint angles. *Med. Biol. Eng. Comput.* **37** (1999) 155-161
11. Fitzpatrick, M., West, J., Maurer, C.: Predicting error in rigid-body point-based registration. *IEEE Trans. Med. Imag.* **17** (1998) 694-702

A Hand-Held Probe for Vibro-Elastography

Hassan Rivaz and Robert Rohling

Department of Electrical and Computer Engineering,
University of British Columbia, Vancouver, Canada
rohling@ece.ubc.ca
www.ece.ubc.ca/~rohling

Abstract. Vibro-elastography is a new medical imaging method that identifies the mechanical properties of tissue by measuring tissue motion in response to a multi-frequency external vibration source. Previous research on vibro-elastography used ultrasound to measure the tissue motion and system identification techniques to identify the tissue properties. This paper describes a hand-held probe with a combined vibration source and ultrasound transducer. The design uses a vibration absorption system to counter-balance the reaction forces from contact with the tissue. Simulations and experiments show a high level of vibration absorption. The first elastograms from the probe are also shown.

1 Introduction

Elastography creates images that depict local tissue stiffness. Such images can be a valuable tool for the diagnosis of disease. For example, prostate and breast cancer are known to exhibit stiff tumors in surrounding softer tissue. Elastography may compliment manual palpation techniques for early detection of the tumors. Static elastography is a well-known technique that uses a quasi-static compression of tissue and simultaneous imaging with ultrasound [5]. Local tissue strain estimates can be obtained by analysis of the ultrasound images (e.g. speckle tracking by cross-correlation of the radiofrequency (RF) signals). The strain is subsequently related to the tissue stiffness.

Vibro-elastography (VE) is a new imaging technique that uses a dynamic excitation source to create tissue motion over a range of frequencies [8]. Ultrasound images are used to capture the tissue motion at multiple locations and time instants. The method then adapts techniques from system identification to estimate the tissue properties from the ultrasound-based strain measurements. In one version of VE, the tissue is modelled as a set of interconnected mass-spring-damper elements. The values of the density-stiffness-viscosity are then identified by solving an over-determined set of equations of motion. In another version of VE, the tissue properties are identified directly from transfer functions between local tissue regions. In either case, the VE images can be shown to have better quality (lower standard deviation to mean values) compared to static elastography, mainly due to the inherent averaging in the approach over various levels and rates of tissue strain. Moreover, measurements of viscosity and density are also possible with the dynamic approach. So far, good results

for stiffness and viscosity measurements have been obtained with fixed or rigidly mounted ultrasound transducers. Widespread clinical acceptance would be aided by the development of a hand-held probe so that a standard ultrasound scanning protocol can be performed. This paper describes a combined ultrasound transducer and vibrator system as the next step in the development of VE.

For a hand-held probe, the reaction forces between the tissue and vibrator act as an external disturbance on the probe assembly and will cause unwanted vibration. These vibrations should be minimized to maintain the comfort of the operator and the accuracy of the VE calculations. A light-weight probe can exhibit a large vibration level because it has a low displacement-force impedance. A heavier probe alleviates some vibration but at the cost of operator fatigue. Therefore, the probe design should use some form of vibration absorption.

Four different methods for reducing the vibration are well known: vibration isolation, passive vibration absorption, semi-active vibration absorption and active vibration absorption. For vibration isolation, a spring-damper combination is placed between the source of excitation force and the system to be isolated (Figure 1a). In this figure, k_i is the stiffness and b_i is the viscosity of the isolating material and F_e is the excitation force. To minimize the transmitted force, the spring and damper values should be selected according to the excitation frequency [7]. For passive vibration absorption, an additional mass-spring-damper trio is attached to the primary system (Figure 1b). In this figure, m_a , k_a and b_a are the absorber mass, stiffness and viscosity respectively. Unlike the previous method, passive vibration absorption adds one degree of freedom (DOF) to the system. To minimize the vibration of the primary system (x_p), the mass, spring and damper values should be selected according to the frequency range of vibration of the primary system [3]. Yet these absorbers can be designed for only a narrow band of excitation frequencies. There is a trade-off between the frequency bandwidth and the absorption level: the wider the bandwidth, the lower the absorption. Semi-active vibration absorption follows the same idea as passive absorption, but uses a tunable spring and damper (Figure 1c). Varying the spring and damper values allows the absorber to be tuned to different excitation frequencies [2]. For active vibration absorption, an actuator is added to the mass-spring-damper trio of a passive absorber (Figure 1d). Active vibration absorption differs from semi-active vibration absorption because it needs more control effort to suppress the vibration, requiring an external energy source. Semi-active and active vibration absorbers have the capability of cancelling the vibration over a wider frequency range than the first two methods.

For VE, an excitation over the range of 10-50Hz is needed, so only active or semi-active vibration absorption methods are suitable. In practice, VE rapidly sweeps this range, and semi-active vibration absorbers have difficulty responding with sufficiently speed. Moreover, the inherent damping produces sub-optimal resonance properties at the frequency of excitation, offering low levels of absorption, especially in the miniature hand-held device.

In this paper an active vibration system is designed and built into the hand-held probe. This paper describes the controller design then the design of the

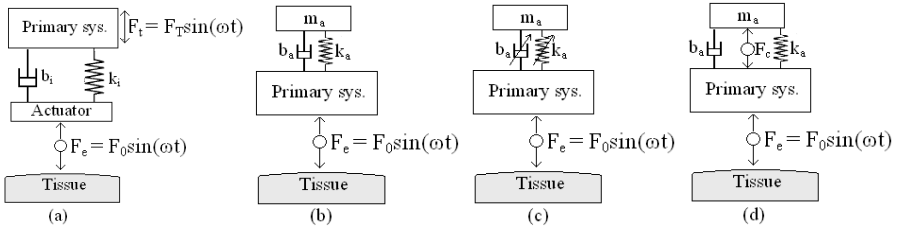


Fig. 1. Vibration reduction methods: (a) Vibration isolation (b) Passive vibration absorption (c) Semi-active vibration absorption (d) Active vibration absorption

overall device. Parameter identification of the absorber properties are then performed to optimize performance. Simulations and experiments are then performed to test the level of vibration absorption and the ability to create images with VE. But first, simulations are used to demonstrate the level of errors in VE that occur without vibration absorption.

2 The Effect of Probe Vibration

Like VE, the tissue is modelled as an n DOF mass-spring-damper model. The tissue is vibrated from one side with the amplitude of 1mm and the position of each mass is calculated at intervals in time. If the probe is stationary, the positions measured from the ultrasound images should be the same as the positions calculated from the simulation. To simulate probe vibration, a portion of the vibration displacement source is added to the simulation results and considered as measurements. The parameters of mass, viscosity and stiffness are then identified from the measurements by the VE equations.

Figure 2 shows the relationship between the amplitude of probe vibration and the relative error in estimating the tissue properties. The error is not the same at all elements/nodes of the modelled tissue, so are averaged to produce a single error value at each amplitude. As shown, there is a significant error introduced by

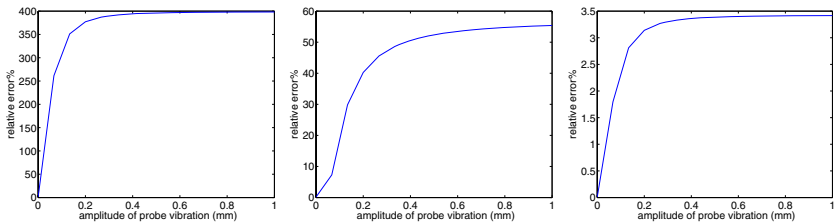


Fig. 2. Simulations of probe vibration on the error in VE. The left, middle and right plots corresponds to the error in the calculated mass, viscosity and stiffness respectively. The tissue is excited at 1mm amplitude over a range of 10 Hz.

the probe vibration. For example, an absorption of 90% of the vibration results in a drop in the error in estimating viscosity from approximately 60% to 20%.

3 Controller Design

The vibration absorber on the probe should be designed to minimize the vibration without requiring a large number of sensors (to minimize costs). A PI controller is chosen here with measurement of only the absorber mass acceleration. The control force of a PI controller with acceleration feedback is

$$F_c = K_p \ddot{x}_a + K_i \dot{x}_a \quad (1)$$

where K_p and K_i are the proportional and integral gain coefficients respectively and x_a is the position of the absorber mass. Having this force as the control command, the equation of motion of the absorber mass is

$$m_a \ddot{x}_a + c_a \dot{x}_a + k_a x_a = K_p \ddot{x}_a + K_i \dot{x}_a \quad (2)$$

Taking the Laplace transform one has $m_a s^2 + c_a s + k = K_p s^2 + K_i s$. By selecting the K_p and K_i such that the roots of the above equation are placed on the imaginary axis at $\pm j\omega$, the absorber system will mimic a resonator at the frequency of ω . This system can cancel the vibrations of the primary system at the frequency ω . Solving the equation with $s = \pm j\omega$, one has

$$K_p = m_a - k_a/\omega^2, \quad K_i = c_a \quad (3)$$

With an electromagnetic actuator, the control command is the applied voltage. Therefore, the dynamics of the electromagnetic actuator must be considered. The block diagram representation of the absorber system with the electromagnetic actuator is provided in Figure 3. In this figure, L and R are the inductance and the DC resistance and K_f and K_b are the force sensitivity and back emf constants of the actuator. To get the velocity data from the acceleration data, a first order Butterworth highpass filter with 0.5 Hz cutoff frequency is used before integration to eliminate the low frequency drift of the accelerometer data.

Again, to have a vibration absorber at the frequency ω , K_p and K_i should be selected such that the roots of the closed loop system are at $\pm j\omega$. Finding the closed loop transfer function of the whole system and forcing it to have poles at $\pm j\omega$, the integrator and proportional gains can be found

$$K_p = \frac{R(k_a/\omega^2 - m_a) - Lc_a}{K_f} \quad (4)$$

$$K_i = \frac{L(k_a - m_a\omega^2) + Rc_a + K_f K_b}{K_f} \quad (5)$$

It is straightforward to show that the highpass filter has a very small effect on K_p and K_i and therefore is neglected in (4) and (5).

4 Hand-Held Probe Design

Figure 3 shows the hand-held device. Parts labelled as (c) and (d) are the primary actuator’s coil and magnet. This actuator vibrates the tissue, while the ultrasound probe (a) images the tissue. The absorber actuator is mounted in an opposite way: its coil (e) is fixed and the magnet (f) moves. Therefore the absorber mass is simply the magnet of the actuator. This choice obviates an additional absorber mass, making the hand-held device lighter. Part (g) is one of the two absorber springs. There is no damper in the absorber; the inevitable coulomb friction between the magnet and the housing and also the rods and the bushings play the role of a damper. Parts (b) and (h) are the linear potentiometer and piezoelectric accelerometer respectively. The accelerometer is mounted on the absorber mass (the magnet) and provides the only feedback for the vibration absorption algorithm. The potentiometer signal is only used to operate the first actuator in closed loop for displacement control. Here a conventional PID controller is used and since the mathematical model of the vibrator is unknown, the coefficients of the PID controller are obtained by the Ziegler-Nicholas method [4].

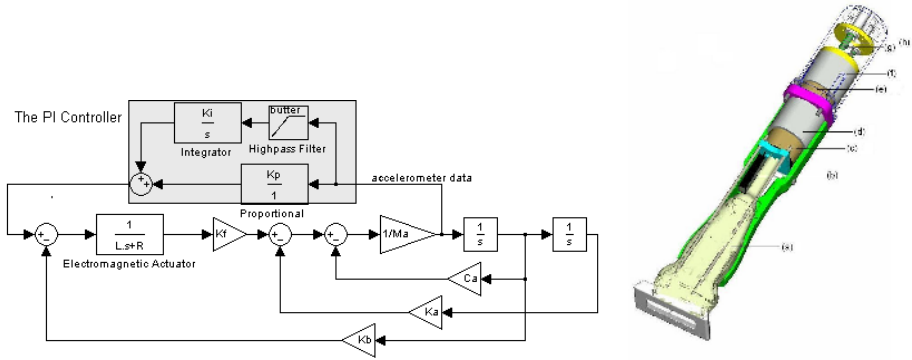


Fig. 3. Left: Controller for vibration absorber. Right: Exploded view of the hand-held probe. (a) ultrasound transducer; (b) potentiometer; (c) and (e) actuator coils; (d) and (f) actuator magnets; (g) absorber spring; (h) accelerometer.

The probe is designed for a desired frequency range of 10-50 Hz to cover a sufficient bandwidth of the tissue response. The values of m_a , k_a and b_a should be chosen from the stability analysis of the absorber. To cover the desired bandwidth, m_a is 180 grams and k_a is 650 N/m (implemented as two serried Spaenaur B-646 springs). The value for c_a can vary between 1 kg/s and 15 kg/s since damping comes from friction.

5 Identification of the Absorber Parameters

After manufacturing the device, the exact values of m_a , k_a and b_a need to be measured. The value of m_a was measured by a scale and is 199 grams. To mea-

sure k_a and b_a , a free vibration test on the absorber mass was done and the acceleration measured by the accelerometer. Assuming that the free vibration is of the form $x(t) = Ae^{-\zeta\omega_n t} \sin(\omega_d t + \phi)$ where $\omega_n = \sqrt{k/m}$ is the natural frequency and ζ is the damping coefficient and ω_d is the frequency of damped oscillation, the two unknowns ω_n and ζ can be found by matching the measured motion to the free vibration equation. This gives $\omega_n = 9.0$ Hz and $\zeta = 0.5$. These translate into $k_a = 636.4$ N/m, $c_{ad} = 11$ Kg/s, and $\omega_d = 7.8$ Hz. The coulomb friction is the major source of damping (plus eddy currents induced by the moving magnet in the housing) and the given c_a value is the result of linearization. Because of its nonlinearity, the measured c_a is valid for the frequency of the test so is labelled c_{ad} . The c_a at the frequency ω is then modelled as[7]

$$c_a = \frac{\omega_d}{\omega} c_{ad} \quad (6)$$

6 Simulation Results

The performance of the proposed active vibration absorber is investigated by simulation. Figure 4 compares the vibration of the probe with and without active absorption. An excitation is applied to the probe held by a simple model of a human arm (stiffness of 100 N/m and damping of 2 kg/s). An attenuation level of 70 dB is achieved with the active vibration absorber after 5 sec.

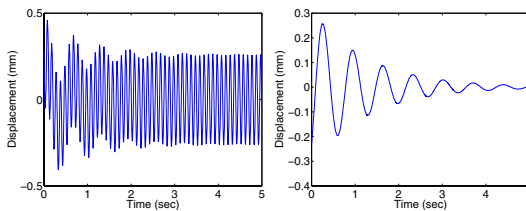


Fig. 4. Simulation of probe motion without and with the vibration absorber (left and right respectively). Tissue is excited at 10Hz with an amplitude of 1 mm.

7 Experimental Results

Several experiments were carried out on the probe to show the effectiveness of the active vibration absorber. A mechanical arm was built to hold the probe in order to perform repeatable and comparable experiments. The arm has a stiffness of 140 N/m, mimicking the stiffness of the human arm holding the probe. Speckle tracking is performed by a window-based cross-correlation algorithm [5]. An OPTOTRAK 3020 (NDI, Waterloo, Canada) tracker with a measurement rate of 400 Hz is used to track the motion of the probe. The OPTOTRAK is only used to investigate the performance of the vibration absorber and is not used in the vibration absorption control. In the first experiment, the probe vibrates a

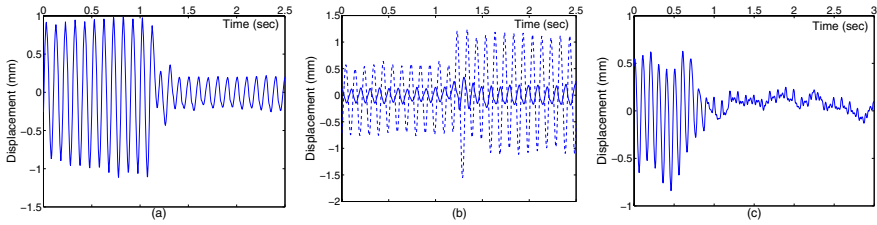


Fig. 5. Experimental results. (a) Probe vibration on a gelatin phantom with a mechanical arm, before and after activating the vibration absorption ($t=1.1$ s). (b) Another test looks at an increase in the excitation amplitude from 2 mm to 3 mm at $t=1.3$ s. Dashed and solid curves show the motion of the probe without and with the active vibration absorber respectively. (c) Probe vibration for experiments with a human operator on real tissue. The vibration absorber is turned on at $t=0.7$ s.

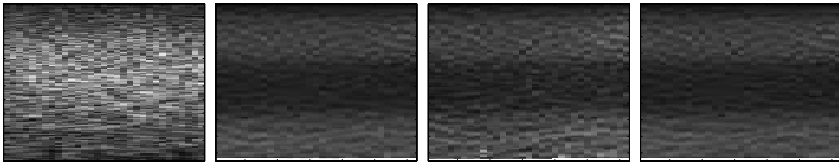


Fig. 6. left to right respectively: B-mode image, best-case example of static elastogram, worst-case static elastogram, and the VE of a three layered phantom

phantom at a frequency of 10 Hz . Figure 5a shows that the amplitude of probe vibration is reduced from 0.9 mm to 0.2 mm when the active vibration absorber is turned on at $t=1.1$ s. To see the performance of the probe to absorb forces with variable amplitude, two experiments with and without active vibration absorption are done. In these experiments, the excitation amplitude is increased from 2 mm to 3 mm at $t=1.3$ s (Figure 5b). The amplitude of vibration is increased from 0.6 mm to 1.1 mm without vibration absorption and from 0.12 mm to 0.17 mm with the vibration absorption. The experiment shows an increase in the attenuation level when the excitation force is increased. In Figure 5c, the probe is held by a human operator on the forearm of a human subject. In this experiment, the vibration absorber decreases the peak to peak motion from 1.28 mm to 0.13 mm. Because the probe is held by hand and applies a slightly uneven pressure to the tissue, a net motion at a low 2 Hz frequency (compared to the 10 Hz vibration frequency) is seen. This motion is easily seen in the frequency domain and excluded from VE. Experiments with a 20 Hz excitation frequency show similar 0.9 mm to 0.2 mm vibration reduction. A B-mode ultrasound image and three real elastograms obtained from a three layered gelatin phantom are shown in Figure 6. The middle layer of the phantom is stiffer than the other layers but exhibits the same acoustic properties. The stiff layer is not visible in the B-mode image but it is detectable in the elastograms. The static elastograms (second and third images) are obtained by applying a 0.01 compression to the

phantom slowly and rapidly acquiring RF data during the compression, to ensure high correlation coefficients between consecutive echo frames. A vibro-elastogram is obtained from the same phantom by vibrating it with a 5 Hz band limited white noise. Static elastograms can be noisy (e.g. third image in Figure 6), but vibro-elastograms are more robust to noise due to their averaging step performed in frequency domain.

8 Conclusions

Despite the uncertainties in the absorber model and the friction model, the vibration absorber performed well. The control law of the active vibration absorber requires only the measurement of the absorber mass acceleration. Also, the control law is decoupled from the primary system and the amplitude of the excitation force. This property is inevitable for the probe because of two main reasons. First, the properties of the primary system can change dramatically since the probe can be held by various operators with various inertia, damping and stiffness properties. Second, the probe vibrates various types of tissues with various mechanical properties so the excitation force can vary considerably. As a final note, the designed PI controller has the advantage of not requiring a double integration from the acceleration data compared to the PD controller proposed in [1]. Since the accelerometer data has a low frequency variable drift, it is impossible to twice integrate the acceleration data because of the accumulation of error. By replacing the bushings with linear bearings and increasing the air gap between the magnet and the housing to decrease the eddy currents, the performance of the vibration absorber will be improved to the simulation results. Given a working probe, work is now underway on clinical applications of VE.

References

1. Elmali, H., Renzulli, M., Olgac, N.: Experimental comparison of delayed resonator and PD controlled vibration absorbers using electromagnetic actuators: *J. of Dynamic Systems, Measurement and Control*. Vol. 122 (2000) 514–520.
2. Jalili, N.: A comparative study and analysis of semi-active vibration-control systems. *J. of Vibration and Acoustics*, Trans. of the ASME, Vol. 124(4) (2002) 593–605.
3. Korenev, B.G., Reznikov, L.M.: *Dynamic vibration absorbers: theory and technical applications*. John Wiley, New York (1993).
4. Ogata, K.: *Modern control engineering*. 4th edn. Prentice Hall, New Jersey (2002).
5. Ophir, J., Alam, S.K., Garra, et al.: Elastography: imaging the elastic properties of soft tissues with ultrasound. *J Med Ultrasound*, Vol. 29 (2002) 155–171.
6. Sun, J.Q., Jolly, M.R., Norris, M.A.: passive, adaptive and active tuned vibration absorbers—a survey. *J of Vibration and Acoustics* Vol. 117 (1995) 134–242
7. Thomson, W.T., Dahleh, M.D.: *Theory of vibration with applications*. 5th edn. Prentice Hall, New Jersey (1998)
8. Turgay, E., MacIntosh, S., Rohling, R., Salcudean S. Parameter identification of tissue lumped models based on sequences of ultrasonic strain images. *Second Intl. Conf. on Ultrasonic Measurement and Imaging of Tissue Elasticity*. Corpus Christi, Texas (2003) 29.

Real-Time Quality Control of Tracked Ultrasound

Emad M. Boctor, Iulian Iordachita, Gabor Fichtinger, and Gregory D. Hager

Engineering Research Center, Johns Hopkins University
eboctor@ieee.org, iordachita@jhu.edu, gabor@cs.jhu.edu,
hager@cs.jhu.edu

Abstract. The overwhelming majority of intra-operative hazard situations in tracked ultrasound (US) systems are attributed to failure of registration between tracking and imaging coordinate frames. We introduce a novel methodology for real-time in-vivo quality control of tracked US systems, in order to capture registration failures during the clinical procedure. In effect, we dynamically recalibrate the tracked US system for rotation, scale factor, and in-plane position offset up to a scale factor. We detect any unexpected change in these parameters through capturing discrepancies in the resulting calibration matrix, thereby assuring quality (accuracy and consistency) of the tracked system. No phantom is used for the recalibration. We perform the task of quality control in the background, transparently to the clinical user while the subject is being scanned. We present the concept, mathematical formulation, and experimental evaluation in-vitro. This new method can play an important role in guaranteeing accurate, consistent, and reliable performance of tracked ultrasound.

1 Background and Significance

Ultrasound (US) imaging has become a widely accepted guidance modality in medical interventions, because it is real-time, safe, and of low cost. Significant research has been dedicated to quantitative tracked ultrasound, involving tracking the US probe in 3D space with respect to a stationary frame of reference. While tracked US originates from interventional applications, it recently has become an indispensable tool in external beam radiation therapy (EBRT) guidance [1] that is expected to become the largest user of tracked ultrasound in the next couple of years. Each year in the United States 65,000 patients are treated for prostate cancer alone. Considering an average of 40 treatment fractions that each patient receives, the total number of procedures accumulates to 2.6 million a year. Tracked US guidance is also applicable to the EBRT of breast cancer, adding about 2 million more cases to the potential market.

Typically, tracking is achieved by rigidly attaching 3D localizers to the US probe. The missing link, however, is the spatial transformation between the US image pixels and the 3D localizers on the probe, which requires calibration. Hence, calibration is ubiquitously present in all systems where ultrasound is used for quantitative image guidance. From our experience, the wide majority of intra-operative hazard situations in tracked US systems are caused by failure of registration between tracking and imaging coordinate frames, thereby manifesting in miscalibration of the tracked US.

The most typical form of errors is false reading of the tracker that occurs quite often in electromagnetic tracking systems due to invisible field distortions caused by metal objects or electromagnetic noise. Another typical problem related to tracking is deformation or physical damage of the tracking body attached to the probe, causing a latent misreading of pose. What makes these problems exceedingly dangerous is that they occur without apparent warning. Among human operator errors, inadvertent changes of lateral image polarity occur quite frequently and transparently to the clinical user. With regular off-line recalibration some of the aforementioned errors can be caught prior to procedure. The process is called *Quality Control (Q/C)*, a mandatory routine in any clinical department. Typically, Q/C is performed annually, monthly, or weekly, which places a heavy financial burden on the department. In addition to increasing patient safety, reduction in Q/C costs clearly is an incentive.

In all current calibration methods, a set of objects (often referred to as phantoms) of known geometrical properties are scanned and then various mathematical procedures are applied to discern the unknown transformation that maximizes the similarity between the US images and the phantom [3-6]. The use of special objects and phantoms, however, is cumbersome and foreign to the operating room where interrupting the procedure for the sake of probe calibration is not practical. What is required therefore is a paradigm shift in calibration technology to phantomless self-calibration that is performed directly on the patient, intra-operatively, in real-time, transparently to the physician.

Generally, full calibration involves six degrees-of-freedom (DOF) of rigid transformation and three dimensional scaling. However, according to a practical observation, the possibility of pure translational error in the calibration matrix is practically low. One atypical scenario would involve the rigidly attached sensor to slide without experiencing any relative rotation with US image reference frame. Thus for the purpose of Q/C, it is sufficient to recalibrate the system for the remaining degrees-of-freedom. This paper presents the concept, mathematical framework, experimental implementation, and in-vitro evaluation of a phantomless real-time method that detects intra-operative failures of the tracked US while recovering the calibration matrix to a 2D scale factor and then achieve full calibration in an additional step.

2 Mathematical Formulation

The key enabler of our self-calibration method is a closed-form mathematical formulation of the problem. Fig. 1 presents the coordinate systems for the mathematical formulation. A_1, A_2 are the transformations of US picture coordinate system (P) with respect to the fixed construction frame (C) at poses 1 and 2, respectively. Note that the actual selection of C is arbitrary and the only requirement is that it must be rigidly fixed during the calibration process. Using A_1 and A_2 , we obtain the transformation between poses 1 and 2, as $A=A_2A_1^{-1}$. At the same time, the transformation between the two poses can be recovered using

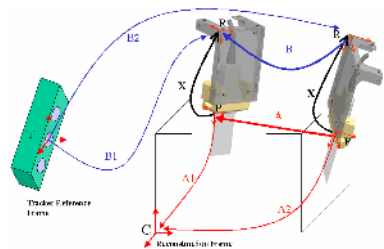


Fig. 1. Coordinate definitions in the closed form $AX=XB$

a calibration phantom or recovered directly by matching the 2D ultrasound images acquired in these poses to a prior 3D model of the phantom object. B_1 and B_2 are readings from the tracker for the sensor frame (R) with respect to tracker reference frame (T), at poses 1 and 2 respectively. The relative pose between the sensor frame (R) at pose 1 and 2 is given by $B = B_2^{-1}B_1$. This yields the following homogeneous matrix equation:

$$\begin{aligned} AX &= XB, \text{ which is :} \\ R_a R_x &= R_x R_b, \text{ and} \\ R_a t_x + D(u_a)\bar{\lambda} &= R_x t_b + t_x \end{aligned} \tag{1}$$

Where A is estimated from images, B is assumed to be known from the external tracking device, and X is the unknown transformation between the US image coordinate system and the sensor frame (R). Expanding this equation yields two separate constraints on translation and rotation, where R_a is the rotation of the US image frame between pose 1 and 2, λ is the unknown scale factor vector that relates the translation vector u_a in voxel space to the US image frame translation vector t_a (usually expressed in mm), such that:

$$t_a = \begin{pmatrix} \lambda_x u_{ax} \\ \lambda_y u_{ay} \\ \lambda_z u_{az} \end{pmatrix} = \begin{pmatrix} u_{ax} & 0 & 0 \\ 0 & u_{ay} & 0 \\ 0 & 0 & u_{az} \end{pmatrix} \cdot \begin{pmatrix} \lambda_x \\ \lambda_y \\ \lambda_z \end{pmatrix} = D(u_a)\bar{\lambda}. \tag{2}$$

It is important to account for the most general case where the scale factor λ (which again converts from voxel space to image space) is not known. This scenario typically occurs when A is recovered by registering the US image to some a priori known model (or phantom) given in voxel space. In the linear formulation of the problem, we use the linear operator vec and the *Kronecker product* (\otimes)[8]. Using the following property of the Kronecker product $vec(CDE) = (C \otimes E^T)vec(D)$, we rewrite Eq. 1 as:

$$\begin{aligned} (R_a \otimes R_b)vec(R_x) &= vec(R_x), \text{ and} \\ (I_3 \otimes t'_b)vec(R_x) + (I_3 - R_a)t_x - D_u \lambda &= 0 \end{aligned}$$

This leads to a homogeneous linear system :

$$\begin{bmatrix} I_9 - R_a \otimes R_b & 0_{9 \times 3} & 0_{9 \times 3} \\ I_3 \otimes t'_b & I_3 - R_a & -D_u \end{bmatrix} \begin{pmatrix} vec(R_x) \\ t_x \\ \lambda_{3 \times 1} \end{pmatrix} = \begin{pmatrix} 0_{9 \times 1} \\ 0_{3 \times 1} \end{pmatrix} \tag{3}$$

The solution for this homogeneous linear system could be given by finding the null space, which is a subspace in R^{15} . Then the unique solution could be extracted from the null space using the unity constraint to the first 9 coefficients representing the R_x . However, a better solution is described in [9] where the system is solved in two steps: first extract the rotation, and then solve for the translation and scale. The complete algebraic analysis for this problem (where the scale factor is assumed to be the same in three directions) is given in [7], where it is proved that two independent motions with non-parallel axes is sufficient to recover a unique solution for $AX=XB$. We have extended this solution method to account for inhomogeneous scale in the three coordinate axes. As discussed in [5], two independent motions with A 's and B 's are sufficient to recover X . Therefore, our present task reduces to recover A 's as we are scanning real tissue and collecting the corresponding B 's from the tracker, and then to obtain the calibration by solving the homogenous linear system in Eq. 3. The

challenge lays in the full recovery of the A 's, including all six DOF. The in-plane motion parameters can be recovered with sub-pixel accuracy in real-time, using speckle based tracking algorithms applied to the B-mode images. The out-of-plane motion parameters are difficult to recover from B-mode and we performed this task only with RF or pseudo-RF data, the latter being decompressed B-mode data as described by Prager *et al.* [13]. Out-of-plane motion, however, is not necessary for our immediate task of Q/C, thus it is omitted from further discussion in this paper.

Translation Motion:

This motion scenario is realized by moving the 2D/3D US probe in translational sweep (without rotation) to collect nearly parallel stack of images and/or series of 3D slabs. Also it can be shown in a panoramic scan where the images can be stitched together without introducing a relative rotation. Given this kind of motion, $Rb = Ra = I_3$, leading to $Rx * t_{bi} = t_{ai}$, where i denotes the motion. Using the property of the Kronecker product used in Eq. 1:

$${}_{3 \times 9} (I_3 \otimes t_{bi}^T) *_{9 \times 1} \text{vec}(Rx) = t_{ai}, \text{ we have}$$

$$t_{ai} = \begin{bmatrix} \lambda_x * u_{ai}(x) \\ \lambda_y * u_{ai}(y) \\ \lambda_z * u_{ai}(z) \end{bmatrix} = \begin{bmatrix} u_{ai}(x) & 0 & 0 \\ 0 & u_{ai}(y) & 0 \\ 0 & 0 & u_{ai}(z) \end{bmatrix} \begin{bmatrix} \lambda_x \\ \lambda_y \\ \lambda_z \end{bmatrix} = D(u_{ai}) \vec{\lambda}$$

Three independent translations are sufficient to obtain a full-rank system. Solving for Rx and the three scale factors:

$${}_{9 \times 9} \begin{bmatrix} (I_3 \otimes t_{b1}^T) \\ (I_3 \otimes t_{b2}^T) \\ (I_3 \otimes t_{b3}^T) \end{bmatrix} *_{9 \times 1} \text{vec}(Rx) = \begin{pmatrix} t_{a1} \\ t_{a2} \\ t_{a3} \end{pmatrix} = \begin{bmatrix} D(u_{a1}) \\ D(u_{a2}) \\ D(u_{a3}) \end{bmatrix} * \begin{pmatrix} \lambda_x \\ \lambda_y \\ \lambda_z \end{pmatrix}$$

$$\text{vec}(Rx) = \frac{1}{\Delta} \underbrace{\begin{bmatrix} I_3 \otimes (t_{b2} \times t_{b3}) & I_3 \otimes (t_{b3} \times t_{b1}) & I_3 \otimes (t_{b1} \times t_{b2}) \end{bmatrix}}_{M^{-1}} * \begin{bmatrix} D(u_{a1}) \\ D(u_{a2}) \\ D(u_{a3}) \end{bmatrix} * \begin{pmatrix} \lambda_x \\ \lambda_y \\ \lambda_z \end{pmatrix}$$

Using Khatri-Rao product [8] which defined as:

$${}_{q \times u} F \circ_{r \times u} G = [F_1 \otimes G_1 \quad F_2 \otimes G_2 \quad \dots \quad F_u \otimes G_u], \text{ and the following fact}$$

$$\text{vec}(AVD) = (D^T \circ A) \text{vecd}(V), \text{ vecd is a diagonal vec operator}$$

$$\text{vec}(Rx) = \frac{1}{\Delta} \left[\begin{pmatrix} D \\ I_3 & I_3 & I_3 \end{pmatrix} \circ (v_1 \ v_1 \ v_1 \ v_2 \ v_2 \ v_2 \ v_3 \ v_3 \ v_3) \right] * \underbrace{\begin{bmatrix} D(u_{a1}) \\ D(u_{a2}) \\ D(u_{a3}) \end{bmatrix} * \begin{pmatrix} \lambda_x \\ \lambda_y \\ \lambda_z \end{pmatrix}}_V$$

$$\text{vec}(Rx) = \frac{1}{\Delta} \text{vec}(A * V * D^T), \text{ from linearity of vec operator}$$

$$Rx = \frac{1}{\Delta} A * V * D^T = \frac{1}{\Delta} \begin{bmatrix} \lambda_x (u_{a1}(x)v_1 + u_{a2}(x)v_2 + u_{a3}(x)v_3) \\ \lambda_y (u_{a1}(y)v_1 + u_{a2}(y)v_2 + u_{a3}(y)v_3) \\ \lambda_z (u_{a1}(z)v_1 + u_{a2}(z)v_2 + u_{a3}(z)v_3) \end{bmatrix}^T$$

We have obtained a closed-form solution that recovers an anisotropically scaled rotation matrix from three independent translations. We can recover the three scale factors by applying the unit constraint on each column vector, and recover the underlying rotation by then rescaling the columns to be unit vectors. This basically proves that three independent motions are sufficient to recover the rotation and the three scale factors. However, the third motion is not even necessary. It can be shown [7] that given two independent motions t_{b1} , t_{b2} , the third constraint comes from the cross-product between t_{b1} , t_{b2} .

In order to map this analysis to our application, several requirements must be considered. The ultrasound machine generates real-time 2D US pixel-map, meaning that we have only two unknown scale factors in x and y , denoted as λ_x and λ_y . With sweeping probe motion we obtain multiple poses that suggests folding the closed form representation into a least squares problem. Starting from the following equation:

$$\begin{aligned}
 Rx * t_{bi} = t_{ai}, \text{ we have} \\
 t_{ai} = \begin{bmatrix} \lambda_x * u_{ai}(x) \\ \lambda_y * u_{ai}(y) \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda_x & 0 & 0 \\ 0 & \lambda_y & 0 \\ 0 & 0 & \lambda_z \end{bmatrix} \begin{bmatrix} u_{ai}(x) \\ u_{ai}(y) \\ 0 \end{bmatrix} = D(\vec{\lambda})u_{ai} \tag{4} \\
 [t_{b1} \ t_{b2} \ \dots \ t_{bn}] = Rx^T D(\vec{\lambda}) [u_{a1} \ u_{a2} \ \dots \ u_{an}]
 \end{aligned}$$

The solution of this equation can be achieved in many ways. One can solve non-linearly for the three rotations and the two scale factors (λ_x , λ_y). Alternatively, one can solve linearly for the nine parameters of the scaled rotation and perform QR factorization with positive scale factor constraints. Or simply, apply the norm constraint on Eq. 4 as follows:

$$\|Rx * t_{bi}\|^2 = \|t_{bi}\|^2 = \|D(\vec{\lambda})u_{ai}\|^2 = u_{ai}^T D(\vec{\lambda})^2 u_{ai}$$

This allows solving for both (λ_x , λ_y). After recalculating the scaled u_{ai} , we obtain two sets of points governed by an SO(3) rotation and recovered with Horn’s method[11].

Planar Motion:

Planar motion is more general as it also allows, in addition to translation, in-plane rotation. This type of sweeping probe motion is commonly applied clinically. With this general motion, one can recover both rotation and anisotropic scale factors as shown before, but it will not yield a full recovery of the position offset t_x . By applying Eq. 3 we obtain:

$$\begin{bmatrix} I_3 - R_{ai} \otimes R_{bi} & 0_{9 \times 3} \\ I_3 \otimes t'_{bi} & -D_{ai} \end{bmatrix} \begin{pmatrix} \text{vec}(R_x) \\ \lambda_{3^{*1}} \end{pmatrix} = \begin{pmatrix} 0_{9 \times 1} \\ 0_{3^{*1}} \\ \vdots \end{pmatrix}, \text{ and} \tag{5}$$

$$\begin{bmatrix} (I_3 - R_{ai}) \\ \vdots \end{bmatrix} t_x = \begin{bmatrix} D_{ai} \vec{\lambda}_{3^{*1}} - (I_3 \otimes t'_{bi}) \text{vec}(R_x) \\ \vdots \end{bmatrix} \tag{6}$$

Note that when $R_b=R_a=I_3$, Eq. 5 becomes similar to the pure translation case in Eq. 4. Also note that Eq. 6 is always under constraint as (I_3-R_a) has rank 2, regardless to the number of in-plane rotations, meaning there is no single solution for t_x and the general solution will have exactly one (the number of unknowns minus the rank) arbitrary scale factor ζ . So the solution can take the form:

$$t_x(\zeta) = t_o + \zeta * t_n \tag{7}$$

Where ζ is the unknown scale factor. t_o is a unique solution in the plane of motion (2-dimensional), since $(I_3 - R_a)$ has rank 2. t_n is the normal to the plane of motion. In our case, if the plane of motion is the US image plane (the x-y plane), t_n may equal $(0 \ 0 \ 1)^T$, which is a unit vector in the z-direction and thus perpendicular to the plane of motion. In our US calibration and Q/C system, we can recover this unknown ζ , if we have a collection of a previously scanned cross-wire tar get from different insonification angles, the idea being is that the cloud of reconstructed cross-wire targets will be the smallest (i.e. standard deviation minimal) for the correct value for ζ .

3 Experimental System

In our experimental prototype (Fig. 2), we used a SONOLINE Antares US scanner (Siemens Medical Solutions USA, Inc. Ultrasound Division, Issaquah, WA), with a Siemens VF 10-5 linear array probe. The patient was replaced with a tissue mimicking agar phantom. The US probe was held against the phantom in a rigid acrylic holder mounted on an adjustable steady arm. The arm had dual purpose: to adjust the spatial position of the tracked US probe over the calibration phantom, and to ensure temporal synchronization between the tracker and the US scanner. Multiple optical markers were attached to the probe holder, which then were tracked by an OPTOTRAK device (Northern Digital Inc.).



Fig. 2. In-vitro experimental setup

The tissue mimicking phantom provides realistic images of fully developed speckles. Its construction is based on a recipe by Fenster *et al.* [10]. Three percent by weight of agar gel (A-7002 Agar, Sigma-Aldrich, St. Louis, MO) was added to distilled water, with three percent by weight 50µm cellulose particles (S-5504 Sigmacell, Sigma-Aldrich), and with seven percent by volume glycerol (W25250, Sigma-Aldrich). The mixture was heated to 92 C°, stirred constantly, gradually cooled to 60 C°, and then poured into a container mold. We also introduced specular features and structures to mimic bone appearance and to allow for testing algorithmic performance under different echogenicity conditions.

Algorithm Workflow

The workflow in the real-time Q/C procedure is described in Fig 3. The *Acquisition Module* receives US video signal and tracker reading, from which it prepares synchronized indexed sequences of images and tracking information. The *Motion Analyzer* sorts out

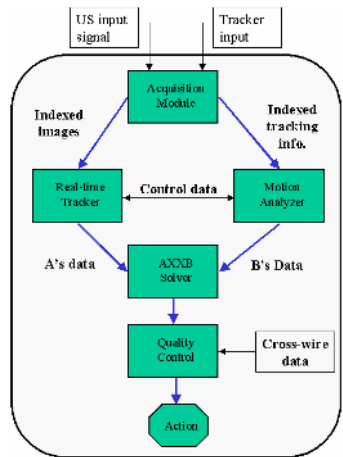


Fig. 3. The Q/C workflow

the types of motions in these sequences and sends control signal for the *Real-time Tracker*, which recovers the A matrices. Finally, the $AX=XB$ solver receives corresponding A 's and B 's data, recovers the X calibration matrix. The *Quality Control* unit analyzes the new calibration and compares it with previous runs. In case of suspected discrepancy, an appropriate *Action* is initiated to deal with the hazard condition. The action could range from generating a warning message to demanding to halt the procedure and fully recalibrate the system. It must be noted that a full recalibration is necessary only if a scale factor is found to be at fault, because the partial calibration recovers all other parameters during the Q/C process. The remaining scale factor can be easily recovered from a pre-scanned geometrical object, such as a cross-wire, as described earlier in relation to Eq. 7.

Real-time Tracker

As mentioned above, the role of *Real-time Tracker* is to recover the A matrices, the motion of the US image in construction frame, as it was described in Fig. 1. What is necessary is to compute the relative motion in pairs of ultrasound images for which the absolute (tracked) motion is known. We accomplish this using direct image registration methods similar to those described in [12]. Specifically, we introduce an intermediate “warped” image representation W defined as

$$W(u, t; p, \alpha) = I(Rot(\alpha) * u + p)$$

where $u=(x,y)^T$ is an image location, P is a translation offset, and α is an interframe rotation. Let $W(t; p, \alpha)$ denote the column vector constructed by stacking the value of W for all possible image locations u . We then compute an estimate of the offset (P_b, α_t) between images at time t and $t+d$ by iterating the following equation:

$$(p_t^k, \alpha_t^k) = (p_t^{k-1}, \alpha_t^{k-1}) + J^{\#}(W(d + t; p_t^{k-1}, \alpha_t^{k-1}) - W(t;0,0))$$

where $J^{\#}$ denotes the pseudo-inverse of the Jacobian matrix of W with respect to P and α . The values of (P_b, α_t) is taken from the previous frame.

4 Experiments and Results

We used the experimental system to collect 5 datasets, altogether containing 20 motions in B-mode, of rectangular view and 8 cm depth. One of the datasets contained 4 motions and was obtained under a faulty condition. We purposely flipped lateral polarity of the B-mode image to simulate a common operator error. Table 1 shows the rotation and scale reported by the Q/C system. In testing the image tracker, we used $d=10$ step size, for

Table 1. Q/C system report on rotation and scale

	Magnitude of rotation (Rodrigues form)	Scale in x and y (mm/pixel)	
Dataset-1	3.163	0.22	0.26
Dataset-2	3.072	0.23	0.26
Dataset-3	2.992	0.23	0.27
Dataset-4	3.008	0.23	0.29
Dataset-5	0.086	0.24	0.27

which images were matched with an RMS gray-scale error of less than 2% of signal value, suggesting a registration error of less than 1/10 pixel [12]. The heart of the Q/C system, as seen in Fig. 3, are the real-time image tracker and $AX=XB$ solver. The image tracker is sensitive to the step d between registered images, while the $AX=XB$ solver is sensitive to the number and type of motions used to recover the calibration parameters. The upper graph in Fig. 4 illustrates the relation between the recovered scale and number of images and the step size d taken in the image registration algorithm, for “bad” (appears in the upper graph for the first 30 images) and “good” (lower graph). Note that number of images can be represented by the scanning time, where 1 second corresponds to 33 frames, sweeping over a certain distance covered with the given scanning velocity. The x-axis represents the number of frames used in the $AX=XB$ solver and the y-axis is the scale ratio in mm/pixel. The upper figure is for λ_x , the lower is for λ_y . Figure 4 also shows convergence for the scale ratio under different image registration steps (small d : 1, 3, and 5; large d : 10, 20, 30, and 40). Note that as we increase the step size, we also introduce a delay equal to the step size before we start estimating a given parameter. (It is because we must wait for the d^{th} image to come.)

We have found that steps bigger than 40 are not reliable to track and predict the motions of the speckle-based images. At the same time, they are considered bad kernels, suggesting that we have to wait for about 2 Sec. to detect a problem. Yet smaller tracking steps need fewer readings to start convergence, due to the small motions they deliver to the $AX=XB$ solver. Note that the convergence for $d= 1, 2,$ and 3 appears after 20 small motions for λ_y and appears after 30 small motions for λ_x , this is mainly due to the type of motions. The intuition behind this is, we can’t estimate a scaling parameter in a direction normal to the direction of motion. Similarly favorable results were obtained for the rotation component, not discussed here due limitation of space. *We can conclude that given the right motion, a kernel of 10 steps ($d=10$) converges in 10-20 steps, meaning 0.3-0.6 sec with a total travel of ~1.5mm (scanning speed @ 3mm/sec).*

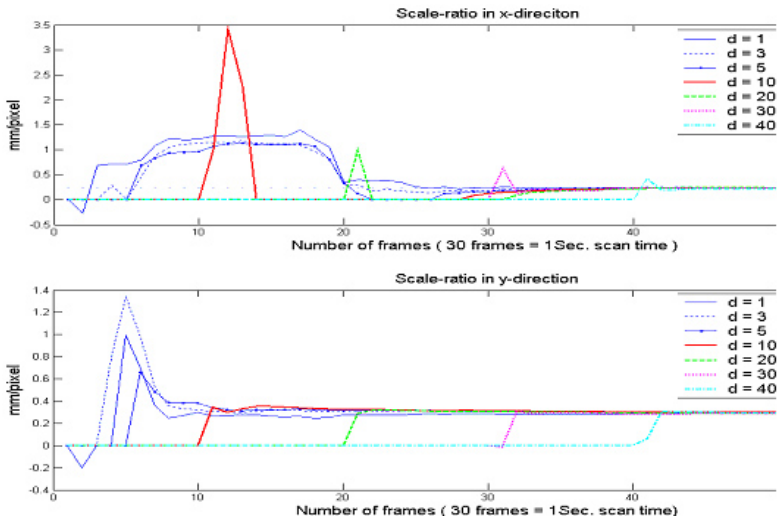


Fig. 4. Recovered scale versus the number of images and step size

5 Conclusion and Future Work

Our Q/C system, that reports the calibration matrix robustly and consistently, recovered the correct calibration parameters under normal working conditions by monitoring the constancy of calibration matrix and did not produce false alarm. It also managed to distinguish the faulty condition through catching the outlier calibration matrix. It is a work in current progress to extend the Q/C framework to estimate the out-of-plane motion of the US probe, thereby providing full calibration in-vivo, real-time, as the patient is being scanned, which will obviate phantom-based calibration. The real-time performance of the Q/C system allows for averaging hundreds of independent calibrations from a single sweep, which in turn promises to retire cross-wire based reconstruction accuracy evaluation and at last make US calibration free from any phantom, whatsoever. The key enabler to this is tuning the *Real-time Tracker* module to coping with out-of-plane motion components in the A matrix, for which speckle decorrelation techniques are under investigation. We will also compare the accuracy of our current partial recalibration against published off-line calibrations. Our present ability to reproduce calibration parameters to a 2D scale factor is most promising in this regard. Last but not least, the sensitivity and specificity of the Q/C system will be examined. A clinical-grade Q/C system must catch all faulty conditions (high sensitivity), yet it must not halt the clinical intervention with producing false alarm (high specificity). Altogether, the combined Q/C and calibration system appears to have high practical utility for clinical departments that use and maintain tracked ultrasound systems. Finally, the authors acknowledge the financial support from the NSF #EEC 9731478 and loan equipment from Siemens Corporate Research.

References

1. Langen KM, Pouliot J, Anezinos C, et al. Evaluation of ultrasound-based prostate localization for image-guided radiotherapy, *Int J Rad. Oncol Biol Phys.* 57(3):635-44.
2. Prager RW, Rohling RN, Gee AH, Berman L. Rapid Calibration for 3-D Freehand Ultrasound, *US in Med. Biol.*, 24(6):855-869, 1998.
3. Pagoulatos N, Haynor DR, Kim Y. A Fast Calibration Method for 3-D Tracking of Ultrasound Images Using a Spatial Localizer, *US in Med. Biol.*, 27(9):1219-1229
4. Rousseau F, Hellier P, Barillot C. A fully automatic calibration procedure for freehand 3D ultrasound, *In IEEE Int. Symp. on Biomedical Imaging*, 2002.
5. Boctor EM, Viswanathan A, Choti MA, Taylor RH, Fichtinger G, Hager G. A Novel Closed Form Solution for Ultrasound Calibration, *In IEEE ISBI* 2004.
6. Muratore DM, Galloway RL. Beam calibration without a phantom for creating a 3-D freehand ultrasound system, *Ultrasound Med Biol.* 2001 Nov;27(11):1557-66.
7. Andreff N, Horaud R, Espiau B. Robot Hand-Eye Calibration Using Structure from Motion, *International J. of Robotics Research*, 20(3), pp 228-248, 2001.
8. Brewer JW. Kronecker Products and Matrix Calculus in System Theory, *IEEE Trans. Circuits and systems*, 25(9) Sep.1978.
9. Tsai RY, Lenz RK. A New Technique for Fully Autonomous and Efficient 3D Robotics Hand/Eye Calibration, *IEEE Trans. On Robotics and Automation*, 5(3):345-358

10. Rickey DW, Picot PA, Christopher DC, Fenster A. A Wall-less Vessel Phantom for Doppler Ultrasound Studies, *US in Med. Biol.*, 21, 1163-1175, 1995.
11. Horn BKP. Closed-form solution of absolute orientation using unit quaternions, *J Opt Soc Am A* 1987;4:629-642.
12. Hager G, Belhumeur P. Efficient region tracking with parametric models of geometry and illumination. *IEEE PAMI Trans.*, 20(10):1125-1139, 1998.
13. Prager RW, Gee AH, Treece GM, Berman LH. Decompression and speckle detection for ultrasound images using the homodyned k-distribution. *Pattern Reco. Letters* 24:705-13.

Fully Truncated Cone-Beam Reconstruction on Pi Lines Using Prior CT*

Krishnakumar Ramamurthi¹, Norbert Strobel²,
Rebecca Fahrig³, and Jerry L. Prince¹

¹ Johns Hopkins University, Electrical and Computer Engineering,
Baltimore, MD 21218, USA

krishna@jhu.edu, prince@jhu.edu

² Siemens Medical Solutions, Stanford University Medical Center,
Stanford, CA 94305, USA

Norbert.Strobel@siemens.com

³ Stanford University, Department of Radiology,
Stanford, CA 94305, USA

fahrig@stanford.edu

Abstract. C-arms are well suited for obtaining cone-beam projections intra-operatively. Due to the compact size of the detector used, the data are usually truncated within the field of view. As a result, direct application of a standard cone-beam reconstruction algorithm gives rise to undesirable artifacts and incorrect values in the reconstructed image volume. When prior information such as a pre-operative CT scan is available, fully truncated cone-beam projections can be used to recover the change within a small region of interest without such artifacts. A method for integrating prior CT is developed using the concept of pi-lines and tested on real flat-panel and simulated cone-beam data.

1 Introduction

Three dimensional image reconstruction from C-arm cone-beam projections has been an area of growing practical interest over the last couple of decades. C-arms can provide rotational cone-beam images that can be used for the purpose of three dimensional reconstruction of anatomy. Filtered backprojection methods, such as Feldkamp's method [1], are among the most popular approaches to cone-beam reconstruction of C-arm images. These methods require a sufficiently large angular sweep, and require that the anatomy being scanned is entirely captured in the detector. In intra-operative applications, such as in hip replacement surgeries, these conditions cannot be practically met, due to patient size, orientation, and other objects such as tools, patient table, etc. that occlude the C-arm's view of the anatomy. These restrictions cause undesirable artifacts within the reconstruction field of view (FOV) thus rendering the reconstructions unsuitable for prognosis or planning. While intra-operative CT scans are quite impractical

* This work is jointly supported by NIH research grants R21-EB003616 (Johns Hopkins University), EB003524 (Stanford University) and Siemens Corporate Research.

for such applications, prior CT scans of the same patient may provide valuable information that can be used to compensate for aforesaid artifacts.

In this paper, we investigate the possibility of using prior CT information in cone-beam reconstruction, from fully truncated C-arm images. Recent developments in fan-beam and cone-beam tomography [2, 3, 4, 5] have shown, by introducing the concept of “Pi-lines”, that a certain degree of truncation (along the width of the detector) is admissible. However, when data is truncated at both ends of the detector these methods do not work. This is known as the *interior problem* and it has been shown that a unique inverse does not exist [6]. Our goal is to show that in certain cases prior CT information can be used to convert the interior problem into a *sub-interior problem*, i.e., a case where the aforementioned Pi-lines technique can be used. Consider for example an application involving the imaging of the pelvic bone during a sacroplasty procedure. Although a 3D pre-operative CT is typically obtained to find the regions of sacral fracture, it would also be helpful to obtain 3D reconstructions during the various steps of the procedure itself. We approach this problem by modeling the effects of the procedure as a change in the anatomy with respect to the prior CT. We show how this change can be recovered within a small region of interest from a semi-circular sweep of heavily truncated cone-beam projections. The x-ray source can be collimated to as narrow a beam as required to cover this region of interest thereby keeping radiational exposure to a minimum.

2 Background

2.1 Classical Framework

Cone-beam reconstruction techniques may be broadly classified based upon the geometry of the source path: (a) Non-planar source paths and (b) Planar source paths. Exact 3D reconstruction is possible only when the source path satisfies Tuy’s [7] condition, according to which every plane passing through any point of interest must intersect the source path non-tangentially at least once. It immediately follows that, only non-planar source paths are capable of satisfying this requirement for an entire volumetric region. Planar source paths can at best satisfy this requirement only for points contained within the plane containing the source path itself. In such cases it is simplest to develop cone-beam algorithms by simply extending existing exact 2D fan-beam reconstruction techniques above and below the mid-plane (the plane containing the source path). Feldkamp’s reconstruction method [1] for volumetric cone-beam reconstruction is an example of this approach. It is an extension of classical exact 2D fan-beam reconstruction algorithm on a circular path.

2.2 Recent Developments

In recent years, exact fan-beam reconstruction techniques have improved progressively by relaxing the requirements on the projection data. While classical fan-beam tomography cannot handle truncated data and requires a minimum

sweep of $\pi + \text{fan-angle}$ (short-scan), Noo et al [8] proposed a method that could work with a very short-scan sweep (less than a short-scan) for exact reconstruction in a region of interest (ROI) within the field of view (FOV). Truncation is still not admissible in this method however. An extension of this method to cone-beam tomography was proposed in [9, 10]. An important breakthrough in handling data truncation in fan-beam tomography was presented in [11] by introducing the idea of a virtual fan-beam re-binning technique. An alternate and emerging approach for handling data truncation is to reconstruct the Hilbert transform of the underlying function on pi lines [2, 3] by backprojecting the first derivative of the projection data. The Hilbert transform is then inverted on these lines to obtain the final reconstruction. These methods were largely inspired by a method introduced in helical CT by Zou et al in 2004 [12]. The idea of pi-lines has also been generalized for the purpose of cone-beam reconstruction from arbitrary source paths [4, 5].

3 PI Lines Theory

For completeness, since this is a relatively new concept, we now briefly describe the idea of cone-beam inversion on pi-lines [4]. For mathematical ease, we use the notion of a virtual detector centered at the origin \mathbf{O} (see Fig. 1). The orthonormal vectors $\hat{\mathbf{u}}, \hat{\mathbf{v}}$ and $\hat{\mathbf{w}}$ define the orientation of the detector width, detector length, and the position of the x-ray source $\mathbf{a}(\lambda)$. The cone-beam projection data is given by a scalar function $g(u, v, \lambda)$. The x-ray source assumes various positions along its trajectory that is parameterized by $\lambda \in A$, a real interval. Without loss of generality it is assumed that the detector is well oriented along these various poses, i.e., $\mathbf{a}'(\lambda)$ the tangent to the source trajectory is always parallel to $\hat{\mathbf{u}}$. A pi-line is defined as any line that connects two points of a connected source trajectory. Figure 1 shows a pi-line $\Pi_{\lambda^* \lambda}$ that contains $\mathbf{a}(\lambda)$ and $\mathbf{a}(\lambda^*)$. The only pi-lines that concerns us are those that intersect the virtual detector, since these are ones that are actually measured. Consider a point $\mathbf{x} \in \Pi_{\lambda^* \lambda}$ and let the underlying x-ray attenuation function at this point be denoted as $f(\mathbf{x})$. Since $\Pi_{\lambda^* \lambda}$ is a measured line, it is safe to assume that $f(\mathbf{x})$ is of finite support say

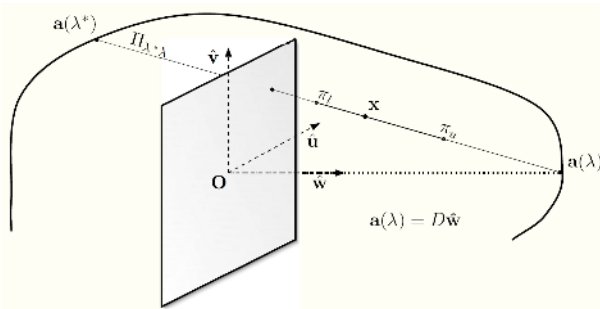


Fig. 1. Differentiated back projection in cone-beam geometry

(π_l, π_u) , on this line. Although truncation is admissible in general, it is strictly required that this section of the pi-line $\Pi_{\lambda^* \lambda}$ be captured completely in all the views from λ to λ^* . Under these assumptions it has been shown in [4] that the Hilbert transform of $f(\mathbf{x})$ along $\Pi_{\lambda^* \lambda}$ can be measured as follows:

$$\mathcal{H}_{\lambda^* \lambda} [f(\mathbf{x})] = \frac{1}{2\pi} \left[\int_{\lambda}^{\lambda^*} \frac{D \|\mathbf{a}'(\lambda)\| g_F(u^*(\lambda, \mathbf{x}), v^*(\lambda, \mathbf{x}), \lambda)}{[(\mathbf{a}(\lambda) - \mathbf{x}) \cdot \hat{\mathbf{w}}]^2} d\lambda + \frac{g(u^*(\lambda, \mathbf{x}), v^*(\lambda, \mathbf{x}), \lambda)}{\|\mathbf{a}(\lambda) - \mathbf{x}\|} \Big|_{\lambda}^{\lambda^*} \right], \tag{1}$$

where $(u^*(\lambda, \mathbf{x}), v^*(\lambda, \mathbf{x}))$ is the point where $\Pi_{\lambda^* \lambda}$ intersects the virtual detector, and

$$g_F(u, v, \lambda) = \frac{\partial}{\partial u} \left[\frac{D}{\sqrt{u^2 + v^2 + D^2}} g(u, v, \lambda) \right]. \tag{2}$$

The Hilbert transform obtained in (1) is directed from $\mathbf{a}(\lambda^*)$ to $\mathbf{a}(\lambda)$. Let t denote a scalar that specifies position on $\Pi_{\lambda^* \lambda}$ along this direction. A complete measure of $\mathcal{H}_{\lambda^* \lambda} [f(t)] \forall t \in (\pi_l, \pi_u)$ can then be obtained. Since the Hilbert transform is shift-invariant, any particular choice for $t = 0$ will do. Under the condition that $f(t)$ vanishes outside the interval (π_l, π_u) , the Hilbert transform can be inverted by,

$$f(t) = \frac{-1}{\sqrt{(t - \pi_l)(\pi_u - t)}} \int_{\pi_l}^{\pi_u} \frac{\sqrt{(t' - \pi_l)(\pi_u - t')} \mathcal{H} [f(t')]}{\pi(t - t')} dt' + C, \tag{3}$$

where C is a constant of integration that can be determined by the knowledge that $f(t) = 0$ outside this interval.

4 Methods

In this section we present the two main contributions of this paper. We first show how to obtain the Hilbert transform on a set of parallel pi-lines from a set of diverging fan-beam (or cone-beam in the mid-plane) projections. We then show how to integrate prior information along these pi-lines from a prior CT. Since the source geometry is planar the method is approximate above and below the mid-plane. For explanatory purposes we use the ‘‘Popeye’’ phantom (see Fig. 2a) defined in [3] to serve as the prior CT ($f_p(\mathbf{x})$). A modified Popeye phantom ($f_m(\mathbf{x})$) is then obtained by adding a long cylinder, and an ellipsoid, changes that are assumed to occur during surgery (Fig. 2b). The goal is to image a small ROI of this modified phantom using fully truncated cone-beam projections obtained over a semi-circular source path.

4.1 Selective Differentiated Back-Projection

As shown in Fig. 2b, the size of the detector is clearly too small to capture the modified Popeye phantom from any view angle ($g(u, v, \lambda)$ is only known $\forall u \in$

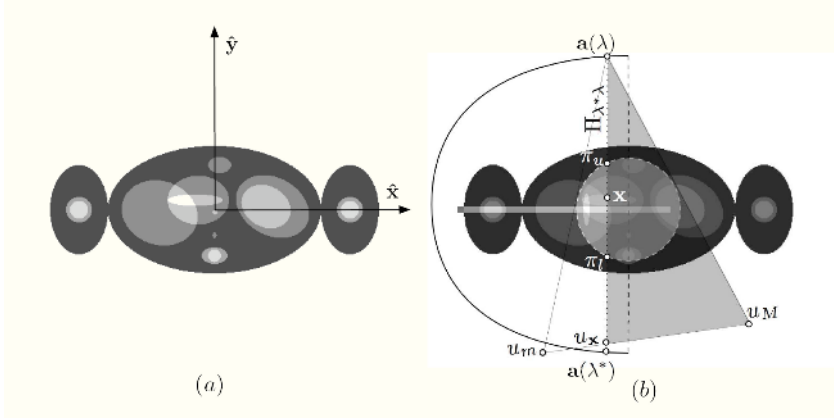


Fig. 2. (a) Popeye phantom serves as prior CT, (b) Selective backprojection to obtain Hilbert transform on parallel lines of modified Popeye phantom

$[u_m, u_M]$). However, the local nature of the derivative operation in (2) allows for accurate calculation of $g_F(u, v, \lambda) \forall u \in (u_m, u_M)$. Since the limits of the integral in (1) are dependent on the reconstruction point \mathbf{x} , a voxel-driven backprojection summation algorithm cannot be directly used. In order to get over this hurdle, we perform a selective differentiated backprojection of the data by defining a modified data function defined as follows:

$$\tilde{g}_F(u, v, \lambda) = g_F(u, v, \lambda) \quad \forall \quad u \begin{cases} u_{\mathbf{x}} < 0 \\ \geq u_{\mathbf{x}} \\ u_{\mathbf{x}} > 0 \end{cases} \quad (4)$$

where $u_{\mathbf{x}}$ corresponds to the column on the detector where the vertical pi-line Π_{λ^*} (see Fig. 2b) containing $\mathbf{a}(\lambda)$, \mathbf{x} , and $\mathbf{a}(\lambda^*)$ intersects it. We backproject this modified projection data over the entire semi-circle to obtain

$$\hat{\mathcal{H}}_y [f_m(\mathbf{x})] = \frac{1}{2\pi} \int_0^\pi \frac{D \|\mathbf{a}'(\lambda)\| \tilde{g}_F(u^*(\lambda, \mathbf{x}), v^*(\lambda, \mathbf{x}), \lambda)}{[(\mathbf{a}(\lambda) - \mathbf{x}) \cdot \hat{\mathbf{w}}]^2} d\lambda \quad (5)$$

Since the limits of the integration in (5) are not position dependent we can process the data using a typical cone-beam backprojection summation algorithm. The bounds in (4) ensure that for points lying on vertical pi-lines, backprojection is effectively computed only over the range defined by the intersection of such a pi-line with the source path. We have also ignored the second term of (1) arguing that its contribution to the summation is significantly less than that from the first term.

4.2 Integration of Prior CT

The selective backprojection of data on vertical pi-lines gives us a good estimate of the Hilbert transform of $f_m(\mathbf{x})$ along the positive y axis. With the knowledge of

$f_p(\mathbf{x})$ along the same pi-lines we can obtain an estimate of the Hilbert transform of the change $c(\mathbf{x}) = f_m(\mathbf{x}) - f_p(\mathbf{x})$ using the fact that the Hilbert transform is a linear transform, i.e.,

$$\hat{\mathcal{H}}_y [c(\mathbf{x})] = \hat{\mathcal{H}}_y [f_m(\mathbf{x})] - \mathcal{H}_y [f_p(\mathbf{x})], \quad (6)$$

where

$$\mathcal{H}_y [f_p(\mathbf{x})] = \mathcal{H}_y [f_p(x, y, z)] = \int_{-\infty}^{+\infty} \frac{f_p(x, y', z)}{\pi(y - y')} dy'. \quad (7)$$

An estimate of $c(\mathbf{x})$ can then be evaluated by using (3) to invert $\hat{\mathcal{H}}_y [c(\mathbf{x})]$ over the domain (π_l, π_u) . This inversion is done over the entire FOV, and then added back to the prior CT to obtain our final reconstruction result. This method works well when $c(\mathbf{x}) = 0$ outside the FOV covered by the fully truncated scan. In this case since the change is a long object oriented along the x axis, it makes sense to reconstruct along vertical pi-lines. In general it would be effective to perform the selective backprojection such that the pi-lines are oriented orthogonal to the longer dimension of the change.

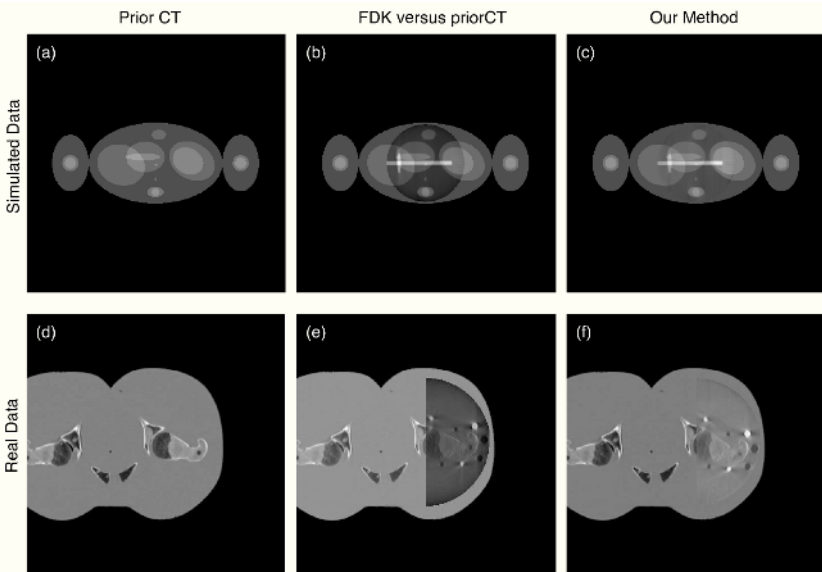
5 Results and Discussion

We conducted two sets of experiments to test the method described in Sect. 4. Using the Popeye phantom and its modified version we began with a simulated experiment. We also conducted a real experiment using an anthropomorphic hip section phantom known as the Alderson Rando Phantom. We first obtained a complete CT scan of the entire hip phantom to serve as the prior CT for our experiment. The phantom consists of several slabs (≈ 1 inch) that are stacked vertically. This allowed for easy modification of the interior section of the phantom. We drilled out 7 cylindrical plugs of varying diameters (3 to 7 mm) from the most central slab (also coinciding with the mid-plane of the sweep). We then replaced three of these with Aluminum plugs. Using a Siemens flat panel ceiling mounted C-arm (Artis dTA), we then obtained well calibrated cone-beam projections of this modified hip phantom over a semi-circular sweep. The poses of the C-arm were accurately represented using projection matrices obtained by scanning an offline calibration phantom. For the purpose of registration we used a standard Levinson-Marquardt optimization routine that searched through a 6 parameter space to find a rigid body transformation that would take the prior CT to the space of the C-arm projections. The cost function was the combined error in the prediction of where some known external markers would be projected to on the image domain. We used two orthogonal views for this purpose. The scan parameters for the two experiments are summarized in Table 1.

The results from the two experiments are shown in Fig. 3. In the case of the real data experiment, it was sufficient to apply our method on pi-lines that spanned only half of the FOV since there was no change outside this region. The results are shown on slices that correspond to the mid-plane of the source geometry. Figures 3a and 3d show instances of the prior CT in each case. For the

Table 1. Scan Parameters

	Simulated Data	Real Data
Angular Range	180°	180°
Angular Sampling	1°	0.4°
Source to Detector	1040 mm	1180 mm
Source to Iso-center	570 mm	785 mm
Number of Rows	256	1024
Number of Columns	256	756
Row Spacing	2.554 mm	0.373 mm
Col Spacing	1.4 mm	0.373 mm
Prior CT grid	(512,512,200)	(512,512,200)
Prior CT size	(640 mm, 640 mm, 250 mm)	(380 mm, 380 mm, 200 mm)

**Fig. 3.** Integration of prior CT information along pi-lines reduces truncation artifacts to a great extent

purpose of comparison, we show the results of applying the very short-scan FBP algorithm on fully truncated projections. We juxtapose these results with the prior CT outside the FOV in order to emphasize the mismatch in reconstructed CT values (see Figs. 3b and 3e). The discrepancy in the values is not due solely to a DC offset but also to an indeterminable cupping attenuation that is typical of the interior problem. This is mainly an effect of the non-local nature of the ramp-filter that is used in such methods. The characteristic strong edge on the boundary of the FOV has been avoided by artificially setting the boundary value of the ramp-filtered projections to zero. Using our method instead, we present the

results of adding the recovered change within the FOV to the prior CT. As can be seen in Figs. 3c and 3f, our method does a remarkably good job of consistently fusing the two reconstructions. The improvement in the reconstructed values of our method is a result of mathematically consistent compensation along pi-lines, and not due to an ad-hoc scaling of values. In the case of real data however, we had to determine a scalar scale in order to match the C-arm data to the prior CT data. This was automatically done by comparing the average values of $\mathcal{H}_y [f_p(\mathbf{x})]$ and $\mathcal{H}_y [f_m(\mathbf{x})]$ within the FOV.

References

1. L.A. Feldkamp, L.C. Davis, and J.W. Kress. Practical cone-beam algorithm. *Optical Society of America*, 1(6), June 1984.
2. Y. Zou, X. Pan, and E. Y. Sidky. Image reconstruction in regions-of-interest from truncated projections in a reduced fan-beam scan. *Physics in Medicine and Biology*, 50:13–27, DEC 2004.
3. F. Noo, R. Clackdoyle, and J. D. Pack. A two-step hilbert transform method for 2D image reconstruction. *Physics in Medicine and Biology*, 49:3903–3923, AUG 2004.
4. J. D. Pack, F. Noo, and R. Clackdoyle. Cone-beam reconstruction using the back-projection of locally filtered projections. *IEEE Transactions on Medical Imaging*, 24(1):70–85, JAN 2005.
5. T. Zhuang, S. Leng, B. E. Nett, and Chen G. Fan-beam and cone-beam image reconstruction via filtering the backprojection image of differentiated projection data. *Physics in Medicine and Biology*, 49:5489–5503, DEC 2004.
6. F. Natterer. *The mathematics of computerized tomography*. Wiley, New York, 1986.
7. H.K. Tuy. An inversion formula for cone-beam reconstruction. *SIAM Journal of Applied Mathematics*, 43:546–552, 1983.
8. Frederic Noo, Michel Defrise, Rolf Clackdoyle, and Hiroyuki Kudo. Image reconstruction from fan-beam projections on less than a short-scan. *Physics in Medicine and Biology*, 47:2525–2546, July 2002.
9. K. Ramamurthi and J. Prince. Tomographic reconstruction for truncated cone beam data using prior CT information. *Sixth International Conference: Medical Imaging Computing and Computer-Assisted Intervention (MICCAI)*, pages 134–141, 2003.
10. F. Noo and H. D.J. Image reconstruction from cone-beam data on a circular short-scan. *SPIE Medical Imaging*, February 2002.
11. R. Clackdoyle, F. Noo, J. Guo, and J. A. Roberts. Quantitative reconstruction from truncated projections in classical tomography. *IEEE Transactions on Nuclear Science*, 51(5):2570–2578, OCT 2004.
12. Y. Zou and X. Pan. Exact image reconstruction on PI-lines from minimum data in helical cone-beam CT. *Physics in Medicine and Biology*, 49:941–959, FEB 2004.

C-arm Calibration - Is it Really Necessary?*

Ameet Jain¹, Ryan Kon¹, Yu Zhou², and Gabor Fichtinger^{1,2}

¹ Department of Computer Science, Johns Hopkins University

² Department of Mechanical Engineering, Johns Hopkins University

Abstract. C-arm fluoroscopy is modelled as a perspective projection, the parameters of which are estimated through a calibration procedure. It has been universally accepted that precise intra-procedural calibration is a prerequisite for accurate quantitative C-arm fluoroscopy guidance. Calibration, however, significantly adds to system complexity, which is a major impediment to clinical practice. We challenge the status quo by questioning the assumption that precise intra-procedural calibration is really necessary. We derived theoretical bounds for the sensitivity of 3D measurements to mis-calibration. Experimental results corroborated the theory in that mis-calibration in the focal spot by as much as 50 *mm* still allows for tracking with an accuracy of 0.5 *mm* in translation and 0.65° in rotation, and such mis-calibration does not impose any additional error on the reconstruction of small objects.

1 Introduction

C-arm fluoroscopy is ubiquitous in general surgery, due to its real-time nature, versatility, and low cost. At the same time, quantitative fluoroscopy has not found a *large scale clinical acceptance*, because of inherent technical difficulties and needs to solve four major problems: (1) C-arm image distortion; (2) Calibration of model parameters; (3) Pose recovery or tracking when multiple images are taken; and (4) Registration to imaging modalities. Some of the prominent works that have tackled the above problems are [1, 2]. The driving application of our research is prostate brachytherapy, where radioactive seeds are required to be precisely placed into the prostate. Quantitative fluoroscopy could enable a significant improvement in the current clinical practice.

If it is known that both image distortion[3] and calibration[4] may vary significantly with pose. Image distortion usually has a consequential contribution to reconstruction error and needs to be compensated. Thus the additional cost of a full online calibration is not substantial. Recently developed advanced intensifier tubes allow for lesser distortion, while modern flat panel detectors obviate distortion correction altogether. This fact brings up the question whether we need to calibrate the C-arm fully at each pose. The question also leads to the broader issue, that even if it is not pose dependent, how accurate does calibration need to be. In spite of the importance of calibration in C-arm fluoroscopy, as far as the authors are aware, there has been no prior work that conducts this analysis. The

* This work has been supported by NIH 1R43CA099374-01 and NSF EEC-9731478.

vision community has a similar problem [5, 6] when cameras are used for visual serving of robots. We do not go into a detailed comparison for lack of space.

In quantitative C-arm fluoroscopy, we typically need to measure the spatial transformation between two objects, such as a vertebra and a bone drill, as compared to the transformation between an object and the C-arm itself. Thus the central intuition of this paper is that *while an incorrect calibration gives erroneous estimates for the absolute transformations, nevertheless it still provides acceptable relative estimates*. The consequence of this conjecture is potentially far reaching, as it can turn fluoroscopy to an affordable quantitative measurement tool in a large family of procedures. It should be however noted that we do not claim that calibration would always be unnecessary, since there are many applications that require high reconstruction accuracy. The decision should always be made case by case, experimentally. In this paper, we build a mathematical framework to formally address this issue and lend credit to the intuition that a loose estimate of the C-arm parameters might suffice in applications where the relative pose of objects is to be measured. In particular, we prove in theory and demonstrate experimentally that intra-operative dosimetry of brachytherapy implants is feasible with an un-calibrated C-arm.

2 Mathematical Framework

C-arm Imaging Model: Geometric aspects of fluoroscopic imaging can be approximated as a perspective projection with five parameters[7]. There are a total of five independent parameters that need to be evaluated by the calibration procedure - the pixel sizes (two) and the focal spot (three). The pixel sizes are fixed and remain unchanged throughout the life of the C-arm, reducing online-calibration to just the focal spot. Though our framework can study sensitivity due to any of the five parameters, we limit ourselves only to the focal spot.

2.1 Model for Reconstruction Space Transformation

As illustrated in Figure 1(a), let A & B (with reference frames F_A & F_B) be the two objects being imaged. The assumptions are: (i) ${}^I F_A, {}^I F_B$ can be computed from the images; (ii) A & B are *not large* in comparison to the focal length; (iii) F_A and F_B are *close by*; and (iv) the quantity of interest is ${}^A F_B = ({}^I F_A)^{-1} {}^I F_B$. Let \bar{f}_1 be the true focal spot and $\bar{f}_2 = (\bar{f}_1 + \bar{D})$ be the mis-calibrated estimate. We claim that even though the absolute locations of the objects are off, their relative transformation might still be accurate.

A transformation is needed that can take the absolute location of an object reconstructed with calibration \bar{f}_1 , and compute its corresponding location with calibration \bar{f}_2 . We claim that the simplest transformation will be a linear affine model \mathcal{T} . The intuition derives from the observation that the image plane is the same in both reconstruction spaces. Thus if P_1 (not in homogenous coordinates) projects to a point p on the image, then it is constrained to be on line \bar{L}_1 in the \bar{f}_1 -space and on \bar{L}_2 in \bar{f}_2 -space. Thus we seek a continuous invertible

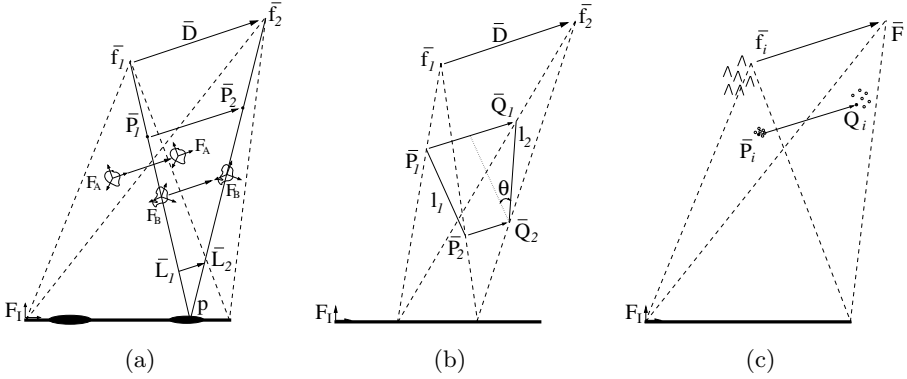


Fig. 1. Mis-calibration (a) shifts all reconstructed objects under an affine transformation; (b) rotates and scales a straight line segment; (c) Pose dependent calibration might be successfully approximated by using the mean value

transformation that projects \bar{L}_1 to \bar{L}_2 . By incorporating the above constraints, \mathcal{T} can be evaluated to be,

$$P_2 = \mathcal{T} \cdot P_1 = \begin{bmatrix} 1 & 0 & D_x/f_{1z} \\ 0 & 1 & D_y/f_{1z} \\ 0 & 0 & 1 + (D_z/f_{1z}) \end{bmatrix} \cdot P_1 = P_1 + (d \cdot Z/f_{1z})\hat{D} \quad (1)$$

where with respect to (wrt) F_I , $\bar{D} = (D_x, D_y, D_z)$; $d = \|\bar{D}\|_2$; $\hat{D} = \bar{D}/d$; $\bar{f}_1 = (f_{1x}, f_{1y}, f_{1z})$; and $P_1 = (X, Y, Z)$. Each point is effectively translated in direction \hat{D} by an amount proportional to its distance from the image. Experiments measuring the correctness of this affine model are available in Section 3. Thus to study sensitivity, it is sufficient to study the properties of \mathcal{T} .

2.2 Changes in Length and Scale

\mathcal{T} preserves the scale along the x, y -axes, but scales the space along the z -axis. Let $P_1(X_1, Y_1, Z_1)$ & $P_2(X_2, Y_2, Z_2)$ be any two points (not necessarily close to each other) in the \bar{f}_1 -space at a distance of l_1 . \mathcal{T} maps them to points Q_1 & Q_2 in the \bar{f}_2 -space at a distance of l_2 (Figure 1 (b)). It can be shown that

$$\|l_2 - l_1\| \leq \frac{d}{f_{1z}}|Z_1 - Z_2| \quad (2)$$

It directly follows from Equation (2) that \mathcal{T} does not alter the length significantly. As an example, a 10 mm calibration error would affect the length of a 30 mm thoracic pedicle screw at an angle of 45° by less than 0.2 mm (focal length ~ 1 m), which is significantly less than the error from other sources. Thus F_A, F_B will not change their relative translation by a factor more than that specified by Equation (2).

2.3 Changes in Absolute Orientation

A change in orientation results from the object having a depth (Figure 1 (c)). It can be shown geometrically that the orientation error is maximal when the vector $\overline{P_1P_2}$ is roughly orthogonal to \bar{D} and is purely in the vertical plane. The amount (θ) and the axis ($\hat{\kappa}$) of rotation, through a series of computations can be shown to be as in Equation (3). The bound on the rotation error is dependent only on origin mis-calibration and not on that in focal length. More importantly it is independent of the height/depth of the object (as far as it is non-planar) and its distance from the image plane. Thus F_A, F_B in Figure 1 will observe the same absolute rotation, in effect not experiencing any relative rotation. Experimental results corroborating this claim are available in Section 3.

$$|\theta| \leq \arcsin \left[\frac{\sqrt{D_x^2 + D_y^2}}{f_{1z}} \right] \sim \frac{\sqrt{D_x^2 + D_y^2}}{f_{1z}} ; \quad \hat{\kappa} = \frac{1}{\sqrt{D_x^2 + D_y^2}} (D_y, -D_x, 0) \quad (3)$$

2.4 Error in Reconstruction of Point Features

In many applications (particularly in brachytherapy), C-arms are used to reconstruct 3D point objects. This is done by obtaining multiple images at varying orientations and then using triangulation to obtain the desired intersection. In ideal circumstances, all the lines would intersect at a unique point. In practice however, calibration (and other) errors lead to non-intersecting lines. We will attempt to bound the error in this symbolic reconstruction of the point. Let point P be imaged from N different poses and reconstructed in a tracker frame F_A , which is stationary wrt P . Let the i^{th} pose have a focal spot error (in frame F_A) of \bar{D}_i . Without errors, each reconstructed line (l_i) would pass through P . It can be shown that due to the calibration error \bar{D}_i , the new line passes through a new point \bar{P}'_A and undergoes a rotation ϕ .

$$\bar{P}'_A \sim \bar{P}_A + [0 \ 0 \ \frac{(\bar{P}_A \cdot \bar{D}_i)}{f_{iz}}]' ; \quad \phi \sim \frac{(\hat{l}_i \cdot \bar{D}_i) \sin\theta_i}{f_{iz}} \quad (4)$$

where θ_i is the angle that l_i makes with the z-axis of F_A . The rotation is fairly small and can be ignored. Thus P_A is at a distance of $(\bar{P}_A \cdot \bar{D}_i) \sin\theta_i / f_{iz}$ from l_i . If Q is the symbolic intersection of all l_i 's, then it can be shown that Q is no further away than $(\frac{d_{max}}{f_z} \sin\theta_{max}) \|P_A\|$ away from any of the lines. Moreover, the reconstruction error (RE) can also be shown to be bounded by

$$RE = \|(\bar{Q} - \bar{P}_A)\| < \frac{\sqrt{2} d_{max}}{f_z} \|\bar{P}_A\| \quad (5)$$

where d_{max} is the maximum amount of mis-calibration and f_z is the minimum focal length. Thus a 10 mm focal length error causes an error less than 0.5 mm for a point at a distance of 35 mm. Note that this is the worst case error analysis and in practice the dot product in Equation (4) mutually cancels positive and negative errors, leading to extremely low reconstruction errors (Section 3).

2.5 The Optimal Choice for Calibration Parameters

Since the focal spot is pose dependant, and the results from Section 2.2, & 2.3 suggest robustness to mis-calibration, choosing a constant calibration for quantitative reconstruction might be viable. In the scenario that the focal spot might vary as much as 10 mm from one pose to another, “*what constant calibration should be chosen to minimize error?*”

Let us assume that we are imaging a point P from N different poses (Figure 1 (c)). Wrt frame F_I , let the i^{th} pose have the focal spot at $\bar{f}_i = (f_{ix}, f_{iy}, f_{iz})$ and the point be at location $P_i = (X_i, Y_i, Z_i)$. Note that we assume: (a) variations in each of $f_{ix}, f_{iy}, f_{iz}, X_i, Y_i, Z_i$ & pose are independent; (b) P_i 's are close to the iso-center, *i.e.* variations in X_i, Y_i, Z_i are not high. We choose a constant value of $\bar{F} = (F_x, F_y, F_z)$ for the focal spot, which will displace the point P_i to $Q_i = \mathcal{T}(\bar{f}_i, F) \cdot P_i$. The aim is to choose an \bar{F} which minimizes the net variation of $\Delta Q_i = Q_i - \mu_Q$. Through a series of computations, it can be shown that

$$\mu_Q = \mu_P + \frac{\mu_z}{\mu_{f_z}}(\bar{F} - \bar{\mu}_f) \tag{6}$$

$$\Delta Q_i = (P_i - \mu_P) + \left[\frac{\Delta Z_i}{\mu_{f_z}} - \frac{\mu_z \Delta f_{iz}}{\mu_{f_z}^2} \right] \bar{F} + \frac{\mu_z \Delta f_{iz}}{\mu_{f_z}^2} \bar{\mu}_f - \frac{\Delta Z_i}{\mu_{f_z}} \bar{\mu}_f - \frac{\mu_z}{\mu_{f_z}} \Delta \bar{f}_i \tag{7}$$

where $\mu_Q, \mu_P, \mu_z, \mu_{f_z}, \bar{\mu}_f$ are the mean values of $Q_j, P_j, Z_j, f_{jz}, \bar{f}_j$; $Z_j = \mu_z + \Delta Z_j$ and likewise for f_{jz}, \bar{f}_j , where $j = 1 \dots N$. In the above calculations, the second order terms either summed to 0 due to the independence of the variables or were too small in comparison. Our choice of \bar{F} should be the one that minimizes the *variance*(ΔQ) = $var(\Delta Q_x) + var(\Delta Q_y) + var(\Delta Q_z)$. It should be noted that F_z scales the whole space, *i.e.* a lower value will decrease the variance, implying that the choice of $F_z = 0$ forces $var(Q_z) = 0$ by forcing all Q_i 's to lie on a plane. Thus $var(Q_z)$ does not provide sufficient constraints for F_z . We will first obtain F_x, F_y by minimizing the variance along x, y -axes (since there is no scaling in these directions), and then will compute F_z . Notice that the first term in Equation (7) is due to the relative movement in P , while the rest is due to an error in the calibration. Since we are interested only in the variance due to mis-calibration, we will ignore the variations in P . Minimizing $var(\Delta Q)$ and enforcing independence of f_{ix}, f_{iy} & f_{iz} gives

$$\bar{F} = \bar{\mu}_f - \frac{\sum_1^N \Delta f_{iz} \Delta \bar{f}_j}{\sum_1^N \Delta f_{iz}^2} \mu_{f_z} = [\mu_{f_x}, \mu_{f_y}, 0]^T \tag{8}$$

As expected, $F_z = 0$ from above. To compute F_z , we need to impose length preserving constraints. Thus if we measure a line segment of length l in each image, use Equation (2) to derive the net length error, the minimization implies

$$F_z = \mu_{f_z} \left(1 - \frac{\sum_1^N \Delta f_{iz}^2}{N \mu_{f_z}^2} \right) \sim \mu_{f_z} \tag{9}$$

Thus $\bar{F} = \bar{\mu}_f$ (the mean), which is fairly intuitive and probably in common practice. Likewise, this particular choice of F_x, F_y is also a length preserving

constraint, *i.e.* it minimizes the error in lengths of line segments. Calibration error in ΔQ_i now reduces to $-\frac{\mu_z}{\mu_{fz}} \Delta \bar{f}_i$, which has a stable mean and low variance. Equation (10) gives a bound on the error when the assumed value of \bar{F} is away from the mean $\bar{\mu}_f$ by a distance d . A 10 mm variation in the focal length ($var \sim 3$ mm), P 's roughly at the iso-center having a depth variation of 100 mm and the assumed calibration unusually away from $\bar{\mu}_f$ by 50 mm still bounds the maximum error by 0.75 mm. Thus large and constant mis-calibration in many applications, might still provide sub-millimetric 3D quantitative measurements.

$$error \leq \frac{\sqrt{d^2 var(Z) + \mu_z^2 var(\|f\|)}}{\mu_{fz}} \quad (10)$$

3 Phantom Experiments and Results

Validity of the Model: Equations (1) & (3) give the translation and rotation transformations as predicted by the affine model, the accuracy of which would furnish the validity of the model. We used the FTRAC fiducial (Figure 3), a small image-based fluoroscope tracking fiducial, which (given the calibration) can track a C-arm with an accuracy of 0.5 mm in translation and 0.65° in rotation [7]. The fiducial was imaged using a Philips Integris V3000 fluoroscope and the true calibration read off the machine display. The images were not corrected for distortion. The pose of the fiducial (wrt to F_I) was first evaluated using the correct calibration, and then with the mis-calibrated parameters. The difference between the pose change predicted by the equations and the one computed using the non-linear pose estimation software, is displayed in Figure 2 (a) as a function of maximum calibration error. Even when mis-calibration is as high as 50 mm, the model can predict the rotation-axis with an accuracy of 4° , amount of rotation under 1° and translation under 1.5 mm. For extreme mis-calibrations the translation error linearly increases, while rotation is still stable. Thus the model seems to predict with an acceptable accuracy.

Accuracy of C-arm Tracking: The FTRAC fiducial was mounted on a 0.02° accuracy rotational turntable, while the fluoroscope was kept stationary. The turntable was rotated by known precise amounts (ground truth) and images were taken. The relative poses were also computed using the pose estimation software. The accuracy in the estimation of C-arm motion is given by the difference between the computed relative pose and the true relative pose. The tracking accuracy is plotted in Figure 2 (b) as a function of mis-calibration. Even a high mis-calibration of 150 mm adds no *additional* error in C-arm motion estimation, fixing the value at 0.45 mm in translation and 0.6° in rotation. An unusually high mis-calibration of 400 mm also only marginally decreases accuracy. Thus, mis-calibration does not increase the error of C-arm tracking .

3D Quantitative Reconstruction using Multiple Images: In addition to tracking a C-arm, it is equally important that multiple objects in the field of view (eg. vertebrae and screws) be reconstructed accurately relative to each

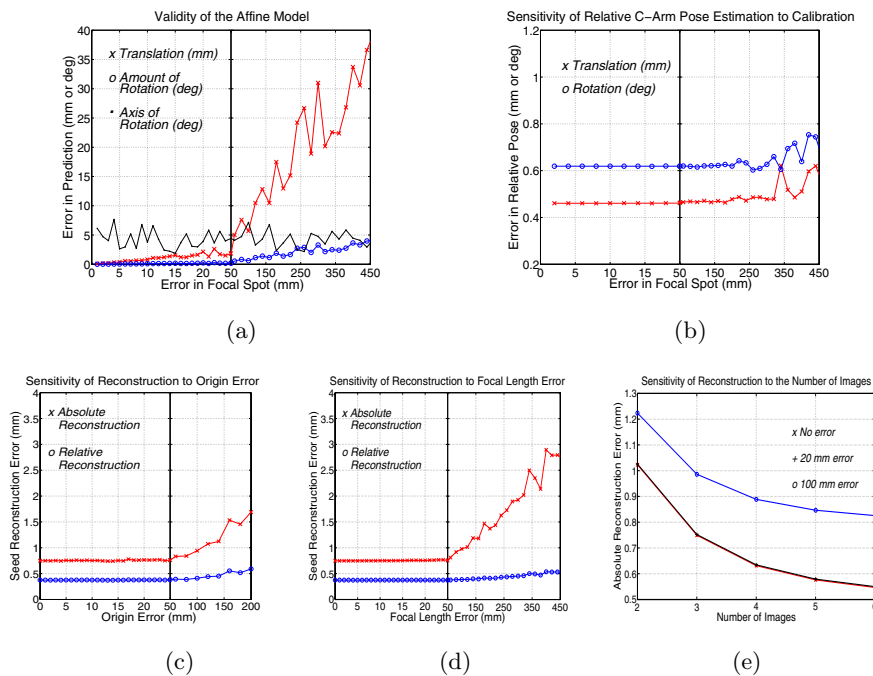


Fig. 2. Note the scale variation in x-axis. (a) An affine transformation is able to predict the movement of 3D objects due to mis-calibration; (b) C-arm tracking is insensitive to mis-calibration; 3D Reconstruction is insensitive to mis-calibration in (c) origin; (d) focal length up to 50 mm, beyond which it starts to linearly drift away from the tracking fiducial. Notice that the shape of the implant (relative err) is barely altered; (e) 3D reconstruction error decreases with an increase in images used.

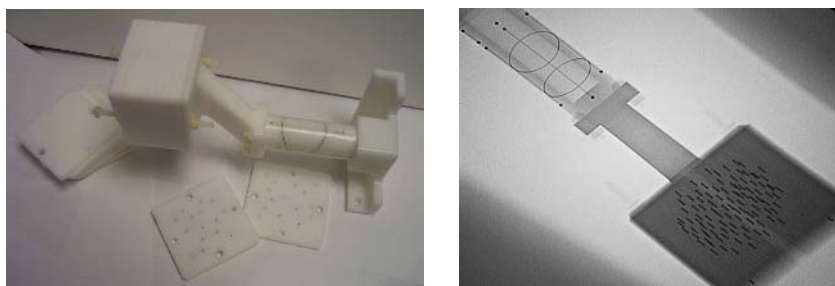


Fig. 3. An image of the seed phantom attached to the FTRAC fiducial (left). The seed phantom can replicate any implant configuration, using the twelve 5 mm slabs each with over a hundred holes. A typical X-ray image of the combination (right).

other. In order to validate our hypothesis that 3D reconstruction might not be sensitive to mis-calibration, we use an accurate acetol phantom (Figure 3) having 100 dummy radioactive seeds, approximating a brachytherapy implant

(Figure 3). The true 3D coordinate of each seed wrt the fiducial is known by rigid attachment. The C-arm is tracked using the FTRAC fiducial and the 3D seed coordinates are computed by triangulation (an algorithm called MARSHAL is used to establish correspondences). The difference between the computed and the true seed location gives us the 3D reconstruction error for each seed (wrt fiducial). The relative reconstruction error removes any consistent shift reflecting any change in shape. These errors are plotted as a function of mis-calibration in Figure 2 (c), (d). The reconstruction error is insensitive to mis-calibration in origin and focal length errors of up to 50 *mm*. The shape of the implant is stable even for large calibration errors. Figure 2 (e) shows a drop in reconstruction error as the number of images increase. Thus mis-calibration does not decrease reconstruction accuracy.

4 Conclusion

We modelled the the effects of mis-calibration on 3D reconstruction as an affine transform, and proved its validity experimentally. We have derived bounds on the amount of scaling, translation and rotation error. For pose dependant calibration, we proved that using the mean calibration minimizes the reconstruction variance. Phantom experiments with a radiographic fiducial indicate that C-arm tracking is insensitive to mis-calibrations. We also showed that mis-calibration up to 50 *mm* adds no additional error in 3D reconstruction of small objects, beyond which the reconstructed objects begin to drift wrt the fiducial, while still retaining the shape. In conclusion, a significant family of quantitative fluoroscopy applications involving localization of small markers can function without cumbersome on-line calibration. A constant loose calibration might suffice.

References

1. Hofstetter, R., Slomczykowski, M., Sati, M., Nolte, L.: Fluoroscopy as an imaging means for computer-assisted surgical navigation. *CAS* **4(2)** (1999) 65–76
2. Yao, J., Taylor, R.H., et al: A c-arm fluoroscopy-guided progressive cut refinement strategy using a surgical robot. *Comput Aided Surg* **5(6)** (2000) 373–90
3. Fahrig, R., et al: Three-dimensional computed tomographic reconstruction using a c-arm mounted XRII: correction of image distortion. *Med Phys.* (**24(7)**) 1097–106
4. Livyatan, H., Yaniv, Z., Joskowicz, L.: Robust automatic c-arm calibration for fluoroscopy-based navigation: A practical approach. In: *MICCAI*. (2002) 60–68
5. Kumar, R., Hanson, A.: Sensitivity of the pose refinement problem to accurate estimation of camera parameters. In: *ICCV90*. (1990) 365–369
6. Malis, E.: Visual servoing invariant to changes in camera intrinsic parameters. *IEEE Transaction on Robotics and Automation* **20** (2004) 72–81
7. Jain, et al: A robust fluoroscope tracking (FTRAC) fiducial. In: *SPIE Medical Imaging; Visualization, Image-Guided Procedures & Display*. (2005) 798–809

Laser Needle Guide for the Sonic Flashlight

David Wang^{1,3, 4}, Bing Wu^{2,3}, and George Stetten^{1,3,4,5}

¹ Department of Biomedical Engineering, Carnegie Mellon University,
Pittsburgh PA 15213, USA

² Department of Psychology, Carnegie Mellon University, Pittsburgh PA 15213, USA

³ Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA

⁴ University of Pittsburgh Medical Center, Pittsburgh PA 15261, USA

⁵ Department of Bioengineering, University of Pittsburgh,
Pittsburgh, PA 15261, USA

Abstract. We have extended the real-time tomographic reflection display of the Sonic Flashlight to a laser guidance system that aims to improve safety and accuracy of needle insertion, especially for deep procedures. This guidance system is fundamentally different from others currently available. Two low-intensity lasers are mounted on opposite sides of a needle aimed parallel to the needle. The needle is placed against a notch in the Sonic Flashlight mirror such that the laser beams reflect off the mirror to create bright red spots on the flat panel display. Due to diffuse reflection from these spots, the virtual image created by the flat panel display contains the spots, identifying the projected destination of the needle at its actual location in the tissue. We have implemented our design and validated its performance, identifying several areas for potential improvement.

1 Introduction

In interventional radiology, it is often necessary to insert a needle outside the plane of an ultrasound scan. Because, in such cases, the tip of the needle is not visible until it reaches the ultrasound scan, a potential exists for the needle to miss the target, requiring multiple needle insertions and unnecessary trauma to the patient. Hence a method to accurately guide the needle to the target in just one attempt would be valuable.

Typically, needle guides are attached to the ultrasound probe and restrict the needle to travel along a specific path within the ultrasound plane. They have been routinely used to perform needle biopsies of various organs, including the liver, kidney, prostate, and breast [1], [2], [3], [4]. The needle pathway is indicated on the monitor by means of guide lines superimposed on the ultrasound image. While steerable in-plane needle guides are currently being developed [5], the needle is still restricted to travel in the scanning plane.

Needle guides have also been developed to operate out of the plane of the ultrasound scan. Two commercial systems are currently available. The PunctSURE™ vascular access imaging system (Inceptio Medical Technologies, L.C., Kaysville, Utah) is a variation on traditional ultrasound systems, presenting real-time cross-sectional and longitudinal B-mode scans simultaneously on the display side by side.

With the vein centered in the cross-sectional scan, the longitudinal ultrasound array is properly oriented parallel to the vein. The needle, when inserted in the plane of the longitudinal scan, can be visualized in its entirety, and no needle guide device is needed with the system.

The second system is the Site-Rite™, (CR Bard, Murray Hill, New Jersey) in which an out-of-plane needle guide attaches to the ultrasound probe, restricting the needle to a pathway that intersects the ultrasound scanning plane at specific depths, ranging from 0.5 cm to 3.5 cm, in 1 cm steps. The choice of depth depends on which of 4 disposable needle guides is attached to the probe. After guiding the needle into the vein, the guide can be separated from the needle, facilitating insertion of a catheter.

Both guidance systems suffer from a lack of perceptual coupling between the act of needle insertion and visual feedback from the ultrasound image, with the display located separately from the transducer. With the PunctSURE, the user can follow the needle trajectory, but must look away from the site of operation in order to do so. Such displaced hand-eye coordination causes attentional shifts and may introduce errors and variability in the operator's performance. The mental imagery involved in locating the target is a demanding cognitive process, subject to error. The Site-Rite permits prediction of the needle trajectory, but restricts the insertion to a fixed number of pre-determined angles, depriving the operator of the ability to perform insertions along an arbitrary path.

Another experimental needle-guidance system uses a 3D ultrasound machine with computer analysis to locate the needle [6]. Others required placing radio-opaque markers on the skin, CT-scanning, 3D segmentation, stereo-cameras, and tracking devices to localize the needle [7]. These systems are cumbersome and require extensive calibration and sophisticated software tools.

We propose to solve these problems by using a new ultrasound guidance system called the Sonic Flashlight (SF), and adapting the needle with two small lasers. As described below, the operator will use the lasers to illuminate targets in a virtual ultrasound image projected directly within the patient by the SF, and thereby "home in" on the target following a straight line. The procedure will occur without looking away from the patient or physically restricting the angle of insertion.

We start by reviewing the concept of Real-Time Tomographic Reflection behind the Sonic Flashlight in Section 2. We then explain our proposed method in Section 3 and present experimental results in Section 4. We conclude with a proposal for future work.

2 Real-Time Tomographic Reflection

Real-Time Tomographic Reflection (RTTR) was separately proposed by Stetten et al. [8], [9], [10], [11] and Masamune et al. [12]. Stetten's RTTR system was developed for real-time visualization of ultrasound. It functions by fixing the relative geometry of the ultrasound transducer, the display, and a half-silvered mirror to produce a virtual image of a tomographic slice (a sector-scan B-mode ultrasound image) at the actual scanned plane within the body (see Fig. 1). Through the half-silvered mirror, the ultrasound image is projected as if it "shines out" from the probe and illuminates the inner tissue, which is no longer occluded by the proximal surface of the skin. For that reason, this implementation of RTTR was named the Sonic Flashlight. Using the Sonic Flashlight, there is no need to track the observer, the patient, or the transducer,

due to the direct registration between the virtual image and the ultrasound beam. Moreover, the patient, the ultrasound image, the surgical instrument, and the operator's hands are merged into one perceptual environment for all observers looking through the half-silvered mirror, facilitating cooperation or training.

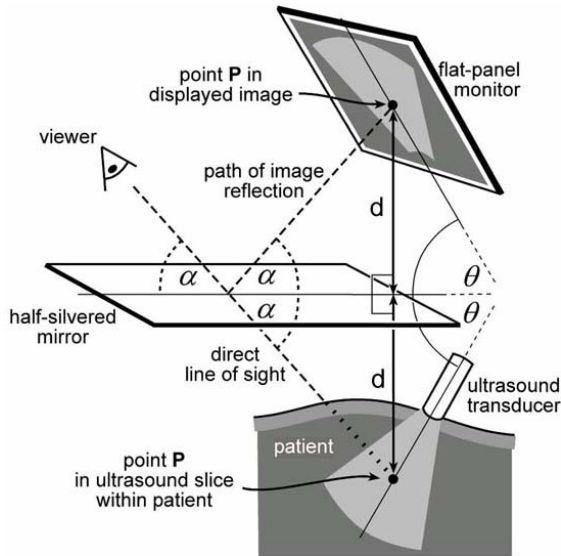


Fig. 1. Optics of Real-Time Tomographic Reflection (RTTR) in the Sonic Flashlight

3 Method

We have applied the concept of RTTR to produce a new method of needle guidance, capitalizing on the fact that the optical geometry of RTTR works as well for light hitting the display as coming from it. If a low-intensity laser is aimed at a target in the virtual image, it can be used to define a straight path for needle insertion. As shown in Figure 2, if the laser beam hits the mirror, it will both reflect and pass through. The part that passes through the mirror will create a light spot on the skin, which shows the point the needle should enter the body. The part of the laser beam that reflects off the mirror will create another light spot on the flat panel monitor at exactly that point in the image displaying the target. As the image on the flat panel monitor reflects off the half-silvered mirror to create a virtual image, a diffuse reflection of the laser spot also shows up in that virtual image, at the actual location within the patient where the needle will intersect the ultrasound scan.

In order to have the laser beams strike as closely as possible to the needle destination, we place the laser generators parallel to, and as close as possible to, the needle. As shown in Figure 3 two lasers are used so that the mid-point of the two spots in the virtual image flank the destination of the needle. Since the lasers must reflect off the half-silvered mirror, the needle is positioned in a small notch cut into the edge of mirror.

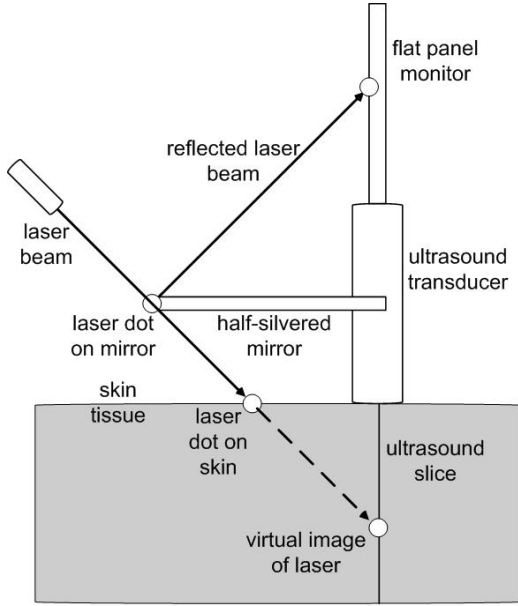


Fig. 2. Laser Needle Guide optics overview (see text)

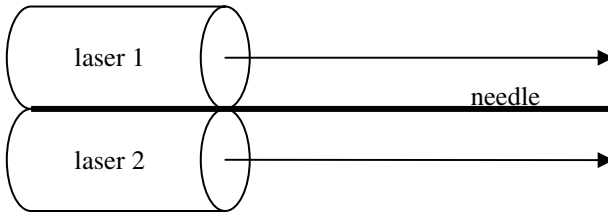


Fig. 3. Positional relationships between the two lasers and the needle

4 Experimental Results

We have tested two implementations of the laser needle guide described in Section 3. In the first implementation, we used an older Sonic Flashlight prototype (Model 4), and used it to guide the insertion of a needle into a water tank to hit a small spherical target mounted in the tank. The water surface in the tank was covered with loose screening, which permitted penetration by the ultrasound beam and the needle. As shown in Figure 4, each of the lasers generates four pairs of bright spots, labeled 1-4 (the spots are red if you are reading a color version of this paper). The needle is difficult to see in the darkness, but was inserted into the tank to hit the target, while maintaining contact with a notch in the edge of the mirror. Following the paths of the laser beams, the four pairs of spots are as follows: The laser beams traverse the lower half of Figure 4 from left to right, striking the mirror (4) and splitting into two beams. The upper (reflected)

beams reach the flat panel monitor (1) while the lower beams penetrate the mirror to product bright spots on the surface of the water covered by white screening (3). The virtual image (2) of the spots on the flat panel monitor (1) accurately flank the target at its actual location in the water tank. The photograph does not convey the strong perceptual depth of these spots felt by the observer. By keeping the lasers aimed on either side of the target, we successfully and easily reached the target with the needle.

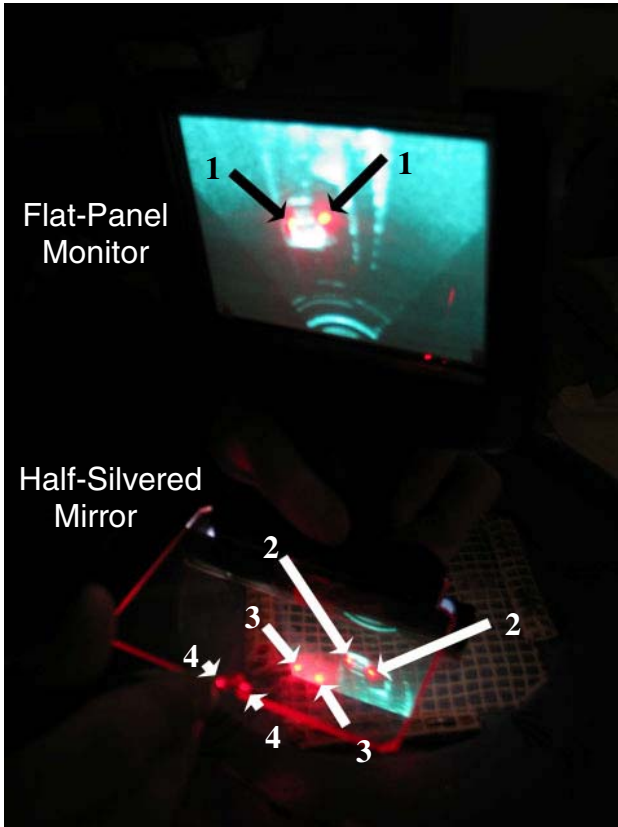


Fig. 4. Using a two-laser needle guide with the Model 4 Sonic Flashlight (see text)

In the second implementation of the laser needle guide, we used a more recent prototype of the Sonic Flashlight (Model 6), which is more compact and produces higher quality ultrasound images than the Model 4. In this case, a gel phantom containing a simulated vein was used as a target. Although a clear photograph of this apparatus in operation proved difficult to obtain, it was nonetheless easy to use. When we penetrated the phantom and pushed the needle toward the target, the needle bent slightly, changing its course. Since the lasers do not bend, the needle became misaligned with the laser. This problem can be easily solved by using a stiffer needle or an unbendable biopsy gun.

4 Conclusions and Future Work

We have demonstrated the feasibility of the two-laser needle guide with the Sonic Flashlight. Whereas the Sonic Flashlight with unaltered needles has shown good accuracy for relatively shallow targets such as veins in the arm, the addition of laser guidance may be appropriate for deeper procedures such as biopsies of the liver or kidney. The longer needles required for deeper procedures would lend themselves nicely to the apparatus, given the requirement that they maintain contact with the edge of the mirror. We are planning more extensive testing by examining human performance, in a special facility equipped with optical tracking.

Several areas for potential improvement could be addressed. First, we would like to eliminate the laser spot visible on the half-silvered mirror itself, due to scattering within it and on the mirror's surface. This spot is of no particular use and may be a potential distraction. The solution is to keep the mirror surface clean so as not to scatter light, and perhaps to find a different type of mirror that minimizes internal scattering. Another area for improvement is the specular surface of the flat panel monitor. Ideally we want the monitor to scatter the laser beam to create a red spot instead of reflecting (or absorbing) it. This could be accomplished by adding an antiglare diffusive surface. Finally, we are considering various schemes for reducing the number of lasers from two to one.

Since lasers are involved in this system, a safety analysis of potential damage to the retina is warranted. A number of paths for the laser light are created by reflections. Although the light attenuates at each of these, potential danger still exists. Our present apparatus uses two Class 3B lasers, which should not be viewed directly or in a specular reflection, but normally will not produce a hazardous diffuse reflection. Therefore the safe operation of the apparatus will depend on eliminating direct paths or specular reflections to the eye of the operator or patient.

References

1. Reid MH. Real-time sonographic needle biopsy guide. *Am J Roentgenol* 1983;140(1):162-3.
2. Yeh HC, Mitty HA, Wolf BS. A simple ultrasound guide for needle puncture. *J Clin Ultrasound* 1976;4(1):53-4.
3. Lindgren PG, Hemmingsson A. Percutaneous nephropylotomy. A new technique. *Acta Radiol Diagn (Stockh)* 1980;21(6):759-61.
4. Lindgren PG. Ultrasonically guided punctures. A modified technique. *Radiology* 1980;137(1 Pt 1):235-7.
5. Han D, Lan Seo Y, Soon Choi C, et al. A steerable guiding device: the new method in ultrasound guidance. *Invest Radiol* 2002;37(11):626-31.
6. O. Chavanon, C. Barbe, J. Troccaz, L. Carrat, and D. Ribuot, C. aand Blin. Accurate guidance for percutaneous access to a specific target in soft tissues. *Journal of Laparoendoscopic and advanced surgical techniques*, 9(3):259-266, 1999.
7. S. Nicolau, A. Garcia, X. Pennec, L. Soler, N. Ayache An Augmented Reality system to guide radio-frequency tumor ablation *Journal of Computer Animation and Virtual World*, 2005, Vol. 16(1), p1-10.
8. Stetten G, "System and Method for Location-Merging of Real-Time Tomographic Slice Images with Human Vision," U.S. Patent no. 6,599,247, issue date, July 29, 2003.

9. Stetten GD, Chib V. Overlaying Ultrasound Images on Direct Vision. *J Ultrasound Med* 2001;20(3):235-40.
10. Stetten G., Chib V., & Tamburo R., "System for Location-Merging Ultrasound Images with Human Vision," *IEEE Proceedings of the Applied Imagery Pattern Recognition (AIPR) Workshop*, Washington D.C., pp. 200-205, 2000.
11. Chang WM, Stetten GD, Lobes Jr. LL, Shelton DM, Tamburo RJ. Guidance of retrobulbar injection with real time tomographic reflection. *J Ultrasound Med* 2002;21(10):1131-5.
12. Masamune K, Fichtinger G, Deguet A, Matsuka D, & Taylor RH: "An Image Overlay System with Enhanced Reality for Percutaneous Therapy Performed Inside CT Scanner", *Fifth International Conference on Medical Image Computing and Computer-Assisted Intervention*, Lecture Notes in Computer Science, 2488, Part 2, pp. 77-84, Springer Verlag, 2002.

Differential Fly-Throughs (DFT): A General Framework for Computing Flight Paths

M. Sabry Hassouna¹, Aly A. Farag¹, and Robert Falk²

¹ Computer Vision & Image Processing Laboratory (CVIP),
University of Louisville, Louisville, Kentucky, 40292
`{msabry, farag}@cvip.uofl.edu`

² Director of medical imaging, Jewish Hospital, Louisville, Kentucky
`robert.falk@jhhs.org`

Abstract. In this paper, we propose a new variational framework based on distance transform and gradient vector flow, to compute flight paths through tubular and non-tubular structures, for virtual endoscopy. The proposed framework propagates two wave fronts of different speeds from a point source voxel, which belongs to the medial curves of the anatomical structure. The first wave traverses the 3D structure with a moderate speed that is a function of the distance field to extract its topology, while the second wave propagates with a higher speed that is a function of the magnitude of the gradient vector flow to extract the flight paths. The motion of the fronts are governed by a nonlinear partial equation, whose solution is computed efficiently using the higher accuracy fast marching level set method (HAFMM). The framework is robust, fully automatic, and computes flight paths that are centered, connected, thin, and less sensitive to boundary noise. We have validated the robustness of the proposed method both quantitatively and qualitatively against synthetic and clinical datasets.

1 Introduction

Virtual endoscopy (VE) is a computer-based alternative to true fiber optic endoscopy (TE) for screening hollow organs. VE is not intended to replace TE, but rather to complement it by providing additional supportive information. For example VE: (1) allows the visualization of neighboring structures outside the screened organ, and hence can assist in the pathology localization, (2) allows viewing in forward and reverse directions, (3) visualize areas that are hard to reach by TE, (4) has the ability to pass high grade stenoses, and finally (5) VE is the only alternative offered to those patients that either refuse TE or are severely ill [1, 2].

The extraction of 3D flight paths or discrete curve skeletons (\mathcal{CS}) of anatomical structures [3–10] is an important component of any VE system. We have recently developed a level set based-framework for computing \mathcal{CS} of 3D tubular and articulated objects [11], which addresses several shortcomings of existing techniques. The key idea is to propagate from a medial voxel wave fronts of different speeds. The first front propagates with a moderate speed to capture the

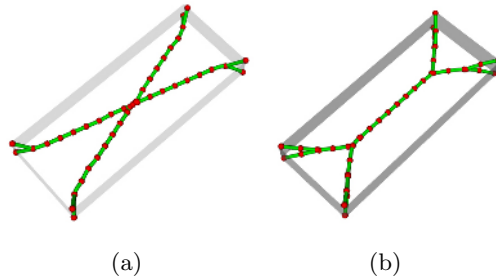


Fig. 1. Computed \mathcal{CS} for non-tubular structure by (a) [12]. (b) Proposed framework.

object topology, while the second one propagates much faster at medial voxels such that \mathcal{CS} intersect the propagating fronts at those voxels of maximum positive curvatures. The framework is slightly modified and quantitatively validated to show that the computed \mathcal{CS} can be used as reliable flight paths through tubular structures for VE [12]. In [12], if the structure's cross section deviates too much from the circle, for example, in case of a severe stenoses, where the cross section is nearly rectangular, the computed paths are the shortest as shown in Figure 1(a) rather than the centered as shown in Figure 1(b), which is generated by the proposed framework of this paper.

In this paper, we extend our recent framework [12] to compute flight paths for both tubular and non-tubular structures by utilizing the magnitude of the gradient vector flow, because it does not form medial surfaces for 3D objects.

2 Methods

2.1 Gradient Vector Flow

The gradient vector flow (GVF) is a bi-directional external force field that moves active contours in highly concave regions [13]. The GVF is the vector field $V(\mathbf{x}) = [u(\mathbf{x}) \ v(\mathbf{x}) \ w(\mathbf{x})]^T$, which minimizes the following energy function,

$$E(V) = \int \int \int \mu |\nabla V(\mathbf{x})|^2 + |\nabla f(\mathbf{x})|^2 |V(\mathbf{x}) - \nabla f(\mathbf{x})|^2 d\mathbf{x} \quad (1)$$

Where $\mathbf{x} = (x, y, z)$, μ is a regularization parameter, and $f(\mathbf{x})$ is an edge map derived from the imaging volume $I(\mathbf{x})$. For a binary volume, $f(\mathbf{x}) = -I(\mathbf{x})$. The interpretation of Eq. (1) is that if $|\nabla f(\mathbf{x})|$ is small, $E(V)$ is dominated by the sum of squares of the partial derivatives of the vector field, yielding a slowly varying field. On the other hand, if $|\nabla f(\mathbf{x})|$ is large, $E(V)$ is dominated by the second term, and is minimized by setting $V = \nabla f(\mathbf{x})$. This produces a vector field $V(\mathbf{x})$ that is nearly equal to the gradient of the edge map $\nabla f(\mathbf{x})$ when it is large and slowly varying in homogeneous regions. $V(\mathbf{x})$ can be computed iteratively by solving the following decoupled partial differential equations in u , v , and w [13].

$$u(\mathbf{x}, t + 1) = u(\mathbf{x}, t) + \Delta t(\mu \nabla^2 u(\mathbf{x}, t) - (u(\mathbf{x}) - f_x(\mathbf{x}))|\nabla f(\mathbf{x})|^2) \tag{2}$$

$$v(\mathbf{x}, t + 1) = v(\mathbf{x}, t) + \Delta t(\mu \nabla^2 v(\mathbf{x}, t) - (v(\mathbf{x}) - f_y(\mathbf{x}))|\nabla f(\mathbf{x})|^2) \tag{3}$$

$$w(\mathbf{x}, t + 1) = w(\mathbf{x}, t) + \Delta t(\mu \nabla^2 w(\mathbf{x}, t) - (w(\mathbf{x}) - f_z(\mathbf{x}))|\nabla f(\mathbf{x})|^2) \tag{4}$$

$$V(x, 0) = \nabla f(\mathbf{x}) \tag{5}$$

The iterative process is guaranteed to converge if

$$\Delta t \leq \frac{\Delta x \Delta y \Delta z}{6\mu} \tag{6}$$

where Δx , Δy , and Δz are the data spacing of a given dataset. In [14], Xu has shown that the GVF possesses as well some medial properties by suggesting two different medialness functions based on the magnitude of the GVF, to measure only how close a point from the skeleton of a 2D shape, rather than computing the skeleton itself. One of the interesting properties of the GVF, $V(\mathbf{x})$ over the distance field $D(\mathbf{x})$ is that it does not form medial surfaces for non-tubular 3D objects [15] because only one boundary voxel contributes to the computation of $D(\mathbf{x})$, while more than one boundary voxels contribute to the computation of $V(\mathbf{x})$ during the diffusion process.

2.2 Single Flight Path Extraction

Consider the *minimum-cost path problem* that finds the path $C(s) : [0, \infty) \rightarrow R^n$ that minimizes the cumulative travel cost from a starting point A to some destination B in R^n . If the cost U is only a function of the location \mathbf{x} in the image domain, the cost function is called isotropic, and the minimum cumulative cost at \mathbf{x} is defined as

$$T(\mathbf{x}) = \min \int_A^B U(C(s)) ds \tag{7}$$

The path that gives the minimum integral is the minimum cost path. The solution of Eq. (7) is a nonlinear partial differential equation known as the *Eikonal equation*[16] Eq. (8), where $F(\mathbf{x}) = 1/U(\mathbf{x})$, and $T(\mathbf{x})$ is the time at which the front crosses \mathbf{x} .

$$|\nabla T(\mathbf{x})|F(\mathbf{x}) = 1.0 \tag{8}$$

In this paper, we solve the *Eikonal* equation using the higher accuracy fast marching method (HAFMM) [17]. Let A and B be medial voxels. Assume that A is a point source P_S that transmits a high speed front Eq. (9), where $\lambda(\mathbf{x})$ is a medial descriptor function that distinguishes medial voxels from others and α controls the curvature of the front at medial voxels. In [12], we showed that the minimum cost path between A and B is the medial curve or centerline connecting them. Since the GVF does not form medial surfaces in 3D, we propose the following medial descriptor function Eq. (10), where the magnitude of GVF goes to zero at medial voxels.

$$F(\mathbf{x}) = e^{\alpha \lambda(\mathbf{x})} \quad \alpha \geq 0 \tag{9}$$

$$\lambda(\mathbf{x}) = \frac{2.0}{1.0 + \|V_n(\mathbf{x})\|^{0.05}} - 1, \quad \|V_n(\mathbf{x})\| = \frac{\|V(\mathbf{x})\| - \|V(\mathbf{x})\|_{min}}{\|V(\mathbf{x})\|_{max} - \|V(\mathbf{x})\|_{min}} \quad (10)$$

The propagating front is monotonically increasing in time; there is only one global minimum over the cumulative cost field T , that is P_S , which has zero travel time. Then, the path between B and A can be found by backtracking from B along the gradient of T until A is reached. The extraction process is the solution of the ordinary differential equation Eq. (11). $C(t)$ traces out the \mathcal{CS} , which is found by solving Eq.(11) using Runge-Kutta of order 2. The error of the method is $O(h^3)$, where h is the integration step. h is set to 1.0.

$$\frac{dC}{dt} = -\frac{\nabla T(\mathbf{x})}{\|\nabla T(\mathbf{x})\|}, \quad C(0) = B \quad (11)$$

For $C_i = [x_i, y_i, z_i]^T$,

$$f(C_i) = -\frac{\nabla T(C_i)}{\|\nabla T(C_i)\|}, \quad k_1 = hf(C_i), \quad C_{i+1} = C_i + hf\left(C_i + \frac{k_1}{2}\right) \quad (12)$$

2.3 Multiple Flight Path Extraction

In order to extract the entire \mathcal{CS} of an anatomical structure, we have to identify the starting voxel (node) of each \mathcal{CS} as well as the merging voxel, if the structure contains a loop. To achieve the goal, we follow the cluster graph (CG) approach [11], which can be summarized as follows: Initially, we compute the normalized distance field $D(\mathbf{x})$ using the HAFMM. Then, we propagate a moderate speed wave from P_S , which results in a new distance field $D_1(\mathbf{x})$. The speed of the front is given by Eq. (13).

$$F(\mathbf{x}) = e^{\beta D(\mathbf{x})} \quad (13)$$

$D_1(\mathbf{x})$ is discretized by computing its integer values. The Discretization converts the structure into a cluster graph, whose root is the cluster containing P_S . Each cluster consists of a set of neighbor voxels with the same code. The CG contains two main types of clusters; Extreme cluster (*Xcluster*), which exists at the tail of the CG and Merging cluster (*Mcluster*), which exists if the structure contains a loop. The point source P_S is any medial voxel that belongs to the \mathcal{CS} of the structure, which is found by searching the CG for the voxel with maximum $D(\mathbf{x})$. If there exists more than one voxels with the same maximum value, we select the one with minimum $\|V(\mathbf{x})\|$. The medial voxel of a cluster is computed similarly to P_S but with searching the cluster rather than the CG.

The proposed framework can be summarized as follows: (1) Construct the gradient vector flow $V(\mathbf{x})$ and then compute $\lambda(\mathbf{x})$ Eq. (10), (2) Construct the minimum distance field $D(\mathbf{x})$, (3) Find the point source P_S , (4) Propagate a moderate speed wave from P_S Eq. (13), discretize the resultant distance field $D_1(\mathbf{x})$, and construct the CG, (5) Identify the extreme and merging nodes, (6) Construct a new distance field $D_2(\mathbf{x})$ from P_S by propagating a fast speed wave Eq. (9), (7) If the object contains loops, extract their \mathcal{CS} as suggested in [12], and finally, (8) Extract those \mathcal{CS} that originate from extreme nodes and ends with P_S or ends on a previously extracted path to prevent overlapped paths.

3 Results and Discussion

We have quantitatively validated the proposed framework against ground truth \mathcal{CS} that are generated analytically and then discretized. Each phantom is created by translating a sphere of a fixed or varying radius along its ground truth \mathcal{CS} . The phantoms are designed to mimic the geometrical and topological properties of anatomical structures such as: (1) high curvature and torsion (e.g., blood vessels), (2) sudden change in the organ's cross section (e.g., colon or aneurysm in vessels), and (3) several branching nodes (e.g., blood vessels and tracheo-bronchial trees). To study the sensitivity of the proposed method to noise, 50% of phantom's boundary voxels are corrupted by additive noise to simulate segmentation error as shown in Figure 2. A quantitative analysis was carried out by computing the amount of overlap, average, and maximum distance between the ground truth and computed \mathcal{CS} for both noise-free and noisy phantoms. The quantitative results are presented in Table 1. Although the amount of overlap is less than 90 %, the average and maximum distance never exceeded 0.42 and 1.41 mm (e.g., 1-2 voxels), respectively. In the presence of noise, the amount of overlap has been decreased by only 8 %, while the average and maximum distance has been increased slightly. To conclude, the computed \mathcal{CS} are always adjacent to the ground truth ones, which is quite acceptable for flight paths in VE.

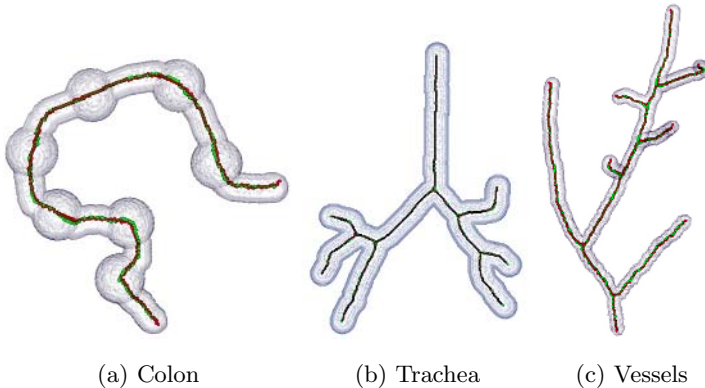


Fig. 2. Noisy synthetic phantoms: (a) Colon. (b) Trachea. (c) Vessels. The voxels of the ground truth, computed, and overlapped \mathcal{CS} are represented by green, red, and black spheres, respectively.

Table 1. Quantitative validation for noise-free and noisy phantoms

Phantom	Colon		Trachea		Vessels	
Size	200 × 357 × 50		200 × 357 × 50		220 × 110 × 210	
State	Noise-Free	Noisy	Noise-Free	Noisy	Noise-Free	Noisy
Percentage of Overlap	72 %	65 %	82 %	76 %	65 %	64 %
Average Distance (mm)	0.33	0.42	0.18	0.19	0.42	0.46
Maximum Distance (mm)	1.41	2.0	1.0	1.41	1.41	2.0

Case Study (Virtual Colonoscopy): The goal of this case study is to bring about one of the important goals of virtual colonoscopy (VC) as a diagnostic tool. The ideal scenario is that the segmented colon maintains the intricate details of the real colon, and that the virtual camera projects views as detailed as those shown in real colonoscopy. If that is achieved, then analysis of projected views can be used for automatic colon scanning against abnormalities, and hence an early detection of colon cancer using VC would be a strong possibility. Our research group is currently pursuing this goal.

In order to illustrate the potential of this research in colonoscopy, the proposed framework has been tested on several CT datasets acquired using Siemens Sensation CT scanner. The dataset volume is $512 \times 512 \times 580$ with voxel size $0.74 \times 0.74 \times 0.75$. All patients have undergone standard cleansing preparations prior to scan. In Figure 3(e-l), we show different polyps captured by the virtual camera for different colons as shown in Figure 3(a-d). The average running time for these large datasets was 9 minutes on a 2.6GHz AMD Linux workstation with 4.0 GB RAM.

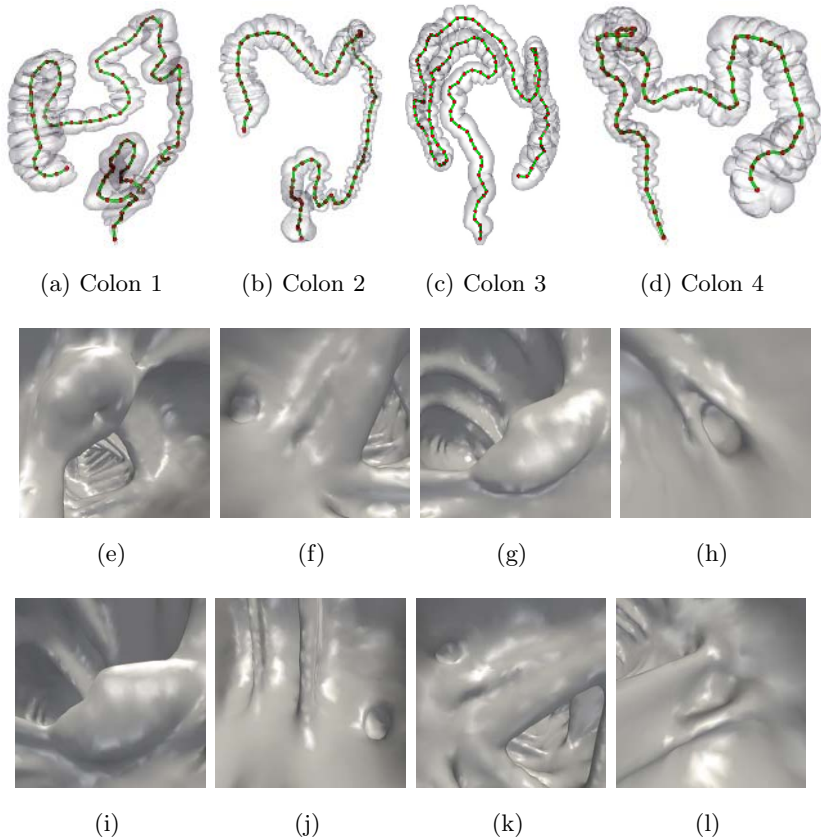


Fig. 3. Virtual Colonoscopy (a-d) Computed \mathcal{CS} for different colon datasets (e-l) Polyp views captured by the virtual camera

The proposed framework is controlled by two main parameters α in Eq.(10), which controls the centeredness of the computed flight path and β in Eq.(13), which controls the generation of the CG. Experimental results showed that the lower bound for α is equal to 15.0, while a suitable range of β is $2.0 < \beta < 4.0$. We automated the framework by setting $\alpha = 20.0$ and $\beta = 3.0$.

We have slightly modified the original GVF to suit the nature of the problem as follows: (1) $f(\mathbf{x}) = I(\mathbf{x})$ such that the vector field points towards the center of the object. (2) The computation of the GVF is restricted to the internal voxels of the structure and hence computationally more efficient than the original GVF. (3) The GVF field is not normalized to maintain the medialness property, otherwise the magnitude of the GVF is unity everywhere. The GVF parameters are set as follows: $\mu = 0.15$, $\Delta t = 0.5$, and the number of iterations is set to 100.

The complexity of the framework in the worst case for n voxels is $O(3n \log n + kn)$. The complexity of computing the distance field using the HAFMM and the GVF is given by $n \log n$ and kn , respectively, where k is the number of iterations.

4 Conclusion and Future Work

In this paper, we have proposed a general framework for computing flight paths through tubular and non-tubular structures using partial differential equations (PDE). Unlike all voxel-based methods, the framework does not require voxel sizes to be isotropic because the solution of the PDE takes into account the data spacing. The computed flight paths are highly centered, connected, topology preserving, and forms a graph at no additional cost.

Most of the processing time of the proposed method is spent on the computation of the GVF because it requires extensive floating point operations. Therefore, in the future, we intend to implement the GVF in the graphical processing unit (GPU) to alleviate its computational complexity.

Acknowledgment. The colonoscopy dataset used in this paper were provided by 3DR Inc., Louisville, Kentucky.

References

1. Baert, A.L., Sartor, K.: Virtual Endoscopy and Related 3D Techniques. Springer (2001)
2. Buthiau, D., Khayat, D.: Virtual Endoscopy. Springer-Verlag (2003)
3. Zhou, Y., Toga, A.W.: Efficient skeletonization of volumetric objects. *IEEE Transactions on Visualization and Computer Graphics* **5** (1999) 196–209
4. Bitter, I., Kaufman, A.E., Sato, M.: Penalized-distance volumetric skeleton algorithm. *IEEE Transactions on Visualization and Computer Graphics* **7** (2001) 195–206
5. Ma, C., Sonka, M.: A fully parallel 3d thinning algorithm and its applications. *Computer Vision and Image Understanding* **64** (1996) 420–433
6. Svensson, S., Nyström, I., di Baja, G.S.: Curve skeletonization of surface-like objects in 3d images guided by voxel classification. *Pattern Recognition Letters* **23** (2002) 1419–1426

7. Deschamps, T.: Curve and Shape Extraction with Minimal Path and Level-Sets techniques - Applications to 3D Medical Imaging. PhD thesis, Université Paris-IX Dauphine (2001)
8. Bouix, S., Siddiqi, K., Tannenbaum, A.: Flux driven fly throughs. In: Computer Vision and Pattern Recognition, CVPR. (2003) 449–454
9. Aylward, S.R., Bullitt, E.: Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction. *IEEE Trans. Medical Imaging* **21** (2002) 61–75
10. Wink, O., Niessen, W., Viergever, M.: Multiscale vessel tracking. *IEEE Trans. Medical Imaging* **23** (2004) 130–133
11. Hassouna, M.S., Farag, A.A.: Robust centerline extraction framework using level sets. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition, CVPR, San Diego, CA (2005) 458–465
12. Hassouna, M.S., Farag, A.A.: Pde-based three dimensional path planning for virtual endoscopy. In: Proc. of Information Processing in Medical Imaging, IPMI, Glenwood springs, Colorado (2005)
13. Xu, C., Prince, J.L.: Gradient vector flow: A new external force for snakes. In: CVPR '97: Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97), Washington, DC, USA (1997) 66–71
14. Xu, C.: Deformable Models with Application to Human Cerebral Cortex Reconstruction from Magnetic Resonance Images. PhD thesis, Department of Electrical and Computer Engineering, Johns Hopkins University (1999)
15. Chang, S., Metaxas, D., Axel, L.: Scan-conversion algorithm for ridge point detection on tubular objects. In: MICCAI, MontrTal, Canada (2003) 158–165
16. Cohen, L., Kimmel, R.: Global minimum for active contour models: A minimal path approach. *International Journal of Computer Vision* **24** (1997) 57–78
17. Sethian, J.: Level Sets Methods and Fast Marching Methods. 2nd edn. Cambridge University Press (1999)

Panoramic Views for Virtual Endoscopy

Bernhard Geiger, Christophe Chefd'hotel, and Sandra Sudarsky

Siemens Corporate Research Inc., Princeton, NJ, 08540, USA

{geiger.bernhard, christophe.chefdhotel, sandra.sudarsky}@siemens.com

Abstract. This paper describes a panoramic projection designed to increase the surface visibility during virtual endoscopies. The proposed projection renders five faces of a cubic viewing space into the plane in a continuous fashion. Using this real-time and interactive visualization technique as a screening method for colon cancer could lead to significantly shorter evaluation time. It avoids having to fly through the colon in both directions and prevents the occlusion of potential polyps behind haustral folds.

1 Introduction

Virtual endoscopy is a non-invasive diagnostic procedure aimed at exploring the inner surface of anatomical structures inside the human body. Using advanced image-processing techniques, 3D models are reconstructed from a series of high-resolution 2D images (e.g. CT or MR). A physician can then automatically or interactively navigate through the 3D virtual model to perform a diagnosis.

One of the most promising uses for virtual endoscopy is the screening of patients for colorectal cancer. This technique mimics conventional colonoscopy with the added advantage that it is less invasive and potentially more desirable as a screening method. An added benefit of virtual colonoscopy is the ability to fly-through the colon in both an antegrade and retrograde direction, significantly increasing the amount of surface displayed. Nevertheless, important sections of the colon often remain obstructed behind haustral folds (see Fig. 1).

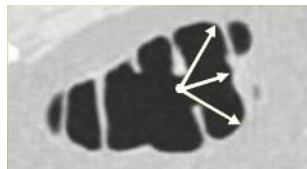


Fig. 1. Limited visibility of a standard virtual camera projection

In this paper we introduce a real-time, interactive visualization technique meant to increase the surface visibility while introducing very little deformation.

2 Previous Work

A number of projections have been proposed in the literature to overcome some of the limitations of the standard perspective volume rendering. These projections seek to display more of the surface of the colon and thereby increase the polyp detection sensitivity.

Serlie et al. [2] propose the use of a cubic viewing space with the virtual camera located in the center of the cube and projecting on each face an image with a 90-degree viewing angle. The cube is then unfolded into a single plane presenting a 360-degree field of view of the colon surface. Off-line animated image sequences are generated in [2] from a number of point samples selected along the central path through the colon. Figure 2 shows a single frame of an animated sequence using this cubing mapping. One of the main drawbacks of this approach is that lesions can be split across several sides of the cube. In addition, the layout is not very natural. It requires to focus on multiple regions of the image to obtain an overall perception of the observed tissues. Similarly, Tiede et al. [6] use the images corresponding to the six faces of the cube to generate precomputed movies known as spherical panoramas.



Fig. 2. Six sided unfolded cube projection

Paik et al. [1] suggest the use of map projections that visualize the entire surface of a viewing sphere. Using Mercator and stereographic projections, the surface of the sphere is transformed onto a flat surface. Unfortunately, the distortion introduced by these projections is too high specially near the poles.

Vilanova et. al. [4, 3] propose a method to visualize the complete inner colon surface by virtually bisecting the colon along its centerline and unfolding it. Although physicians are unfamiliar with the resulting view, it provides them with a complete overview of the colon in a single image. However, unfolding introduces geometric distortions that can make lesions difficult to identify.

3 Our Approach

We propose two panoramic endoscopic projections as variations of the unfolded cubic view. They are designed to obtain a continuous projection, a large field of

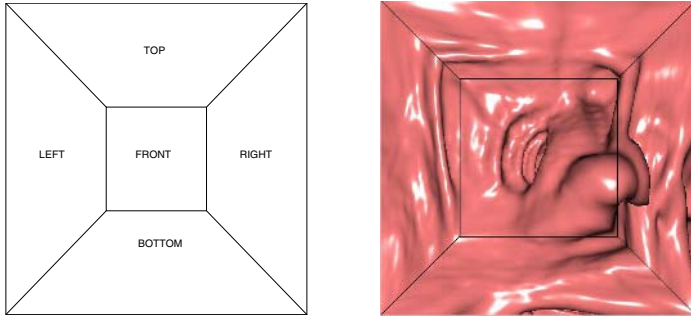


Fig. 3. Panoramic projection of the unfolded cube onto a square

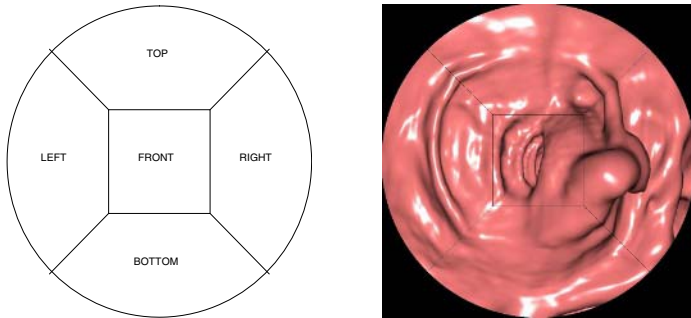


Fig. 4. Panoramic projection of the unfolded cube onto a disk

view, and limited distortion. As in [2], we place the virtual camera in the center of a cubic space and project the images into a single plane using a 90-degree viewing angle. However, instead of using a straightforward unfolding, the front image is mapped into a rectangular window at the center of the plane, the back projection is removed, while the other four images are projected into trapezoidal shaped windows arranged as shown in Figure 3. This approach removes the discontinuities found in [2] and at the same time makes better use of the screen space.

A fast raycast algorithm is used to render each frame in real-time, providing a fully interactive navigation. Figure 3 displays a single frame during a fly-through sequence. The only distortion introduced arises from mapping the cubic walls into trapezoidal views. In order to minimize these distortions, we propose a second projection in which the front face is mapped into a square while the other four faces are mapped around it into a disk (see Fig. 4).

These projections are designed to guarantee a smooth transition between the five faces. A careful comparison between these projections and the original mapping shown in Figure 2 reveals some distortion closer to the disk periphery. The deformation lessens towards the center of the disk, displaying the frontal view without any deformation whatsoever.

The size of the frontal view can be adjusted as desired. In the limiting case, when the frontal view covers the entire window, both of these panoramic projections are equivalent to the standard projection with a single virtual camera looking straight ahead.

4 Projections

Each endoscopic view is obtained by casting rays from the center of a virtual endoscope positioned inside the volume. The intensity of a point in the endoscopic view is an arbitrary function of the volume intensity values along the corresponding ray. From a given endoscope position, each ray is uniquely characterized by a direction vector. Hence, we model a panoramic view as a vector field (a collection of direction vectors).

In this section, we describe the direction fields corresponding to the proposed panoramic views. These vectors are computed for a fixed endoscope position (we assume the endoscope is placed at the origin of the volume coordinate system and points in the z direction). Arbitrary location and orientation of the endoscope can be taken into account by translation and rotation of the vector field.

The projection is a two-step process. First we find for each point in the panoramic view the corresponding point in one of the faces of the unfolded cube. Then, we find the corresponding direction vector. We focus here on the first step, while the second step is briefly described in Section 4.3.

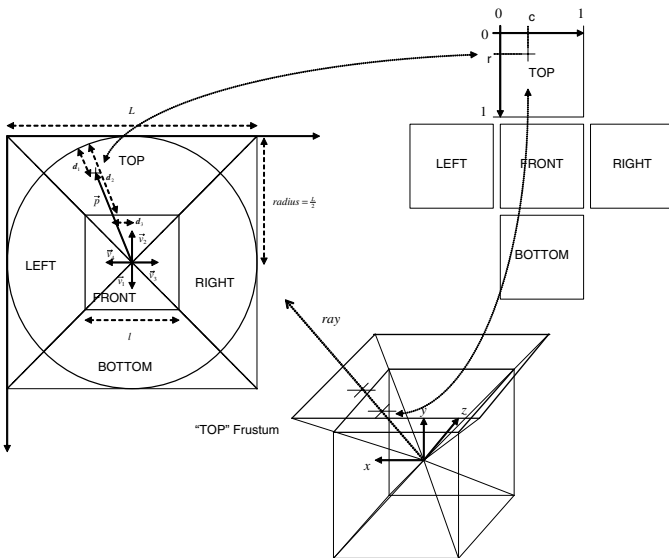


Fig. 5. Determining the coordinates (r, c) on the unfolded cube corresponding to a vector \vec{p} on the disk

4.1 Continuous Panoramic Projection of the Unfolded Cube Onto a Disc

The endoscopic view is modeled as a disk inscribed in a square of side length L (i.e. the disk has a radius equal to $L/2$). Any point inside this square is characterized by a position vector \vec{p} with respect to its center (see Fig. 5). The inner square of side length l corresponds to the front view of the unfolded cube. The remaining sections of the disk contain transformed versions of the top, left, right, and bottom views of the unfolded cube.

We introduce four unit vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$. We denote $\vec{v}(i)$ the i -th component of a vector \vec{v} and $\|\vec{v}\|$ its norm. The expression $|x|$ denotes the absolute value of a scalar x . The operator \cdot denotes the scalar product of two vectors. The scalar product of the vector \vec{p} with \vec{v}_i determines the face (bottom, top, right, left, $i = 1, 2, 3, 4$ respectively) used during the mapping.

We attach to each face of the unfolded cube a coordinate system (where $(0, 0)$ is the upper-left corner of a face and the side length is 1). On a given face, points are located by a set of coordinates (r, c) (see Fig. 5). The mapping between the disk and the cube's faces is given by the following algorithm:

```

If  $\|\vec{p}\| > radius$  Then set the intensity value to zero
Else (the point is inside the disk)
    If  $\max(|\vec{p}(1)|, |\vec{p}(2)|) < \frac{l}{2}$  Then (the point is in the front view)
         $r = 0.5 + \frac{\vec{p}(1)}{l}$ 
         $c = 0.5 + \frac{\vec{p}(2)}{l}$ 
    Else
         $\vec{n} = \frac{\vec{p}}{\|\vec{p}\|}$ 
        If  $\vec{n} \cdot \vec{v}_i > \frac{\sqrt{2}}{2}$  Then
             $cos\_angle = \vec{n} \cdot \vec{v}_i$ 
             $norm\_p\_inside\_front = \frac{1}{2 \cdot cos\_angle}$ 
             $r = \frac{radius - \|\vec{p}\|}{radius - norm\_p\_inside\_front} = \frac{d_1}{d_2}$  (see Fig. 5)
             $c = 0.5 + \frac{\vec{n}(2)}{2 \cdot cos\_angle}$  (note:  $\frac{d_3}{l} = \left| \frac{\vec{n}(2)}{2 \cdot cos\_angle} \right|$ )
        End
    End
End

```

4.2 Continuous Panoramic Projection of the Unfolded Cube Onto a Square

This view is a variant of the previous approach, where the unfolded cube is mapped onto a square. There is no region left unused, but image distortions are more significant along the diagonals of the endoscopic view (see Fig. 6). The main difference from the computational point of view is the calculation of the r coordinate given below:

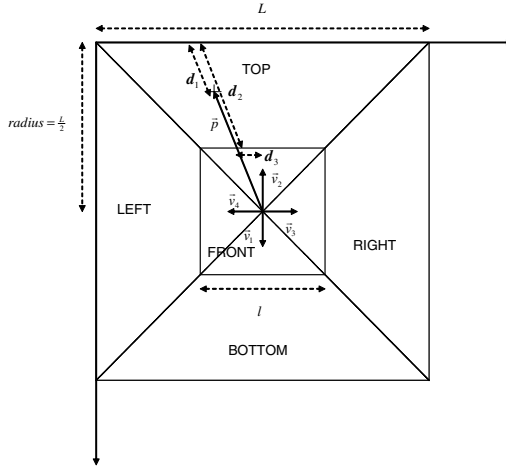


Fig. 6. Determining the coordinates (r, c) on the unfolded cube corresponding to a vector \vec{p} on the square

$$r = \frac{(radius / \cos_angle) - \|\vec{p}\|}{(radius / \cos_angle) - norm_p_inside_front} = \frac{d_1}{d_2}.$$

4.3 Finding Direction Vectors from Unfolded Cube Coordinates

To each face of the unfolded cube corresponds a frustum (see Fig. 5 and Fig. 7). A frustum is characterized by a front plane (a square of edge length l_{front} placed at distance d_{front} of the endoscope location) and a back plane (a square of edge length l_{back} placed at distance d_{back} of the endoscope location).

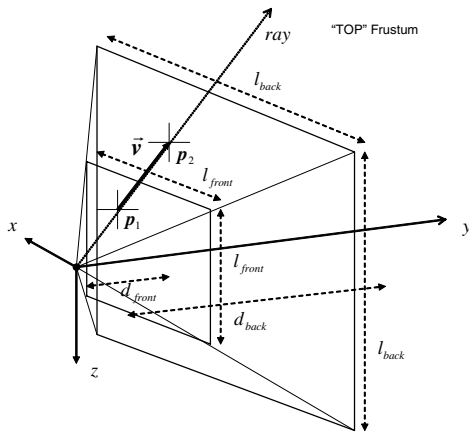


Fig. 7. Finding the direction vector \vec{v} of the ray corresponding to the location (r, c) on the cube faces

The direction vector $\vec{v} = p_2 - p_1$ of the ray corresponding to the location (r, c) on one of the cube faces is given by:

$$p_1 = \begin{bmatrix} \frac{l_{front}}{2} - c \cdot l_{front} \\ d_{front} \\ -\frac{l_{front}}{2} + r \cdot l_{front} \end{bmatrix} \text{ and } p_2 = \begin{bmatrix} \frac{l_{back}}{2} - c \cdot l_{back} \\ d_{back} \\ -\frac{l_{back}}{2} + r \cdot l_{back} \end{bmatrix}.$$

Most of the previous operations can be combined to obtain a more compact and efficient implementation. The symmetry of the vector field can also be exploited to reduce the computational cost.

5 Results

Figure 8 shows a side by side comparison of the amount of the colon surface visualized using a standard camera projection and the proposed panoramic projections. Voxels shown in red correspond to the areas not visible after a one-way flight through the virtual colon. Since the panoramic projections show simultaneously a forward, left, right, up and down views, we get nearly 100% of surface coverage with one single pass [5]. To achieve better coverage, a conventional virtual navigation typically executes an additional fly-through in the reverse direction. The added flight not only increases the reading time as some parts of the surface are examined twice, but even then some areas remain hidden behind haustral folds. The panoramic projections therefore could be used to speed up the 3D reading as we can cut the fly-time in half.

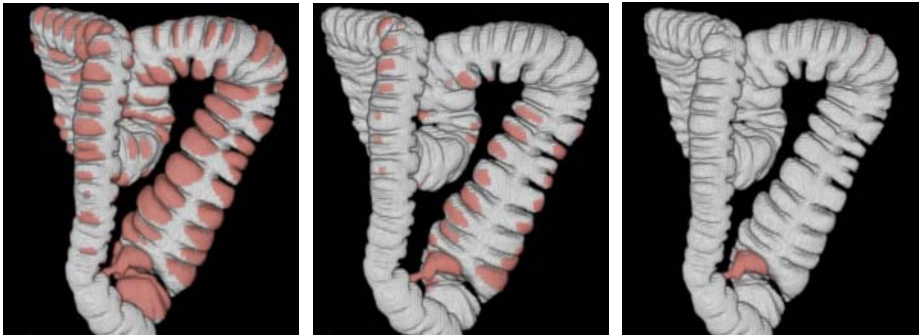


Fig. 8. Global view of the colon after a one-way flight using a standard perspective projection (left), a two-way flight using a standard projection (middle), and a one-way flight using a panoramic projection (right)

Our implementation allows a real-time fly through and, as discussed before, the geometry of the projection limits distortion. User studies are underway to compare the detection rate, speed and general acceptance of the proposed projections. Preliminary results seem very positive, as features tend to stay in view

for longer periods of time. The proposed projections are not restrictive to the colon and therefore can potentially be used to display the inner surface of any virtual organ.

References

1. Paik, D., Beaulieu, C., Jeffrey, R., Karadi, C., Napel, S.: Visualization Modes for CT Colonography Using Cylindrical and Planar Map Projections. *J. Comput. Assist. Tomography*, **24**:179-188, 2000.
2. Serlie, I., Vos, F., van Gelder, R., Stoker, J., Truyen, R., Gerritsen, F., Nio, Y., Post, F.: Improved Visualization in Virtual Colonoscopy Using Image-Based Rendering, *Proceedings of IEEE TCVG Symposium on Visualization*, 137–146, 2001.
3. Vilanova Bartrolí, A., Gröller, E.: Geometric Modelling for Virtual Colon Unfolding, in *Geometric Modeling for Scientific Visualization*, 453–468, Springer-Verlag, 2003.
4. Vilanova Bartrolí, A., Wegenkittl, R., König, A., Gröller, E.: Nonlinear Virtual Colon Unfolding. *IEEE Visualization Conference Proceedings*, 411–418, 2001.
5. Wax, M., Kreeger, K., Anderson, J.: Endoscopic View in Virtual Colonoscopy: Achieving Complete Surface Visualization, *RSNA*, 2001.
6. Tiede, U., Sternberg-Gospos, N., Steiner, P., Höhne, K. H., Virtual Endoscopy using Spherical QuickTime-VR Panorama Views, in J. Westwood et al. (eds.): *Medicine meets Virtual Reality, Studies in Health Technology and Informatics 85*, IOS Press, Amsterdam, 2002, 523-528.

Toward Automatic Computer Aided Dental X-ray Analysis Using Level Set Method

Shuo Li¹, Thomas Fevens¹, Adam Krzyżak¹, Chao Jin¹, and Song Li²

¹ Medical Imaging Group,

Department of Software Engineering and Computer Science,
Concordia University, Montréal, Québec, Canada

{shuo_li, fevens, krzyzak, chao_jin}@cs.concordia.ca

² School of Stomatology, Anhui Medical University, Hefei, Anhui, P. R. China
xlisong@sohu.com

Abstract. A Computer Aided Dental X-rays Analysis (CADXA) framework is proposed to semi-automatically detect areas of bone loss and root decay in digital dental X-rays. In this framework, first, a new proposed competitive coupled level set method is proposed to segment the image into three pathologically meaningful regions using two coupled level set functions. Tailored for the dental clinical environment, the segmentation stage uses a trained support vector machine (SVM) classifier to provide initial contours. Then, based on the segmentation results, an analysis scheme is applied. First, the scheme builds an uncertainty map from which those areas with bone loss will be automatically detected. Secondly, the scheme employs a method based on the SVM and the average intensity profile to isolate the teeth and detect root decay. Experimental results show that our proposed framework is able to automatically detect the areas of bone loss and, when given the orientation of the teeth, it is able to automatically detect the root decay with a seriousness level marked for diagnosis.

1 Introduction

The past few years has seen a great increase in the usage of digital dental X-rays in dental practices in North America. Their adoption is motivated by the fact that digital X-rays systems are more sensitive than X-ray film systems while allowing for up to a 90% reduction in a patient's exposure to X-rays compared to traditional systems. With the higher resolution digital X-rays of sections of the teeth, or of the entire jaw, comes the possibility of more accurate diagnoses.

Dental X-rays play an important role in detecting such pathological problems such as periodontitis, chronic periapical periodontitis and bone loss that cannot be seen during a visual examination. Periodontitis is a dental disorder that results from the progression of gingivitis, involving the inflammation and infection of the ligaments and bones that support the teeth. Early detection of bone loss and root decay is very important since often they can be remedied by dental procedures, such as a root canal, for example. Without early treatment, bone loss may lead to loss of teeth or erosion of the jaw bone. In the dental X-ray film

based environment, the dentist does the analysis of the X-ray in their head. With the increasingly wide use of digital radiography, the dental X-rays are accessible electronically which makes computer aided dental X-rays analysis possible.

However, although dental X-rays are widely used, it is a challenging task to do automatic, or even semi-automatic, computer aided dental X-rays analysis. Compared with other types of images, dental X-rays analysis is a challenging problem for classical image processing methods due to the following characteristics: (1) poor image modalities: noise, low contrast, and sample artifacts; (2) very complicated topology; and (3) there may not be clear lines of demarcation between regions of interest, which is especially true for dental X-rays since problem teeth tend to have very complicated structures and are normally coupled with healthy teeth. Therefore dental X-rays are normally inspected by a dentist. Although efficient, human inspection requires specialized training which is increasingly expensive. In addition, human inspection gives a subjective judgment which may vary from person to person, and, as such, does not give a quantitative measurement. Inspection results could be affected by many factors, such as fatigue and distraction by other features in the image, for example. Also, some early bone loss may not be visible to the human eye. Early detection of bone loss and root decay is very important since often they can be remedied by dental procedures, such as a root canal, for example. Without early treatment, bone loss may lead to loss of teeth or erosion of the jaw bone. All these issues indicate a need for effective automatic dental X-rays analysis.

In this paper, we report on innovative work on computer aided dental X-rays analysis which semi-automatically provides indications to the dentist aid in finding bone loss and root decay, which are the primary reasons that X-rays are taken in many countries. For this implementation, we will be dealing primarily with X-rays which are close-up views of a few individual teeth, a common type of dental X-rays, although the approach developed here can be adapted to larger scale dental X-rays. Compared with panoramic dental X-rays which include the entire jaw region, close-up images taken for these purposes are more challenging since the orientation of the teeth may not be fixed and problem areas are either complicated, or easily overlooked. To the best of our knowledge, we are the first group working towards automatic computer aided dental X-rays diagnosis for the detection of bone loss and root decay. This paper reports on our preliminary results towards this goal.

2 Proposed Framework

As shown in Fig. 1, the framework consists of two phases: segmentation and analysis. First, we employ a new proposed competitive level set segmentation method to segment the image into three regions using two coupled level set functions. Based on the segmentation results, an analysis scheme is applied. The scheme first builds an uncertainty map which is then used to automatically mark any areas of bone loss. Subsequently, an average intensity profile based method is employed to isolate the teeth and detect possible root decay. Finally the estimated seriousness level of the root decay will be marked.

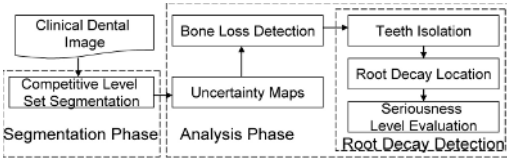


Fig. 1. Framework diagram

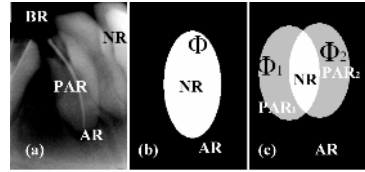


Fig. 2. Region modelling of dental image (a), one level set segmentation (b) and competitive level set (c)

2.1 Variational Level Set Segmentation

In this implementation, we propose a new variational level set segmentation method driven by pathological energy modelling. The modelling explicitly incorporates regions of problems as part of the modelling, so the identification of such areas would be an automatic product of the segmentation. The level set method segments the dental X-rays into three regions: Normal Region (NR), Potential Abnormal Region (PAR), Abnormal and Background Region (ABR) using two competitive level set functions.

Competitive Level Set Segmentation. With an evolving curve C , one level set function divides the image (u) into two parts: Normal Region Ω_{NR} (“+” region) and Abnormal Region Ω_{AR} (“-” region) as shown in Fig. 2(b). The energy functional is given by

$$E(\Phi) = \beta_1 \int_{\Omega_{NR}} \frac{(u - c_{NR})^2}{\sigma_{NR}^2} dx dy + \beta_2 \int_{\Omega_{AR}} \frac{(u - c_{AR})^2}{\sigma_{AR}^2} dx dy, \tag{1}$$

where c_i is the mean grey value of the Ω_i , σ_i is the variance and β_i is a constant.

However for the diagnosis of pathological problems, potential problems areas which might be between normal and abnormal regions are of particular interest. Therefore we propose a competitive coupled level set model for two level set functions to segment the image into three regions. As shown in the Fig. 2(a), a X-ray image (u_0) can be divided into four regions of interest: the Normal Region (Ω_{NR}), the Potential Abnormal Region (Ω_{PAR}), the Abnormal Region (Ω_{AR}) and the Background Region (Ω_{BR}). Since Ω_{AR} and Ω_{BR} is not separable in terms of intensity values, so in the segmentation, we take Ω_{AR} and Ω_{BR} to be a single region: the Abnormal and Background Region (Ω_{ABR}). The energy functional for the two coupled level set functions (Φ_1 and Φ_2) can be modeled as:

$$E(\Phi_1, \Phi_2) = \lambda_1 \int_{\Omega_{NR}} \frac{(u - c_{NR})^2}{\sigma_{NR}^2} dx dy + \lambda_3 \int_{\Omega_{ABR}} \frac{(u - c_{ABR})^2}{\sigma_{ABR}^2} dx dy + \lambda_2 \int_{\Omega_{PAR}} \text{Min}\left(\frac{(u - c_{PAR1})^2}{\sigma_{PAR1}^2}, \frac{(u - c_{PAR2})^2}{\sigma_{PAR2}^2}\right) dx dy \tag{2}$$

where the function $\text{Min}(x, y)$ returns the smaller value of x and y , and λ_i is a constant. The modelling is in the same spirit as multiphase modelling in [1].

Using competitive coupled level set functions, if both level set functions classify an area as a normal region, we take it as a normal region Ω_{NR} ; if both level

set functions classify a region as an abnormal and background region, we take it as an abnormal and background region Ω_{ABR} ; however if only one of the level set functions segments a region as a normal region, the region will be taken as a potential abnormal region Ω_{PAR} . Compared to three coupled level set functions, two coupled level sets is able to achieve faster and more robust segmentation as described in [1].

To achieve fast and robust segmentation, a hybrid coupled level sets functional that combines minimal variance (Eq. 2), the optimal edge integrator [2] and the geodesic active contour model [3] is used:

$$E = E(\Phi_1, \Phi_2) - \gamma_1 E_{LAP} + \gamma_2 E_{GAC}, \tag{3}$$

where γ_i are constants.

The geodesic active contour (E_{GAC}) and edge functional (E_{LAP}) are defined in Eq. 4. The edge functional was proposed in [2] where the authors show that a Laplacian edge detector Δu provides optimal edge integration with regards to a very natural geometric functional.

$$E_{GAC}(C) = \iint g(C) dx dy \tag{4}$$

$$E_{LAP}(C) = \int_C < \nabla, \mathbf{n} > ds + \iint_{\Omega_C} K_u |\nabla u| dx dy.$$

Here K_u is the mean curvature of the level set function, \mathbf{n} is the unit vector normal to the curve and ds is the arc length of curve C . Function $g(x, y)$ is an inverse edge indicator function introduced in [4], and defined as $g(x, y) = \alpha^2 / (\alpha^2 + |\nabla u|^2)$, where α is a constant and ∇ is the gradient operator. The level set functions Φ_i are derived from the functional in Eq. 3 as shown:

$$\frac{\partial \Phi_1}{\partial t} = \delta_\epsilon(\Phi_1) \left[\gamma_2 \operatorname{div} \left(g \frac{\nabla \Phi_1}{|\nabla \Phi_1|} \right) - \frac{(u - c_{NR})^2}{\sigma_{NR}^2} H(\Phi_2) - \frac{(u - c_{PAR})^2}{\sigma_{PAR}^2} (1 - 2H(\Phi_2)) \right. \tag{5}$$

$$\left. + \frac{(u - c_{ABR})^2}{\sigma_{ABR}^2} (1 - H(\Phi_2)) - \gamma_1 u_{\xi\xi} \right],$$

$$\frac{\partial \Phi_2}{\partial t} = \delta_\epsilon(\Phi_2) \left[\gamma_2 \operatorname{div} \left(g \frac{\nabla \Phi_2}{|\nabla \Phi_2|} \right) - \frac{(u - c_{NR})^2}{\sigma_{NR}^2} H(\Phi_1) + \frac{2(u - c_{PAR})^2}{\sigma_{PAR}^2} H(\Phi_1) \right. \tag{6}$$

$$\left. + \frac{(u - c_{ABR})^2}{\sigma_{ABR}^2} (1 - H(\Phi_1)) - \gamma_1 u_{\xi\xi} \right].$$

Here, $H(\cdot)$ is the Heaviside function, $\operatorname{div}(\cdot)$ is the divergence operator, and $u_{\xi\xi} = \Delta u - K_u |\nabla u|$.

Although we only apply competitive level set segmentation to dental X-rays, the segmentation method can be extended to X-rays and some types of CT images for three region segmentation using two level set functions.

Segmentation Phase. To apply the level set method in the clinical environment, we adapt the segmentation framework proposed by Li *et al.* [4, 5] which uses a trained SVM to provide a good initial contour for the level set method which greatly speeds up convergence of the coupled level set functions. Following the same principle, we use an SVM to provide initial contours for two coupled

level set functions. The purpose is not only to speed up segmentation convergence, but also to use competitive couple level set functions to find regions of interest, as discussed in section 2.1. The segmentation phase has two stages: a training stage and a clinical segmentation stage.

During the training stage, manually chosen representative images are segmented by hierarchical level set region detection using the Chan and Vese level set method [6]. In the hierarchical level set region detection, first a level set function is used to separate Ω_{ABR} from the rest of the image (Ω_{AR} and Ω_{BR}). Then another level set function is used to separate Ω_{AR} and Ω_{BR} . Then these results are used to train an SVM classifier.

During the clinical segmentation stage, dental X-rays are first classified by the trained SVM. The classifier is able to classify three regions (Ω_{ABR} , Ω_{NR} and Ω_{PAR}) based only on intensity which may not be accurate. Then two coupled level set functions are used to further segment the images. For Φ_1 , we set classified Ω_{NR} as the “+” region and rest of the image as the “-” region; for Φ_2 , we set classified Ω_{NR} and Ω_{PAR} region as the “+” region and rest of the image as the “-” region. Although SVM is only able to give a coarse segmentation, it provides a very good and competitive initial contours for two coupled level set functions. The final segmentation will be obtained by evolution of these two level set curves.

2.2 Analysis Phase

The analysis phase contains three steps: uncertainty map building, bone loss detection and root decay detection. The first two are fully automatic. The only manual input is the image orientation during root decay detection. This requirement is not difficult to accommodate in the dental clinical environment.

Uncertainty Maps. First for each image, an uncertainty map is built based on following uncertainty measurement:

$$\psi = \frac{(u-c_{NR})\tau_1+(u-c_{PAR})(\tau_2-2\tau_1)+(u-c_{ABR})(1-H(\phi_1))(1-H(\phi_2))}{\sigma_{NR}\tau_1+\sigma_{PAR}(\tau_2-2\tau_1)+\sigma_{ABR}(1-H(\phi_1))(1-H(\phi_2))}$$

where $\tau_1 = H(\phi_1)H(\phi_2)$ and $\tau_2 = H(\phi_1) + H(\phi_2)$.

Bone Loss Detection. Areas of bone loss will generally occur in those regions of high uncertainty. Therefore, we mark these areas with different levels of emphasis according to the uncertainty measurement and region segmented.

Although the uncertainty map is an objective uncertainty measure, it fails to provide direct visual cues. To achieve the visual assistance, the RGB channels of the image are used to couple the intensity values of the image with the degree of uncertainty at each pixel. For all regions, the G channel is used to represent the intensity value of each pixel of the original dental X-ray image. The uncertainty values are nonlinearly (For this implementation, users can interactively choose the polynomial functions to get best visualization results.) scaled to the range -255 to 255. To differentiate between the three regions, we apply the following schemes to R and B channel for pixels in each region: 1) For Ω_{ABR} , B channel is

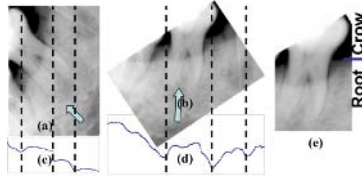


Fig. 3. Teeth isolation. (a) Original image. (b) Rotated image. (c) Integrated intensity. (d) Average intensity profile. (e) Two parts of a tooth: crown and root.

set to 0. The R channel is set to the uncertainty value if it is negative; otherwise the R channel is set to 0. 2) For Ω_{PAR} , both the R and B channels are set to 200 to emphasize this region. 3) For Ω_{NR} , both the R and B channels are set to the uncertainty value if it is negative; otherwise, both channels are set to 0.

Root Decay Detection. Root decay detection consists of three steps: tooth isolation, root decay location and seriousness level evaluation. This is a semi-automatic process in which the orientation of the teeth is supplied manually.

Teeth isolation: As suggested by Jain *et al.* [7], since the teeth usually yield higher intensity values than the jaws and other tissues, the gap of teeth will have a very low value on the integrated intensity value profile which are the sums of the intensities of pixels along the vertical direction. However, unlike dental X-ray images used for a forensic purpose which can be assumed to have certain orientation, clinical dental X-rays used to detect root decay, etc., could have any orientation. As shown in the Fig. 3(a) and (c), if the orientation varies, this profile method will not able to obtain the correct isolation. Therefore as a preliminary step, we assume that the orientation is given. Then we rotate the image according to the given orientation so that the teeth are aligned in a consistent direction. After rotation, instead of using an integrated intensity value, we use the average of the intensity value (the integrated intensity value divided by the number of pixels) as shown in Fig. 3(b) and (d). After rotation, we use an SVM to judge where are the gaps in the teeth using a 1D window based feature vector from average intensity profile for training and classification.

Root decay location and seriousness level evaluation: Three types of regions are considered to be root decay if they are found at the root of teeth with the following seriousness levels, in order of most serious to least serious: 1. advise level: if the Ω_{ABR} is found at the root of the tooth; 2. warning level: if the Ω_{PAR} of any uncertainty is found at the root of the tooth; 3. attention level: if a high uncertainty area of Ω_{NR} is found at the root of the tooth.

3 Experimental Results

3.1 Segmentation

Figs. 4,5 and 6 show results of competitive level set segmentation. Since regions of problems are incorporated as part of the modelling, the identification of

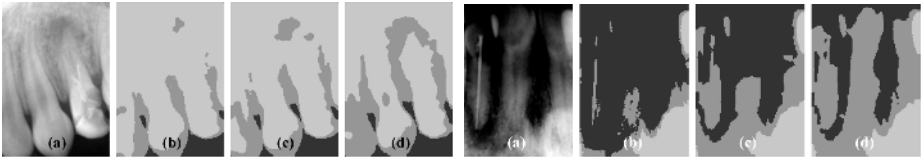


Fig. 4. Segmentation Results. (a) Original image. (b) Initial condition provided by image. (b) Iteration 0 provided by SVM. (c) SVM. (c) Iteration 20. (d) Iteration 60. **Fig. 5.** Segmentation Results. (a) Original image. (b) Iteration 40. (c) Iteration 80. (d) Iteration 80.

such areas would be an automatic product of the segmentation. The pathological meaningful segmentation can be used at the next stage of analysis. In additional, as shown in Fig. 4 and 5, although the SVM only gives a rough approximate segmentation, it is able to provide a good competitive initial contour for two level set functions, Φ_1 and Φ_2 , which accelerates the segmentation. Indeed, the competitive level set segmentation is robust to the placement of the initial contours so that even when the initial contour does not closely correspond to the final segmentation, the level set functions can still achieve an accurate segmentation as shown in Fig. 5.

3.2 Analysis

Based on the uncertainty map and segmentation results, the analysis is able to indicated some possible areas in the image as shown in Fig. 6(c) in which those bone loss areas are emphasized by the color channel scheme. The scheme provides more direct visual cues which will greatly reduce the possibility that those areas, the area pointed to by an arrow in Fig. 6(c) for example, might be overlooked.

Figs. 6, 7 and 8 show results of teeth root decay detection. In these examples, the segmentation and uncertainty map successfully locate the area of bone loss and, after being given the orientation of the image, the area of root decay can be automatically detected. Fig. 6 shows an interesting result of automatic root decay detection. In this example, the system automatically detected three root

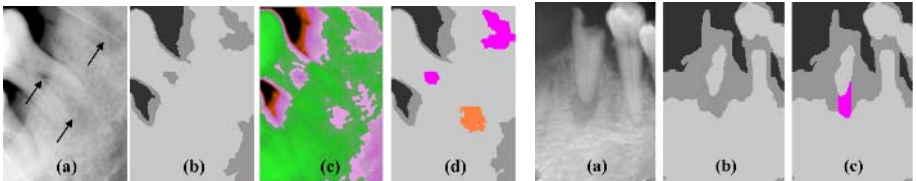


Fig. 6. Segmentation Results. (a) Original image. (b) Segmentation results. (c) Bone loss area marked with color channel method. (d) Root decay detected. **Fig. 7.** Root Decay Detection Results. (a) Segmentation Results. (b) Root decay detected. (c) Root decay detected.

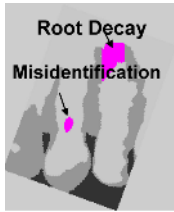


Fig. 8. Root decay detection of Fig. 6

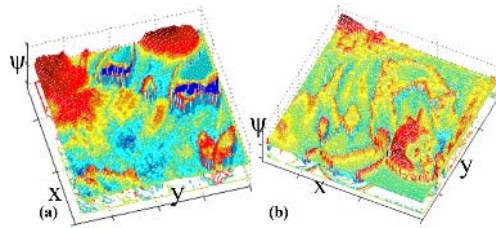


Fig. 9. Uncertainty maps of (a) Fig. 4 and (b) Fig. 6

problems with the orientation given. Two of them are warning level (pink area) and one is attention level (orange area) as shown in Fig. 6(d). Fig. 9 shows two of resulting uncertainty maps. The map is able to give a general idea for the problem areas. And more importantly, it provides a digitized uncertainty measurement. Altogether 50 dental X-rays with different levels of bone loss and root decay are used to test the proposed framework and the results validated by dentist. The experimental results show that the proposed framework is able to help to find all the areas of bone loss. For the root decay detection, the proposed framework is able to find all of them. But due to complexity of the dental X-rays, there are still 4 misidentifications of root decay. One of these misidentifications are indicated by arrows in Fig. 8.

4 Summary and Conclusions

Leveraging the transition towards a totally electronic format for dental X-rays, a framework to semi-automatically detect areas of bone loss and root decay in these images is proposed. The system is designed in particular for the dental clinical environment through the use of a classifier to ensure accurate segmentation of these images. Utilizing these segmentations and a computed uncertainty map, regions of pathological abnormality are emphasized for the dentist's attention. Experimental results indicate that the system correctly identifies such problem areas with a few misidentifications. To the best of our knowledge, this is the first work on semi-automatic dental X-ray computer aided diagnosis.

References

1. L. Vese and T. Chan, "A multiphase level set framework for image segmentation using the mumford and shah model," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271–293, 2002.
2. R. Kimmel and A. M. Bruckstein, "Regularized Laplacian zero crossings as optimal edge integrators," *International Journal of Computer Vision*, vol. 53, pp. 225–243, July 2003.
3. V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, vol. 22, pp. 61–79, 1997.

4. S. Li, T. Fevens, and A. Krzyżak, “Image segmentation adapted for clinical settings by combining pattern classification and level sets,” in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, (St-Malo, France), pp. 160–167, 2004.
5. S. Li, T. Fevens, and A. Krzyżak, “An SVM based framework for autonomous volumetric medical image segmentation using hierarchical and coupled level sets,” in *Computer Aided Radiology and Surgery (CARS)*, (Chicago, USA), pp. 207–212, 2004.
6. T. Chan and L. Vese, “Active contour model without edges,” *IEEE Trans. on Image Processing*, vol. 24, pp. 266–277, 2001.
7. A. K. Jain and H. Chen, “Matching of dental x-ray images for human identification,” *Pattern Recognition*, vol. 37, no. 7, pp. 1519–1532, 2004.

Exploiting Temporal Information in Functional Magnetic Resonance Imaging Brain Data*

Lei Zhang^{1,2}, Dimitris Samaras¹, Dardo Tomasi², Nelly Alia-Klein²,
Lisa Cottone², Andreana Leskovjan², Nora Volkow², and Rita Goldstein²

¹ Department of Computer Science, SUNY at Stony Brook, NY, USA

² Medical Department, Brookhaven National Laboratory, NY, USA

Abstract. Functional Magnetic Resonance Imaging (fMRI) has enabled scientists to look into the active human brain, leading to a flood of new data, thus encouraging the development of new data analysis methods. In this paper, we contribute a comprehensive framework for spatial and temporal exploration of fMRI data, and apply it to a challenging case study: separating drug addicted subjects from healthy non-drug-using controls. To our knowledge, this is the first time that learning on fMRI data is performed explicitly on temporal information for classification in such applications. Experimental results demonstrate that, by selecting discriminative features, group classification can be successfully performed on our case study although training data are exceptionally high dimensional, sparse and noisy fMRI sequences. The classification performance can be significantly improved by incorporating temporal information into machine learning. Both statistical and neuroscientific validation of the method's generalization ability are provided. We demonstrate that incorporation of computer science principles into functional neuroimaging clinical studies, facilitates deduction about the behavioral probes from the brain activation data, thus providing a valid tool that incorporates objective brain imaging data into clinical classification of psychopathologies and identification of genetic vulnerabilities.

1 Introduction

Functional Magnetic Resonance Imaging (fMRI) has enabled scientists to look into the active human brain by providing sequences of 3D brain images. This has revealed exciting insights into the spatial and temporal changes underlying a broad range of brain functions, including basic functions such as how we see, feel,

* We thank Steve Berry, B.A., for help with preliminary data analyses, F. Telang, E.C. Caparelli, L. Chang, T. Ernst and N.K. Squires for helpful discussions; This study was supported by grants from the National Institute on Drug Abuse (to NDV: DA06891-06; and to RZG: 1K23 DA15517-01), Laboratory Directed Research and Development from U.S. Department of Energy (OBER), NARSAD Young Investigator Award, SB/BNL seed grant (79/1025459), National Institute on Alcohol Abuse and Alcoholism (AA/ODO9481-04), ONDCP, and General Clinical Research Center (5-MO1-RR-10710).

move, understand. Concomitantly, this new instrumentation has led to a flood of new data: a twenty-minute fMRI session with a single human subject produces a series of 3D brain images each containing approximately 150,000 voxels, collected once a second (or two), yielding tens of millions of data observations. Thus, developing appropriate data analysis methods is needed for truly comprehensive exploration of this ample volume of data. We suggest that through incorporation of computer data analysis principles into functional neuroimaging studies we will be able to identify unique patterns of variability in brain states and deduce about the behavioral probes from the brain activation data (in contrast to the reverse: deducing about brain activation data from behavioral probes). We further propose that this interscientific incorporation may provide a valid tool where objective brain imaging data are used for clinical purpose of classification of psychopathologies and identification of genetic vulnerabilities.

Functional Magnetic Resonance Imaging (fMRI) [1][2] is based on the increase in blood flow to the local vasculature that accompanies neural activity in the brain, so that human cortical functions can be observed without the use of exogenous contrast agents. To date, the analyses and interpretation of fMRI data that are most commonly employed by neuroscientists depend on the cognitive-behavioral probes that are developed to tap regional brain function. Thus, brain responses are a-priori labeled based on the putative underlying task condition (e.g., regions involved in reward vs. regions involved in punishment) and are then used to separate a priori defined groups of subjects. A variety of machine learning methods have also been used for exploratory analysis of fMRI data[3][4][5]. In recent research[6][7][8][9], machine learning methods have been applied for fMRI data analysis but only statistical maps [8][9] or the mean of each fMRI time interval[7] are used while temporal information has yet to be fully employed. Discarding temporal information results in more manageable data sizes; however, the cost of such information loss is still unclear.

In this paper, we consider a different classification problem: separating different groups of human subjects based on the observed fMRI time sequences. We contribute a comprehensive framework of spatially and temporally exploring fMRI data, and apply it to a challenging case study: separating drug addicted subjects from healthy non-drug-using controls based on their observed fMRI time sequences. This learning problem is challenging for a number of reasons: 1) oversized dimensionality of the fMRI BOLD sequences; 2) undersized data space; 3) increased inter-subject variability and intra-subject variability: even for the same person, activations are different from trial to trial due to brain-behavior complexity; and 4) decreased between group experimental heterogeneity. Fig. 1 shows time sequences of one voxel in three human brains.

To our knowledge, this is the first time that machine learning is performed on the temporal information of fMRI data for classification purposes. In this work, we aim to answer the following questions: 1) given the difficulties inherent in this type of the data, what are the most discriminative features for this classification problem? 2) Will temporal/function information help us with classification? We explore fMRI data in two different representations: 3D brain

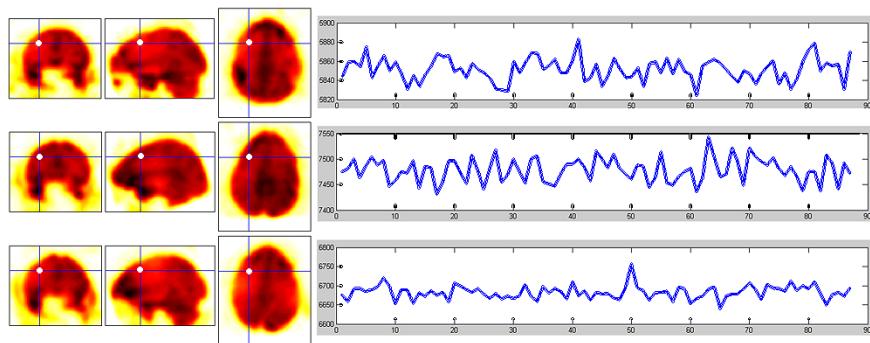


Fig. 1. In each row, the left three images show three slides of the a 3D fMRI scan in different views with a white point representing the location of one voxel, followed by a time sequence of this voxel across the 87 task-sequence time points

“video” and a set of voxel-specific time series. Dimensionality reduction and feature selection techniques are proposed and examined in each representation. Experimental results demonstrate that group classification is improved by selecting discriminative features and incorporating fMRI temporal information into a machine learning framework. The automatically selected features prove reliable, i.e., stable in cross-task validation. Furthermore, these blindly selected features prove valid, i.e., they include voxels in regions previously implicated in the successful performance of the core cognitive-behavioral task. Thus, our approach is generalizable, as tested statistically and integrated in a wider neuroscience context. For example, the anterior cingulate gyrus has been assumed to be essential for error detection, conflict resolution, and in emotional regulation. The emergence of anterior cingulate voxels, blindly selected through our analyses, confirms its their central role in underlying the core characteristics of addiction and its relevance to performing this specific sustained attention task.

Drug addiction is a complex disorder characterized by compromised inhibitory control. Individuals with compromised mechanisms of control are difficult to identify unless they are directly subjected to challenging conditions. Solving this problem is essential because patterns of variability in brain states may be unique to a certain psychopathology and could be therefore used for improving diagnosis and prevention efforts (e.g. diagnosis of drug addiction, prevention of relapse or craving). In addition, the development of this “*clinical machine learning framework*” can be applied to further our understanding of other human disorders and states such as those impacting insight and awareness, that similarly to drug addiction are currently identified based mostly on subjective criteria and self-report.

2 Methodology

Acquisition of fMRI data: In our experiments, data were collected to study the neuropsychological problem of loss of sensitivity to the relative value of

money in cocaine users[10]. MRI studies were performed on a 4T Varian scanner and all stimuli were presented using LCD-goggles connected to a PC. The human participants pressed a button or refrained from pressing based on a picture shown to them. They received a monetary reward if they performed correctly. Specifically, three runs were repeated twice (T1, T2, T3; and T1R, T2R, T3R) and in each run, there were three monetary conditions (high money, low money, no money) and a baseline condition where a fixation cross was shown on the screen; the order of monetary conditions was pseudo-randomized and identical for all participants. Participants were informed about the monetary condition by a 3-sec instruction slide, which visually presented the stimuli: or \$0.45, \$0.01 or \$0.00. The feedback for correct responses in each condition consisted of the respective numeral designating the amount of money the subject has earned if correct or the symbol (X) otherwise. To simulate real-life motivational salience, subjects could gain up to \$50 depending on their performance on this task. 16 cocaine dependent individuals, 18-55 years of age, in good health, were matched with 12 non-drug-using controls on sex, race, education and general intellectual functioning.

In this work, we use Statistical Parametric Mapping (SPM)[11] to preprocess (realignment, normalization/registration and smoothing) the fMRI sequences.

Methodology: In this paper, we aim to separate the drug-addicted subjects from controls by applying machine learning methods to observed fMRI sequences. Following [7], the classification problem can be represented: $f : \langle fMRI - sequence(t1, t2) \rangle \rightarrow [DrugAddicted|Control]$ where fMRI-sequence($t1, t2$) is the sequence of fMRI images collected during the contiguous time interval $[t1, t2]$. The input data is an extremely high dimensional feature vector, consisting of hundreds of thousands of features (87 scans per sequence and $53 \times 63 \times 46$ voxels per scan). In order to evaluate the discriminative power of the extracted features, we first perform classification using a simple Euclidean method that uses the Euclidean distance on the original fMRI sequences. This is computationally expensive due to the high dimensionality and yields inferior results due to the large number of non-discriminative features. Feature selection is essential to achieve accurate classification when only a small number of data are available [12], hence we explore a variety of approaches for dimensionality reduction and feature selection. We group our approaches into two categories based on the different views of the fMRI sequences:

2.1 fMRI: 3D Brain “Video”

The 4D input fMRI data ($53 \times 63 \times 46 \times 87$) can be treated as sequences of 3D images. Previous work[13] demonstrated that, by traversing the 3D space of fMRI images using the Hilbert space-filling curve[14], a 3D image can be linearly mapped into 1D space(153594×1). A space filling curve defines a continuous path in a multidimensional grid, visiting each point exactly once and never crossing itself. Hilbert space-filling curve has been proven optimal in preserving the locality and clustering properties of data[15]. The Hilbert space-filling proceeds

in a recursive manner, following the same rotation and reflection pattern at each vertex of the basic curve (details in [16]). By using this technique, the original input fMRI sequences can be mapped into 2D spaces (153594×87) where each 3D fMRI scan is linearly mapped into 1D space. Hence, two groups of dimensionality reduction and feature selection techniques applied in the signal analysis domain can be applied: those based on global information, such as Singular Value Decomposition (SVD)[17]; and those based on local information, such as the Discrete Fourier Transform (DFT)[18]. Due to the dimensionality of the input data, we chose to use local dimensionality reduction techniques.

Dimensionality Reduction and Feature Selection: After mapping each of the 3D scans into 1D space, the input fMRI sequences are transformed into 2D matrices with one spatial dimension and one temporal dimension.

We applied two sets of signal analysis approaches for dimensionality reduction:

1. *1D Transformation:* by performing 1D DFT or DWT on the spatial dimension, the original input data can be represented by a small number of time series.
2. *2D Transformation:* by performing 2D DFT or DWT on the 2D matrices directly, the original input data can be represented by a small 2D feature matrix.

After dimensionality reduction, we can perform classification either by: i) employing the whole reduced feature space for classification or, ii) learning the most discriminative features for classification. As expected, experimental results in Sec. 3 demonstrate the need for learning discriminative features.

2.2 fMRI: A Set of Voxel-Specific Time Series

In Sec. 2.1, we treat each fMRI sequence as a time series of 3D images. Alternatively, each fMRI sequence can also be thought of as a combination of N time series where N is the number of the voxels. In this section, we examine voxel-based feature selection methods for group classification. The most common approach for feature selection[19] is to select those features that best discriminate the target classes: given the goal of learning a target classification function, one common approach to feature selection is to rank the features by their mutual information with respect to the class variable, then to select the n highest scoring features. In our experiments, the information of each voxel is represented by a time sequence of intensities representing BOLD brain activations. Distance (dissimilarity) between voxel intensities is straightforward to define. However, there are different ways to measure this distance when the features are voxel time series. We examine various time series analysis techniques. More specifically, given training data Q and C , for the i th voxel, we propose to compute and examine the distance between two time series $q_{1..T}$ and $c_{1..T}$ by:

1. *Raw_Dist:* Euclidean Distance Metric directly on the two time series of the voxel: $D(q, c) = \sum_{t=1}^T (q_t - c_t)^2$;
2. *Norm_Dist:* Euclidean Distance Metric on normalized time series, where $X' = (X - \bar{X})/\sigma(X)$ with \bar{X} and $\sigma(X)$ the mean and standard deviation;

- 3. *Mean_Dist*: Distance of the mean values of the two series: $D(q, c) = \|\bar{q} - \bar{c}\|$;
- 4. *Var_Dist*: Distance of the variances of the two series: $D(q, c) = \|\sigma^2(q) - \sigma^2(c)\|$;

In this paper, we choose to use a simple kth-Nearest Neighbor(KNN)[19] retrieval scheme to evaluate classification using different sets of selected features.

3 Experiments and Results

In our data collection, there are totally 6 runs: T1, T2, T3, T1R, T2R and T3R with the three latter repeating the three former, grouped into 3 data sets *T1, T2, T3*. The first set of experiments is to evaluate classification performance in a "leave-one-out" cross validation procedure. Each of the *K* human subjects was used as a test subject while training on the remaining *K* - 1 subjects, and the mean classification rate of the *K* experiments is reported.

For comparison purposes, in Table 1, we report the classification performance of the plain Euclidean method(Sec.2) which uses the computationally expensive Euclidean distance on the original fMRI sequences and yields classification results that are close to random. Table 1 also presents the classification performance with the first 30 Fourier coefficients by applying DFT to the mapped matrices using Hilbert space-filling curve. The results of DWT(double the number of features) are similar to the results reported in Table 1. After dimensionality reduction, most of the features are still non-discriminative and a feature selection step can significantly improve classification although standard global dimensionality reduction techniques such as SVD, which also perform feature selection, cannot be applied.

Table 2 reports classification rates for voxel-specific time series analysis in each data set individually. It also reports cross validation using the features selected from voxel-based methods. Contrast map based classification rates in [9] are listed in the last row for comparison since the creation of contrast maps can be thought of a dimensionality reduction process using statistical inferences along

Table 1. 3D Brain "Video" Analysis: classification rates of the simple Euclidean method on the original fMRI sequences are close to random due to the large number of non-discriminative features. Classification using features computed from signal-based methods (DFT). The "ALL" column shows classification rates using all the first 30 Fourier coefficients and the "Discrim" column reports rates using 5 selected discriminative coefficients only. Experimental results demonstrate that in the locally reduced feature space, appropriate features can significantly improve classification.

	Euclidean Distance		1D DFT		2D DFT	
	1NN	5NN	ALL	Discrim	ALL	Discrim
T1	51.92%	53.84%	65.38%	84.62%	63.46%	84.62%
T2	51.92%	51.92%	61.54%	80.77%	63.46%	82.69%
T3	50%	52.08%	60.42%	72.92%	62.50%	68.75%

Table 2. Voxel-Specific Time Series Analysis: classification rates using features selected from voxel-based methods. Experimental results in each individual data set demonstrate that classification performance can be significantly improved by incorporating temporal information into learning. Classification rates of using Contrast maps *Contrasts* are the best reported in [9] that use the kNN classifier. Experimental results across data sets demonstrate that the features selected using temporal information (*Raw_Dist*) stably perform better across data sets.

Training Testing	T1			T2			T3		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
<i>Raw_Dist</i>	94.2%	90.4%	87.5%	92.3%	96.2%	91.7%	86.5%	88.5%	93.8%
<i>Norm_Dist</i>	84.6%	75.0%	70.8%	78.9%	90.4%	72.9%	71.2%	75.0%	85.4%
<i>Mean_Dist</i>	88.5%	76.9%	70.8%	80.8%	90.4%	75.0%	75.0%	76.9%	85.4%
<i>Var_Dist</i>	86.5%	82.7%	72.9%	80.8%	86.5%	75.0%	78.9%	78.9%	83.3%
<i>Contrasts</i>	88.5%	N/A	N/A	N/A	86.8%	N/A	N/A	N/A	85.7%

the temporal axis. Experimental results on each individual data set demonstrate that classification performance can be significantly improved by incorporating temporal information into learning. Classification rates of *Raw_Dist* outperformed other methods. In time series analysis, temporal distortion is a common problem. Dynamic Time Warping (DTW) has been successfully applied in many time series analysis problems [20] but is inapplicable here due to dimensionality. Luckily the nature of the data prevents significant temporal distortion (since each 87 point sequence can be divided into 6 short sub-sequences). For validation purposes, we performed a set of DTW experiments on 5 selected features. Experimental results demonstrate that DTW gives similar classification results as the Euclidean distance method.

In the three data sets, subjects are performing exactly the same task whereas the sequence of the monetary conditions is different. We address the generalization question by examining classification performance when selected features from one data set are applied onto other data sets. Table 2 reports the cross validation results. The features that are automatically selected using temporal information (*Raw_Dist*), prove very stable in cross-task validation. Hence, by selecting the most discriminative features using temporal information, group classification can be successfully performed.

We also examined the selected features under a neuro-scientific context: the selected most discriminative voxels cluster into two prefrontal brain regions: the middle frontal gyrus (dorsolateral prefrontal cortex) and the anterior cingulate gyrus (ACG). Those regions are known to be involved in sustained attention/working memory and in the processing of salient stimuli/inhibitory control, respectively. It is very intriguing that these blindly selected voxels represent the two core functions of the delayed forced-choice task used in this study. Further, their location in the prefrontal cortex lends support to our working hypothesis that this region is crucially involved in the underlying core characteristics of drug addiction. Specifically, the involvement of the ACG, a region which has been previously implicated in drug intoxication and craving [21], lends support

to the dysfunction in drug addiction of this corticolimbic region (see [22] where a hypofunctionality of the ACG to a GO/NO-GO task was reported in cocaine users compared to controls).

4 Conclusions and Future Work

We have demonstrated that, by selecting discriminative features, group classification can be successfully performed on a challenging case study although training data are fMRI sequences that are exceptionally high dimensional, sparse and noisy. We have also shown that classification rates can be significantly improved by incorporating temporal information into machine learning analysis of such data. To our knowledge, this is the first time that the temporal/functional information of the fMRI data is explicitly explored for machine learning classification purposes. This comprehensive framework of exploring spatial and temporal information of fMRI data for classification problems can be extended to many other fMRI analysis applications. Our analyses provide an additional method for validation of a regional-functionality specificity, however external validation using a lesioned sample would still be necessary to confirm a particular region's role in a specific function (i.e., loss of a certain function in individuals lacking a specific region).

Since feature selection is the key for pattern recognition problems, especially when only a small number of data are available, as in most human subject research, one of our future research directions is to explore efficient global dimensionality reduction techniques[23] that can be applied on extremely high dimensional training data and examine more sophisticated classifiers. Another future research direction is to apply grouping/clustering as a preprocessing step to reduce dimensionality of the raw data and smooth noise. Finally, connectivity and interactivity information play important roles in brain activation patterns. For example, in the voxel-based feature selection part, voxels are considered to be independent, however, this is not accurate for the human brain. After further validation with other data sets (additional subjects with addiction or other psychopathology), we aim to explore the connectivity and interactivity between voxels to reveal more discriminative brain activation patterns.

References

1. Kwonget et al., K.: Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc Natl Acad Sci USA* (1992) 5675–5679
2. Bandettini, P., Jesmanowicz, A., Wong, E., Hyde, J.: Processing strategies for time-course data sets in functional mri of the human brain. (*Magn Reson Med*)
3. Goutte, C., e.a.: On clustering fmri time series. (Technical Report IMM-REP-1998-11)
4. Penny, W.: Mixture models with adaptive spatial priors. In: *Concepts and Methods in NeuroImaging workshop, NIPS*. (2001)
5. LaConte, S., Strother, S., Cherkassky, V., Hu, X.: Predicting motor tasks in fmri data using support vector machines. In: *ISMRM*. (2003)

6. Wang, X., Hutchinson, R., Mitchell, T.: Training fmri classifiers to detect cognitive states across multiple human subjects. In: NIPS03. (2003)
7. Mitchell, T., Hutchinson, R., Niculescu, R., Pereira, F., Wang, X., Just, M., Newman, S.: Learning to decode cognitive states from brain images. (Machine Learning)
8. Ford, J., Farid, H., Makedon, F., Flashman, L., McAllister, T., Megalooikonomou, V., Saykin, A.: Patient classification of fmri activation maps. In: MICCAI. (2003)
9. Zhang, L., Samaras, D., Tomasi, D., Volkow, N., Goldstein, R.: Machine learning for clinical diagnosis from functional magnetic resonance imaging. In: CVPR. (2005)
10. Goldstein et al., R.: A modified role for the orbitofrontal cortex in attribution of salience to monetary reward in cocaine addiction: an fmri study at 4t. In: Human Brain Mapping Conference. (2004)
11. Friston, K., Holmes, A., Worsley, K., et al.: Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping* (1995) 2:189–210
12. Levi, K., Weiss, Y.: Learning object detection from a small number of examples: The importance of good features. In: CVPR. (2004)
13. Wang, Q., Kontos, D., Li, G., Megalooikonomou, V.: Application of time series techniques to data mining and analysis of spatial patterns in 3d images. In: ICASSP. (2004) 525–528
14. Hilbert, D.: Uber die stetige abbildung einer linie auf flachenstuck. *Math. Annln.* (1891) 459–460
15. Moon, B., Jagadish, H., Faloutsos, C., Saltz, J.: Analysis of the clustering properties of the hilbert space-filling curve. *IEEE Transactions on Knowledge and Data Engineering* (2001) 124–141
16. J.G.Griffiths: An algorithm for displaying a class of space-filling curves. *Software Practice and Experience* **16** (1986) 403–411
17. Keogh, E., Chakrabarti, K., Pazzani, M., Mehrotra, S.: Dimensionality reduction for fast similarity search in large time series databases. *Journal of Knowledge and Information Systems* (2000)
18. Agrawal, R., Faloutsos, C., Swami, A.: Efficient similarity search in sequence databases. In: Proc. of the 4thConference on Foundations of Data Organization and Algorithms. (1993)
19. Mitchell, T.: *Machine Learning*. McGraw-Hill (1997)
20. Keogh, E.: Data mining and machine learning in time series databases. In: Tutorial in ICML. (2004)
21. Goldstein, R., Volkow, N.: Drug addiction and its underlying neurobiological basis: Neuroimaging evidence for the involvement of the frontal cortex. (*American Journal of Psychiatry*) 1642–1652
22. J.N.Kaufman, T.J.Ross, E.A.Stein, H.Garavan: Cingulate hypoactivity in cocaine users during a go-nogo task as revealed by event-related functional magnetic resonance imaging. *The Journal of Neuroscience* (2003)
23. Ye, J.: Generalized low rank approximations of matrices. In: ICML. (2004) 887–894

Model-Based Analysis of Local Shape for Lesion Detection in CT Scans

Paulo R.S. Mendonça¹, Rahul Bhotika¹, Saad A. Sirohey², Wesley D. Turner¹,
James V. Miller¹, and Ricardo S. Avila³

¹ GE Global Research, One Research Circle, Niskayuna, NY 12309, USA
{mendonca, bhotika, turner, millerjv}@research.ge.com

² GE Healthcare, 3200 N Grandview Blvd, Waukesha, WI, 53188, USA
saad.sirohey@ge.com

³ Kitware Inc., 28 Corporate Drive, Suite 204, Clifton Park, NY 12065, USA
rick.avila@kitware.com

Abstract. Thin-slice computer tomography provides high-resolution images that facilitate the diagnosis of early-stage lung cancer. However, the sheer size of the CT volumes introduces variability in radiological readings, driving the need for automated detection systems. The main contribution of this paper is a technique for combining geometric and intensity models with the analysis of local curvature for detecting pulmonary lesions in CT. The local shape at each voxel is represented via the principal curvatures of its associated isosurface without explicitly extracting the isosurface. The comparison of these curvatures to values derived from analytical shape models is then used to label the voxel as belonging to particular anatomical structures, e.g., nodules or vessels. The algorithm was evaluated on 242 CT exams with expert-determined ground truth. The performance of the algorithm is quantified by free-response receiver-operator characteristic curves, as well as by its potential for improvement in radiologist sensitivity.

1 Introduction

State-of-the-art computer tomography (CT) scanners, with spatial resolutions of less than a millimeter, routinely provide images of smaller and smaller nodules, characteristic of early-stage lung cancer [1]. However, these gains in spatial resolution have led to an explosion in the sizes of the image volumes that a radiologist has to review, resulting in significant variability in radiological readings [2]. Figure 1 shows examples of nodules detected by two expert radiologists but missed by a third expert.

Computer-aided detection (CAD) systems have been developed to aid radiologists in reading CT exams. Brown et al [3] use simple shape- and intensity-based features of segmented regions with a fuzzy classifier. The data consisted only of selected 2 cm thick cross-sections of the lung, favoring the 3D segmentation technique upon which the algorithm is dependent and avoiding the difficult apex region. McCulloch et al [4] obtained encouraging results with a Bayesian classifier — 70% sensitivity at 8 false positives per case operating on noisy low-dose screening data. Their method uses 2D segmentation to generate candidates and the number of false positives grows with the use of thinner slices in CT. Paik et al [5] use a geometric model and learn parameters from training data. They report results on only 8 CT volumes and on nodules with

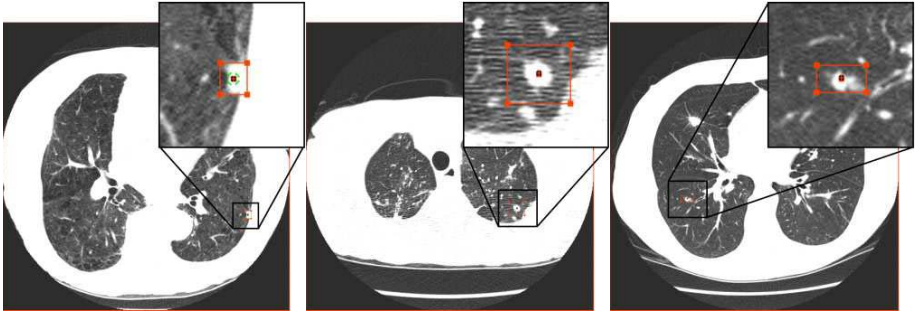


Fig. 1. Three examples of structures missed by an expert radiologist but independently classified as nodules by two other experts. These abnormalities were successfully detected by the algorithm proposed here.

diameters above 6 mm, thus no solid conclusions can be drawn. Farag et al [6] report numbers as high as 82.3% sensitivity with 9.2% false positive rate. However, the majority of their nodules are calcified and therefore clinically irrelevant [7] and over 80% of their ground truth consists of nodules above 12 mm which are easy to detect. Others have used local structure information alone to discriminate between particular shapes [8, 9, 10]. However, in all of these methods the thresholds for classification are either set empirically or learned from data.

This paper introduces a technique for the shape analysis of 3D images and demonstrates its application to the detection of lung nodules in CT exams. The method is based on a fully 3D algorithm that does not depend on segmenting or grouping the relevant anatomical structures or any edge detection method. The proposed algorithm combines geometric and intensity models with the eigenanalysis of the *curvature tensor* in order to identify pulmonary nodules in CT. All parameters and thresholds are derived from these models, eliminating the need for training data. Furthermore, no model fitting or optimization is performed. All operations are local and therefore the proposed method can be interpreted as a filter for highlighting nodule-like structures in the image.

2 Differential Operators on Volume Images

A *volume image* I is defined as a twice-differentiable (\mathcal{C}^2) mapping from a compact domain $\mathcal{V} \subset \mathbb{R}^3$ into \mathbb{R} . For any given k , the equation

$$I(\mathbf{x}) = k \quad (1)$$

defines an *isosurface* $M_k \subset \mathcal{V}$ at the points \mathbf{x} satisfying (1) and $\nabla I(\mathbf{x}) \neq \mathbf{0}$ [11]. A common descriptor for the local structure of an image I is its Hessian \mathbf{H} [8]. Although the eigenvalues of \mathbf{H} provide an intuitive measure of local structure, they do not capture the true underlying shape, which is more accurately described by the curvature of the isosurfaces defined by (1). For example, consider an isotropic Gaussian intensity profile (see Fig. 2). Hessian-based shape measures [8] would incorrectly signal the presence of a cylinder at the inflection points of the profile, while the principal curvatures would correctly flag the entire structure as spherical.

2.1 The Curvature Tensor

It can be shown that, at the point \mathbf{x} , the principal directions of the isosurface given by (1) can be obtained directly from the implicit function by solving the eigenproblem

$$\min_{\hat{\mathbf{v}}} / \max_{\hat{\mathbf{v}}} \frac{-\hat{\mathbf{v}}^T \mathbf{N}^T \mathbf{H} \mathbf{N} \hat{\mathbf{v}}}{\|\nabla I\|}, \tag{2}$$

subject to $\|\hat{\mathbf{v}}\| = 1$,

where \mathbf{N} is the 3×2 matrix of the null space of ∇I . The principal directions \mathbf{v}_1 and \mathbf{v}_2 are given by $\mathbf{v}_{1,2} = \mathbf{N} \hat{\mathbf{v}}_{1,2}$, where $\hat{\mathbf{v}}_{1,2}$ are the solutions of (2), and the corresponding principal curvatures κ_1 and κ_2 are the eigenvalues of the 2×2 matrix $-\mathbf{N}^T \mathbf{H} \mathbf{N} / \|\nabla I\|$, with $\kappa_1 \leq \kappa_2$. The matrix $\mathbf{C} = -\mathbf{N}^T \mathbf{H} \mathbf{N} / \|\nabla I\|$ is herein defined as the *curvature tensor* of the volume image.

Yoshida et al [9] present a method to compute κ_1 and κ_2 for an implicit surface. Their technique also estimates curvatures directly from an implicit function. However, it requires the isosurfaces to be amenable to a local parameterization via Monge patches, which cannot be achieved everywhere on the surface [12]. The solution from (2) circumvents this problem and also avoids the rotation step required by Vos et al [10].

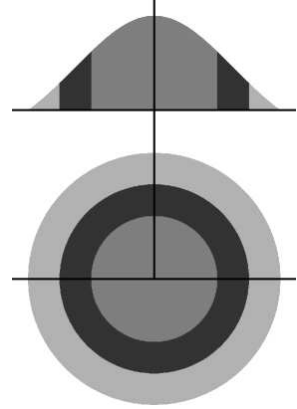


Fig. 2. Cross-section and top view of Hessian responses on a 2D Gaussian profile. The Hessian response is spherical in the innermost circle, cylindrical in the dark band containing the inflection point, and void in the outer band.

2.2 Response of the Curvature Tensor

At points \mathbf{x} for which the isosurface defined by (1) is well-approximated by a spherical patch with surface normal pointing outward, κ_1 and κ_2 satisfy $\kappa_1 > 0$ and $\kappa_2 > 0$. For cylindrical patches with an outward normal, $\kappa_1 \approx 0$ and $\kappa_2 > 0$. Finally, saddle points have $\kappa_1 < 0$ and $\kappa_2 > 0$. This suggests a methodology to discriminate between spherical and cylindrical regions in volume images, such as nodules and vessels in CT scans of the lung. The *sphericalness* and *cylindricalness measures* $b(\mathbf{x})$ and $c(\mathbf{x})$ for the curvature tensor at the isosurface intersecting the point \mathbf{x} are defined as

$$b(\mathbf{x}) = \begin{cases} \frac{\kappa_1}{\kappa_2} & \text{if } \kappa_1, \kappa_2 > 0, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad c(\mathbf{x}) = \begin{cases} 1 - \frac{\kappa_1}{\kappa_2} & \text{if } \kappa_2 > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

3 Anatomical Modeling

The eigenvalues of the curvature tensor can be used to quantify the sphericalness or cylindricalness of a voxel in a CT volume by thresholding the values attained by (3) at that voxel. An important contribution of this work is the use of a *model-based* approach to set these thresholds from generative models for the relevant anatomical structures.

3.1 Anatomical Modeling of Local Shape

The eigenvalues κ_1 and κ_2 of the curvature tensor are independent of any transformation $I'(\mathbf{x}) = aI(\mathbf{x}) + b$ applied to the image volume $I(\mathbf{x})$. In fact, κ_1 and κ_2 at a voxel \mathbf{x} depend only on the shape of the associated isosurface, not on its isovalue. Therefore one only needs to take into account the local shapes of the relevant anatomical structures, not the exact form of their intensity profiles.

Radiologists routinely model pulmonary nodules as spherical objects when attempting to measure quantities such as volume and growth rate [13]. In this work, the local shapes of isosurfaces of relevant 3D anatomical structures — nodules and vessels — are approximated using ellipsoidal and toroidal surface patches.

Ellipsoidal model. Under the appropriate curvature-preserving mapping the implicit equation of an ellipsoid \mathcal{E} is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \tag{4}$$

where $a \leq b \leq c$ are the lengths of the ellipsoid's semi-axes, as shown in Fig. 3(a). The principal curvatures $\kappa_1^\mathcal{E}$ and $\kappa_2^\mathcal{E}$ take on minimum values

$$\kappa_{1,\min}^\mathcal{E} = \frac{a}{c^2} \text{ and } \kappa_{2,\min}^\mathcal{E} = \frac{a}{b^2}. \tag{5}$$

Furthermore, the minimum and maximum values of the ratio $\kappa_r^\mathcal{E} = \kappa_1^\mathcal{E}/\kappa_2^\mathcal{E}$, a component of the measures defined in (3), are given by

$$\kappa_{r,\min}^\mathcal{E} = \frac{a^2}{c^2}, \kappa_{r,\max}^\mathcal{E} = \max\left(\frac{a^2}{b^2}, \frac{b^2}{c^2}\right). \tag{6}$$

Toroidal model. A torus \mathcal{T} with small radius r and large radius R , as shown in Fig. 3(b), can be parameterized as

$$\mathcal{T} : \begin{cases} x = (R + r \cos \psi) \cos \theta \\ y = (R + r \cos \psi) \sin \theta \\ z = r \sin \psi \end{cases} \tag{7}$$

with $(\theta, \psi) \in (-\pi, \pi]^2$. Its principal curvatures $\kappa_1^\mathcal{T}$ and $\kappa_2^\mathcal{T}$ take on values

$$\kappa_1^\mathcal{T} = \frac{\cos \psi}{R + r \cos \psi} \text{ and } \kappa_2^\mathcal{T} = \frac{1}{r}, \tag{8}$$

and the minimum and maximum values of the ratio $\kappa_r^\mathcal{T} = \kappa_1^\mathcal{T}/\kappa_2^\mathcal{T}$ can be directly computed from (8) by varying ψ .

The analysis for the bounds in κ_1 , κ_2 , and κ_r presented in this section can be directly used to set an acceptable range for κ_1 and κ_2 and thresholds for (3). In the case of the ellipsoidal model, the range of values for a , b , and c can be set based on the range of

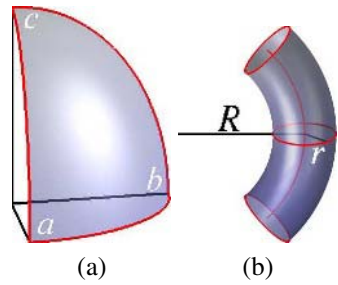


Fig. 3. The local shape of the isosurfaces of nodules is approximated by ellipsoidal patches (a) and that of vessels is approximated by toroidal patches (b)

targeted nodule sizes. Given a , b , and c , we can set thresholds for $\kappa_{1,\min}^{\mathcal{E}}$, $\kappa_{2,\min}^{\mathcal{E}}$, and $\kappa_{r,\min}^{\mathcal{E}}$. The lower bound on the aspect ratio for the ellipsoidal model was set at $1/2$, by assuming that nodules are roughly round. As for vessels, one typically has $R \gg r$, and thus $\kappa_{1,\min}^{\mathcal{T}} \approx 0$. However, the surface of vessel junctions can be modeled as highly bending tori, which would result in $\kappa_{1,\min}^{\mathcal{T}} \ll 0$.

4 Application to Lung Nodule Detection

The algorithm was initially implemented using the open source Insight Toolkit (ITK) [14], which provides a general framework for processing n -dimensional medical images and includes basic libraries for image processing, segmentation and registration.

The volume image was first smoothed by convolution with a Gaussian kernel to reduce the effect of noise. Then the lung volume was automatically extracted via region growing to provide a region of interest (ROI) for subsequent operations. The eigenvalues of the curvature tensor were computed at every voxel in the ROI, and thresholds derived from the anatomical models were used to label each voxel as spherical, cylindrical, or neither.

A major source of false positives for the Hessian are vessel junctions, since they display large second derivatives in all directions. Junction isosurfaces, on the other hand, are clearly non-spherical, as seen in Fig. 4. However, they have sharp peaks that cause the curvature tensor to generate a thin streak of spurious spherical responses along the centerline of the vessels. These spurious responses can be easily removed by morphological closing of the surrounding cylindrical responses.

The responses defined at individual voxels are sufficient for presentation to radiologists as a tool for highlighting spherical regions, e. g., by overlaying the responses on the original CT volume, as shown in Fig. 5. To quantify the algorithm's performance neighboring voxel responses were grouped and counted as a single detection. Experimental results and validation are presented in the next section.

The algorithm was re-implemented with several optimizations, including specialized data structures for memory access and morphological operations. On a 500-slice

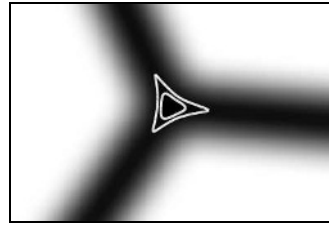


Fig. 4. Isosurfaces around the center of a junction display sharp peaks, with strong spherical responses along the centerline of the surrounding vessels



Fig. 5. A maximum intensity projection image showing spherical responses overlaid on the original CT data

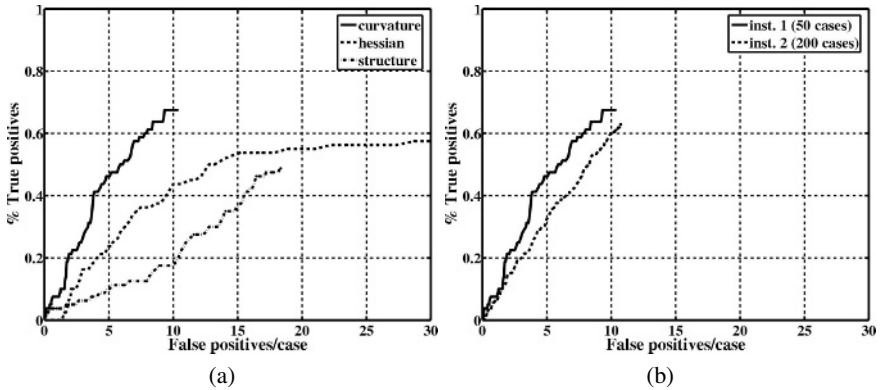


Fig. 6. (a) FROC curves showing the performance of the proposed algorithm compared to two other commonly-used methods. (b) FROC curves showing the algorithm's performance on datasets from two different institutions with different acquisition protocols.

dataset with slice thickness of 0.625 mm, the optimized version ran in 10 seconds on off-the-shelf hardware with dual 3.06 GHz Xeon processors and 2 GB of memory. The algorithm works with physical coordinates, using the pixel and slice spacing information contained in the CT images. For example, the standard deviation σ of the smoothing kernel and the radii for the morphological structuring elements were specified in millimeters and selected based on the anatomical models and clinical protocol. The smoothing kernel has $\sigma = 2$ mm, targeted at detecting nodules of diameter 4 mm.

5 Experimental Results

The algorithm was validated on datasets from two different institutions, totaling 242 exams. The dataset from institution 1 consisted of 50 low-dose CT exams of high-risk subjects acquired at 40 mAs and 120 kVp, with slice thickness of 1.25 mm. The exams were independently read by three expert radiologists. Ground truth was defined as all nodules marked by two or more radiologists and consisted of 109 non-calcified solid nodules with diameter greater than 3 mm of which 80 had diameter greater than 4 mm. The detections of the algorithm were considered true positives (TP) if they overlapped with a ground truth nodule, and false positives (FP) otherwise. Free response receiver-operator characteristic (FROC) curves were produced by sweeping the detections generated by the algorithm ranked in decreasing order of average spherical response.

For the nodules 4 mm and above in diameter, the algorithm achieved a sensitivity of 67.5% at 9.3 FP/case. These results were compared with those of two other well-known and oft-used local shape descriptors: the Hessian [8] and the structure tensor [15]. At the same rate of 9.3 FP/case, the Hessian and the structure tensor achieved sensitivities of only 40.0% and 17.5%, respectively. The FROC curves for all three techniques are shown in Fig. 6(a), demonstrating the superior performance of the proposed technique. In another experiment, nodules with diameter between 3 mm and 4 mm were also included, since 3 mm is the smallest diameter for which a clinical protocol has

been defined [7]. In this experiment, the proposed algorithm achieved a sensitivity of 70.6% at 25.6 FP/case compared to sensitivities of 46.8% for the Hessian and 47.5% for the structure tensor at the same FP rate (FROC curve not shown).

Table 1. The table shows, for each radiologist, the number of nodules missed, the radiologist sensitivity, the number of missed nodules detected by the algorithm, the algorithm’s false positive rate, the sensitivity improvement, the combined sensitivity of the radiologist and the algorithm, and the average sensitivity improvement weighted by the number of nodules on the ground truth for each radiologist. The difference in the algorithm’s false positive rate is due to the different ground truth for each radiologist.

	radiol.	missed nod.	initial sensit.	added detect.	FP/case	sensit. improv.	final sensit.
inst. 1	A	6	86.4%	3	10.0	6.8%	93.2%
	B	6	86.4%	4	10.0	9.1%	95.4%
	C	30	55.9%	19	9.5	27.9%	83.8%
inst. 2	D	18	85.4%	4	10.8	3.3%	88.6%
	E	19	84.7%	16	10.8	12.9%	97.6%
	F	68	60.7%	36	10.5	20.8%	81.5%
avg. improv.: 14.2%							

ologist’s reads. Using a leave-one-out scheme the ground truth for each radiologist was defined as only those nodules marked by the other two radiologists from the same institution. For institution 1, there were 38 nodules that were marked by all three radiologists. For radiologist A, 6 nodules were considered “missed,” i. e., marked only by B and C. Therefore, the ground truth defined by radiologists B and C consisted of 44 nodules (38 + 6), 38 of which were marked by A, yielding an initial sensitivity of 86.4%. The algorithm added three more detections, which, when combined with those of A, resulted in a final sensitivity of 93.2%, at a cost of 10 false positives per case. This procedure was repeated for radiologists B and C, and for the radiologists from institution 2. The corresponding sensitivity improvements are shown in Table 1. The overall average sensitivity improvement was 14.2%. Figure 1 shows examples of nodules detected by the algorithm but not marked by one radiologist.

6 Conclusions

This paper introduces a technique for detecting pulmonary nodules in CT volume images of the lung that is based on combining geometric and intensity models with the eigenanalysis of the curvature tensor. The method does not require either segmentation of the relevant anatomical structures or sophisticated classifiers and can be viewed as a filter that highlights specific anatomical shapes. Furthermore, the algorithm does not

The dataset from institution 2 consisted of 192 low-dose CT scans acquired using a different protocol (20 mAs, 140 kVp, with slice thickness of 2.5 mm). Ground truth was defined as before and consisted of 210 non-calcified solid nodules with diameter above 4 mm. The algorithm achieved a sensitivity of 62.9% at 10.3 FP/case. As seen in Fig. 6(b), there is only a minor degradation in the algorithm’s performance when compared to the results on the previous dataset, even though that was acquired at half the slice thickness.

To estimate the improvement in radiological sensitivity the algorithm’s detections were combined with each radiologist’s reads.

require any training data as values of all parameters and thresholds are derived analytically from the models. The method is robust to changes in scanning protocols, including slice thickness, as demonstrated by the algorithm's performance on multiple databases with radiologist-provided ground truth.

References

1. Mulshine, J.L., Smith, R.A.: Lung cancer 2: Screening and early diagnosis of lung cancer. *Thorax* **57** (2002) 1071–1078
2. McCulloch, C.C., Yankelevitz, D., Henschke, C., Patel, S., Kazerooni, E., Sirohey, S.: Reader variability and computer aided detection of suspicious lesions in low-dose CT lung screening exams. *Radiology* **226** (2003) 37A
3. Brown, M.S., Goldin, J.G., Suh, R.D., McNitt-Gray, M.F., Sayre, J.W., Aberle, D.R.: Lung micronodules: Automated method for detection at thin-section CT — initial experience. *Radiology* **226** (2003) 256–262
4. McCulloch, C.C., Kaucic, R.A., Mendonça, P.R.S., Walter, D.J., Avila, R.S.: Model-based detection of lung nodules in computed tomography exams. *Academic Radiology* **11** (2004) 258–266
5. Paik, D.S., Beaulieu, C.F., Rubin, G.D., Acar, B., Jeffrey, Jr., R.B., Yee, J., Dey, J., Napel, S.: Surface normal overlap: A computer-aided detection algorithm with application to colonic polyps and lung nodules in helical CT. *IEEE Trans. Medical Imaging* **23** (2004) 661–675
6. Farag, A.A., El-Baz, A., Gimel'farb, G.G., Falk, R., Hushek, S.G.: Automatic detection and recognition of lung abnormalities in helical CT images using deformable templates. In Barillot, C., Haynor, D.R., Hellier, P., eds.: *Medical Image Computing and Computer-Assisted Intervention*. Number 3217 in *Lecture Notes in Computer Science*, Saint-Malo, France (2004) 856–864
7. ELCAP: International early cancer action program — Protocol (2003) <http://icscreen.med.cornell.edu/ielcap.pdf>.
8. Sato, Y., Westin, C., Bhalerao, A., Nakajima, S., Shiraga, N., Tamura, S., Kikinis, R.: Tissue classification based on 3D local intensity structures for volume rendering. *IEEE Trans. Visualization and Computer Graphics*. **6** (2000) 160–180
9. Yoshida, H., Näppi, J.: Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps. *IEEE Trans. Medical Imaging* **20** (2001) 1261–1274
10. Vos, F.M., Serlie, I.W.O., van Gelder, R.E., Post, F.H., Truyen, R., Gerritsen, F.A., Stoker, J., Vossepoel, A.M.: A new visualization method for virtual colonoscopy. In Niessen, W.J., Viergever, M.A., eds.: *Medical Image Computing and Computer-Assisted Intervention*. Number 2208 in *Lecture Notes in Computer Science*, Berlin, Springer-Verlag (2001) 645–654
11. O'Neill, B.: *Elementary Differential Geometry*. Academic Press, New York (1966)
12. Forsyth, D.A.: Shape from texture and integrability. In: *Proc. 8th Int. Conf. on Computer Vision*. Volume II., Vancouver, Canada (2001) 447–452
13. Yankelevitz, D.F., Reeves, A.P., Kostis, W.J., Zhao, B., Henschke, C.I.: Small pulmonary nodules: Volumetrically determined growth rates based on CT evaluation. *Radiology* **217** (2000) 251–256
14. Ibáñez, L., Schroeder, W., Ng, L., Cates, J.: *The ITK Software Guide*. Kitware Inc. (2003)
15. Krissian, K., Malandain, G., Ayache, N., Vaillant, R., Troussset, Y.: Model based detection of tubular structures in 3D images. *Computer Vision and Image Understanding* **80** (2000) 130–171

Development of a Navigation-Based CAD System for Colon

Masahiro Oda^{1,*}, Takayuki Kitasaka¹, Yuichiro Hayashi¹, Kensaku Mori¹,
Yasuhito Suenaga¹, and Jun-ichiro Toriwaki²

¹ Graduate School of Information Science, Nagoya University,
Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8603, Japan
moda@suenaga.m.is.nagoya-u.ac.jp
{kitasaka, kensaku, suenaga}@is.nagoya-u.ac.jp

² School of Life System Science and Technology, Chukyo University,
101 Tokodachi, Kaizu-cho, Toyota, Aichi, 470-0393, Japan
jtoriwak@life.chukyo-u.ac.jp

Abstract. We propose a navigation-based computer aided diagnosis (CAD) system for the colon. When diagnosing the colon using virtual colonoscopy (VC), a physician makes a diagnosis by navigating (flying-through) the colon. However, the viewpoints and the viewing directions must be changed many times because the colon is a very long and winding organ with many folds. This is a time-consuming task for physicians. We propose a new *navigation-based* CAD system for the colon providing virtual unfolded (VU) views, which enables physicians to observe a large area of the colonic wall at a glance. This system generates VU, VC, and CT slice views that are perfectly synchronized. Polyp candidates, which are detected automatically, are overlaid on them. We applied the system to abdominal CT images. The experimental results showed that the system effectively generates VU views for observing colon regions.

1 Introduction

In Japan, the number of patients suffering from colonic cancer is increasing. Colonic cancer can be cured completely if it is found in its early stage, i.e., colonic polyps. However, if found in its latter stages, a complete cure is difficult due to metastasis. Therefore, screening of the colon for early detection of colonic polyps is important. Colonoscopy, which is mainly performed in the clinical field, is physically and mentally painful for patients. Furthermore, some problems are

* The authors thank Dr. Masanori Tomikawa of Kyushu University, Masaki Mori, and Takeshi Hagiwara of Sapporo General Hospital, Japan, for providing CT images and for their useful comments from a medical perspective. The authors also wish to thank our colleagues for their suggestions and advices. Parts of this research were supported by the Grant-In-Aid for Scientific Research from the Ministry of Education, the 21st-Century COE program, a Grant-In-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology, Japan Society for Promotion of Science, and a Grant-In-Aid for Cancer Research from the Ministry of Health and Welfare.

that a physician needs much skill and experience to operate the colonoscope, and it is time-consuming task due to the long and winding shape of the colon, which has many folds.

In the clinical field, virtual colonoscopy (VC) is a new diagnostic tool that is considered less-invasive, due to development of imaging devices such as multi detector-row (MD) CT scanners. VC provides a virtual environment of the colon based on 3D CT images of a patient and visualizes the inside of the colon everywhere a physician wants to observe. The CT images on a computer can be regarded as the virtual human body of a real patient. A physician can diagnose the virtual human body while doing a fly-through the inside of it. We call this diagnostic scheme *navigation diagnosis*. However, because the colon has a complicated shape, a physician has to change viewpoints and viewing directions many times. Thus, VC is still a time-consuming task. Therefore, a computer aided diagnosis (CAD) system must be developed for the colon that has various functions to assist physicians such as automated polyp detection and “easy to observe” visualization, e.g., unfolded views of the colon.

Automated polyp detection has been researched by several groups [1, 2, 3, 4, 5]. They used curvature or intensity distribution information of colonic polyps. Each method developed by them can detect polyps with high sensitivity, but false positives also detected. Ways to reduce the false positives are being researched now. Another important function for CAD is visualization. Even if CAD could accurately detect polyp candidates, final decisions should be made by physicians no matter how they use CAD. Therefore, we should also think about how to efficiently visualize the colon. One possible solution is unfolding the colon.

Several studies have examined the virtual unfolding of the colon [6, 7, 8]. The unfolding must (1) be less distorted, (2) give a real-time rendering of the unfolded views and (3) unfold using arbitrary cutting. Our system focuses on requirements (2) and (3) to make diagnoses using CAD efficient.

The CAD system we propose here has automated polyp detection and real-time visualization of unfolded views of the colon. A special characteristic of our system is that three kinds of views are completely synchronized, i.e., ordinary CT slice view, VC view and virtual unfolded (VU) view. For example, if a physician finds a suspicious area in a VU view, he/she can seamlessly check it on both the VC and CT slice views.

In Section 2, we describe the navigation-based CAD system we developed. Experimental results of polyp detection and virtual unfolding are shown in Section 3. A brief discussion is in Section 4.

2 Navigation-Based CAD System for the Colon

2.1 Overview

For a CAD system for the colon, functions that assist physicians such as automated polyp detection, navigation to suspicious areas, and “easy to observe” visualization to reduce the inspection time are important. Because physicians diagnose patients’ CT data by flying through their virtual bodies, we call a CAD

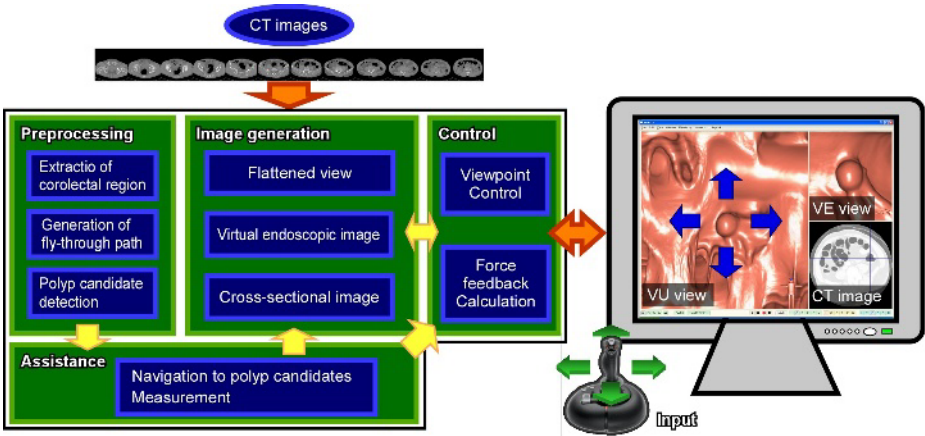


Fig. 1. Overview of navigation-based CAD system

system that uses this technique a *navigation-based CAD system*. Our system consists of four parts: (a) preprocessing, (b) image generation, (c) system control, and (d) observation assistance. An overview of the system is shown in Fig. 1.

The preprocessing part generates a central line of the colon for navigation and unfolding. Polyp candidates are also detected here. The image generation part generates VC, VU, and CT slice views. These views are synchronized by the control part. The system control part also allows physicians to navigate to polyp candidates detected in the preprocessing part. Of course, physicians can fly-through inside the colon freely using intuitive input devices such as joysticks. We explain each part below.

2.2 Preprocessing

Central line generation. First, we extract a colonic lumen region using a region-growing method. Then, we generate a central line from the lumen region.

(1) **Extraction of colonic lumen region.** A colonic lumen region is segmented using a region-growing method (Fig. 2 (a)). CT values of a lumen are defined by -1000 to -700 H.U..

(2) **Thinning and spurious branch elimination.** A thinning algorithm [9] is applied to extract the central line of a colonic lumen region. A thinning result is shown in Fig. 2 (b). As you see, many spurious branches remain after the thinning process. Those spurious branches are eliminated by calculating the minimal path between the farthest two points on the thinned figure (Fig. 2 (c)).

(3) **Central line generation using spline interpolation.** A central line is generated using cubic spline interpolation as follows. First, control points \mathbf{P}_k ($k = 0, \dots, N$) are allocated at intervals of 20 voxels on the thinned figure. Then, the central line is obtained using cubic spline interpolation from \mathbf{P}_k (Fig. 2 (d)). Points on the central line are denoted as \mathbf{p}_i ($i = 0, \dots, M$).

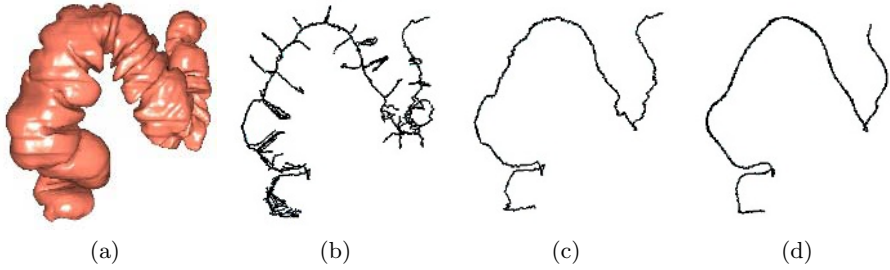


Fig. 2. Generation of central line of colon. (a) Extracted lumen region, (b) result of thinning, (c) result of spurious branch elimination, and (d) generated central line.

Polyp candidate detection. Polyp candidates are detected using a detection method [5]. Convex regions on the colonic wall are detected as polyp candidates. The shape index and curvedness, which are calculated from curvatures on the colonic wall, are used for shape classification. See Kitasaka et al. [5] for details.

2.3 Image Generation

In this part, VC, VU, and CT slice views are generated. The VU view is generated by controlling ray-casting directions in the volume rendering.

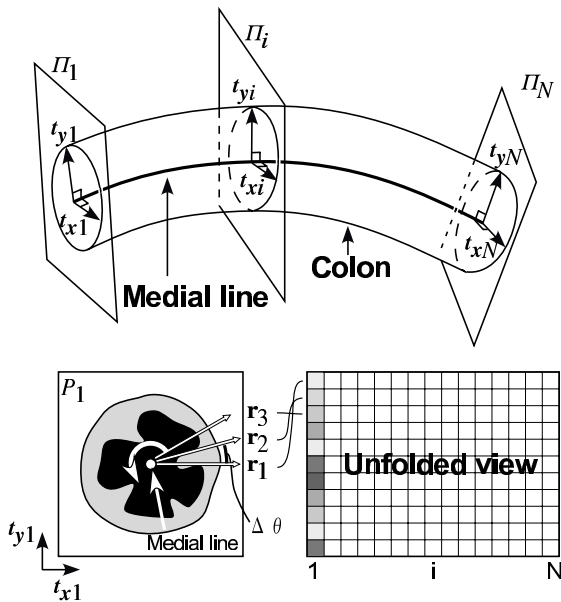


Fig. 3. Schematic illustration of unfolding process

Virtual unfolding view. For each point on the central line \mathbf{p}_i , a ray is casted along a direction vertical to the tangent, \mathbf{t}_i , of the central line at the voxel. Let \mathbf{t}_{x_i} be a vector perpendicular to \mathbf{t}_i , and \mathbf{t}_{y_i} be the vector perpendicular to both \mathbf{t}_i and \mathbf{t}_{x_i} . Then, casting directions $\mathbf{r}_j^i (0 \leq j \leq m; m = 2\pi/\Delta\theta)$ at \mathbf{p}_i are defined by

$$\mathbf{r}_j^i = \mathbf{t}_{x_i} \cos(j\Delta\theta) + \mathbf{t}_{y_i} \sin(j\Delta\theta). \quad (1)$$

That is, a VU view is generated by casting rays on an orthogonal plane, Π_i , toward the colonic wall side at each point, \mathbf{p}_i , on the central line (Fig. 3). Because this unfolding is equivalent to a volumetric deformation, we can observe not only the surface of the colonic wall but also blood vessels beyond the wall on a VU view by adjusting rendering parameters (color, opacity, etc.).

2.4 System Control

This system uses a joystick as an intuitive input device. When the stick is inclined right and left, the camera moves forward and backward along the central line. When the stick is inclined back and forth, the VU view is rotated along the central line; i.e., the cutting line of the colon is modifiable.

2.5 Observation Assistance

This part is to assist observation and diagnosis by synchronizing views and coloring polyp candidates detected automatically and unobserved regions.

Synchronized views. Let \mathbf{p}_i be the point on the central line at the center of a VU view. VU, VC, and CT slice views are synchronized as follows. In the VC view, the viewpoint is the same as \mathbf{p}_i , and the view direction is set as $\mathbf{r}_{m/2}^i$ (the vector that points to the center of the VU view). The CT slice that contains \mathbf{p}_i is displayed.

Colored polyp candidates and unobserved regions. Polyp candidates are overlaid on VU views and VC and a CT slice views. Because of the complicated shape of the colon, polyps may be overlooked. Thus, our system detects unobserved regions [10]. Physicians are shown which parts they have not observed yet. Polyp candidates can be navigated to.

3 Experiments

We implemented the navigation-based CAD system on a PC platform (CPU: Intel Xeon 3.6GHz, Memory: 2GByte). Four cases of the colonic wall were unfolded by the system. The acquisition parameters of the CT data are: 512×512 pixels, 0.586 - 0.781 mm/pixel, 333 - 370 slices, 1.0 - 2.5 mm slice thickness, and 1.0 - 1.25 mm reconstruction intervals.

The results of the central line extraction are shown in Fig. 4. Examples of unfolding are shown in Fig. 5. System overviews are shown in Fig. 6. The top,

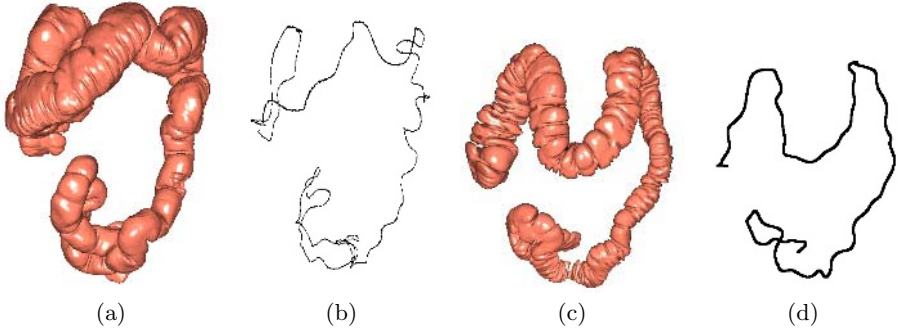


Fig. 4. Results of central line generation. (a) Extracted lumen of Case 2, (b) central line generated from (a), (c) extracted lumen of Case 3, and (d) central line generated from (c).

Table 1. Ray intersection percentages

Case	Ray intersection (%)
1	24.56
2	33.34
3	19.41
4	13.74

left-bottom, and right-bottom views are the VU, VC, and CT slice views. Polyp candidates detected by the system are colored blue. In Fig. 6 (b), blood vessels beyond the colonic wall can be observed. During VU view generation, some rays cast from the central line intersect each other before they hit the colonic wall. The percentages of rays that intersect are shown in Table 1. The rendering speed, including the unfolding process, was approximately 0.1 second per frame. The rendering resolutions of VU and VC views are 256×256 and 256×256 .

4 Discussion

Our system can unfold the colon in real-time. As shown in Fig. 5, haustra are clearly distinguished in VU views. However, we found that areas near sharp bends are distorted. Rays intersected in such areas. As shown in Table 1, rays intersected in all cases. As a consequence, a polyp may appear more than once on the VU view, or a part of a polyp may switch its position with another part of the colonic wall. To reduce this distortion, ray intersection must be reduced. This subject will be addressed in future work.

As shown in Fig. 6, physicians can easily find polyp candidates in the VU view and verify them on both the VC and CT slice views. Those three views are completely synchronized in real-time. This allows physicians to feel free to observe, i.e., comfortably navigate the colon. Furthermore, blood vessels beyond the colonic wall can be observed in the VU view. The unfolding process used in

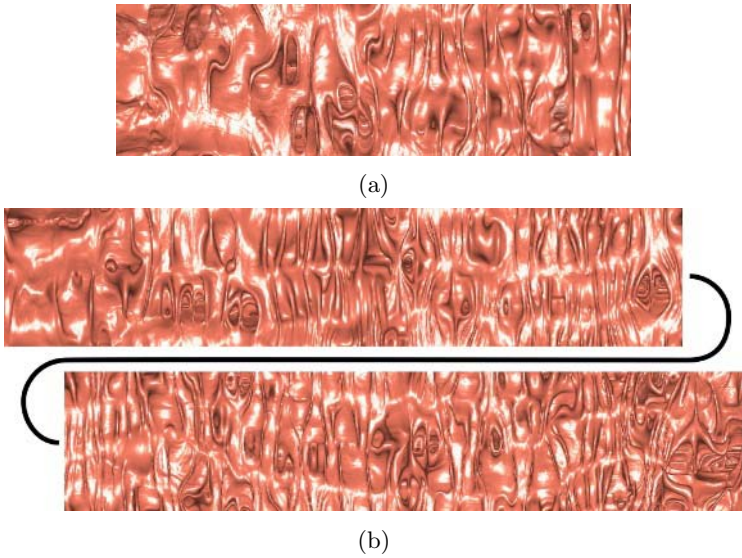


Fig. 5. Results of unfolding of colon. (a) Case 1 and (b) Case 3.

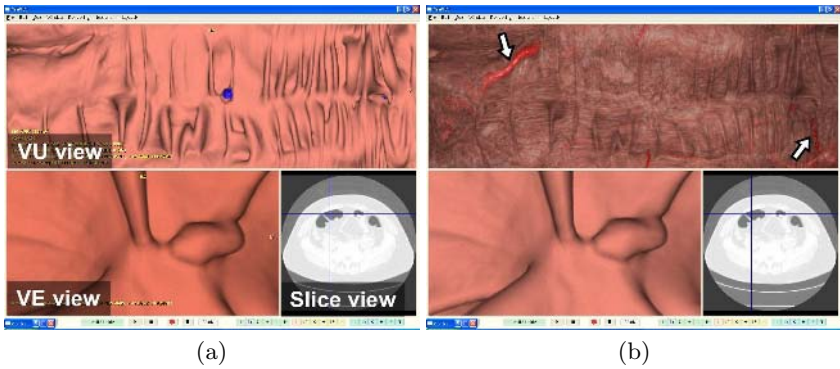


Fig. 6. Example of polyp candidate detection and its display in VU view for Case 3. Blood vessels are observed in VU view by changing transparency of colonic wall, indicated by arrows.

the system volumetrically deforms the colon simply by controlling the directions of ray-casting, so we can observe not only the surface of the colonic wall but also beyond the wall by changing the rendering parameters.

5 Conclusion

We proposed a navigation-based CAD system for the colon. This system has functions to assist physicians; automated polyp detection, virtual unfolding, and

synchronized VU, VC, and CT slice views. We applied the navigation-based CAD system to four cases of 3D abdominal CT data. By comparing synchronized VU and VC views, we observed the entire colonic wall efficiently. Future work includes further experiments using larger number of CT data, evaluations by physicians, and reduction of distortion in VU views.

References

1. R. M. Summers and H. Yoshida, "Future directions of CT colonography: computer-aided diagnosis," In: Dachman AH, ed. Atlas of virtual colonoscopy. New York: Springer-Verlag, 2002.
2. J. Nappi and H. Yoshida, "Automated detection of polyps with CT colonography: evaluation of volumetric features for reduction of false-positive findings," *Academic Radiology*, 9, pp. 386-397, 2002.
3. Anna K. Jerebko, James D. Malley, Marek Franaszek, and Ronald M. Summers, "Multiple Neural Network Classification Scheme for Detection of Colonic Polyps in CT Colonography Data Sets," *Academic Radiology*, 10, pp. 154-160, 2003.
4. H. Yoshida and A. H. Dachman, "CAD techniques, challenges, and controversies in computed tomographic colonography," *Abdominal Imaging*, 30, pp. 26-41, 2005.
5. T. Kitasaka, K. Mori, T. Kimura, Y. Hayashi, Y. Suenaga, and J. Toriwaki, "A method for detecting colonic polyps using curve fitting from 3D abdominal CT images," *Proc. of SPIE Medical Imaging 2005*, 5746-44, 2005. (in print)
6. Ge Wang, Elizabeth G. McFarland, Bruce P. Brown, and Michael W. Vannier, "GI Tract Unraveling with Curved Cross Sections," *IEEE Trans. on Medical Imaging*, 17, 2, pp. 318-322, 1998.
7. S. Haker, S. Angenent, A. Tannenbaum, and R. kikinis, "Nondistorting Flattening Maps and the 3-D Visualization of Colon CT Images," *IEEE Trans. on Medical Imaging*, 19, 7, pp. 665-670, 2000.
8. A. V. Bartroli, R. Wegenkittl, A. Koenig, E. Groeller, and E. Sorantin, "Virtual Colon Flattening," *Proc. of IEEE TCVC Symposium on visualization*, 2001.
9. T. Saito and J. Toriwaki, "New algorithms for n-dimensional Euclidean distance transformation," *Pattern Recognition*, 27, 11, pp. 1551-1565, 1994.
10. Y. Hayashi, K. Mori, Y. Suenaga, J. Hasegawa, and J. Toriwaki, "A method for detecting undisplayed regions in virtual colonoscopy and its application to quantitative evaluation of fly-through methods," *Academic Radiology*, 10, 12, pp. 1380-1391, 2003.

A Prediction Framework for Cardiac Resynchronization Therapy Via 4D Cardiac Motion Analysis

Heng Huang¹, Li Shen³, Rong Zhang¹, Fillia Makedon¹,
Bruce Hettleman², and Justin Pearlman^{1,2}

¹ Department of Computer Science, Dartmouth College, Hanover, NH 03755

² Department of Cardiology, Dartmouth Medical School, Lebanon, NH 03756

³ Computer and Information Science Department, Univ. of Mass. Dartmouth, MA 02747
hh@cs.dartmouth.edu

Abstract. We propose a novel framework to predict pacing sites in the left ventricle (LV) of a heart and its result can be used to assist pacemaker implantation and programming in cardiac resynchronization therapy (CRT), a widely adopted therapy for heart failure patients. In a traditional CRT device deployment, pacing sites are selected without quantitative prediction. That runs the risk of suboptimal benefits. In this work, the spherical harmonic (SPHARM) description is employed to model the ventricular surfaces and a novel SPHARM-based surface correspondence approach is proposed to capture the ventricular wall motion. A hierarchical agglomerative clustering technique is applied to the time series of regional wall thickness to identify candidate pacing sites. Using clinical MRI data in our experiments, we demonstrate that the proposed framework can not only effectively identify suitable pacing sites, but also distinguish patients from normal subjects perfectly to help medical diagnosis and prognosis.

1 Introduction

Heart failure, also called congestive heart failure, is a major health problem that continues to increase in prevalence. It is a disorder in which the heart loses its ability to pump blood efficiently. Low cardiac output resulting from heart failure may cause the body's organ systems to fail. As one important part of the problems, the walls of the left ventricle (LV) are unable to contract synchronously.

Over the past decade, investigators [2] have established the feasibility of placing multiple pacing leads of pacemaker to improve the activation synchrony (sameness of activation time) of LV and biventricle. Based on these studies, a promising therapeutic option, called cardiac resynchronization therapy (CRT), has been proposed as an alternative treatment in patients with severe, drug-refractory heart failure. It is aimed at correcting contraction delays that result in different regions of the heart not working optimally in concert [1].

Although clinical trials have confirmed that CRT improved clinical symptoms, increased exercise capacity, and led to cessation or even reversal of chronic chamber remodeling, a significant minority seem not to benefit. [3] There are a lot of potential explanations for the CRT failure cases, and improper surgical placement of the LV lead is one of the most important reasons. Improvements in the methods of identification of likely responsive implantation placements are needed.

The initial CRT device utilization incorporates right ventricle and LV pacing sites, but a right ventricle pacing site is not required for hemodynamic benefit in many patients [4]. Since LV pacing alone has the almost similar benefit with biventricular pacing, it is used more popular in the CRT system. Thus, in this paper, we focus on the CRT with LV pacing sites.

The principal goal of the present study is to efficiently predict the optimal LV pacing sites that should be stimulated by electrical impulses of pacemaker and provide the corresponding parameters (timing delay, *etc.*) to help the programmable device. In order to identify the optimal pacing sites, the mechanical dyssynchrony is directly analyzed by the spatio-temporal modeling. Based on the fact that ventricular wall thickening and motion reflect activation, we build an integrated framework to estimate the most effective places for implanting the pacemaker to achieve maximized CRT benefit. Given a stack of cardiac MRI, both endocardium and epicardium are reconstructed, and the optimal correspondences between them are established by minimizing the surface Euclidean distance. LV wall motion is described by the three dimensional (3D) wall thickness change that is computed using the reconstructed LV surfaces. After applying the hierarchical clustering method on a time series of wall thickness measurements, we can find candidate pacing sites with abnormal local motion. Our experiments also show that this study can be used to distinguish patients and normal subjects and judge the disease degree.

2 Methods

2.1 Spatio-Temporal LV Motion Modeling

In order to quantify the ventricular mechanical asynchrony or synchrony that can directly help determine optimal treatment, we develop our spatio-temporal model to describe a temporal sequence of wall thickness changing during a heart cycle.

Surface reconstruction. We reconstruct both endocardium and epicardium of the LV by using the spherical harmonic (SPHARM) method, which was introduced by Brechbühler, Gerig and Kübler [5] for modeling any simply connected 3D object. The object surface is parameterized as $\mathbf{v}(\theta, \phi) = (x(\theta, \phi), y(\theta, \phi), z(\theta, \phi))^T$ using a pair of spherical coordinates (θ, ϕ) , where the parameterization aims to preserve the area and minimize the angle distortion. Thus, $\mathbf{v}(\theta, \phi)$ becomes a vector of three spherical functions that can be expanded using spherical harmonics $Y_l^m(\theta, \phi)$ as follows,

$$\mathbf{v}(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \mathbf{c}_l^m Y_l^m(\theta, \phi),$$

where $\mathbf{c}_l^m = (c_{lx}^m, c_{ly}^m, c_{lz}^m)^T$ are expansion coefficients that can be used to reconstruct the object surface. SPHARM has been used by Gerig and Styner in many medical imaging applications (*e.g.*, shape analysis of brain structures [6, 7, 8]). It has also been used for shape modeling and functional analysis of cardiac MRI [9]. Since SPHARM provides an implicit correspondence between surfaces of 3D objects, it is suitable to be used to analyze the LV wall motion during heart cycle.

In our cardiac MRI data sets, each MRI sequence holds seventeen temporal phases per heartbeat. Since the LV deformation is exhibited by the thickness change of the wall between endocardium and epicardium, we use 17 SPHARM reconstructed surface pairs (including both endocardium and epicardium) to describe the LV contraction and dilation during a whole heart cycle.

Surface correspondence. In order to measure the wall thickness at each surface location as well as compare thickness changes between different time points, a registration step is necessary for aligning all the reconstructed epicardial surfaces together. Given two SPHARM models, we establish their surface correspondence by minimizing the Euclidean distances between their corresponding surface locations. Formally, for two surfaces given by $\mathbf{v}_1(s)$ and $\mathbf{v}_2(s)$, their distance $D(\mathbf{v}_1, \mathbf{v}_2)$ is defined as [7]

$$D(\mathbf{v}_1, \mathbf{v}_2) = \left(\oint \|\mathbf{v}_1(s) - \mathbf{v}_2(s)\|^2 ds \right)^{1/2} = \left(\sum_{f \in \{x, y, z\}} \sum_{l=0}^L \sum_{m=-l}^l (c_{lf_1}^m - c_{lf_2}^m)^2 \right)^{1/2}.$$

The epicardial surface in the first time phase (diastolic phase in our MRI data) during heartbeat is used as the template. For any other epicardial surface in the same sequence, we align it to the template by rotating its parameter net [10] so that the surface distances $D(\mathbf{v}_1, \mathbf{v}_i)$, ($i = 2, \dots, 17$) between them is minimized. Given an aligned surface sequence, we use the same method to align the endocardium to the epicardium in the same timing phase.

Wall motion series. Because the wall thickness change of LV directly shows the wall motion in 3D space during a heart cycle, we can use it as the wall motion descriptor. In this study, we observe that the distance between the corresponding points (*i.e.*, with the same (θ, ϕ)) on the endocardium and epicardium surfaces, can be directly used to approximate the wall thickness, since we have already minimized this distance in the previous step. In addition, the underlying equal area parameterization implies the correspondence relationships between any pair points on these two surfaces are reasonable and effective.

After that, we create the wall motion series that includes the thickness values for each corresponding surface location at each time phase during a heart cycle, from end-diastolic phase to next end-diastolic one. Since we are only interested in the LV wall motion, we ignore the points appearing on the top of reconstructed surfaces. Even if only one point whose thickness value in the wall motion series appears on the top of its surface, the whole motion series is discarded.

2.2 Clustering Based Pacing Sites Search

After the above steps, a set of motion series are used to present the LV wall contraction. Given a pair of (θ, ϕ) , we use $\mathbf{w}(\theta, \phi) = \{w_1(\theta, \phi), w_2(\theta, \phi), \dots, w_n(\theta, \phi)\}$ to denote its corresponding wall motion series, where $w_i(\theta, \phi)$ is the wall thickness value of wall motion phase i corresponding to the parametrized point (θ, ϕ) on the epicardium.

Similarity measurement. Formally, given two wall motion series $\mathbf{w}(\theta_x, \phi_x)$ and $\mathbf{w}(\theta_y, \phi_y)$, we employ the following formulae to measure the distance or dissimilarity between them:

$$d_{corr}(\mathbf{w}(\theta_x, \phi_x), \mathbf{w}(\theta_y, \phi_y)) = 1 - r(\mathbf{w}(\theta_x, \phi_x), \mathbf{w}(\theta_y, \phi_y)) = 1 - \left[\sum_{i=1}^n \left(\frac{w_i(\theta_x, \phi_x) - w_{mean}(\theta_x, \phi_x)}{\sigma_x} \right) \left(\frac{w_i(\theta_y, \phi_y) - w_{mean}(\theta_y, \phi_y)}{\sigma_y} \right) \right] / n,$$

where $\sigma = \sqrt{(\sum_{i=1}^n (w_i(\theta, \phi) - w_{mean}(\theta, \phi))^2) / n}$. $r(\mathbf{w}(\theta_x, \phi_x), \mathbf{w}(\theta_y, \phi_y))$ is the Pearson correlation coefficient of two wall motion series, $w_{mean}(\theta, \phi)$ is the mean of wall motion series, and σ is the standard deviation of $\mathbf{w}(\theta, \phi)$. The Pearson correlation coefficient is always between -1 and 1, and we normalize distance function as $d_{corr}/2$ (the result will change from 0 to 1) in our experiments.

Hierarchical clustering. By combining or clustering similar wall motion series we can identify groups of wall motion series that are the main trend of LV contraction and dilation for different locations in the 3D space. To group similar wall motion series together, we employ hierarchical agglomerative clustering approach [11], which is a bottom-up clustering method where clusters can have sub-clusters.

The hierarchical clustering result, a dendrogram, is a binary tree (see Fig. 1 for an example) in which each data point corresponds to a leaf nodes, and distance from the root to a subtree indicates the similarity of subtrees—highly similar nodes or subtrees have joining points that are farther from the root.

Sweep line method. We move the horizontal sweep-line from top to bottom in the dendrogram result (for example, the “sweep-line 1” in Fig. 1) to get the abnormal clusters (small clusters) that have a large dissimilarity to the main cluster. Note that the pacemaker system uses electrical impulses to adjust the sites whose contraction characteristics are considerably different from other sites’. Thus, hierarchical clustering results can help us to find these location candidates for installing the pacing leads.

2.3 Pacing Site Pre-filtering

Cross correlation. In order to set the electrical impulses in a pacemaker system, a technician still needs to know the timing delay value between the pacing position and a common position. Hence we use cross correlation method to acquire such a value between the wall motion series. For two wall motion series $\mathbf{w}(\theta_x, \phi_x)$ and $\mathbf{w}(\theta_y, \phi_y)$, the correlation function of two wall motion series is defined as:

$$cc_{xy}(t) = \mathbf{w}(\theta_x, \phi_x) \star \mathbf{w}(\theta_y, \phi_y) = \sum_{m=1}^n w_m(\theta_x, \phi_x) w_{m+t}(\theta_y, \phi_y)$$

where “ \star ” is the correlation operator, and $t = 0, 1, \dots, n - 1$. If $t = t_0$ satisfies $cc_{xy}(t_0) = \max(cc_{xy}(t))$ for $t \in [0, n - 1]$, then the wall motion series $\mathbf{w}(\theta_x, \phi_x)$ shifts t_0 to get the maximum correlation with the wall motion series $\mathbf{w}(\theta_y, \phi_y)$. Thus, t_0 is the timing shift (or delay). The timing period between two neighbouring phases can be calculated using the heartbeat velocity. Thus, the timing delay can be calculated by: *timing delay* = $t_0 \times (\text{a heartbeat period} / \text{the number of phases})$.

Pacing sites selection. We introduce a filtering step on the pacing site candidates list, because a few of them don't have contraction timing delay to the normal activation. After picking up the site candidates, there is a single big cluster in the dendrogram, called the main cluster (see Fig. 1 for a marked sample main cluster). The wall motion series (average motion series) of the main cluster is regarded as the normal wall motion variation of the LV, for example the square-line in Fig. 2 and Fig. 3. Using the contraction timing delay between pacing site candidates and main cluster, we filter out the site candidates without contraction delay.

In the implantation, a doctor still needs to test the lead to see whether a candidate location is suitable for pacing, because the pacing lead cannot be placed into some regions of left ventricle (such condition normally is created by epicardial scar or unacceptable phrenic nerve stimulation, *etc.*). Based on the dendrogram result, we provide the location candidates for implanting and they are rated by the distances from the main cluster, which is described below.

3 Results and Discussions

We have implemented our pacing site prediction framework using Matlab 6.5. To show the effectiveness of this framework, we use cardiac MRI data from 20 patients in our experiments, where half of them have heart failure problems. These experiments are conducted on a PC with a 2.40GHz CPU and 512 MB main memory. Note that the patients are diagnosed by specialized physicians, and these diagnostic results are used to validate our results in the experiments.

For convenience, we allocate a number to each wall motion series. From apex to basis of the LV, 1 ~ 96 are used to mark the points of wall motion series level by level. Therefore, the points represented by consecutive numbers are in the neighbour locations on the surface, and the points with small numbers should be close to apex and the points with large numbers should be close to the basis of the LV.

Fig. 1 shows the hierarchical clustering result of a patient with heart failure problem. The dendrogram consists of a main cluster and several other small ones. The locations corresponding to the motion series in those small clusters are selected as the candidate pacing sites. Note that a single small cluster may include multiple regions on the LV, since the different regions may have similar motion behaviors. In Fig. 1, {92, 93} (close to the basis of LV) is the top-priority option for resynchronization therapy, and the next pacing candidates that should be considered are {77, 78, 79} and {30, 31}.

Since the distance function used by us cannot discriminate the timing delay between wall motion series, pre-filtering step should be executed here. Fig. 2 and Fig. 3 show the pacing sites filtering step. In Fig. 2, the curve with square tags is the average motion series of the main cluster in Fig. 1 and the curve with circle tags is the average motion series of region {92, 93}. Because there is no timing delay between the main cluster and this region, it is filtered out, although their average wall motion series is very different from the main cluster's. Regions {77, 78, 79} and {30, 31} still remain in the candidate list, since obvious timing delays are observed in Fig. 3. After filtering step, our results can be used for pacemaker implantation. As we mentioned before, the pacing lead cannot be placed into some particular regions of the LV. The physician will test the pacing

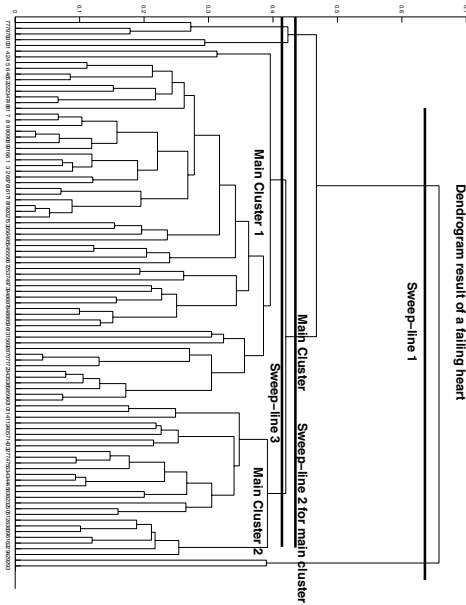


Fig. 1. Dendrogram result of a failing heart. The x -axis label represents the number of wall motion series. The y -axis label corresponds to the distance between clusters. The dendrogram is cut into clusters by the “sweep-line 3”.

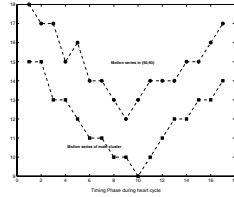


Fig. 2. There is no obvious timing delay between the average wall motion series of the main cluster (square-curve) and the motion series of region $\{92, 93\}$ (circle-curve)

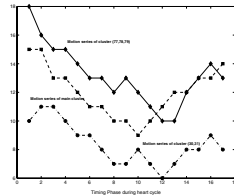


Fig. 3. There is a contraction timing delay between the main cluster and regions $\{77, 78, 79\}$ (diamond-curve) and $\{30, 31\}$ (circle-curve)

lead on candidate pacing sites according to the suggested site ordering until they find a suitable region for fixing the tip of pacing lead. If the list is empty and a suitable site isn’t found, we will continue to select a lower value sweep-line in the dendrogram result, for example, the “sweep-line 2” and “sweep-line 3” in Fig. 1. Because the candidates list includes locations with notable asynchronous contraction and timing delay, the optimal resynchronization therapy can be obtained after adding electrical pulse into these candidates. These sites are potentially good candidates to implant the pacemaker for a more efficient CRT. Furthermore, in some clinical cases, a physician may want to use multiple sites in left ventricular pacing for cardiac resynchronization, and they can select additional locations from the candidates list.

We test our methods on the MRI data of both normal and failing hearts. The dendrogram results of the normal hearts are very different from the failing ones. In the normal heart dendrogram, see Fig. 4, if the value of sweep-line is selected as ≥ 0.3 (the distance between clusters), we obtain only one single main cluster without any other small clusters. This matches our intuition, since the wall motion of a normal heart tends to be synchronous and so the motion difference on different surface locations are very small. Thus our analysis may be useful in identifying patient candidates for helping diagnosis. After obtaining twenty dendrograms for all subjects, for each single case, we move the sweep-line from top to bottom until the result contains exactly two clusters. Fig. 5

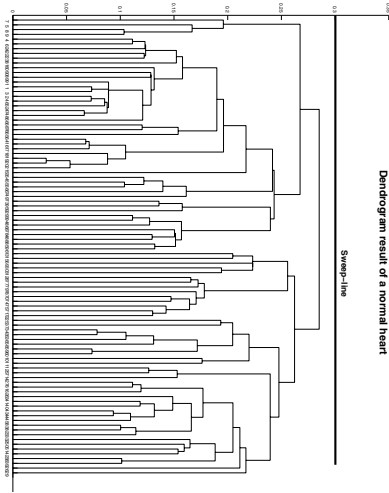


Fig. 4. Dendrogram result of a normal heart. The x -axis label represents the number of wall motion series. The y -axis label corresponds to the distance between clusters.

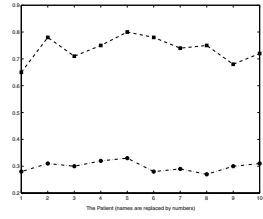


Fig. 5. The cutoff value of the sweep-line

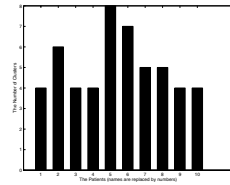


Fig. 6. The number of small clusters extracted from the abnormal subjects

summarizes the final values of these sweep-lines, sorted in two groups – one group holds 10 low values, and the other one holds 10 high values. The clinical diagnosis indicates that all low value points (circles) correspond to normal hearts and all high value ones (squares) correspond to failing hearts. Note that there is a big gap between these two lines, and so such hierarchical clustering results can actually separate subjects with heart failure from normal subjects. This observation is helpful for heart failure diagnosis and prognosis. The value $0.4 \sim 0.5$ seems to be a good threshold for the sweep-line to distinguish failing hearts from normal hearts in our data.

In the failing heart data set, we move the sweep-line to extract all small clusters (this sweep-line may separate the main cluster into two or more main clusters). “Sweep-line 3” in Fig. 1 is such an example and in this case we have five clusters: $\{92, 93\}$, $\{77, 78, 79\}$, $\{30, 31\}$, main cluster 1 and main cluster 2. Fig. 6 shows the number of clusters gotten from these 10 abnormal subjects, without counting the main cluster. In this result, the dendrogram result of patient #5 has more small clusters and the patient also has worse heart failure symptom. It is often the case that, the more severe heart failure the patient has, the more asynchronous wall contraction the LV performs. Our results reveal such a reasonable relationship between the number of pacing site candidates and the degree of patients’ symptom.

4 Conclusion

In this paper, we propose a new pacing sites prediction framework that is based on spatio-temporal analysis of cardiac motion patterns and hierarchical clustering method.

It can automatically generate the candidate site list to help a physician to localize the pacing areas. Blinded analysis of clinical MRI data demonstrates that our approach can not only identify pacing sites in the LV for assisting pacemaker implantation in CRT, but also be used to help medical diagnosis of heart failure. Since the bi-ventricular pacing is useful for some clinical cases, our further study will be carried out on the prediction of its pacing sites. This further study will be combined with our current prediction system to improve the accuracy and feasibility of estimated results.

References

1. Khaykin, Y., Saad, E.B., Wilkoff, B.L.: Pacing in heart failure: the benefit of resynchronization. *Cleve. Clin. J. Med.* **70(10)** (2003) 841–865
2. Trautmann, S.I., Kloss, M., Auricchio, A.: Cardiac resynchronization therapy. *Curr. Cardiol. Rep.* **4(5)** (2002) 371–378
3. Auricchio, A., Stellbrink, C., Sack, S., *et al.*: Long-term clinical effect of hemodynamically optimized cardiac resynchronization therapy in patients with heart failure and ventricular conduction delay. *J. Am. Coll. Cardiol.* **39(12)** (2002) 2026–2033.
4. Stevenson, W.G., Sweeney, M.O.: Single Site LV Pacing for Cardiac Resynchronization. *Circulation.* **90(5)** (2004) 483–484
5. Brechbühler, Ch., Gerig, G., and Kübler, O.: Parametrization of closed surfaces for 3D shape description. *Computer Vision and Image Understanding.* **61(2)** (1995) 154–170
6. Styner, M., Gerig G.: Three-Dimensional Medial Shape Representation Incorporating Object Variability. *Proc. IEEE Conf. Computer Vision and Pattern Recognition.* (2002) 651–656
7. Gerig, G., and Styner, M.: Shape versus Size: Improved Understanding of the Morphology of Brain Structures. *International Conf. on Medical Image Computing and Computer Assisted Intervention.* LNCS **2208** (2001) 24–32
8. Styner, M., Lieberman, A.J., Pantazis, D., and Gerig, G.: Boundary and Medial Shape Analysis of the Hippocampus in Schizophrenia. *Med. Im. Ana. MEDIA.* **8(3)** (2004) 197–203.
9. Huang, H., Shen, L., Ford, J., Makedon, F., Zhang, R., *et al.*: Functional analysis of cardiac MR images using SPHARM modeling. *Proceedings of the SPIE* **5747** (2005) 1384–1391.
10. Huang, H., Shen, L., Zhang, R., Makedon, F., Hettleman, B., Pearlman, J.: Surface Alignment of 3D Spherical Harmonic Models: Application to Cardiac MRI Analysis. *International Conf. on Medical Image Computing and Computer Assisted Intervention.* (2005)
11. Alpaydin, E.: *Introduction to Machine Learning.* The MIT Press, (2004)

Segmentation and Size Measurement of Polyps in CT Colonography

J.J. Dijkers^{1,2}, C. van Wijk¹, F.M. Vos^{1,3}, J. Florie³, Y.C. Nio³,
H.W. Venema^{3,4}, R. Truyen², and L.J. van Vliet¹

¹ Quantitative Imaging Group, Department of Imaging Science and Technology, Faculty of Applied Sciences, Delft University of Technology, Delft, The Netherlands

L.J.vanVliet@tnw.tudelft.nl <http://www.qi.tnw.tudelft.nl/>

² Philips Medical Systems, Best, The Netherlands

³ Department of Radiology, Academic Medical Center, Amsterdam, The Netherlands

⁴ Department of Medical Physics, Academic Medical Center, Amsterdam, The Netherlands

Abstract. Virtual colonoscopy is a relatively new method for the detection of colonic polyps. Their size, which is measured from reformatted CT images, mainly determines diagnosis. We present an automatic method for measuring the polyp size. The method is based on a robust segmentation method that grows a surface patch over the entire polyp surface starting from a seed. Projection of the patch points along the polyp axis yields a 2D point set to which we fit an ellipse. The long axis of the ellipse denotes the size of the polyp. We evaluate our method by comparing the automated size measurement with those of two radiologists using scans of a colon phantom. We give data for inter-observer and intra-observer variability of radiologists and our method as well as the accuracy and precision.

1 Introduction

Colorectal cancer is one of the most commonly diagnosed types of cancer. Specifically, the American Cancer Society predicts 145,000 new cases and 56,000 deaths from colorectal cancer for 2005 [1]. Polyps are a well-known precursor to such carcinoma. Not surprisingly, it has been shown that early removal of polyps ensures a decrease in incidence [7].

In recent years, CT colonography has been proposed as a noninvasive alternative to traditional polyp detection by colonoscopy [9, 5]. In CT colonography, the colon structure is often visualized from an endoluminal perspective by means of surface or volume rendering. Recently, methods have been proposed to support the inspection by a computer aided detection (CAD) system indicating suspect locations [8, 4]. The size of a detected polyp is an important aspect for diagnosis and decision making. It is generally accepted that polyps with diameter ≤ 5 mm require no direct further action, whereas larger polyps should be removed via colonoscopy. Typically, the size of polyps is measured in colonoscopy by comparison with an open biopsy forceps. In CT colonography, it is usually measured in reformatted images, in which the largest polyp diameter is selected for size

measurement. However, polyp sizes thus measured by human experts can show significant inter- and intra-observer variability.

Clearly, an automated method is needed to enable more accurate measurement of polyp size. As a side effect, such a procedure is also useful in CAD algorithms. Automated polyp detection is usually based on sophisticated pattern recognition techniques that take into account many features measured on tentatively selected candidates (e.g. size, area, average shape index etcetera). Proper segmentation is crucial to perform reliable feature measurement.

The existing methods for colonic polyp segmentation (such as Summers et al. [3,2] and Yoshida et al. [6]) are especially designed to work directly on the 3D CT data. Such an approach is hindered by not operating on a specifically defined region of interest c.q. the colon surface. Hence, segmentation of polyps which are by definition protrusions of the colon surface is not a trivial task.

In this paper we present a new method for semi-automatic segmentation of polyp-like structures. Additionally, a technique is described to automatically measure polyp sizes using this algorithm. Our method assumes that the colon surface has been identified as a region of interest. Moreover, it is asserted that a candidate location has been identified; in our system by a vertex detection step based on the measured shape index [10]. We will compare the size measurement by our algorithm with that of physicians in a set of phantom objects (in which the size is known a priori).

2 Method

The description below assumes that the colon wall is described by a triangular mesh. However, the basic ideas of the method are not restricted to a mesh based surface representation of the colon wall; they can as well be implemented to work on a voxel based model. An additional advantage of our method is that the segmentation requires no user input, that is all parameters are drawn from the underlying data.

Outline of segmentation procedure. Ideally, a polyp could be described as a rather spherical, symmetric mound on a background shape (see e.g. Figure 2a). One could intuitively delineate a polyp by the inflection points on both sides. However, these points may not be easily identifiable due to the curvature of the background shape (think of a fold).

Hence, we model a polyp to have a symmetry axis that goes through the center point (P_c) in which the apical surface normals converge, and the mean position (P_m) calculated from the polyps surface points. The edge of the polyp is defined by the points at which the surface normals tend to deflect from the center point (we will formalize this below).

Initially, a single position or a small patch indicates a point on the polyp candidate [10]. Since the center and mean points may not be robustly determined from such a seed patch, the polyp segmentation procedure is set up as an iterative process. During each cycle of this process neighboring vertices are added if certain

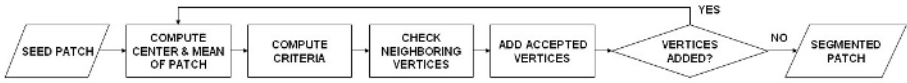


Fig. 1. Schematic overview of segmentation procedure

criteria are met. The process terminates when no more points are added. An overview of the procedure is shown in figure 1.

Computing the center and mean points. As depicted in Figure 2a, the surface normals on the polyp apex tend to converge in a center point. This point (P_c) is found by minimizing the sum of the distances (d_i) to all normals (\mathbf{n}_i). The surface normals are calculated by Gaussian derivatives the underlying 3D CT data at a scale of 2mm. This scale was determined experimentally such that no polyps are missed. The distances can be computed according to:

$$d_i = \|\mathbf{n}_i \times (P_i - P_c)\| / \|\mathbf{n}_i\| \tag{1}$$

where P_i is a point on the patch and \times denotes the vector outer product. Additionally, a mean point (P_m) is associated with a patch. The position of the mean is simply computed by averaging the positions of all vertices: $P_m = \frac{1}{N} \sum P_i$. The mean and the center points define a centerline (dashed in Figure 2b). Henceforth it is called the polyp axis.

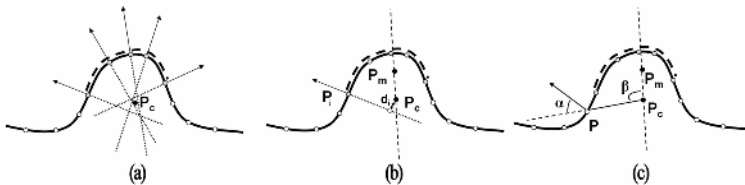


Fig. 2. Schematic representation of a patch (dashed curve) on the colon wall. Figure (a) shows how convergent normals define a center point; figure (b) shows how the minimized distance d_i is defined for surface point P_i ; figure (c) shows how the angles α and β are defined.

Adding points to a seed patch Points are to be added to a seed patch until the local surface normal tends to deviate from P_c . To formalize the stopping criterion, consider first a sphere on a flat background. Let us define α as the angle between the line from the center point (P_c) to the vertex (P_i) and the normal at the position of the vertex (see Figure 2c). Clearly, on top of the polyp α is small (exactly zero on a spherical cap, see Figure 3). The angle α increases while moving to periphery of the polyp. Right outside the polyp the angle is given by (compare with Figure 3):

$$\alpha_{edge} = \arccos \left[\frac{P_{edge} - P_c \cdot \mathbf{n}}{\|P_{edge} - P_c\| \cdot \|\mathbf{n}\|} \right] \tag{2}$$

in which P_{edge} is defined as in Figure 3 and \mathbf{n} is the normal at point P_{edge} . We assume that the ideal threshold-value lies somewhere between these extreme values (respectively 0 and α_{edge}). The required midway point is closely approximated by the angle calculated via (compare with Figure 3):

$$\alpha_{mid} = \arccos [(P_m - P_c) / R] \quad (3)$$

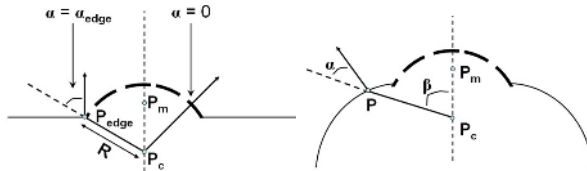


Fig. 3. Schematic representation of a polyp (dashed curve) on a flat background (left) and on a fold (right)

Thus, $\alpha < \alpha_{mid}$ yields a safe stopping criterion for adding neighboring vertices to a polyp on a flat background. On a fold, however, the angle α remains small (see Fig. 3). Let us define β as the angle between the polyp axis and the line between the vertex and the center point (as in Figure 2c):

$$\beta = \arccos \left[\frac{(P - P_c) \cdot (P_m - P_c)}{\|P - P_c\| \cdot \|P_m - P_c\|} \right] \quad (4)$$

At the edge of the polyp β is given by $\beta_{edge} = \alpha_{edge}$. Typically, β continues to increase while moving onto the fold. Consequently, $\beta < \beta_{edge}$ yields a logical stopping criterion for a polyp on a fold. It should be noticed that the two posed criteria are mutually exclusive: the sidepoints of a polyp on a fold do not meet the criterion of $\alpha < \alpha_{mid}$. On the other hand, points besides a polyp on flat background do not fulfill $\beta < \beta_{edge}$. Also, the angles α_{mid} and β_{edge} are both dependent on the shape of a polyp. Flatter polyps tend to have lower values for α_{mid} and β_{edge} than more protruding polyps. In other words, the threshold values automatically depend on the polyp shape.

All vertices neighboring a seed patch that match the conditions are accepted and added at once to yield a new seed. Consequently, the outcome does not depend on the order in which points are processed. Clearly, if none of the vertices match the criteria, no points are added and the current patch is considered the final, segmented polyp. Otherwise, all steps are iterated.

Automated size measurement. The size measurements for polyps are based upon the segmented patches. The edges of these patches are projected along the polyp axis onto a plane. An ellipse is fitted to these points in 2D space by computation of the first and second order moments. This is in accordance with the current medical practice in the Academic Medical Center where the polyp size is characterized by its largest diameter.

3 Experiments and Results

The performance of the method was assessed by comparing the automated size measurement with those of radiologists using scans of a colon phantom. We have looked into several aspects to test our approach:

- Inter-observer variability of radiologists
- Intra-observer variability of radiologists and our method
- Accuracy and precision of the radiologists and our method

Experimental data. All data was acquired using a Mx8000 multislice CT-scanner (Philips Medical Systems, Best, the Netherlands) using the same scanning protocol for all scans (scan parameters: 120 kV, 100 mAs, 4 x 2.5 mm collimation, pitch 1.25, standard reconstruction filter, and a 180o interpolation algorithm).

The phantom consisted of a lucite cylinder into which fabricated polyps were inserted. At first, the phantom contained 10 hemispherical lucite objects of various sizes, and 4 objects with reduced height (2 mm) in order to mimic flat lesions. Subsequently, 8 asymmetric objects from plasticine were inserted in the phantom (maximum width 6-19 mm). The size of all objects (see Table 1) was measured by sliding calipers. The phantom was placed in a cylinder, 34 cm in diameter that was filled with water to arrive at a signal to noise ratio comparable to that in patient data. The two phantoms with lucite and plasticine polyps respectively were scanned twice: in the axial plane of the cylinder, and an orthogonal plane (see Figure 5).

Table 1. Dimensions of the phantom objects; of all lucite objects two specimens were used

Object	Lucite Objects (lengthxheight) [mm]	Plasticine Objects (lengthxheight) [mm]
1	10.0 x 5.0	19 x 9
2	10.0 x 2.5	17 x 8
3	8.0 x 4.0	14 x 10
4	8.0 x 2.0	14 x 8
5	6.0 x 3.0	12 x 8
6	5.0 x 2.5	11 x 11
7	4.0 x 2.0	11 x 5
8	-	6 x 5

The size of all objects was measured in the CT scans by two radiologists and by our automated method. The radiologists measured the objects in multiplanar reformatted CT images. Each object was measured twice, along the main axes as perceived by the physician. The largest value was taken as the polyp size. The automatic measurements were done as described previously. For that purpose, an arbitrary seed point was manually indicated somewhere on the polyp surface.

Inter-observer variability of the radiologists. Figure 4a,d contains the graphs displaying the measurements of one radiologist against those of the other.

Clearly, radiologist A tends to measure larger diameters compared to radiologist B. The average difference of their measurements was 1.2mm for the lucite objects and 3.1mm for the plasticine objects. The standard deviation of the absolute difference was 0.7mm for the lucite objects and 2.7mm for the plasticine objects.

Intra-observer variability of the radiologists and the automatic method.

The intra-observer variability is assessed by the difference in size measured in the axial scan versus the measurement on the same object in the orthogonal scan. It must be conceded that this involves two different scans of the same object. We assume, however, that it allows for a good approximation of the intra-observer variability.

Figure 4b,e shows the measurements of the observers in one scan versus the measurement in the other. Apparently, neither the radiologists nor the automatic method shows a bias. The average absolute difference between the two measurements on the lucite objects was 1.0mm and 0.8mm for the two radiologists and 0.5mm for the automatic technique. The standard deviation of the absolute difference was 0.9, 1.0 and 0.5mm respectively. For the plasticine objects the average absolute differences were 2.6, 3.2 and 1.1mm and the standard deviations 2.2, 2.1 and 1.1mm respectively.

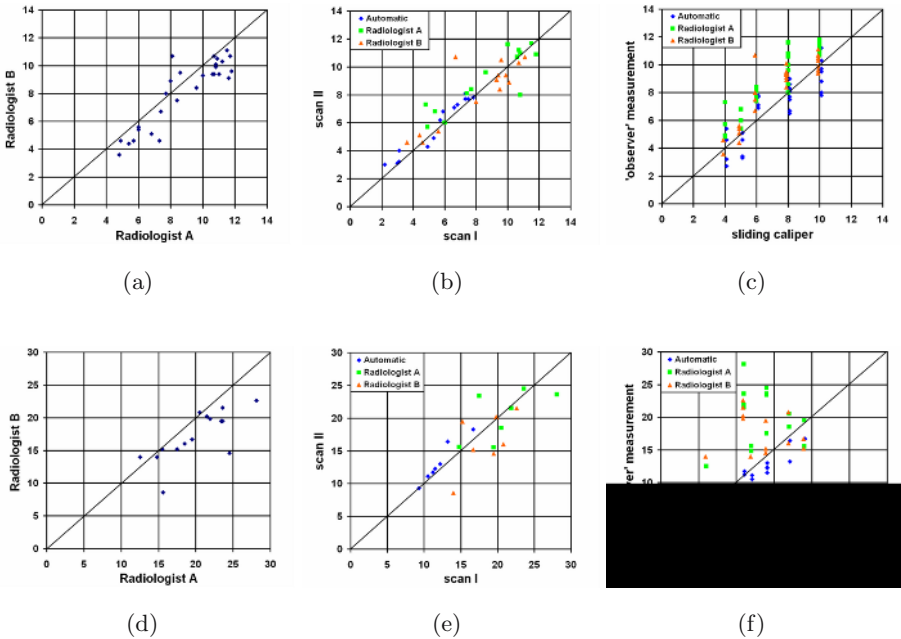


Fig. 4. First column: Inter-observer variability in polyp size for lucite (a) and plasticine (d) objects. Middle column: Intra-observer variability in polyp size for lucite (b) and plasticine (e) objects. Right column: Accuracy and precision of size measurement for lucite (c) and plasticine (f) objects.

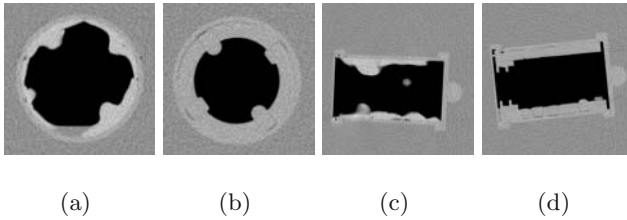


Fig. 5. Single slice from each scan. From a-d: lucite polyps axially, plasticine polyps axially, plasticine polyps orthogonally and lucite polyps orthogonally

Accuracy and precision. The accuracy and precision of the observers is defined by comparison to sliding calipers (see Figure 4c,e). The measurements of both radiologists appear to be slightly biased. This can be explained by the procedure of always selecting the larger of two measurements as the size of the polyp. The bias in the automatic method is less pronounced, but not completely absent. It can be explained by notifying that all points (except due to noise) on a segmented hemispherical polyp surface are projected inside a circle with the diameter of the polyp. An ellipse fitted through the contour points yields a small underestimation of the true size. Clearly, one might correct for all these biases in a calibration step. Specifically noticeable, is the higher precision of the automatic method on the plasticine objects, indicated by the smaller spread of values around the line of identity.

As shown in table 2, the automatic method shows a smaller systematic error than the radiologists. There is no significant difference in the precision (std. dev.) for the (symmetric) lucite objects between the automatic system and the radiologists. However, for the irregular plasticine objects the precision of the automatic system remains the same, whereas the precision of the radiologists decreases significantly.

Table 2. Mean difference and standard deviation of difference between observers and sliding calipers measurements for the lucite objects and the plasticine objects

	Lucite		Plasticine	
	mean diff. [mm]	std. dev. [mm]	mean diff. [mm]	std. dev. [mm]
Rad. A	1.6	0.9	7	5
Rad. B	0.8	1.0	4	5
Automatic	-0.2	1.2	-1	1.2

4 Conclusions and Future Work

The size of a colonographically detected polyp is important for diagnosis and decision making. The size measurement by human observers is generally considered to be imprecise and inaccurate. In this paper we presented a method for the automatic segmentation of polyp-like structures. The polyp size was automatically derived from the segmentation result. It was shown that our algorithm

yields a smaller bias than the measurements from radiologists: on average 1mm or less for the automatic method and between 1 and 7mm for the radiologists, depending on the irregularity of the object. Even more important, the algorithm is consistent irrespective of the polyp shape. As opposed to that, the radiologists show a four times larger variation for the irregularly shaped objects. It is this irregularity which occurs in practice.

A good polyp segmentation algorithm is also useful for automatic polyp detection algorithms. It allows for extraction of features such as volume, surface area, average grey-value etcetera. Such features may improve the specificity of CAD algorithms.

References

1. *Colorectal Cancer Facts & Figures Special Edition 2005*. American Cancer Society, 2005.
2. Jerebko et al. Polyp segmentation method for ct colonography computer aided detection. *Medical Imaging 2003*, pages 359–169, 2003.
3. Summers RM et al. Automated polyp detection and ct colonography: feasibility assessment in a human population. *Radiology*, 219:51–59, 2001.
4. Yoshida H. and NSppi J. Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps. *IEEE Trans Med Imaging*, 20(12):1261–1274, 2002.
5. Serlie I., Vos F.M., van Gelder R.E., Stoker J., Truyen R., Gerritsen F.A., Nio Y., and Post F.H. Improved visualization in virtual colonoscopy using image-based rendering. Data Visualization 2001, Proceedings Joint Eurographics - IEEE TCVG Symposium on Visualization, pages 137–146, 2001.
6. NSppi J. and Yoshida H. Feature-guided analysis for reduction of false positives in cad of polyps for computed tomographic colonography. *Med. Phys.*, 30(7):1592–1601, 2003.
7. Thoeni R.F. and Laufer I. *Polyps and cancer, Textbook of Gastrointestinal Radiology*. Philadelphia: W.B. Saunders, 1994.
8. Summers RM. Challenges for computer-aided diagnosis for ct colonography. *Abdominal Imaging*, 27:268–274, 2002.
9. Hong L. Muraki S., Kaufman A., and He T. Virtual voyage: interactive navigation in the human colon. ACM SIGGRAPH Conf, pages 27–34, 1997.
10. van Wijk C, Truyen R, van Gelder RE, van Vliet LJ, and Vos FM. On normalized convolution to measure curvature features for automatic polyp detection. MICCAI 2004, part I:200-208, pages 200–208, 2004.

Quantitative Nodule Detection in Low Dose Chest CT Scans: New Template Modeling and Evaluation for CAD System Design

Aly A. Farag¹, Ayman El-Baz¹, Georgy Gimelfarb²,
Mohamed Abou El-Ghar³, and Tarek Eldiasty³

¹ CVIP Lab., University of Louisville, Louisville, KY 40292
{farag, elbaz}@cvip.louisville.edu, <http://www.cvip.louisville.edu>

² Department of Computer Science, Tamaki Campus,
University of Auckland, Auckland, New Zealand
g.gimelfarb@auckland.ac.nz

³ Mansoura University, Urology and Nephrology Center, Mansoura, Egypt

Abstract. Automatic diagnosis of lung nodules for early detection of lung cancer is the goal of a number of screening studies worldwide. With the improvements in resolution and scanning time of low dose chest CT scanners, nodule detection and identification is continuously improving. In this paper we describe the latest improvements introduced by our group in automatic detection of lung nodules. We introduce a new template for nodule detection using level sets which describes various physical nodules irrespective of shape, size and distribution of gray levels. The template parameters are estimated automatically from the segmented data (after the first two steps of our CAD system for automatic nodule detection) – no a priori learning of the parameters density function is needed. We show quantitatively that this template modeling approach drastically reduces the number of false positives in the nodule detection (the third step of our CAD system for automatic nodule detection), thus improving the overall accuracy of CAD systems. We compare the performance of this approach with other approaches in the literature and with respect to human experts. The impact of the new template model includes: 1) flexibility with respect to nodule topology – thus *various* nodules can be detected simultaneously by the *same* technique; 2) automatic parameter estimation of the nodule models using the gray level information of the segmented data; and 3) the ability to provide exhaustive search for all the possible nodules in the scan without excessive processing time – this provides an enhanced accuracy of the CAD system without increase in the overall diagnosis time.

1 Introduction

Lung cancer screening is a major research activity worldwide. The goal is to design a CAD system that will contribute to early diagnosis of lung cancer, a major cause of death in various parts of the world. The work of our group uses helical low dose thin slice (2 mm – 2.5mm) chest CT scanning (LDCT)

which provides very high spatial, temporal, and contrast resolution of anatomic structures. Automated detection of lung nodules in thoracic CT scans is an important clinical challenge, especially because manual analysis by a radiologist is time-consuming and may result in missed nodules. Furthermore, the amount of image data to be analyzed continues to increase. We will not attempt an exhaustive search on this subject because of the enormous size of literature on early diagnosis of lung cancer, thus we will limit our scope to the image analysis aspect of the CAD research on this subject. Specifically, we will focus on the automatic detection and identification of lung nodules from the LDCT scanning. Only sample studies will be referred to below as they pertain to our work.

Most of the CAD work in lung cancer screening involves two-stage detection of lung nodules, such that initial candidate nodules are first selected and then the false ones, called false positive nodules (FPNs), are partially eliminated while preserving the true positive nodules (TPNs). For example, the conformal nodule filtering [1] or unsharp masking [2] enhance nodules and suppress other structures at the first stage in order to separate the candidates from the background by simple thresholding. To improve the separation, the background trend is corrected within image regions of interest [3]. A series of 3D cylindrical and spherical filters are used to detect small lung nodules from high resolution CT images [4]. Circular nodule candidates are detected by template matching [2, 5, 6] or other pattern recognition techniques such as fuzzy clustering [7], linear discriminant functions [8], rule-based classification [9], and patient-specific a priori model [10]. Also, cylindrical vascular models are used along with spherical nodular ones to amplify the template matching [11].

The FPNs are excluded at the second stage by feature extraction and classification [12, 13]. Such features as circularity, size, contrast [12], or local curvature [13] are extracted by morphological techniques, and artificial neural networks (ANN) are frequently used as post-classifiers [14].

Our CAD system detects the nodules in LDCT images in three main steps: 1) segmentation of the raw scanning information to isolate the lung tissues from the rest of the structures in the chest cavity [Figs. 1(a) and (b)]; 2) extraction of the 3D anatomic structures (e.g., blood vessels, bronchioles, alveoli, etc., and possible abnormalities) from the already segmented lung tissues [Fig. 1(c)]; and 3) identification of the nodules by isolating the true nodules from other extracted

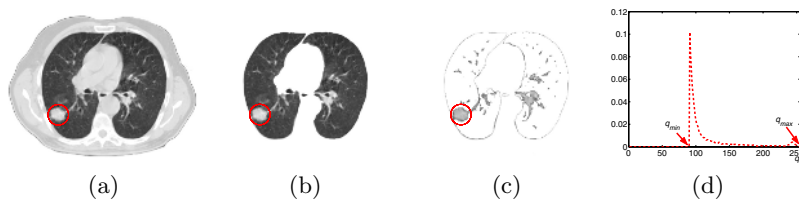


Fig. 1. First two segmentation steps: (a) an initial LDCT slice, (b) separated lung regions, (c) extracted objects (a nodule candidate is encircled), and (d) an empirical marginal gray level distribution for the objects (here, the gray range is [98, 255])

structures. The first two steps considerably reduce the search space by using segmentation algorithms based on representing each CT slice as a sample of a Markov–Gibbs random field of region labels and gray levels. Details of the algorithms are presented in [15].

In this paper we focus on the third step that detects and classifies the nodules among the extracted 3D structures. Both the nodules and normal objects have almost the same marginal gray level distributions similar to that in Fig. 1(d) for all the extracted objects in Fig. 1(c). Therefore, segmentation based solely on gray level distribution (e.g. thresholding) will not work; we need to include geometrical/shape information in this process. The approach proposed in this paper includes a 3D deformable nodule prototype combined with a central-symmetric 3D intensity model of the nodules. The model closely approximates an empirical marginal probability distribution of image intensities in the real nodules of different size and is analytically identified from the empirical distribution.

2 Detecting Lung Nodules with Deformable Prototypes

The detection step extracts, by shape and intensities, and classifies the nodule candidates among all the 3D objects selected at the second segmentation stage.

2.1 Deformable Prototype of a Candidate Nodule

To extract the nodule candidates among the already selected objects like those in Fig. 1(c), we use the deformable prototypes generated by level sets [16] that have become a powerful segmentation tool in recent years. The evolving prototype’s surface at time instant t° is a propagating zero-level front $\phi(x, y, z, t^\circ) = 0$ of a certain 4D scalar function $\phi(x, y, z, t)$ of 3D Cartesian co-ordinates (x, y, z) and time t . Changes of ϕ in continuous time are given by the partial differential equation:

$$\frac{\partial \phi(x, y, z, t)}{\partial t} + F(x, y, z) |\nabla \phi(x, y, z, t)| = 0 \quad (1)$$

where $F(x, y, z)$ is a velocity function and $\nabla = \left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right]^\top$. The scalar velocity function controlling the front evolution depends on local geometric properties, e.g. a local curvature, $k(x, y, z)$, of the front, and on local input data parameters, e.g. a 3D gradient, $\nabla I(x, y, z)$, of the segmented 3D image I .

In practice, the difference relationship replaces Eq. (1) and each next value $\phi(x, y, z, t_{n+1})$ relates to the current one $\phi(x, y, z, t_n)$ at respective time instants t_{n+1} and t_n such that $t_{n+1} - t_n = \Delta t$; $n = 0, 1, \dots$, as follows: $\phi(x, y, z, t_{n+1}) = \phi(x, y, z, t_n) - \Delta t \cdot F(x, y, z) |\nabla \phi(x, y, z, t_n)|$. The velocity function F plays a major role in the propagation process. Among variants in [17, 18], we have chosen $F(x, y, z) = -h(x, y, z)(1 + \varepsilon k(x, y, z))$ where $h(x, y, z)$ and ε are a local consistency term and a smoothing factor, respectively. Since the level set for a segmented 3D image I can always be initialized inside an object, an appropriate consistency term to evolve faster to the object boundary can be as follows:

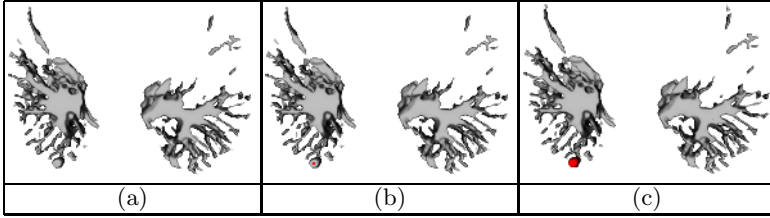


Fig. 2. Part of the separated 3D lung objects (a), the initialized level set (b) indicated by red color, and the finally extracted potential nodule candidate (c)

$h(x, y, z) = (1 + |\nabla I(x, y, z)|)^{-1}$. To keep the level set front from propagating through blood vessels to which the nodules may be connected, we apply a low-pass filter after each propagation step n . Figure 2 shows results of extracting a potential nodule candidate with the deformable prototype. To check whether the extracted object is really a nodule candidate, we should measure similarity between grayscale patterns in the extracted part of the initial 3D image and the intensity prototype of the nodule of that shape.

2.2 Similarity Measure for Grayscale Nodule Prototypes

Analysis of abnormalities in real 3D LDCT slices suggests that gray levels in central cross-sections of a solid-shape 3D nodule or in a solid-shape 2D nodule roughly follow a central-symmetric Gaussian spatial pattern such that the large central intensity gradually decreases towards the boundary. Moreover, the marginal gray level distributions for all 3D objects separated from the lung tissues at the second segmentation stage (e.g. arteries, veins, bronchi, or nodules of different size) are very similar to each other. The 3D Gaussian intensity pattern in each grayscale nodule prototype ensures that the marginal gray level distribution closely approximates the empirical one for each real nodule in the LDCT data.

Let the prototype be a central-symmetric 3D Gaussian of radius R with the maximum intensity q_{\max} in the center so that the gray level $q(r)$ at any location (x, y, z) at radius $r = (x^2 + y^2 + z^2)^{\frac{1}{2}}$ with respect to the center $(0, 0, 0)$ is given by the obvious relationship:

$$q(r) = q_{\max} \exp\left(-\left(\frac{r}{\rho}\right)^2\right); \quad 0 \leq r \leq R \quad (2)$$

The scatter parameter ρ in Eq. (2) specifies how fast the signals decrease towards the boundary of the prototype. The maximum gray level, $q_{\max} = q(0)$, and the minimum one, $q_{\min} = q(R)$, on the boundary of the spherical Gaussian prototype of the radius R uniquely determine this parameter as follows:

$$\rho = R (\ln q_{\max} - \ln q_{\min})^{-\frac{1}{2}} \quad (3)$$

Because all the prototype's points with a fixed gray value q in the continuous interval $[q_{\min}, q_{\max}]$ are located at the spherical surface of the radius $r(q) = \rho(\ln q_{\max} - \ln q)^{\frac{1}{2}}$, their density is proportional to the surface area $4\pi r^2(q)$. Therefore, the marginal probability density function for such a prototype is $\psi(q) = \gamma r^2(q)$ where γ is the normalizing factor such that $\int_{q_{\min}}^{q_{\max}} \psi(q) dq = 1$. It is easily shown that this function has the following closed form:

$$\psi(q|q_{\min}, q_{\max}) = \frac{\ln q_{\max} - \ln q}{q_{\max} - q_{\min} (1 + \ln q_{\max} - \ln q_{\min})} \quad (4)$$

The gray level parameters q_{\max} and q_{\min} are estimated from the empirical marginal distribution for each segmented 3D object, e.g. $q_{\max} = 255$ and $q_{\min} = 98$ for the objects extracted in Fig. 1.

To evaluate similarity, the gray level nodule prototype is centered at the centroid of the volume extracted with the deformable prototype.

2.3 Lung Nodule Detection Algorithm

1. Separate lung regions from a given CT scan using the segmentation algorithms in [15]; see an example in Fig. 1(b).
2. Separate arteries, veins, bronchi, bronchioles, and lung nodules (if they exist) from the above lung regions using the same segmentation algorithms in [15]; see an example in Fig. 1(c).
3. From the empirical marginal gray level distribution for the objects separated at Step 2 calculate q_{\min} and q_{\max} ; see an example in Fig. 1(d).
4. Stack all the voxels separated at Step 2.
5. Pop-up a top voxel from the stack as a seed for the deformable prototype and let this latter propagate until reaching a steady state indicating that the voxel set \mathbf{U} enclosed by the final prototype constitutes an extracted object.
6. Calculate the centroid for the voxel set \mathbf{U} extracted at the previous step; find the maximum, R_{\max} , and the minimum, R_{\min} , radii from the centroid of the boundary of that set; find the average radius $R = (R_{\min} + R_{\max})/2$, and estimate the scatter parameter ρ from Eq. (3).
7. Use Eq. (2) to assign the prototype gray levels $N_{x,y,z}$ for each extracted voxel $(x, y, z) \in \mathbf{U}$.
8. Use the normalized cross-correlation $\text{Corr}_{\mathbf{C},\mathbf{N}}$ between the actual extracted object $\mathbf{C} = [C_{x,y,z} : (x, y, z) \in \mathbf{U}]$ and its gray level nodule prototype $\mathbf{N} = [N_{x,y,z} : (x, y, z) \in \mathbf{U}]$ as the similarity measure (see Fig. 3).
9. If $\text{Corr}_{\mathbf{C},\mathbf{N}} \geq \tau$ where τ is a pre-selected similarity threshold (in our experiments below we set it to $\tau = 0.85$ on the empirical basis), then classify the extracted object as the potential nodule candidate.
10. Remove all the voxels of the extracted object from the stack.
11. If the stack is empty then stop, otherwise go to Step 5.

To reduce the error rate, the initially selected potential candidates are post-classified just as in [19] to distinguish between the false (FPNs) and true (TPNs)

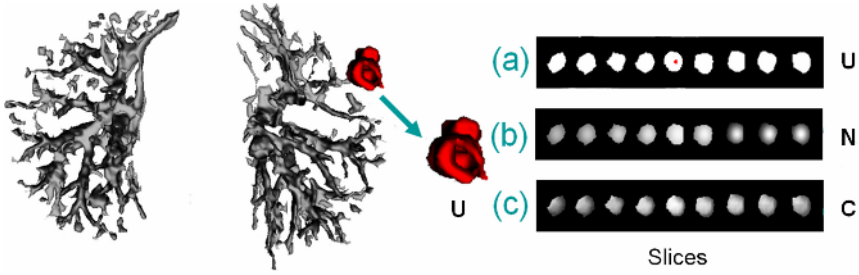


Fig. 3. Detected nodule candidate: (a) slices of the extracted voxel set U , (b) the gray level prototype N , and (c) the actual gray levels C ; the correlation $\text{Corr}_{C,N} = 0.886$

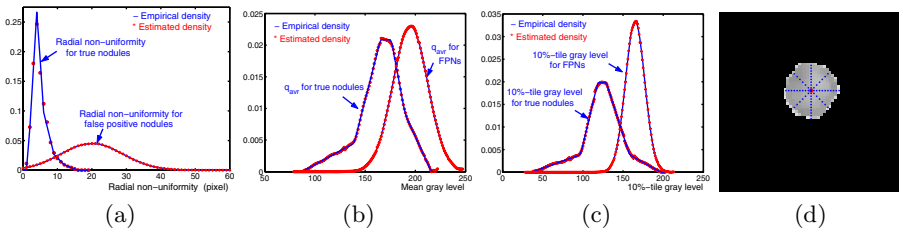


Fig. 4. Estimated and empirical distributions for radial non-uniformity (a), mean gray level (b), and 10%-tile gray level (c), and the radii d_θ in 8 directions θ from the centroid to the boundary that specify the radial non-uniformity $\max_\theta(d_\theta) - \min_\theta(d_\theta)$

positive nodules. In contrast to [19], probability distributions of each feature required by the classifier are accurately estimated with linear combinations of Gaussians (LCGs) using our algorithms in [15]. The LCGs approximate the empirical distributions for a training set of the nodules. Figure 4 shows the empirical and estimated distributions of each feature for both TPNs and FPNs.

3 Experimental Results and Conclusions

The proposed algorithm was tested on the same LDCT scans of 50 screened subjects. Among them, 16 subjects had abnormalities in their CT scans and 34 subjects were normal (this classification was validated by two radiologists). The chest CT data used in this paper were obtained from Mansoura University, Urology and Nephrology Center, Radiology department, Mansoura, Egypt as follows: 1) The scanner: Multidetector scanner (Light speed plus; GE), 2) Scan mode: Helical, 3) Slice thickness: 2.5mm, 4) Field of view: large, 5) K.V.: 120, 6) M.A.: 200, 7) Exposure time: 20-25 sec., and 8) Window level: -500 & length 1500. Our approach extracted 113 potential nodule candidates out of the true 119 nodules and 14 FPNs. The post-classification has reduced the number of FPNs to 4 but simultaneously rejected two true nodules. Thus, the final detection rate of the TPNs was 93.3% (111 out of 119) with the FPNs rate of 3.36%. Both the

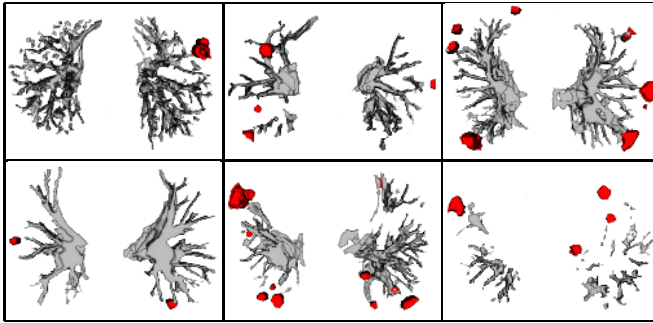


Fig. 5. Large candidate nodules (shown in the red colour) detected with our approach

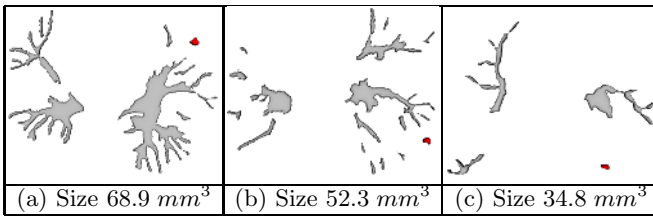


Fig. 6. Small candidate nodules (shown in the red colour) detected with our approach

Table 1. Recognition of different abnormalities with our present algorithm, the previous one in [19], and the algorithm in [5] w.r.t. the “ground truth” by two radiologists (the corresponding numbers are bold, bold italic, italic, and regular, respectively)

Type of the nodules	True detected nodules before removing FPNs	False detected nodules before removing FPNs	True detected nodules after removing FPNs	False detected nodules after removing FPNs
Lung wall	28 : 28 : <i>14</i> : 29	3 : 8 : <i>86</i>	28 : 27 : <i>13</i> : 29	1 : 2 : <i>17</i>
Calcified	49 : 46 : <i>31</i> : 49	2 : 4 : <i>35</i>	48 : 46 : <i>31</i> : 49	0 : 1 : <i>9</i>
Non-calcified	17 : 12 : <i>14</i> : 18	3 : 5 : <i>25</i>	17 : 12 : <i>14</i> : 18	1 : 3 : <i>14</i>
Small	19 : 17 : <i>10</i> : 23	6 : 25 : <i>34</i>	18 : 15 : <i>9</i> : 23	2 : 5 : <i>12</i>
Total TPN rate			93.3% : 82.3% : <i>56.3%</i>	
Total FPN rate			3.4% : 9.2% : <i>43.6%</i>	

rates are notably better than 82.3% and 9.2% in [15], respectively. Figure 5 shows several large lung nodules detected by our approach, and a few small detected TPNs are depicted in Fig. 6.

Table 1 presents the above results in detail and also compares them to the algorithm in [5], the only work known to us on detecting the lung nodules from the LDCT scans using template matching. It uses fixed circular, semicircular and approximate spherical templates with parameters which are *manually* derived from a training data set. This algorithm, after the post-classification, gives the

overall TPN rate of 56.3% (67 out of 119) with the FPN rate of 43.6%. It is felt that this algorithm fails to detect a large number of the true nodules because of the fixed-size templates and manual specification of their gray level patterns. At times these patterns change from one LDCT slice to another depending on the scanned cross section and internal organs that appear in that cross section. In our present algorithm, these patterns are analytically adjusted to each extracted shape by applying Eq. (3) to each CT slice.

These experiments show that our new deformable level-set prototype with the analytically modeled standard intensity pattern detects more than 90% of the true lung abnormalities. The overall processing time for the data set of size $1005 \times 804 \times 186$ is 6, 18, and 16 min. for the proposed approach and the algorithms in [19, 5], respectively. But it is still difficult to accurately detect very small lung nodules similar to bronchi and bronchioles. Our future work is focused on features that distinguish between the small lung nodules and normal objects. Also, we are going to analyze the accuracy of our CAD system with respect to the ground truth on a much larger number of the real LDCT scans.

References

1. S. C. B., M. T. Freedman, J. S. Lin, and S. K. Mun, "Automatic lung nodule detection using profile matching and back-propagation neural network techniques," *J. Digital Imaging*, Vol. 6, pp. 48-54, 1993.
2. F. Mao, W. Qian, J. Gaviria, and L. Clarke, "Fragmentary window filtering for multiscale lung nodule detection," *Academic Radiology*, Vol. 5, pp. 306-311, 1998.
3. X. Xu, S. Katsuragawa, K. Ashizawa, H. MacMahon, and K. Doi, "Analysis of image features of histograms of edge gradient for false positive reduction in lung nodule detection in chest radiographs," *Proc. SPIE*, Vol. 3338, pp. 318-326, 1998.
4. S. Chang, H. Emoto, D. N. Metaxas, and L. Axe, "Pulmonary Micronodule detection from 3D Chest CT," *Proc. MICCAI, Saint-Malo, France, Sept. 26-29, 2004*, pp. 821-828, 2004.
5. Y. Lee, T. Hara, H. Fujita, S. Itoh, and T. Ishigaki, "Automated Detection of Pulmonary Nodules in Helical CT Images Based on an Improved Template-Matching Technique," *IEEE Trans. on Medical Imaging*, Vol. 20, pp. 595-604, 2001.
6. A. El-Baz, A. A. Farag, R. Falk, and R. La Rocca, "Automatic Identification of Lung Abnormalities in Chest Spiral CT Scans," *International Conference on Acoustics, Speech, and Signal Processing, ICASSP-03, April 6-10, 2003*, pp. 261-264.
7. K. Kanazawa, Y. Kawata, N. Niki, H. Satoh, H. Ohmatsu, R. Kakinuma, M. Kaneko, N. Moriyama, and K. Eguchi, "Computer-aided diagnosis for pulmonary nodules based on helical CT images," *Computerized Medical Imaging and Graphics*, vol. 22, pp. 157-167, 1998.
8. Y. Kawata, N. Niki, H. Ohmatsu, M. Kusumoto, R. Kakinuma, K. Mori, H. Nishiyama, K. Eguchi, M. Kaneko, N. Moriyama, "Computer-Aided Diagnosis of Pulmonary Nodules Using Three-Dimensional Thoracic CT Images," *Proc. MICCAI, Utrecht, Netherlands, October 14-17, 2001*, pp. 1393-1394, 2001.
9. R. M. Betke and J. P. Ko, "Detection of Pulmonary Nodules on CT and Volumetric Assessment of Change over Time," *Proc. MICCAI, Cambridge, UK, 1999*, pp. 245-252, 1999.

10. M. S. Brown, M. F. McNitt-Gray, J. G. Goldin, R. D. Suh, J. W. Sayre, and D. R. Aberle "Patient-Specific Models for Lung Nodule Detection and Surveillance in CT Images," *IEEE Trans. on Medical Imaging*, vol. 20, pp. 1242–1250, 2001.
11. H. Takizawa, K. Shigemoto, S. Yamamoto, T. Matsumoto, Y. Tateno, T. Iinuma and M. Matsumoto, "A Recognition Method of Lung Nodule Shadows in X-Ray CT Images Using 3D Object Models," *Int. J. of Image and Graphics*, vol. 3, pp. 533-545, 2003.
12. X. Xu et al., S. Katsuragawa, K. Ashizawa, H. MacMahon, and K. Doi, "Analysis of image features of histograms of edge gradient for false positive reduction in lung nodule detection in chest radiographs," *Proc. SPIE*, Vol. 3338, pp. 318-326, 1998.
13. M. J. Carreira, D. Cabello, M. G. Penedo, and J. M. Pardo, "Computer aided lung nodule detection in chest radiography," *Proc. Conf. Image Analysis Applications and Computer Graphics*, Berlin, pp. 331–338, 1995.
14. S. C. B. Lo, S. L. A. Lou, J. S. Lin, M. T. Freedman, M. V. Chien, and S. K. Mun, "Artificial convolution neural network techniques and applications for lung nodule detection," *IEEE Trans. on Medical Imaging*, vol. 14, 1995, pp. 711-718.
15. G. Gimel'farb, A. A. Farag, and A. El-Baz, "Expectation-Maximization for a linear combination of Gaussians", in *Proc. IAPR Int. Conf. Pattern Recognition (ICPR 2004)*, Cambridge, UK, August 23-26, 2004, vol.2, pp. 422–425, 2004.
16. J.A. Sethian, *Level Set Methods and Fast Marching Methods*, Cambridge University Press, USA, 1999.
17. J. Gomes and O. Faugeras, *Reconciling distance functions and level-sets*. Technical Report 3666, INRIA, April 1999.
18. N. Paragios and R. Deriche, "Unifying boundary and region-based information for geodesic active tracking," *Proc. CVPR, Fort Collins, Colorado, June 1999*, Vol. 2, pp. 300–305, 1999.
19. A. A. Farag, A. El-Baz, G. Gimel'farb, R. Falk, and S. G. Hushek, "Automatic detection and recognition of lung abnormalities in helical CT images using deformable templates," *Proc. MICCAI, Saint-Malo, France, Sept. 26–29, 2004*, pp. 856-864, 2004.

Graph Embedding to Improve Supervised Classification and Novel Class Detection: Application to Prostate Cancer

Anant Madabhushi¹, Jianbo Shi², Mark Rosen²,
John E. Tomaszewski², and Michael D. Feldman²

Rutgers University, Piscataway, NJ 08854,
University of Pennsylvania, Philadelphia, PA 19104
`anantm@rci.rutgers.edu`

Abstract. Recently there has been a great deal of interest in algorithms for constructing low-dimensional feature-space embeddings of high dimensional data sets in order to visualize inter- and intra-class relationships. In this paper we present a novel application of graph embedding in improving the accuracy of supervised classification schemes, especially in cases where object class labels cannot be reliably ascertained. By refining the initial training set of class labels we seek to improve the prior class distributions and thus classification accuracy. We also present a novel way of visualizing the class embeddings which makes it easy to appreciate inter-class relationships and to infer the presence of new classes which were not part of the original classification. We demonstrate the utility of the method in detecting prostatic adenocarcinoma from high-resolution MRI.

1 Introduction

The aim of embedding algorithms is to construct low-dimensional feature-space embeddings of high-dimensional data sets [1–4]. The low-dimensional representation is easier to visualize and helps provide easily interpretable representations of intra-class relationships, so that objects that are closer to one another in the high dimensional ambient space are mapped to nearby points in the output embedding. Recently researchers have begun exploring the use of embedding for solving different problems. Dhillon [1] employed embedding for visually understanding the similarity of different classes, distance between class clusters, and to evaluate the coherence of each of the class clusters. Iwata *et al.* [2] described a parametric embedding method to provide insight into classifier behavior. Euclidean embedding of co-occurrence data has also been successfully applied to classifying text databases [3] and for detecting unusual activity [4]. In this paper we demonstrate a novel application of graph embedding in (i) improving the accuracy of supervised classification tasks and (ii) for identifying *novel classes*, i.e. classes not included in the original classification.

In [5] we presented a computer-aided detection (CAD) methodology for detecting prostatic adenocarcinoma from high resolution MRI, which in several

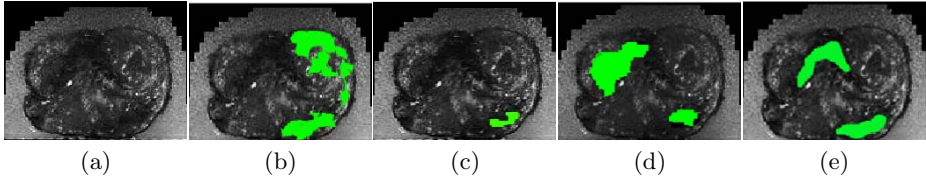


Fig. 1. (a) Original MR image of the prostate, (b) ground truth for tumor (in green) determined manually from the corresponding histology [5]. Three expert segmentations (Fig. 1(c)-(e)) based on visual inspection of Fig. 1(a) without the accompanying histologic information. Note the low levels of inter-expert agreement.

instances outperformed trained experts. It was found that the false positive errors due to CAD were on account of,

- Errors in tumor ground truth labels on MRI, since the tumor labels were established by manually registering the corresponding histologic and MR slices (both MR and histologic slices being of different slice thickness) and due to the difficulty in identifying cancer (Fig. 1).
- The presence of objects with characteristics between tumor and non-tumor (e.g. pre-cancerous lesions). Since the system is not trained to recognize these novel classes, the classifier *forces* these objects into one of the original classes, contributing to false positives.

In order to detect novel classes, we need to first eliminate true outliers due to human errors from the training set. The implications of outlier removal from the training set are two fold.

- (1) It can significantly improve the accuracy of the original classification, and
- (2) It ensures that objects that now lie in the overlap between the object classes after outlier removal, truly represent the novel classes.

We borrow a graph embedding technique used in the computer vision domain [4] for improving classification accuracy and for novel class detection. However, while in [4] both the object and its representative class are co-embedded into the low-dimensional space, in our case, the embedding algorithm takes as input the *a-posteriori* likelihoods of objects belonging to the tumor class. Note that graph embedding differs from data reduction techniques like PCA [7] in that the relationship between adjacent objects in the higher dimensional space is preserved in the co-embedded lower dimensional space. While we have focused on one specific CAD application in this paper [5], we emphasize that our methods are applicable to most supervised or semi-supervised classification tasks, especially those in which class labels cannot be reliably ascertained.

This paper is organized as follows. Section 2 provides a brief description of our methodology, and a detailed description of the individual modules is given in Section 3. In Section 4 we present the results of quantitative evaluation of our methodology on a CAD system for prostate cancer. Concluding remarks are presented in Section 5.

2 System Overview for Prostate Cancer Detection

Fig. 2 shows the main modules and pathways comprising the system. Fig. 3 shows the results of graph embedding on a high resolution MR study of the prostate (Fig. 3(a)). Fig. 3(c) is a map of the posterior likelihood of every voxel belonging to the tumor class; the posterior likelihood being derived from the prior distribution (dashed line in Fig. 3(f)), obtained with the initial set of tumor class labels (Fig. 3(b)) and Fig. 3(e) shows the corresponding *probability* image using the refined prior distribution after graph embedding (solid line in Fig. 3(f)). The plot of graph embedding (Fig. 3(d)) shows considerable overlap (ellipse 3) between the tumor (red circles) and non-tumor (black dots) classes. Using the refined probability map in Fig. 3(e), the resultant embedding (Fig. 3(f)) shows a clear separation between the two classes (ellipses 1, 2). The increased class separation is also reflected in the increased image contrast of Fig. 3(e) over Fig. 3(c). Fig. 3(g) shows a novel way of visualizing the graph embeddings in Fig. 3(f), with objects that are adjacent in the embedding space being assigned similar colors. Objects that lie in the overlap of the class clusters after outlier removal (ellipse 3 in Fig. 3(f)) correspond to the apparent false positive area (marked as FP) in Fig. 3(g). This region is actually inflammation induced by atrophy (confirmed via the histology slice in Fig. 3(h)).

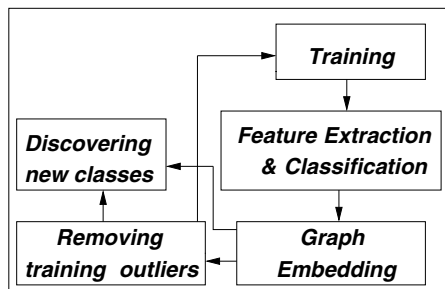


Fig. 2. Training distributions for individual features are generated using existing class labels, and each voxel assigned a posterior likelihood of being tumor. Graph embedding on the posterior likelihoods is used to remove training outliers and (i) improve the prior distributions and (ii) identify new classes.

3 Methodology

3.1 Notation

We represent a 3D image or *scene* by a pair $\mathcal{C} = (C, g)$, where C is a finite 3-dimensional rectangular array of voxels, and g is a function that assigns an integer intensity value $g(c)$ for each voxel $c \in C$. The feature scenes $\mathcal{F}^i = (C, f^i)$ are obtained by application of K different feature operators, for $1 \leq i \leq K$. The tumor class is denoted by ω_t and S_{ω_t} denotes the *true ground truth* set, such that for any voxel $d \in S_{\omega_t}$, $d \hookrightarrow \omega_t$ where \hookrightarrow denotes the “belongs to” relationship. \hat{S}_{ω_t} is the surrogate of ground truth S_{ω_t} obtained by experts by visually registering the MR and the histologic slices [5]. $\hat{S}_{\omega_t}^T \subset \hat{S}_{\omega_t}$ is the training set used for generating the prior distributions $\hat{p}(f^i | c \hookrightarrow \omega_t)$ for each feature f^i . Given $\hat{p}(f^i | c \hookrightarrow \omega_t)$, the *a-posteriori* probability that voxel $c \hookrightarrow$

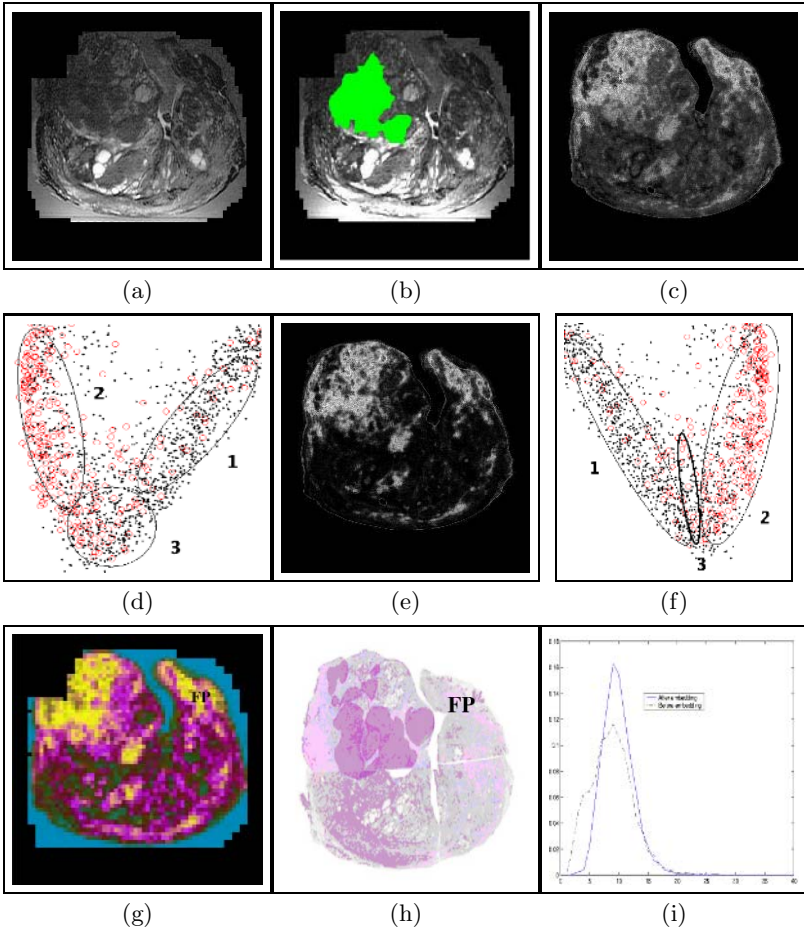


Fig. 3. (a) Original MR scene \mathcal{C} , (b) surrogate of ground truth (in green) for cancer (\hat{S}_{ω_t}) superposed on (a), (c) combined likelihood scene showing tumor class probability before outlier refinement via embedding, (d) graph embedding of tumor/non-tumor class likelihoods in (c), (e) combined likelihood scene showing tumor class probabilities after outlier removal, (f) graph embedding of tumor/non-tumor class likelihoods in (e), (g) RGB representation of graph embeddings in (f), and (h) the histology slice corresponding to the MR slice in (a). Note the greater contrast between intensities in (e) compared to (c), reflecting the increased separation between the tumor and non-tumor clusters after outlier removal. This is also reflected in the overlap of the tumor (red circles) and non-tumor (black dots) clusters in the embedding plot before outlier removal (ellipse 3 in (d)) and the more distinct separation of the two clusters after outlier removal (ellipse 3 in (f)), constitute the intermediate class (between tumor and non-tumor). Also note the tighter envelope of the prior distribution of feature f^i (3(i)) after embedding (solid line) compared to before (dashed line). The embedding scene in 3(g) also reveals that an apparent false positive area (FP on 3(g)) actually corresponds to a new object class not included in the original classification (inflammation induced by atrophy, confirmed via the histology slice (h)).

ω_t for f^i is given as $\hat{P}(c \leftrightarrow \omega_t | f^i)$. $\hat{P}(c \leftrightarrow \omega_t | \mathbf{f})$, for $\mathbf{f} = [f^i | i \in \{1, \dots, K\}]$, is the combined posterior likelihood obtained by combining $\hat{P}(c \leftrightarrow \omega_t | f^i)$, for $1 \leq i \leq K$. $\tilde{p}(f^i | c \leftrightarrow \omega_t)$, $\tilde{P}(c \leftrightarrow \omega_t | f^i)$, and $\tilde{P}(c \leftrightarrow \omega_t | \mathbf{f})$ denote the corresponding prior, posterior, and combined posterior likelihoods obtained after refinement by embedding. $\hat{\mathcal{L}} = (C, \hat{l})$ denotes the combined likelihood scene (Fig. 3(d)), such that for $c \in C$, $\hat{l}(c) = \hat{P}(c \leftrightarrow \omega_t | \mathbf{f})$. $\tilde{\mathcal{L}} = (C, \tilde{l})$, where for $c \in C$, $\tilde{l}(c) = \tilde{P}(c \leftrightarrow \omega_t | \mathbf{f})$, similarly denotes the corresponding likelihood scene (Fig. 3(e)) after refinement by graph embedding.

3.2 Feature Extraction and Classification

A total of 35 3D texture feature scenes $\mathcal{F}^i = (C, f^i)$, for $1 \leq i \leq 35$, are obtained from the MR scene \mathcal{C} . The extracted features include 7 first order statistical features at two scales, 8 Haralick features at two scales, 2 gradient features, and 18 Gabor features corresponding to 6 different scales and 3 different orientations. A more detailed description of the feature likelihood extraction methods has been previously presented in [5]. The *a-posteriori* likelihoods $\hat{P}(c \leftrightarrow \omega_j | f^i)$ for each feature f^i can be computed using Bayes Theorem [6] as, $\hat{P}(c \leftrightarrow \omega_j | f^i) = \hat{P}(c \leftrightarrow \omega_j) \frac{\hat{p}(f^i | c \leftrightarrow \omega_j)}{\hat{p}(f^i)}$, where $\hat{P}(c \leftrightarrow \omega_j)$ is the *a-priori* probability of observing the class ω_j , $\hat{p}(f^i) = \sum_{j=1}^{\mathcal{B}} \hat{p}(f^i | c \leftrightarrow \omega_j) \hat{P}(c \leftrightarrow \omega_j)$, where \mathcal{B} refers to the number of classes. The combined posterior likelihood $\hat{P}(c \leftrightarrow \omega_j | \mathbf{f})$, for $\mathbf{f} = [f^i | i \in \{1, \dots, K\}]$, can be obtained from $\hat{P}(c \leftrightarrow \omega_j | f^i)$, by using any of the various feature ensemble methods, e.g. ensemble averaging, GEM [5], majority voting.

3.3 Graph Embedding for Analyzing Class Relationships

Our aim is to find a placement (embedding) vector $\hat{X}(c)$ for each voxel $c \in C$ and the tumor class ω_t such that the distance between c and class ω_t is monotonically related to the *a-posteriori* probability $\hat{P}(c \leftrightarrow \omega | \mathbf{f})$ in a low-dimensional space [2]. Hence if voxels $c, d \in C$ both belong to class ω_t , then $[\hat{X}(c) - \hat{X}(d)]^2$ should be small. To compute the optimal embedding, we first define a confusion matrix W representing the similarity between any two objects $c, d \in C$ in a high dimensional feature space.

$$W(c, d) = e^{-\|\hat{P}(c \leftrightarrow \omega_t | \mathbf{f}) - \hat{P}(d \leftrightarrow \omega_t | \mathbf{f})\|} \in \mathcal{R}^{|C| \times |C|} \tag{1}$$

Computing the embedding is equivalent to optimization of the following function,

$$E_W(\hat{X}) = \frac{\sum_{(c,d) \in C} W(c, d) (\hat{X}(c) - \hat{X}(d))^2}{\sigma_{\hat{X}}^2} \tag{2}$$

Expanding the numerator of (2) we get $2\hat{X}^T(D - W)\hat{X}$, where $D(c, d) \in \mathcal{R}^{|C| \times |C|}$ is a diagonal matrix with $D(c, c) = \sum_d W(c, d)$. Using the fact that

$$\sigma_{\hat{X}}^2 = \sum_{c \in C} \hat{X}^2(c) \hat{P}(c \leftrightarrow \omega_t | \mathbf{f}) - \left(\sum_{c \in C} \hat{X}(c) \hat{P}(c \leftrightarrow \omega_t | \mathbf{f}) \right)^2, \tag{3}$$

it can be shown that $\hat{P}(c \hookrightarrow \omega_t | \mathbf{f}) \approx \frac{1}{\gamma} D(c, c)$, where $\gamma = |C| - 1$, and $|C|$ represents the cardinality of set C . Centering the embedding around zero (i.e. $\hat{X}^T \gamma = 0$), we get $\sigma_{\hat{X}}^2 = \frac{1}{\gamma} \hat{X}^T D \hat{X}$. Putting all these together we can rewrite (2) as,

$$E_W(\hat{X}) = 2\gamma \frac{\hat{X}^T (D - W) \hat{X}}{\hat{X}^T D \hat{X}}. \tag{4}$$

The global energy minimum of this function is achieved by the eigenvector corresponding to the second smallest eigenvalue of,

$$(D - W) \hat{X} = \lambda D \hat{X}. \tag{5}$$

For voxel $c \in C$, the embedding $\hat{X}(c)$ contains the coordinates of c in the embedding space and is given as, $\hat{X}(c) = [\hat{e}_a(c) | a \in \{1, 2, \dots, \beta\}]$, where $\hat{e}_a(c)$, are the eigen values associated with c .

3.4 Improving Training Distributions by Refining Ground Truth

In several classification tasks (especially in medical imaging), S_{ω_t} , the set of *true ground truth* class labels is not available. For the CAD problem tackled in this work, only an approximation of the ground truth (\hat{S}_{ω_t}) is available, so that there exist objects $d \in \hat{S}_{\omega_t}$ which do not belong to class ω_t . Consequently the prior distributions $\hat{p}(f^i | c \hookrightarrow \omega_t)$, for $1 \leq i \leq K$, and the posterior probabilities $\hat{P}(c \hookrightarrow \omega_t | f^i)$ reflect the errors in \hat{S}_{ω_t} , since $\hat{p}(f^i | c \hookrightarrow \omega_t)$ is generated from a training set $\hat{S}_{\omega_t}^T \subset \hat{S}_{\omega_t}$. Clearly a more accurate estimate (\tilde{S}_{ω_t}) of S_{ω_t} would result in more accurate prior distributions $\tilde{p}(f^i | c \hookrightarrow \omega_t)$, for $1 \leq i \leq K$, and consequently a more accurate posterior likelihoods $\tilde{P}(c \hookrightarrow \omega_t | f^i)$. To obtain \tilde{S}_{ω_t} we proceed as follows,

- (1) The embedding of all voxels $c \in C$, $\hat{X}(C)$ is determined.
- (2) The K -means algorithm is applied on the embedding coordinates $\hat{X}(C)$ to cluster objects $c \in C$ into Z disjoint partitions $\{\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_Z\}$.
- (3) We obtain the union of those disjoint partitions \mathcal{P}_z , for $1 \leq z \leq Z$, sizes of which are above a pre-determined threshold θ . The rationale behind this is that outliers will be partitioned into small sets. \tilde{S}_{ω_t} is then obtained as,

$$\tilde{S}_{\omega_t} = \hat{S}_{\omega_t} \bigcap \left[\bigcup_z \mathcal{P}_z \right], \text{ where } |\mathcal{P}_z| \geq \theta, \text{ for } z \in \{1, 2, \dots, Z\}. \tag{6}$$

The intuition behind Equation 6 is that we only consider objects in \hat{S}_{ω_t} for inclusion into \tilde{S}_{ω_t} . This avoids inclusion of potentially new outliers. Note that, since this procedure is only for the training step, we are not concerned with including every object in class ω_t into \tilde{S}_{ω_t} . Instead, our aim is to ensure as far as possible that for every object $c \in \hat{S}_{\omega_t}$, $c \hookrightarrow \omega_t$.

- (4) New *a-priori* distributions $\tilde{p}(f^i | c \hookrightarrow \omega_t)$, for $1 \leq i \leq K$, are then generated from training set $\tilde{S}_{\omega_t}^T \subset \tilde{S}_{\omega_t}$ and the new posterior likelihoods $\tilde{P}(c \hookrightarrow \omega_t | f^i)$ and combined likelihood $\tilde{P}(c \hookrightarrow \omega_t | \mathbf{f})$, for $\mathbf{f} = [f^i | i \in \{1, \dots, K\}]$, are computed.

Fig. 3(c), (e) correspond to the *likelihood scenes* ($\hat{\mathcal{L}}, \tilde{\mathcal{L}}$) obtained from distributions $\hat{p}(f^i|c \rightarrow \omega_t)$ and $\tilde{p}(f^i|c \rightarrow \omega_t)$ respectively. The intensity at every voxel $c \in C$ in Fig. 3(c), (e) is given by the *a-posteriori* likelihoods $\hat{P}(c \rightarrow \omega_t|\mathbf{f})$ and $\tilde{P}(c \rightarrow \omega_t|\mathbf{f})$, for $\mathbf{f} = [f^i|i \in \{1, \dots, K\}]$, respectively. While Fig. 3(e) is almost a bi-level image, suggesting distinct separation between the tumor and non-tumor classes, Fig. 3(c) is more fuzzy, indicating considerable overlap between the two classes. This is reflected in the plot of class embeddings $\hat{X}(C)$ obtained from $\hat{P}(c \rightarrow \omega_t|\mathbf{f})$ in which considerable overlap (ellipse 3) exists between the two classes (Fig. 3(d)), while in the plot of $\tilde{X}(C)$, the graph embedding obtained from $\tilde{P}(c \rightarrow \omega_t|\mathbf{f})$ (Fig. 3(f)), there is a more distinct separation of class clusters.

3.5 Discovering Novel Classes

Even after removing outliers from the ground truth, there exist objects that occupy the transition between tumor and non-tumor clusters (observe ellipse 3 in Fig. 3(f)), suggesting that the characteristics of these objects are between that of the tumor and benign classes. In Fig. 3(g) is shown a novel way of visualizing and identifying objects from these intermediate classes. Since $\tilde{X}(c)$ contains the embedding coordinates of voxel c , we can represent $\tilde{X}(C)$, the embedding over scene C , as a RGB image in which the value at voxel c is given by the three principal eigen values associated with c . Objects that are adjacent to each other in the embedding space have a similar color (Fig. 3(g)). The apparent false positive area (labeled as FP on Fig. 3(g)), on inspecting the corresponding histology slice (Fig. 3(h)) was found to be inflammation induced by atrophy on account of a prior needle insertion. This new class had not been considered in our original two class classification paradigm.

3.6 Algorithm

- For each scene we compute the corresponding feature scenes for each feature f^i .
- Prior distributions $\hat{p}(f^i|c \rightarrow \omega_t)$ for each feature f^i for class ω_t are obtained using training set $\hat{S}_{\omega_t}^T \subset \hat{S}_{\omega_t}$.
- Bayes Theorem [6] is used to compute posterior likelihoods $\hat{P}(c \rightarrow \omega_t|f^i)$, for $1 \leq i \leq K$. Combined likelihood $\hat{P}(c \rightarrow \omega_t|\mathbf{f})$, for $\mathbf{f} = [f^i|i \in \{1, \dots, K\}]$ is then computed from $\hat{P}(c \rightarrow \omega_t|f^i)$ using any standard ensemble method.
- Confusion matrix W is computed for $c, d \in C$ as $W(c, d) = e^{-|\hat{P}(c \rightarrow \omega_t|\mathbf{f}) - \hat{P}(d \rightarrow \omega_t|\mathbf{f})|} \in \mathcal{R}^{|C| \times |C|}$. Solve for the smallest eigen vectors of $(D - W)\hat{X} = \lambda D\hat{X}$ where the rows of the eigen vectors are the coordinates for the object c in the embedding space $\hat{X}(C)$.
- Partition $\hat{X}(C)$ into disjoint regions \mathcal{P}_z , for $1 \leq z \leq Z$, and compute new set of tumor class objects $\tilde{S}_{\omega_t} = \hat{S}_{\omega_t} \cap [\bigcup_z \mathcal{P}_z]$, where $|\mathcal{P}_z| \geq \theta$.
- Generate new prior distributions $\tilde{p}(f^i|c \rightarrow \omega_t)$, for $1 \leq i \leq K$, from new training set $\tilde{S}_{\omega_t}^T \subset \tilde{S}_{\omega_t}$ and compute new posterior likelihoods $\tilde{P}(c \rightarrow \omega_t|f^i)$ and combined posterior likelihood $\tilde{P}(c \rightarrow \omega_t|\mathbf{f})$, for $\mathbf{f} = [f^i|i \in \{1, \dots, K\}]$.

4 Evaluating CAD Accuracy for Prostate Cancer on MRI

The likelihood scene $\hat{\mathcal{L}}$ is thresholded to obtain binary scene $\hat{\mathcal{L}}_B = (C, \hat{l}_B)$ so that for $c \in C$, $\hat{l}_B(c)=1$ iff $\hat{l}(c) \geq \delta$, where δ is a pre-determined threshold. $\tilde{\mathcal{L}}_B$ is similarly obtained. $\hat{\mathcal{L}}_B$ and $\tilde{\mathcal{L}}_B$ are then compared with \hat{S}_{ω_t} and \tilde{S}_{ω_t} respectively to determine Sensitivity and Specificity values for different values of δ . Receiver operating characteristic (ROC) curves (plot of Sensitivity versus 100-Specificity) provide a means of comparing the performance of detection tasks. A larger area under the ROC curve implies higher accuracy. A total of 33 MR images of the prostate were used for quantitatively comparing $\tilde{\mathcal{L}}$ and $\hat{\mathcal{L}}$ for different values of δ . Fig. 4(a) and (b) show the ROC curves for $\hat{\mathcal{L}}$ (dashed line) and $\tilde{\mathcal{L}}$ (solid line) for two different feature combination methods (ensemble averaging and majority voting) using 5 and 10 training samples respectively. The accuracy of $\tilde{\mathcal{L}}$ was found to be significantly higher compared to $\hat{\mathcal{L}}$ for both classification methods and different sets of training samples, as borne out by the larger area under the ROC curves in Fig. 4(a) and (b). All differences were found to be statistically significant.

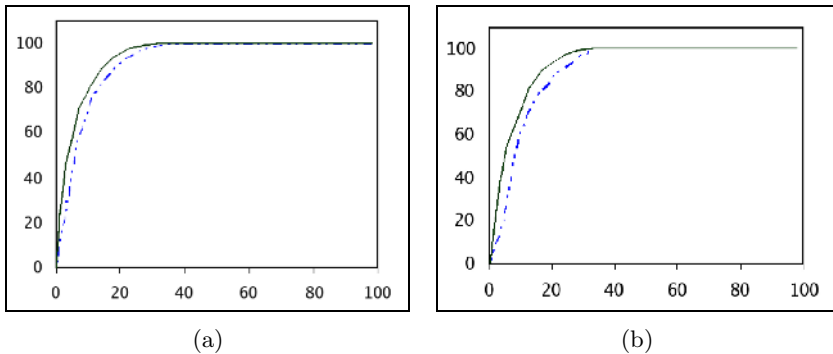


Fig. 4. ROC analysis of $\hat{\mathcal{L}}$ (dashed line) and $\tilde{\mathcal{L}}$ (solid line) using (a) ensemble averaging and 5 training samples, and (b) majority voting and 10 training samples

5 Concluding Remarks

In this paper we have presented a novel application of graph embedding in (i) improving the accuracy of supervised classification schemes, especially in cases where object class labels cannot be reliably ascertained, and (ii) for identifying novel classes of objects not present in the original classification. We have successfully employed this method to improve the accuracy of a CAD system for detecting prostate cancer from high resolution MR images. We were also able to identify a new class (inflammation due to atrophy). The method could be similarly used to detect pre-cancerous lesions, the presence of which has significant clinical implications.

References

1. I. Dhillon, D. Modha, W. Spangler, “Class Visualization of high-dimensional data with applications”, *Computational Statistics & Data Analysis*, 2002, vol. 41, pp. 59-90.
2. T. Iwata, K. Saito, *et al.*, “Parametric Embedding for Class Visualization”, *NIPS*, 2004.
3. A. Globerson, G. Chechik, *et al.*, “Euclidean Embedding of Co-occurrence Data”, *NIPS*, 2004.
4. H. Zhong, J. Shi and M. Visontai, Detecting Unusual Activity in Video, *CVPR*, 2004.
5. A. Madabhushi, M. Feldman, D. Metaxas, J. Tomaszeweski, D. Chute, Automated Detection of Prostatic Adenocarcinoma from High Resolution *in vitro* prostate MR studies, *IEEE Trans. Med. Imag.*, Accepted.
6. R. Duda, P. Hart, *Pattern Classification and Scene Analysis*, New York Wiley, 1973.
7. T. Joliffe, *Principal Component Analysis*, *Springer-Verlag*, 1986.

Quantification of Emphysema Severity by Histogram Analysis of CT Scans

Paulo R.S. Mendonça¹, Dirk R. Padfield¹, James C. Ross¹, James V. Miller¹, Sandeep Dutta², and Sardar Mal Gautham²

¹ GE Research Center, Niskayuna NY 12309, USA

² GE Healthcare, Waukesha WI 53188, USA

Abstract. Emphysema is characterized by the destruction and over distension of lung tissue, which manifest on high resolution computer tomography (CT) images as regions of low attenuation. Typically, it is diagnosed by clinical symptoms, physical examination, pulmonary function tests, and X-ray and CT imaging. In this paper we discuss a quantitative imaging approach to analyze emphysema which employs low-level segmentations of CT images that partition the data into perceptually relevant regions. We constructed multi-dimensional histograms of feature values computed over the image segmentation. For each region in the segmentation, we derive a rich set of feature measurements. While we can use any combination of physical and geometric features, we found that limiting the scope to two features – the mean attenuation across a region and the region area – is effective. The subject histogram is compared to a set of canonical histograms representative of various stages of emphysema using the Earth Mover's Distance metric. Disease severity is assigned based on which canonical histogram is most similar to the subject histogram. Experimental results with 81 cases of emphysema at different stages of disease progression show good agreement against the reading of an expert radiologist.

1 Introduction

Emphysema is a widespread chronic respiratory disorder, characterized by the loss of lung recoil resulting from the deterioration of lung tissue, which leads to physical limitations. The destruction and over-distension of lung tissue in emphysema are manifest on high resolution computer tomography (CT) images as regions of low attenuation, as seen in Fig. 1. While physicians are trained to identify such regions, it is often difficult to accurately quantify the extent of the disease severity, and both intra- and inter-operator variability is common [1]. Thus a variety of computer aided methods have been developed to automatically detect and quantify the extent of disease severity.

An early technique for measuring the amount of tissue degradation in the lung parenchyma is the Müller density index [2]. This technique calculates the percentage of the pixels in the lung field that have low attenuation. A global threshold defines what is considered low attenuation. This threshold must be set in a manner that takes into account variations in attenuation due to differences in scanners, scanning protocols, patient size, and the patients' ability to inflate their lungs.

While the Müller density index provides insight into the percentage of the lung that may be degraded by the disease, the interpretation is based strictly on the percentage of

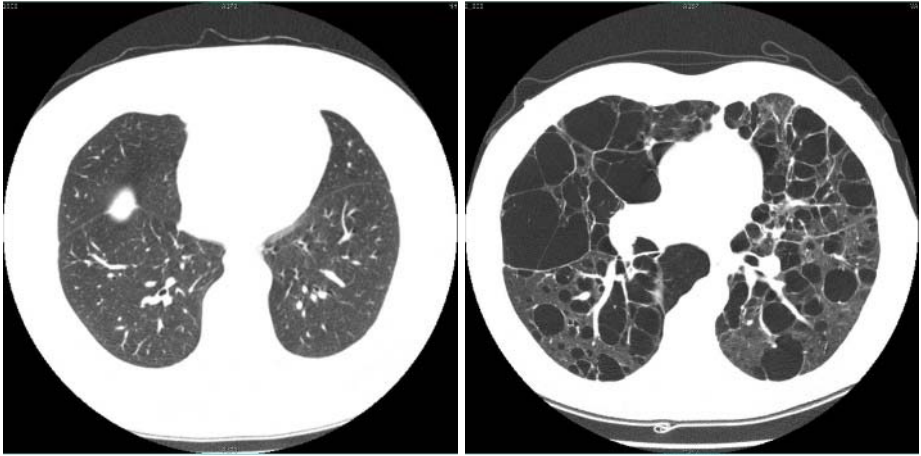


Fig. 1. The image on the left shows a CT scan of a healthy subject, which can be compared to the image of a severely emphysematous subject, shown on the right. The two diagnosis were established by and expert pulmonologist.

low attenuation pixels and uses no information on the spatial distribution or grouping of the low attenuation pixels. Another method of analyzing the degree of lung degradation uses the fractal dimension of the low attenuation areas. The cumulative size distribution of the low attenuation areas has been shown to follow a power law relationship - with healthy subjects and emphysematous subjects having markedly different cumulative size distributions [3].

Coxson provides another extension to the Müller density index [4]. Here, the CT pixel values are converted to local measures of specific volume. This facilitates an interpretation of the volume of gas per mass of tissue in the lung parenchyma. Since emphysema manifests as a degradation of tissue (loss of mass) and a loss of elastic recoil (gain of volume), subjects with emphysema have higher volume of gas per unit mass ratios than healthy subjects.

Finally, emphysema has been analyzed using standard statistical pattern analysis approaches. The adaptive multiple feature method (AMFM) of Uppaluri [5] divides an image into a regular grid of regions and generates a series of derived measurements or features for each grid square. The features generated include local histograms, co-occurrence matrices, run length frequencies, and fractal attributes. Disease severity is determined using a Bayesian classifier on this large feature space. A main disadvantage of this approach is that the classifier must be manually trained with an expert assigning a classification to each grid square in each image in the training set. While the previous techniques looked at the percentage and distribution of low attenuation regions, the AMFM approach utilizes a richer feature set, in essence using more of the information present in the CT data.

The quantification approach presented here incorporates some of the previous technologies, extends others, and introduces new approaches for analyzing emphysema. In the following sections we will describe our classification technique in more detail and

provide quantitative results. Section 2 describes the segmentation and feature selection approach. Section 3 describes the metrics used for histogram comparison, and section 4 provides results computed on a dataset of CT volumes. Finally, in section 5 we will draw conclusions.

2 Feature Selection

In the problem of quantifying emphysema severity, a good deal is known about the data being analyzed, both in terms of anatomic characteristics of the lung and the physics of the scanner. This knowledge is incorporated in our feature extraction model, which includes both signal and shape models, to achieve the highest detection rates. These feature extraction methods are similar to those presented in [6].

In CT imagery, the signal models represent the blurring (point spread function) and noise characteristics inherent to the imaging system. These models permit the accurate detection of primitive events in the image such as intensity peaks and discontinuities. These primitives form the basis for our method's shape representation.

The first stage of shape representation consists of step-edge detection using the Canny edge detector [7]. To account for noise, each image slice is convolved with a two-dimensional Gaussian kernel in which the scale of the kernel can be optimally computed from the point spread function of the scanner [8].

After edge detection we seek to capture the appearance of features in the image in terms of their intensity and shape. This representation is formed by grouping the low-level features detected at the signal level. The edges from the preceding signal modeling stage are grouped into closed regions by tracing along the edges and following a path perpendicular to the edge gradient until a closed structure can be formed. The result is a connected network of vertices and edges delimiting adjacent image regions within

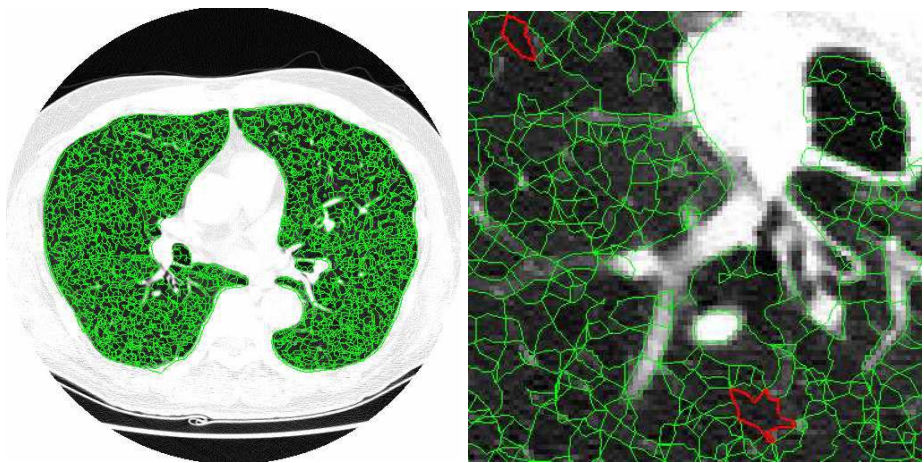


Fig. 2. Left: CT slice showing segmentation of the lung field. Right: close-up of the segmentation showing individual regions over which statistics can be computed.

which the voxel intensities are homogeneous [9]. The boundaries of these regions correspond to underlying tissue morphology, so that, for example, a large emphysematous area is represented by a single segmented region. Figure 2 illustrates a segmented CT slice and shows individual regions over which statistics can be computed.

In our application, we extract a variety of features for each region including intensity, area, number of edges, outer perimeter, diameter, intensity variance, min and max intensity, and many others. Of these, the first two, intensity and area, were chosen for histogram generation because they provide a logical characterization of disease severity. That is, as the severity of emphysema progresses, the lung tissue degrades, leading to larger and darker tissue patches.

3 Histogram Comparison

Using the low-level segmentation of the lung region, we construct two-dimensional histograms of area and mean intensity and compare these histograms to canonical histograms for various stages of the disease. These histograms are rich in information content. From these histograms, one can extract the percentage of the lung parenchyma occupied by low attenuation regions as in the Müller density index or analyze the size distribution of the low attenuation regions as in the fractal dimension approach. But the information in the histograms is richer than either of those approaches since it includes the size distributions for each attenuation value. Thus, in one representation you have the size distributions for the healthy parenchyma regions, the severely diseased regions, and the moderately diseased regions. The histograms are more relevant than the features extracted in the AMFM method because they are aligned with image events and not with an artificial grid.

To assess disease severity, we compare a subject's two-dimensional histogram to a set of canonical histograms. Each canonical histogram represents a different disease severity level, and each is derived from patient scans assessed by an expert pulmonologist, by averaging a set of histograms belonging to a given disease category as established by the expert. In order to compare the canonical histograms with the feature histogram of a new case several techniques can be employed, such as L_p metrics, chi-square distance or Kullback-Leibler divergence. Motivated by the results in [10], we use the Earth's Movers Distance (EMD) as our metric. Intuitively, the EMD between two histograms is the minimum amount work needed to "reshape" one histogram into the other, normalized by the total amount of mass moved to perform the reshaping. The EMD depends both on the amount of mass that has to be moved as well as on the distance over which the transport occurs. In the situation when the distance between histogram bins is a metric [11] and the histograms have the same mass, the EMD is itself a metric. This is an important property in the application tackled here, since the triangular inequality prevents a given histogram from being simultaneously close to two others if those are themselves far apart. This should avoid the idiosyncratic situation in which the disease level of a subject is simultaneously deemed as close to high and low, as long as there is enough separation between these two classes themselves. In the current work the distance between bins was chosen to be the Euclidean distance between the bin locations.

4 Experimental Results

We generated two-dimensional histograms of region intensity and area for 81 CT data sets. Each data set was acquired with a GE scanner with 120 kVp, 1 mm slice thickness, 20 mm interslice distance, 200 mA tube current, 0.5 sec rotation, and x and y pixel spacing ranging from 0.547 mm to 0.813 mm. An expert pulmonologist examined and classified each case as one of five emphysema levels ranging from 0 (no signs of emphysema) to 4 (severe emphysema). The data sets were subgrouped as follows: 16 level 0 cases, 25 level 1 cases, 19 level 2 cases, 9 level 3 cases, and 12 level 4 cases.

For each case we constructed 400×400 two-dimensional histograms as described in section 2. Intensity ranged from -1024 HU to -220 HU with a bin size of 2.01 HU. Area ranged from 0 mm^2 to 338 mm^2 with a bin size of 0.8452 mm^2 . Figure 3 shows area and intensity histogram marginals for a level 0 case and a level 4 case. Though the area marginals are quite similar to one another, the difference in the intensity marginals is clear.

We employed a bootstrapping technique to classify each of the 81 cases. That is, we generated canonical histograms for each severity level while withholding the case to be tested. Table 1 is a confusion matrix showing classification results. The rows indicate the subject cases, and the columns indicate the level at which the case was classified. If considering perfect classification, randomly choosing emphysema levels for each case yields an accuracy of 20%. Table 1 indicates that 62% of the level 0 cases, 20% of the level 1 cases, 21% of the level 2 cases, 33% of the level 3 cases, and 67% of the level 4 cases were perfectly classified, with a weighted kappa statistic of 0.468 (standard error of 0.074) [12, 13].

It is interesting to observe that although the classification rates for levels 1, 2, and 3 are somewhat low, they tend to be misclassified as a similar emphysema level. Therefore, if we instead consider a correct classification to be exact agreement or off-by-one agreement, results improve substantially. This is not an unreasonable approach given that cases with similar levels of severity have similar intensity and spatial characteris-

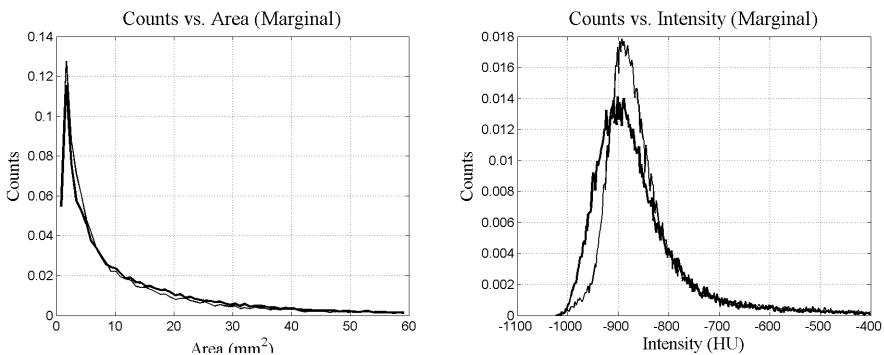


Fig. 3. Left: Area marginal histograms for a level 0 case and a level 4 case (bold). Right: Intensity marginal histograms for a level 0 case and a level 4 case (bold).

Table 1. Classification Confusion Matrix

	Level 0	Level 1	Level 2	Level 3	Level 4
Level 0	10	2	2	0	2
Level 1	9	5	9	2	0
Level 2	3	4	4	7	1
Level 3	1	1	1	3	3
Level 4	0	0	0	4	8

tics. In this case, random classification is 52%. Table 1 indicates that 75% of the level 0 cases, 92% of the level 1 cases, 79% of the level 2 cases, 78% of the level 3 cases, and 100% of the level 4 cases were classified correctly.

5 Conclusion

Emphysema is a serious disease leading to extensive physical restrictions for many individuals. In this paper we demonstrated a completely automated method of measuring emphysema severity. It is based on the observation that emphysema manifests in CT as large regions of low attenuation, properties captured on a two-dimensional joint histogram of the area and intensity of segmented regions in the lung parenchyma. We showed that our algorithm yields results in good agreement with manual classifications of an expert pulmonologist.

The automation of severity quantification addresses the issue of human intra- and inter-observer variability. Inter-observer agreement on emphysema scores for radiologists with varying degrees of expertise was reported in [1], with kappa values in the range of 0.431 to 0.589. The weighted kappa value for our algorithm (0.468) compares well with the range of values given in that work.

Although the results reported here are encouraging, improvements can be gained by employing more sophisticated classification schemes. In particular, we should consider a method that takes the variance of the histograms into account. Additionally, we chose features (mean intensity and area) based on clinical significance. However, because we compute a rich set of features for each segmented region, it would be interesting to validate this choice using a statistical feature selection technique. In principle three-dimensional segmentations would provide a more faithful representation of the true anatomical scenario. However, the imaging protocol may preclude accurate 3D segmentations owing to large interslice distances.

References

1. Bankier, A.A., De Maertelaer, V., Keyzer, C., Gevenois, P.A.: Pulmonary emphysema: Subjective visual grading versus objective quantification with macroscopic morphometry and thin-section ct densitometry. *Radiology* **211** (1999) 851–858
2. Muller, Staples, Miller, Abboud: Density mask: An objective method to quantitate emphysema using computed tomography. *Chest* **94** (1988) 782–787

3. Mishima, M.: Fractal analysis of emphysema in X-ray, CT and simulation. *Medical Imaging Technology* **18** (2000) 179–186
4. Coxson, Rogers, Whittall, D'Yachkova, Pare, Sciruba, Hogg: A quantification of the lung surface area in emphysema using computed tomography. *American Journal of Respiratory and Critical Care in Medicine* **159** (1999) 851–856
5. Uppaluri, R., Hoffman, E.A., Sonka, M., Hartley, P.G., Hunninghake, G.W., McLennan, G.: Computer recognition of regional lung disease patterns. *American Journal Respiratory and Critical Care in Medicine* **160** (1999) 648–654
6. McCulloch, C.C., Kaucik, R.A., Mendonça, P.R.S., Walter, D.J., Avila, R.S.: Model-based detection of lung nodules in computed tomography exams. *Academic Radiology* **11** (2004) 258–266
7. Canny, J.F.: A computational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intell.* **8** (1986) 679–698
8. Elder, J.H., Zucker, S.W.: Local scale control for edge detection and blur estimation. *IEEE Trans. Pattern Analysis and Machine Intell.* **20** (1998) 699–716
9. Rothwell, C., Mundy, J., Hoffman, W., Nguyen, V.D.: Driving vision by topology. In: *IEEE International Symposium on Computer Vision*. (1995) 395–400
10. Rubner, Y., Tomasi, C., Guibas, L.J.: The earth mover's distance as a metric for image retrieval. *Int. Journal of Computer Vision* **40** (2000) 99–121
11. Sutherland, W.A.: *Introduction to Metric and Topological Spaces*. Clarendon Press, Oxford, UK (1975) 1998 reprint.
12. Cohen, J.: A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* **20** (1960) 37–46
13. Cohen, J.: Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin* **70** (1968) 213–220

Efficient Learning by Combining Confidence-Rated Classifiers to Incorporate Unlabeled Medical Data

Weijun He¹, Xiaolei Huang¹, Dimitris Metaxas¹, and Xiaoyou Ying²

¹ Center for Computational Biomedicine Imaging and Modeling,
Division of Computer and Information Sciences, Rutgers University, NJ, USA

{weijunhe, xiaolei, dnm}@cs.rutgers.edu

² Sanofi-aventis, Bridgewater, NJ, USA

Xiaoyou.Ying@sanofi-aventis.com

Abstract. In this paper, we propose a new dynamic learning framework that requires a small amount of labeled data in the beginning, then incrementally discovers informative unlabeled data to be hand-labeled and incorporates them into the training set to improve learning performance. This approach has great potential to reduce the training expense in many medical image analysis applications. The main contributions lie in a new strategy to combine confidence-rated classifiers learned on different feature sets and a robust way to evaluate the “informativeness” of each unlabeled example. Our framework is applied to the problem of classifying microscopic cell images. The experimental results show that 1) our strategy is more effective than simply multiplying the predicted probabilities, 2) the error rate of high-confidence predictions is much lower than the average error rate, and 3) hand-labeling informative examples with low-confidence predictions improves performance efficiently and the performance difference from hand-labeling all unlabeled data is very small.

1 Introduction

In many learning algorithms in medical image analysis, the labeling of training data is often done manually. This process is quite time-consuming since a large set of training data is usually required. However, not all labeled data have the same level of effectiveness in improving a classifier. As in Support Vector Machines [1], only those “support vectors” that are located near the boundaries of different classes are the informative data that affect the final classifier. Hence if we can discover this type of “support vectors” in the unlabeled data, then we need only label these discovered informative data, include them in the training set and re-train the classifier. In this way, the amount of data to be labeled is greatly reduced without sacrificing the learning performance. In our approach we consider the confidence-rated classifiers that can predict a probability distribution over the labels for an example since the probability distribution enables us to determine the “informativeness” of the example.

A single confidence-rated classifier, however, is often insufficient because in many medical images, multiple sets of features have very different characteristics and can not be effectively combined in a single classifier. For instance, image features are often grouped into different categories such as shape and texture. These feature sets have independent bases, and simply concatenating them into a single feature vector produces a complex, unstructured feature space that can potentially degrade learning and classification performance. To tackle this problem, in this paper we train separate confidence-rated classifiers on each category of features and then combine the predictions using Bayes rule, assuming conditional independence between classifiers trained on different feature sets. The classical voting classification algorithms, such as Bagging [2, 3] and AdaBoost [4, 5], are successful in improving the accuracy by combining multiple weak classifiers. Bauer and Kohavi [6] gave an empirical comparison of voting classification algorithms. However, these voting classification algorithms are generally applied to classifiers that just assign a label (not a probability) to an instance. Schapire and Singer [7] proposed new boosting algorithms using confidence-rated predictions, however, their extension to multi-class classification problems is not so straightforward. The new approach proposed in this paper for combining multiple confidence-rated classifiers based on Bayes rule efficiently addresses these problems.

Since our combining rule produces probability distributions over all labels for an unlabeled example, the predicted probabilities can be used to determine the “informativeness” of the example. Examples with high-confidence predictions are less informative than those with low-confidence predictions in improving the classifier. Hence it is more efficient to hand-label only those examples with low-confidence predictions. A classical method in the literature that improves learning by using unlabeled data is the co-training method [8, 9]. The basic idea is to organize the features of training examples into two different feature sets, and learn a separate classifier on each feature set. There are two assumptions in co-training. First, the two feature sets are redundant but not completely correlated. Second, each feature set would be sufficient for learning if enough data were available. Under these assumptions, the high-confidence predictions of one classifier on new unlabeled examples are expected to generate informative examples to enrich the training set of the other. However, these formal assumptions may not hold in many medical image applications that tend to have high complexity and dimensionality. In this paper, instead of trusting that each feature set is sufficient for learning, we determine the high-confidence predictions for new data by combining the opinions of all classifiers based on different feature sets. Our approach can be applied to multi-class classification problems directly.

2 Data Description and Preprocessing

The data we use are microscopic cell images. Each image consists of lots of cells in different developmental stages. The goal is to classify the cells into different stages and count the number of cells in each developmental stage. This problem has wide applications in the pharmaceutical industry for therapy evaluation.

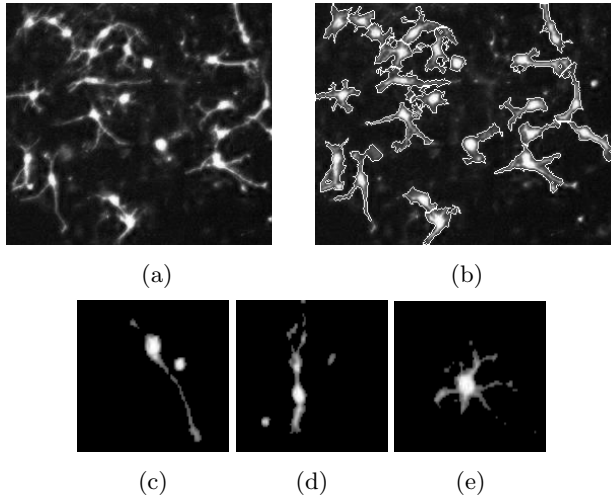


Fig. 1. (a) Cell image; (b) Segmented cell image; (c,d,e) are cells in developmental stage 0, 1 and 2 respectively

We first segment out the individual cells of interest. Since the cell nuclei are usually the brightest and cover a certain amount of area, we locate all cell nuclei by thresholding at a high intensity value and applying connected component analysis. Then we locate the regions occupied by all cells by thresholding the original image at a low intensity value. We apply dilation on all cell nuclei in the cell regions simultaneously until all cells are fully segmented. Finally we extract useful features from each cell for the purpose of classification. There are two categories of features: 1) binary image features, including roundness, eccentricity, solidity, extent and the diameter of a circle with the same area as the region; 2) gray image features, including mean and standard deviation of the gray level intensity. In the training phase, we first label a small set of cells from different developmental stages by hand, and include them in the initial training data set. The labels denote the different developmental stages. There are three developmental stages in our experiments (Fig. 1): 0 - beginning; 1 - immature; 2 - mature.

Because the hand-labeling phase is tedious and time consuming, the initial labeled training set we can acquire is limited. Our methodology is to tackle this problem by strategically adding unlabeled data into the training data set based on the evaluation of the unlabeled data using a confidence-rated classification mechanism. The proposed mechanism is described in the next sections.

3 Learning Framework

Our learning framework is outlined by the flowchart in Fig. 2. In the framework, we train multiple confidence-rated classifiers on separate groups of features. When presented with unlabeled data, each classifier produces a confidence

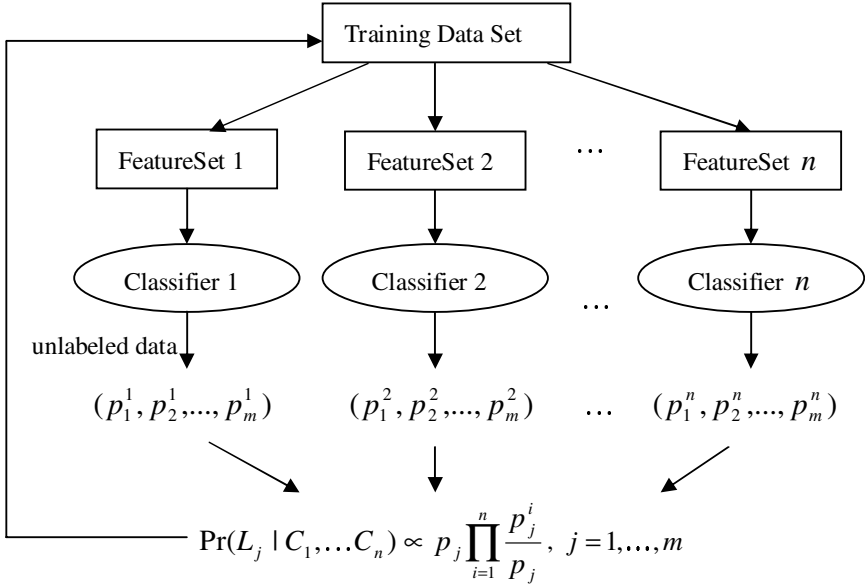


Fig. 2. Learning framework

rate of its own prediction. Turney et al. [10] has argued that opinions from independent modules should be combined multiplicatively. We extend this idea to combine predictions of classifiers trained on different feature sets.

3.1 Combining Confidence-Rated Classifiers Using Bayes Rule

Our combining approach is to apply Bayes rule to evaluate the final confidence rate based on the confidence rates given by the multiple, independent classifiers. Suppose we have n classifiers, m labels, and classifier i is trained on the feature set i . p_j^i ($i = 1, \dots, n, j = 1, \dots, m$) denotes the probability that classifier i assigns label j to an input unlabeled example. L_j denotes that the label of an example is j . Let $C_i = (p_1^i, p_2^i, \dots, p_m^i), p_j = Pr(L_j)$. Using Bayes rule, we have

$$Pr(L_j | C_1, C_2, \dots, C_n) = \frac{Pr(C_1, C_2, \dots, C_n | L_j) Pr(L_j)}{Pr(C_1, C_2, \dots, C_n)} \tag{1}$$

Assuming conditional independency of C_1, C_2, \dots, C_n , we have

$$Pr(L_j | C_1, C_2, \dots, C_n) \propto Pr(C_1 | L_j) Pr(C_2 | L_j) \dots Pr(C_n | L_j) Pr(L_j) \tag{2}$$

Since $Pr(C_i | L_j) = \frac{Pr(L_j | C_i) Pr(C_i)}{Pr(L_j)}$, Eq. 2 can be rewritten as

$$Pr(L_j | C_1, C_2, \dots, C_n) \propto Pr(L_j) \prod_{i=1}^n \frac{Pr(L_j | C_i)}{Pr(L_j)} \tag{3}$$

Table 1. The difference between our approach with naïve Bayes classifier

	Known	$Pr(L_j O_1, \dots, O_q) \propto$
naïve Bayes classifier	$Pr(O_i L_j)$	$Pr(L_j) \prod_{i=1}^q Pr(O_i L_j)$
Our approach	$Pr(L_j C_i)$	$Pr(L_j) \prod_{i=1}^n \frac{Pr(L_j C_i)}{Pr(L_j)}$

Using our abbreviated notation, Eq. 3 can be simplified as

$$Pr(L_j|C_1, C_2, \dots, C_n) \propto p_j \prod_{i=1}^n \frac{p_j^i}{p_j} \tag{4}$$

The above formula intuitively says that if the predicted probability p_j^i is greater than (equal to, less than) the prior probability p_j , then this prediction will increase (not affect, decrease) the final probability given to label j . The following theorem relates our approach to naïve Bayes classifier.

Theorem 1. *If Classifier $i, i = 1, \dots, n$ are themselves naïve Bayes classifiers using disjoint feature sets, then the combined classifier is a naïve Bayes classifier using all features.*

Proof. Suppose the feature sets are $\phi = \{\phi_i, i = 1, \dots, n\}$ and each classifier C_i is a naïve Bayes classifier based on the feature set ϕ_i . Then we have:

$$\begin{aligned}
 Pr(L_j|C_i) &\propto Pr(L_j) \prod_{\alpha \in \phi_i} Pr(\alpha|L_j) \tag{5} \\
 Pr(L_j|C_1, C_2, \dots, C_n) &\propto Pr(L_j) \prod_{i=1}^n \frac{Pr(L_j|C_i)}{Pr(L_j)} \\
 &\propto Pr(L_j) \prod_{i=1}^n \prod_{\alpha \in \phi_i} Pr(\alpha|L_j) \\
 &\propto Pr(L_j) \prod_{\alpha \in \phi} Pr(\alpha|L_j) \tag{6}
 \end{aligned}$$

Hence each classifier can be viewed as a mapping from a set of concrete features to a single abstract feature (a probability distribution over all labels). Our combining approach then predicts the final classification using Bayes rule on these abstract features. Denote the observation on feature $i, i = 1, \dots, q$ as O_i . In table 1, we compare our combining approach with the naïve Bayes classifier. Our combining approach does not assume any particular class-conditional density model as the naïve Bayes classifier does for the continuous variables. Instead base classifiers trained on different feature sets are applied to generate probability distributions over the labels for an example.

3.2 Exploring the Unlabeled Data

Once we have acquired the combined probability distribution over an unlabeled example, we can use it to determine the “informativeness” of this example.

The intuition behind our approach is that not all unlabeled data have equal effectiveness in improving the classifier. For example, in SVM only the support vectors are used in determining the final classifier. If we know those support vectors, it would suffice to label only those data and train the classifier on them. So our strategy is to find those potential “support vectors”, and present them only for hand-labeling. In this way, the amount of human efforts needed to acquire a large labeled training set is greatly reduced.

We notice that the support vectors are near the boundaries between two classes. And the classifier does not predict well their labels. So, the probabilities given by the classifier can be used to discover those informative unlabeled data. If the predicted probability that one unlabeled example belongs to a certain class is high, we include this example along with the predicted label directly into the training set. However, these data may only enlarge the training set without helping much to improve the classifier. On the other hand, if the probability of an unlabeled example belonging to any class is below some threshold, the current classifier is uncertain about the label of this example. Therefore, this type of data is most probably lying around the boundary between two classes. Hand-labeling these data that the current classifier is uncertain about and adding them into the training set will most efficiently improve the classifier. In this way, by quantitatively evaluating the relationship between the unlabeled data and the current classifier, we only need to label those “most profitable” unlabeled data without sacrificing much in performance.

4 Experiments

We tested our algorithm by collecting our data from 40 microscopic cell images, each containing about 50 cells of all developmental stages after segmentation. We classify the cells into 3 developmental stages: beginning, immature and mature. For each segmented cell, we extract two separate sets of features: one related to shape and geometry using the thresholded binary image, and the other related to intensity statistics based on the gray-level image. The multi-class logistic regression [11] classifier is applied on the binary and gray feature sets separately and we get two probability distributions for each unlabeled cell example. However our learning framework does not assume any specific base learning method. We choose logistic regression since it can predict a probability distribution and can be applied to a wide variety of situations as long as the difference between the logarithms of the class-conditional density function is linear in the variables.

In the first experiment, we compare our Bayesian combination approach which takes into account the prior probability of each class with multiplicative combination that does not consider the prior probabilities. The label we assign to an example is the label with the highest predicted probability. In table 2, we can see that combining the predictions without unlabeled data using Bayes rule is better than just simply multiplying the probabilities without considering the prior probability of each class. In each run, the prior probabilities are estimated from the training set (randomly selected 35 labeled examples). We expect that the performance gain would be greater if more feature sets were combined.

Table 2. Prediction accuracy without unlabeled data (average over 20 runs)

Binary features	Gray features	Multiplicative combination	Bayesian combination
81.1%	69.7%	79.3%	81.7%

Table 3. Prediction accuracy (a) with just initial 35 training data; (b) with hand-labeled “informative” unlabeled data, along with the number of such data; (c) with all 100 unlabeled data hand-labeled

runs	(a)	(b)	(c)
1	78.6%	91.1% (24)	92.9%
2	78.6%	83.9% (16)	83.9%
3	80.3%	82.1% (30)	83.9%
4	82.1%	83.9% (31)	85.7%
5	82.1%	87.5% (14)	87.5%

In the second experiment, we examine whether the confidently predicted unlabeled data are really correctly labeled and compute the proportion of this type of data among all unlabeled data. The result depends on the threshold applied on the combined probability. In this experiment, if the predicted probability of one example belonging to one class is higher than 90%, then we treat this example as a confidently predicted example. Over 20 runs, the percentage of the confidently labeled data over all unlabeled data is 60.9%. The average prediction accuracy for those confidently labeled data is 92.3%, which is much higher than the average prediction accuracy (81.7%). However if a higher labeling accuracy for the added unlabeled data is required, we need to increase the threshold, which will decrease the percentage of confidently labeled data. So there is a tradeoff here and the choice of the threshold depends on the application.

Finally, we show that hand-labeling a few informative examples with low-confidence predictions efficiently improves performance, and that the performance difference is small between hand-labeling the few informative examples and hand-labeling all unlabeled data. Similarly we need to set a threshold. In this experiment, if the maximal predicted probability of one example belonging to any class is lower than 80%, then we treat this example as an ambiguous example to the current classifier and we need to label it by hand. In table 3, we can see that, over five runs, the numbers of such ambiguous (i.e. “informative”) unlabeled examples are 24, 16, 30, 31 and 14, which are much less than the total number of unlabeled examples (100). By labeling only this reduced number of unlabeled data, however, we achieve a performance that is comparable to that by labeling all unlabeled examples.

5 Discussions and Conclusions

In this paper, we have presented a Bayesian strategy for combining confidence-rated predictions of classifiers trained on different feature sets. Our method gen-

erates a probability distribution over the labels for an unlabeled example. We utilize these probability distributions to filter out two groups of unlabeled data. One group is the confidently labeled data. We add them directly into the training set. Compared to the co-training method, our approach combines the opinions from different classifiers to ensure that the self-labeled data are correct with very high probability. The other group of the filtered unlabeled data includes those potentially informative examples for whose labels the current classifier is uncertain. By hand-labeling only these informative data, we achieve comparable performance with hand-labeling all data. This results in greatly reduced training expense. Therefore the training phase of our method is not static, but dynamic.

References

1. Cristianini, N., Shawe-Taylor, J.: Support Vector Machines and other kernel-based methods. Cambridge University Press (2000)
2. Breiman, L.: Bagging predictors. *Machine Learning* **26** (1996) 123–140
3. Quinlan, J.R.: Bagging, boosting, and C4.5. In: Proc. AAAI-96 Fourteenth National Conf. on Artificial Intelligence. (1996) 725–730
4. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences* **55** (1997) 119–139
5. Schapire, R.E.: A brief introduction to boosting. In: Proc. of 16th Int'l Joint Conf. on Artificial Intelligence. (1999) 1401–1406
6. Bauer, E., Kohavi, R.: An empirical comparison of voting classification problems: Bagging, boosting and variants. *Machine Learning* **36** (1999) 105–142
7. Schapire, R.E., Singer, Y.: Improved boosting algorithms using confidence-rated predictions. *Machine Learning* **37** (1999) 297–336
8. Blum, A., Mitchell, T.: Combining labeled and unlabeled data with co-training. In: Proc. of the 1998 Conf. on Computational Learning Theory. (1998) 92–100
9. Levin, A., Viola, P., Freund, Y.: Unsupervised improvement of visual detectors using co-training. In: Proc. of the Int'l Conf. on Computer Vision. (2003) 626–633
10. Turney, P., Littman, M., Bigham, J., Shnayder, V.: Combining independent modules to solve multiple-choice synonym and analogy problems. In: Proc. of the Int'l Conf. on Recent Advances in Natural Language Processing. (2003) 482–489
11. Anderson, J.A.: Logistic discrimination. In Krishnaiah, P.R., Kanal, L.N., eds.: *Handbook of Statistics 2*. (1982) 169–191

Mosaicing of Confocal Microscopic *In Vivo* Soft Tissue Video Sequences

Tom Vercauteren^{1,2}, Aymeric Perchant²,
Xavier Pennec¹, and Nicholas Ayache¹

¹ Projet Epidaure, INRIA Sophia-Antipolis, France

² Mauna Kea Technologies, 9 rue d'Enghien Paris, France

Abstract. Fibered confocal microscopy allows *in vivo* and *in situ* imaging with cellular resolution. The potentiality of this imaging modality is extended in this work by using video mosaicing techniques. Two novelties are introduced. A robust estimator based on statistics for Riemannian manifolds is developed to find a globally consistent mapping of the input frames to a common coordinate system. A mosaicing framework using an efficient scattered data fitting method is proposed in order to take into account the non-rigid deformations and the irregular sampling implied by *in vivo* fibered confocal microscopy. Results on 50 images of a live mouse colon demonstrate the effectiveness of the proposed method.

1 Introduction

Fibered confocal microscopy (FCM) is a potential tool for *in vivo* and *in situ* optical biopsy [1]. FCM is based on the principle of confocal microscopy which is the ability to reject light from out-of-focus planes and provide a clear in-focus image of a thin section within the sample. This optical sectioning property is what makes the confocal microscope ideal for imaging thick biological samples. The adaptation of a confocal microscope for *in vivo* and *in situ* imaging can be viewed as replacing a microscope objective by a probe of adequate length and diameter in order to be able to perform *in situ* imaging. For such purpose, a fiber bundle is used as the link between the scanning device and the microscope objective. After image processing of the FCM raw output, the available information is composed of a video sequence irregularly sampled in the space domain, each sampling point corresponding to a fiber center [1].

This imaging modality unveils the cellular structure of the observed tissue. The goal of this work is to enhance the possibilities offered by FCM by using image sequence mosaicing techniques in order to widen the field of view (FOV). Several possible applications are targeted. First of all, the rendering of wide-field micro-architectural information on a single image will help experts to interpret the acquired data. This representation will also make quantitative and statistical analysis possible on a wide field of view. Moreover, mosaicing for microscopic images is a mean of filling the gap between scales and allows multiscale information fusion for probe positioning and multi-modality fusion.

Each frame of the input sequence is modeled as a deformed partial view of a ground truth 2D scene. The displacement of the fiber bundle probe across the tissue is described by a rigid motion. Due to the interaction of the contact probe with the soft tissue, a small non-rigid deformation appears on each input frame. Because of those non-linear deformations, and the irregular sampling of the input frames, classical video mosaicing techniques need to be adapted.

In Section 2, our mosaicing framework is described. The first main contribution presented in Section 3 is the use of Riemannian statistics to get a robust estimate of a set of mean rigid transformations from pairwise registrations results. The second main contribution is proposed in Section 4 where we develop a mosaicing framework using an efficient scattered data fitting method well-suited for non-linear deformations and irregularly sampled inputs. Finally real experiments described in Section 5 demonstrate the effectiveness of our approach and show the significant image improvements obtained by using non-linear deformations.

2 Problem Formulation and Mosaicing Method

The goal of many existing mosaicing algorithms is to estimate the reference-to-frame mappings and use these estimates to construct the mosaic [2]. Small residual misregistrations are then of little importance because the mosaic is reconstructed by segmenting the field into disjoint regions that use a single source image for the reconstruction [3,4]. Since our input frames are rather noisy, we would like to use all the available information to recover an approximation of the true underlying scene. We will therefore estimate the frame-to-reference transformations (instead of the usual reference-to-frame) and consider all the input sampling points as shifted sampling points of the mosaic. This has several advantages for our problem. First of all, this is really adapted to irregularly sampled input frames because we will always use the original sampling points and never interpolate the input data. This approach is also more consistent with a model of noise appearing on the observed frames rather than on the underlying truth. Finally in this framework, it will be possible to get a mosaic at a higher resolution than the input frames. The drawback is that we need an accurate estimate of the unknown transformations.

Let I be the unknown underlying truth and I_n be the observed frames. Our algorithm makes use of the following observation model,

$$I_n(p) = I(f_n(p)) + \epsilon_n(p), \quad \forall p \in \Omega_n, \quad (1)$$

where $\epsilon_n(p)$ is a noise term, Ω_n is the coordinate system associated with the n^{th} input frame and $f_n : \Omega_n \rightarrow \Omega$ is the unknown frame-to-reference mapping composed of a large rigid mapping r_n and a small non-rigid deformation b_n ,

$$f_n(p) = b_n \circ r_n(p). \quad (2)$$

By making the reasonable assumption that consecutive frames are overlapping, an initial estimate of the global rigid transformations can be obtained by

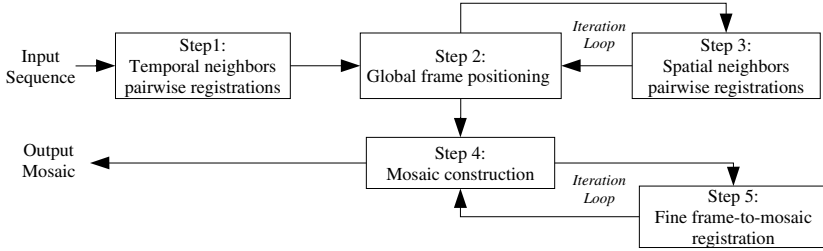


Fig. 1. Block diagram of the mosaicing algorithm

using an image registration technique to estimate the motion between the consecutive frames. Global alignment is then obtained by composing the local motions. This initial estimate suffers from a well-known accumulation of error problem. Our algorithm depicted in Fig. 1 iteratively refines the global positioning by adding new pairwise rigid registration results to estimate the global parameters. Once a consistent set of rigid transformations is found, the algorithm constructs an initial mosaic by mapping all observed sampling points into the common reference coordinate system and using an efficient scattered data fitting technique on this point cloud. The residual non-rigid deformations are finally taken into account by iteratively registering an input frame to the mosaic and updating the mosaic based on the new estimate of the frame-to-mosaic mapping.

3 Global Frame Positioning from Pairwise Registrations

The first task of our algorithm is to register pairs of overlapping input images under a rigid transformation assumption. For that purpose, we use a classical registration framework based on a similarity criterion optimization but any other technique (e.g. block matching framework [5], Mellin transform [3], feature-based registration [6] etc.) can be used. Let $r_{j,i}^{(\text{obs})} : \Omega_i \rightarrow \Omega_j$ be the pairwise rigid registration result between input frames i and j . This result is considered as a noisy observation of $r_j^{-1} \circ r_i$. Based on the set of all available observations, our algorithm looks for a globally consistent estimate of the global parameters $[r_1, \dots, r_N]$. This problem is addressed in [3] where a least-square solution is given when linear transformations are considered. This technique cannot readily be adapted to rigid transformation. In [7], the authors propose a more general approach. Some chosen corner points are transformed through r_i and $r_j \circ r_{j,i}^{(\text{obs})}$. The squared distance between the transformed points added to a regularization term is then minimized. These techniques are sensitive to outliers, and are either tailored to a specific type of transformation or need a somewhat ad hoc choice of points. In this paper statistics for Riemannian manifolds are used to provide a globally consistent and robust estimate of the global rigid transformation.

The computational cost of registering all input frames pairs is prohibitive and not all pairs of input frames are overlapping. It is therefore necessary to choose which pairs could provide informative registration results. For that purpose, we

chose the topology refinement approach proposed in [7]. An initial guess of the global parameters $[r_1, \dots, r_N]$ is obtained by registering the consecutive frames, the algorithm then iteratively chooses a next pair of input frames to register (thus providing a new observation $r_{j,i}^{(\text{obs})}$) and updates the global parameters estimation. As we only consider the pairwise registration results as noisy observations, we need many of them. In order to minimize the computational cost of those numerous registrations, we use a multiresolution registration technique using a Gaussian image pyramid that stops at a coarse level of the optimization.

Given the set of all available pairwise registration results Θ , we need to estimate the true transformations. A sound choice is to consider a least-square approach. However the space of rigid transformations is not a vector space but rather a Lie group that can be considered as a Riemannian manifold. Classical notions using distances are therefore not trivial to generalize. In what follows, we provide an extension of the Mahalanobis distance for Riemannian manifolds [8] and propose an optimization algorithm to find the least-square estimate. Let r be an element of the 2D rigid transformations Riemannian manifold. From the theory of Riemannian manifolds, the vector $\mathcal{V}(r)$ representing r by its angle $\alpha(r) \in [-\pi, \pi]$ and translation components $t_x(r)$, $t_y(r)$ can be considered as a mapping of the transformation r onto the vector space defined by the tangent space at the identity point Id of the manifold. This tangent space will be denoted as Id-tangent space. By using the canonical left-invariant Riemannian metric for rigid transformations, the distance of r to the identity is defined as the norm of the representation $\mathcal{V}(r)$ in the Id-tangent space, $\text{dist}(r, \text{Id}) = \|\mathcal{V}(r)\|$. The distance between two transformations is given by

$$\text{dist}(r_a, r_b) = \text{dist}(r_b^{-1} \circ r_a, \text{Id}) = \|\mathcal{V}(r_b^{-1} \circ r_a)\|. \tag{3}$$

Using this distance, it is possible to define a generalized mean for a random transformation r , the Fréchet mean $E_f[r] = \arg \min_r E[\text{dist}(r_f, r)^2]$. If e is a random error whose Fréchet mean is the identity, its covariance matrix is simply defined as $\Sigma_{ee} = E[\mathcal{V}(e)\mathcal{V}(e)^T]$. The squared Mahalanobis distance between e and the identity is given by $\mu^2(e, \text{Id}) = \mathcal{V}(e)^T \Sigma_{ee}^{-1} \mathcal{V}(e)$. We now have the tools to derive the global parameters estimator. The observation model is given by $r_{j,i}^{(\text{obs})} = r_j^{-1} \circ r_i \circ e_{j,i}^{(\text{obs})}$, where $e_{j,i}^{(\text{obs})}$ is a random error whose Fréchet mean is assumed to be the identity and whose covariance matrix is Σ_{ee} . An estimate of $[r_1, \dots, r_N]$ is given by the set of transformations that minimizes the total Mahalanobis distance:

$$[\hat{r}_1, \dots, \hat{r}_N] = \arg \min_{[r_1, \dots, r_N]} \sum_{(i,j) \in \Theta} \mu^2(e_{j,i}^{(\text{obs})}, \text{Id}) \tag{4}$$

This equation does not admit a closed form solution but an efficient optimization method can be designed by a simple modification of a usual non-linear least square optimizer such as the Gauss-Newton descent. The Riemannian structure of the transformation space is taken into account by adapting the intrinsic geodesic gradient descent in [8]. The idea is to walk towards an optimum by a series of steps taken along a geodesic of the manifold rather than walking in the

tangent vector space. Let $r^{(t)}$ be an estimate of a rigid transformation r at step t , the intrinsic geodesic walking is achieved by finding a direction $\delta r^{(t)}$ and a step length λ for the following update equation in the Id-tangent space:

$$\mathcal{V}(r^{(t+1)}) = \mathcal{V}(r^{(t)} \circ \lambda \delta r^{(t)}). \quad (5)$$

By contrast, if a usual optimization routine on the Id-tangent space is used, we get a walking direction $\Delta \mathcal{V}^{(t)}$ such that the update would be

$$\mathcal{V}(r^{(t+1)}) = \mathcal{V}(r^{(t)}) + \lambda \Delta \mathcal{V}^{(t)}. \quad (6)$$

Using (6) directly can be problematic because we are not assured to remain on the manifold. It is however possible to combine the power of intrinsic geodesic walking and the ease of use of classical optimization routine by mapping a walking direction found in the Id-tangent space onto the manifold. For that purpose, a first order Taylor expansion of (5) around the identity is used:

$$\mathcal{V}(r^{(t)} \circ \lambda \delta r^{(t)}) = \mathcal{V}(r^{(t)}) + \lambda JL(r^{(t)}) \cdot \mathcal{V}(\delta r^{(t)}) + O(\lambda^2), \quad (7)$$

where $JL(r) = \partial \mathcal{V}(r \circ s) / \partial \mathcal{V}(s)|_{s=\text{Id}}$. By identifying (6) and (7), we see that $\Delta \mathcal{V}^{(t)} = JL(r^{(t)}) \cdot \mathcal{V}(\delta r^{(t)})$. The walking direction in the manifold is thus

$$\mathcal{V}(\delta r^{(t)}) = JL(r^{(t)})^{-1} \cdot \Delta \mathcal{V}^{(t)}. \quad (8)$$

Within this general framework, several improvements can easily be added. The terms of the cost function can be weighted by some confidence measure, robust estimation techniques such as M-estimators can be used to discard outliers, the noise variance can be re-estimated based on the measurements etc. We are now able to get robust and globally consistent estimates of the rigid transformations which we use as initial estimates of the complete global transformations $[f_1, \dots, f_N]$. The resulting mosaics in Fig. 2(a) show an accurate global positioning of the input frames that is robust to erroneous pairwise registrations.

4 Frame to Mosaic Fine Registration

Once a globally consistent mosaic has been constructed, it is possible to make a fine multi-image registration by iteratively registering each input frame to the mosaic and updating the mosaic. The mosaicing problem can be written as an optimization problem over the unknown underlying image I and the unknown transformations $[f_1, \dots, f_N]$ of the following multi-image criterion, $\mathcal{S}(f_1, \dots, f_N, I) = \sum_{n=1}^N S(I_n, I \circ f_n)$, where $S(I_a, I_b)$ is a usual pairwise similarity criterion between the two images I_a and I_b . With this framework our mosaic refinement procedure can be seen as an alternate optimization scheme.

We divide the fine frame-to-mosaic registrations into two loops of increasing model complexity. First we refine the global rigid mappings. Then, in order to account for the small non-rigid deformations, the residual deformation fields $[b_1, \dots, b_N]$ are modeled by using B-splines tensor products on a predefined grid.

This framework can easily be extended to use any other non-rigid registration methods using landmarks-based schemes or more accurate deformation models.

This iterative mosaic refinement scheme requires a new mosaic construction at each iteration. It is therefore necessary to use a very efficient reconstruction algorithm. Furthermore, since we want to register input frames with the mosaic, the reconstruction needs to be smooth enough for the registration not to be trapped in a local minimum but detailed enough for the registration to be accurate. Once an estimate \hat{f}_n of f_n is available, we get a point cloud composed of all transformed sampling points from all the input frames

$$\{(p_k, i_k)\} = \{(\hat{f}_n(p), I_n(p)) | p \in \Omega_n^s, n \in [0, \dots, N]\}, \tag{9}$$

where Ω_n^s is the set of sampling points in the input frame n . The usual algorithms for scattered data approximation do not simultaneously meet the requirements of efficiency, and control over the smoothness of the approximation. In the sequel we develop our second main contribution which is an efficient scattered data fitting algorithm that allow a control over the smoothness of the reconstruction. The main idea is to get an approximation of the underlying function by using a method close to Shepard interpolations. The value associated with a point in Ω is a weighted average of the nearby sampled values,

$$\hat{I}(p) = \sum_k w_k(p) i_k = \sum_k \frac{h_k(p)}{\sum_l h_l(p)} i_k. \tag{10}$$

The usual choice is to take weights that are the inverse of the distance, $h_k(p) = \text{dist}(p, p_k)^{-1}$. In such a case we get a true interpolation [9]. An approximation is obtained if a bounded weighting function $h_k(p)$ is chosen. We choose a Gaussian weight $h_k(p) = G(p - p_k) \propto \exp(-\|p - p_k\|^2 / 2\sigma_a^2)$ and thus (10) can be rewritten as

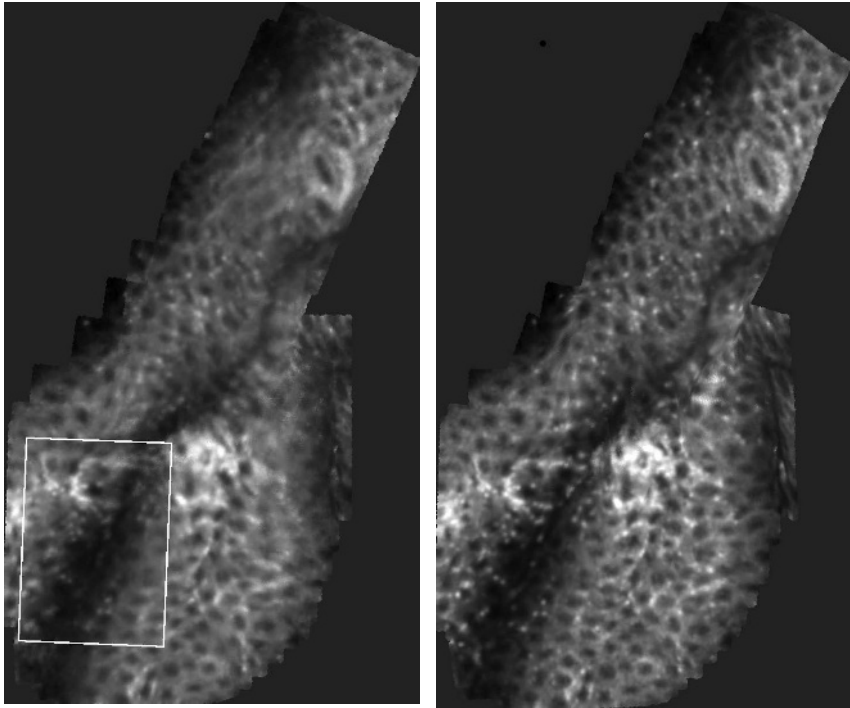
$$\hat{I}(p) = \frac{\sum_k i_k G(p - p_k)}{\sum_k G(p - p_k)} = \frac{[G \star \sum_k i_k \delta_{p_k}](p)}{[G \star \sum_k \delta_{p_k}](p)}, \tag{11}$$

where δ_{p_k} is a Dirac distribution centered at p_k . This scattered data approximation technique requires only two Gaussian filtering and one division and is thus very efficient. The smoothness is controlled by the variance σ_a^2 of the Gaussian kernel. Thanks to this approximation method, we obtain a mosaic that is sharper and smoother than the original input frames as demonstrated in Section 5.

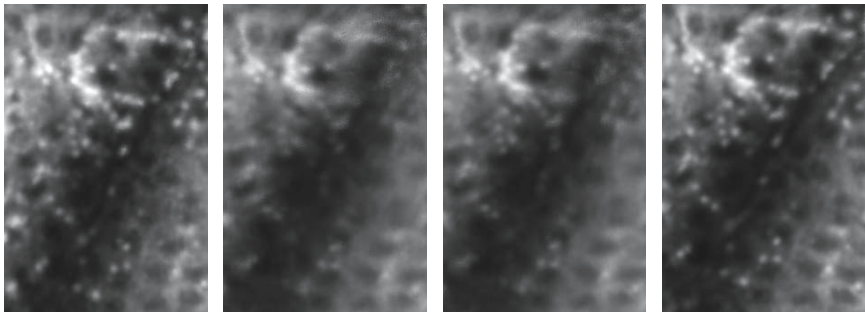
5 Results

In the field of colon cancer research, the development of methods to enable reliable and early detection of tumors is a major goal. In the colon, changes in crypt morphology are known to be early indicators of cancer development. The crypts that undergo these morphological changes are referred to as Aberrant Crypt Foci (ACF) and they can develop into cancerous lesions. Compared to the standard methods of ACF screening, fluorescence FCM enables the operator to see the lesions in real-time and to make an almost immediate evaluation.

However, in many cases the limited field of view restricts the confidence that the operator has in the ACF counting. By offering an extended field of view, mosaicing techniques can be an answer to this restriction.



(a) Mosaic using global frame positioning. FOV: $863\mu\text{m} \times 1462\mu\text{m}$. (b) B-spline registered mosaic. FOV: $862\mu\text{m} \times 1469\mu\text{m}$.



(c) 23^{rd} original input frame. (d) Sequential reg. (no global align.) (e) Global positioning. (f) Fine non rigid frame to mos. reg.

Fig. 2. Mosaic of 50 live mouse colon images (Fluorescence FCM). The zone corresponding to the 23^{rd} input frame is detailed for different steps of the algorithm. FOV one input: $303\mu\text{m} \times 425\mu\text{m}$. Images are courtesy of Danijela Vignjevic, Sylvie Robine, Daniel Louvard, Institut Curie, Paris, France.

The effectiveness of the proposed algorithm is shown on a sequence that has been acquired *in-vivo* on a mouse colon stained by acriflavine at 0,001%. The mouse was treated with azoxymethane (AOM) to induce a colon cancer. As shown in fig.2(b), our algorithm allows for a simultaneous visualization of normal crypts and ACFs.

The global frame positioning mosaicing took approximately 1 min on a 2GHz P4 and 12 min if the non-rigid deformations are compensated. The imaged tissue is really soft and non-linear deformations occur. Figure 2(b) illustrates the gain we obtain by taking into account those non-rigid deformations. Some details are lost if we only use rigid deformations and appear again on our final mosaic. The results shown here prove the feasibility of mosaicing for *in-vivo* soft-tissue microscopy. Current work is dedicated to the validation of the proposed approach.

6 Conclusion

The problem of video mosaicing for *in-vivo* soft tissue confocal microscopy has been explored in this paper. A fully automatic robust approach based on Riemannian statistics and efficient scattered data fitting techniques was proposed. The results shown for two different imaging modalities are promising and encourage the application of the proposed method for qualitative and quantitative studies on the mosaics.

References

1. G. Goualher, A. Perchant, M. Genet, C. Cave, B. Viellerobe, F. Berier, B. Abrat, and N. Ayache, "Towards optical biopsies with an integrated fibered confocal fluorescence microscope," in *Proc. of MICCAI'04*, 2004, pp. 761–768.
2. M. Irani, P. Anandan, and S. Hsu, "Mosaic based representations of video sequences and their applications," in *Proc. ICCV'95*, June 1995, pp. 605–611.
3. J. Davis, "Mosaics of scenes with moving objects," in *Proc. CVPR'98*, 1998, pp. 354–360.
4. S. Peleg, B. Rousso, A. Rav-Acha, and A. Zomet, "Mosaicing on adaptive manifolds," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, no. 10, pp. 1144–1154, Oct. 2000.
5. S. Ourselin, A. Roche, S. Prima, and N. Ayache, "Block matching: A general framework to improve robustness of rigid registration of medical images," in *Proc. MICCAI'00*, 2000, pp. 557–566.
6. A. Can, C. V. Stewart, B. Roysam, and H. L. Tanenbaum, "A feature-based technique for joint linear estimation of high-order image-to-mosaic transformations: Mosaicing the curved human retina," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 24, no. 3, pp. 412–419, Mar. 2004.
7. H. S. Sawhney, S. Hsu, and R. Kumar, "Robust video mosaicing through topology inference and local to global alignment," in *Proc. ECCV'98*, vol. 2, 1998, pp. 103–119.
8. X. Pennec, "Probabilities and statistics on Riemannian manifolds: Basic tools for geometric measurements," in *Proc. NSIP'99*, vol. 1, 1999, pp. 194–198.
9. I. Amidror, "Scattered data interpolation methods for electronic imaging systems: a survey," *J. Electron. Imaging*, vol. 11, no. 2, pp. 157–176, Apr. 2002.

Segmentation and 3D Reconstruction of Microtubules in Total Internal Reflection Fluorescence Microscopy (TIRFM)

Stathis Hadjidemetriou¹, Derek Toomre², and James S. Duncan¹

¹ Departments of Diagnostic Radiology and Biomedical Engineering,
{stathis, duncan}@noodle.med.yale.edu

² Department of Cell Biology, Yale School of Medicine, New Haven, CT 06520
derek.toomre@yale.edu

Abstract. The interaction of the microtubules with the cell cortex plays numerous critical roles in a cell. For instance, it directs vesicle delivery, and modulates membrane adhesions pivotal for cell movement as well as mitosis. Abnormal function of the microtubules is involved in cancer. An effective method to observe microtubule function adjacent to the cortex is TIRFM. To date most analysis of TIRFM images has been done by visual inspection and manual tracing. In this work we have developed a method to automatically process TIRFM images of microtubules so as to enable high throughput quantitative studies. The microtubules are extracted in terms of consecutive segments. The segments are described via Hamilton-Jacobi equations. Subsequently, the algorithm performs a limited reconstruction of the microtubules in 3D. Last, we evaluate our method with phantom as well as real TIRFM images of living cells.

1 Introduction

Microtubules are biopolymers of the cytoplasm. They are composed of tubulin and have a diameter of ≈ 25 nm [1]. One of their ends interacts with the cell cortex and cell adhesions. Those interactions are pivotal for mitosis, cell migration, and active vesicle delivery [1]. In mammalian cells the other end is anchored at an organelle positioned next to the nucleus. The assembly formed provides structural stability.

Microtubules also play critical roles for cells in pathological states such as cancer. Thus, microtubules have been identified as an important target for anti-cancer drugs. Chemical entities such as the taxanes inhibit the polymerization of microtubules. Therefore, they stabilize the microtubule assembly and retard cancer growth. Several novel taxane derivatives are in active clinical development. It is essential to elucidate their specific effects on the microtubule assembly. Current in vivo studies are done by observation or manually [1,2,3]. Automation can enable higher throughput screening of new microtubule drugs. The latter with the goal of improving the therapeutic index with reduced toxicity.

The main objectives of analyzing microtubule data from TIRFM are to quantify their position and dynamics relative to the cortex. To this end the microtubules must be segmented, reconstructed in 3D, and their motion must be

tracked. Also, the state of the assembly must be expressed concisely. In this work we present an algorithm that addresses the problems of 2D microtubule segmentation, 3D reconstruction, and state measurement.

Some common image enhancement techniques for tubular biomedical structures have been linear filtering and morphological operations [4]. The image segmentation of such structures has been implemented sequentially using local information [5,4]. Tubular structures have also been segmented as minimal global paths over intensity for colon in virtual endoscopy [6]. Similarly, global minimal paths have been used to extract white matter fibers in diffusion tensor imaging [7]. Typically, global approaches to segmentation of a tubular structure require the specification of both end points of that structure [6,7]. Some segmentation algorithms have been designed for specific microscopy techniques [8,9]. However, an approach for automatic processing of TIRFM images of microtubules has not been evident in the literature.

2 Methods

2.1 TIRFM Image Formation and Requirements for Processing

The substrate was living cells. They were transfected with tubulin fused with green fluorescent protein. TIRFM exploits the properties of an induced evanescent wave from a totally internally reflected laser light. It selectively illuminates and excites fluorophores in a very thin slice, $\approx 100\text{ nm}$, immediately adjacent to a glass-cell interface [10,2,11]. Specifically, a beam of light travelling in a medium of high refractive index, such as glass with $n_g \approx 1.51$, enters a medium of lower refractive index, such as the cytosol of a cell with $n_c \approx 1.37$, beyond a certain critical angle a_c . The light is totally internally reflected and an evanescent field is induced in the cell. The exponential falloff of the evanescent field intensity does not allow fluorophores farther away from the interface to be excited. The in-vivo imaging of microtubules with TIRFM is drawn in figure 1.

A TIRFM image can be represented by the map $I : D \rightarrow \mathbb{R}$, where $D \in \mathbb{R}^2$ is the domain of dimensions (x, y) . The image intensity of a microtubule point at depth $z(x, y)$ above the interface is [10,2]:

$$I_z(x, y) = I_{max} \exp^{-(z(x,y)-z_{min})/d_p}, \quad (1)$$

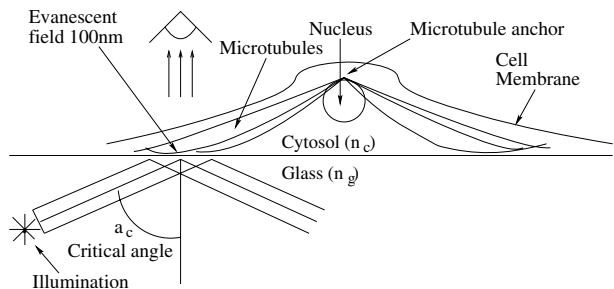


Fig. 1. TIRFM excites fluorescent microtubules that lie exclusively in a thin slice $\approx 100\text{ nm}$ immediately adjacent to the glass.

where d_p is the penetration depth, I_{max} is the maximum image intensity, and z_{min} is the corresponding depth that is an experimental constant with $z_{min} \approx 10 \text{ nm}$. The spatial resolution is $\approx 135 \text{ nm}/\text{pixel}$. This technique yields high-contrast images of surface events and allows the imaging of the interaction of the microtubules with the cell cortex. That interaction mediates key cellular processes such as endocytosis, exocytosis, and protein signaling [10,2,11].

The algorithm extracts a microtubule from its tip alone. The other end is not clearly distinguishable in TIRFM. The algorithm is able to extract the microtubule centerline, is robust to noise, and is robust to microtubule proximity. It can also resolve microtubule intersections. The fluorescence of isolated microtubules is used to obtain information about their 3D position above the interface.

2.2 Cumulative Cost Map for Microtubule Segment

A microtubule segment in domain D of length L can be represented by the curve $\mathcal{M}(s) : [0, L] \rightarrow D$, where s is the arc-length parameter. A microtubule segment has two boundary points $\mathcal{M}(0) = \mathbf{p}_0$ and $\mathcal{M}(L) = \mathbf{p}_1$. The set of all curves with these boundary points is $\mathcal{A}_{\mathbf{p}_0, \mathbf{p}_1}$. Every pixel at $\mathbf{p} = (x, y)$ is associated with cost Q and the set of curves is associated with the cost functional $E(\mathcal{A}) = \int_0^L Q(\mathcal{M}(s), \mathcal{M}'(s)) ds$. We find the curve of minimum E over the set $\mathcal{A}_{\mathbf{p}_0, \mathbf{p}_1}$. To this end we compute the minimum cost E over the sets of intermediate curves $\mathcal{A}_{\mathbf{p}_0, \mathbf{p}}$. The result is a map of minimum cost values $U_0(\mathbf{p}) : D \rightarrow \mathbb{R}^+$ that starts at \mathbf{p}_0 . That is, $U_0(\mathbf{p}) = \min_{\mathcal{M}(L)=\mathbf{p}} \{ \int_{\mathcal{M}} Q ds \}$. Differentiating this relation gives the Hamilton-Jacobi equation:

$$\|\nabla U_0\| = Q \quad (2)$$

that can be solved to obtain U_0 .

The pixel cost Q is the product of two factors. The first is proportional to the reverse intensity $Q_1 = 1 - I(\mathbf{p})/I_{max}$. It is isotropic and favors microtubule fluorescence. The second factor, Q_2 , is anisotropic and favors microtubule centerline. It uses directional parameters associated with every pixel, namely vectors \mathbf{n}_1 and \mathbf{n}_2 , as well as scalars r_1 and r_2 . The pixel cost is non-negative and given by:

$$Q(\mathbf{p}, \nabla U_0) = \left(1 - \frac{I(\mathbf{p})}{I_{max}} \right) \times \left(\frac{1}{\|\nabla U_0\|} \sum_{i=1}^{i=2} \left| \frac{\mathbf{n}_i \cdot \nabla U_0}{r_i + \epsilon} \right| \right), \quad (3)$$

where ϵ is a very small regularizing constant. The directional pixel parameters of Q are derived from the eigenvector decomposition of the Hessian. The vectors \mathbf{n}_1 and \mathbf{n}_2 are set to the eigenvectors of the smaller eigenvalue and larger eigenvalue magnitudes, respectively. The scalars r_1 and r_2 are set to the larger and smaller eigenvalue magnitudes, respectively.

2.3 Solution for a Cumulative Map and Curve Extraction

The solution $U_0(\mathbf{p})$ of equation (2) can be expressed in terms of consecutive level sets [6]. A level set $U_0(\mathbf{p}) = t$ is the closed curve $\mathcal{C}(v, t) : \mathbb{R}^1 \rightarrow \mathbb{R}^2$, where v is

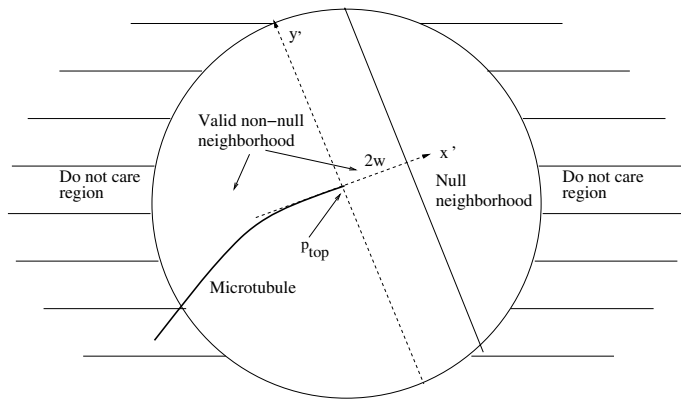
the arc-length parameter. The family $\mathcal{C}(v, t)$ over t is computed with the time evolution equation $\frac{\partial \mathcal{C}(v, t)}{\partial t} = \frac{1}{Q} \mathbf{d}(v, t)$, where $\mathbf{d}(v, t)$ is the normal to $\mathcal{C}(v, t)$ [6]. It is initialized with a curve $\mathcal{C}(\cdot, 0)$ around \mathbf{p}_0 . The numerical solution visits each pixel once [12]. The retrieval of the pixels is done efficiently with a priority heap [12] and their costs Q in equation (3) are minimized. This minimization is simple since Q is a convex function of its parameters with a rhombus boundary [13,12]. The minimum is found over directions \mathbf{d} originating from the eight-connected neighborhood. Subsequently, we extract \mathcal{M} by starting from \mathbf{p}_1 and proceeding along $\frac{\partial \mathcal{M}}{\partial s} = -\nabla U_0$. The negative discrete gradient direction of U_0 is followed to arrive at $p_0 \in \mathcal{C}(\cdot, 0)$.

2.4 Microtubule Segmentation in 2D

A microtubule is extracted in terms of consecutive segments. Each segment is computed based on its neighborhood and the remaining image domain is not considered. The neighborhood is formed using a local coordinate system (x^l, y^l) with origin O^l . The origin for the extraction of the first segment is the microtubule tip, \mathbf{p}_{tip} . The valid and non-null neighborhood is within the circle $\|(x^l, y^l)\| \leq 4w$ and a surrounding annulus is nulled. The origin O^l for the extraction of all subsequent segments is the most recently segmented microtubule point \mathbf{p}_{top} . The valid neighborhood is within the circle $\|(x^l, y^l)\| \leq 4w$. The x^l -axis is the microtubule tangent. The null level set is in front of \mathbf{p}_{top} , where $x^l > 2w$ and w is the microtubule width. The coordinate system is shown in figure 2. Backtracking from \mathbf{p}_{tip} or \mathbf{p}_{top} to the null level set gives curves μ_i .

To determine whether a curve μ_i is indeed a microtubule segment we compute a contrast measure across its axis. We form three zones of width w , an inner zone along the curve axis, and two surrounding outer zones. We compute the average intensities g_{in}^i of the inner zone and g_{out}^i of the two outer zones. The ratio $c^i = g_{in}^i / g_{out}^i$ gives a contrast measure. If $c^i > 1$, the curve μ_i is appended to the microtubule \mathcal{M}_{tot} , otherwise the extraction of the microtubule ends. The algorithm is repeated independently for each microtubule tip specified. A micro-

Fig. 2. A coordinate system centered at the most recently segmented point, p_{top} , of the microtubule.



tubule consists of $\frac{L_{tot}}{2w}$ segments, where L_{tot} is the total microtubule length. The cost of extracting one segment from a neighborhood of size N is $O(N \log N)$, where $N \leq \pi(4w)^2$. Thus, the total segmentation cost is $O(rL_{tot}w \log w)$, where r is the number of microtubule tips.

2.5 Partial Microtubule Reconstruction in 3D

The microtubule segments close to their tips are distinct, adjacent to the glass, and have a higher signal to noise ratio. That part of the microtubules is used to perform a partial 3D reconstruction. The TIRFM formula in equation (1) is inverted to give the depth coordinate $z(x, y) = z_{min} + d_p \log(I_{max}/I_z(x, y))$ relative to z_{min} [2,10]. Subsequently, the 3D data points are mean normalized and subjected to eigenvector analysis. The eigenvector of the largest eigenvalue is the microtubule tip tangent. Finally, statistics of these measures over all the microtubules are computed.

3 Experiments on Microtubule Extraction

We tested the sensitivity of the algorithm using parameterized sets of phantom images of microtubules with natural shapes. The phantom data was generated based on equation (1) with $d_p = 15 \text{ nm}$ and $z_{min} = 0$. To model the degradation caused by TIRFM the microtubules were first smoothed with a Gaussian filter of standard deviation equal to two pixels and then corrupted with Gaussian noise. The ratio of the maximum intensity of the phantom microtubules to the standard deviation of the Gaussian noise gives the signal to noise ratio (SNR). The noise for the phantom images is $SNR = 100$ unless otherwise stated. The cross section of a microtubule is assumed to be Gaussian. The width within one standard deviation of the peak is the effective microtubule width. The phantom microtubules have $w \approx 5$. The size of the phantom images is 150×150 .

To evaluate the algorithm we measured the total sum of the intensities corresponding to the segmented microtubules as well as the average error in the elevation estimates. Both were plotted as a function of the phantom set parameter in figure 4. We first tested the sensitivity of microtubule segmentation using a phantom set of images with noise in the range $SNR = 5 - 50$. An example is in figures 3 (a-c). In figure 4 (a) the 2D segmentation is robust to noise. The elevation estimate in figure 4 (b), however, is more sensitive.

The sensitivity of the algorithm with respect to the proximity between neighboring microtubules was also tested. We used images containing five phantom microtubules each. The distance between consecutive microtubules is $\delta = 5 - 14$ pixels. In figure 3 (d) that distance is 6 pixels. The biggest part of the microtubules was segmented successfully as shown in figure 3 (e). The 2D segmentation errors in figure 4 (c) are large only for $\delta \leq w$. The elevation errors in figure 4 (d) are not affected by proximity because the initial microtubule segments have a high SNR.

Subsequently, we used sinusoidal microtubules to test the sensitivity of the algorithm to curvature. The amplitudes of the microtubules are in the range $A =$

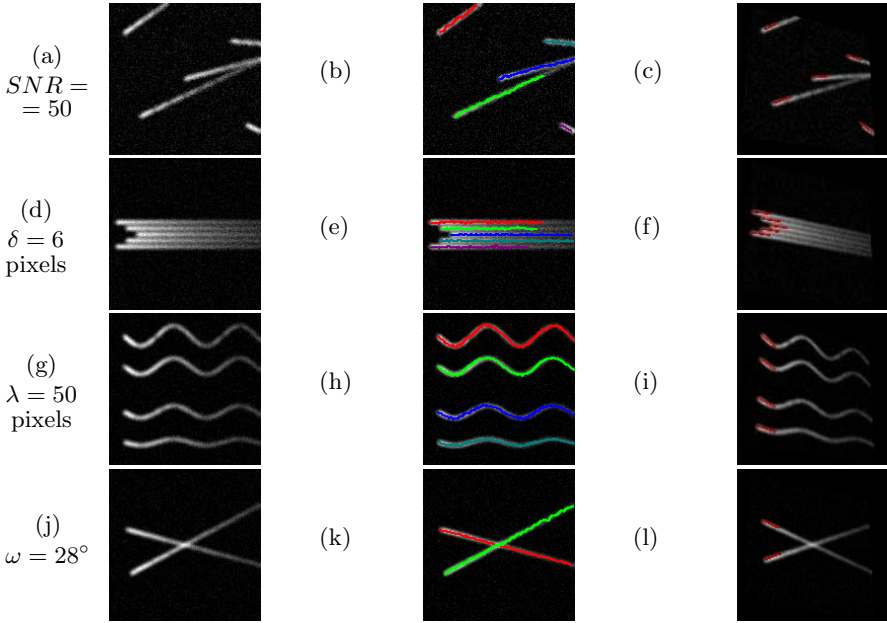


Fig. 3. In the left column are the phantom images. In the middle column are the 2D segmentations in color. The right column shows the 3D tip tangents in red.

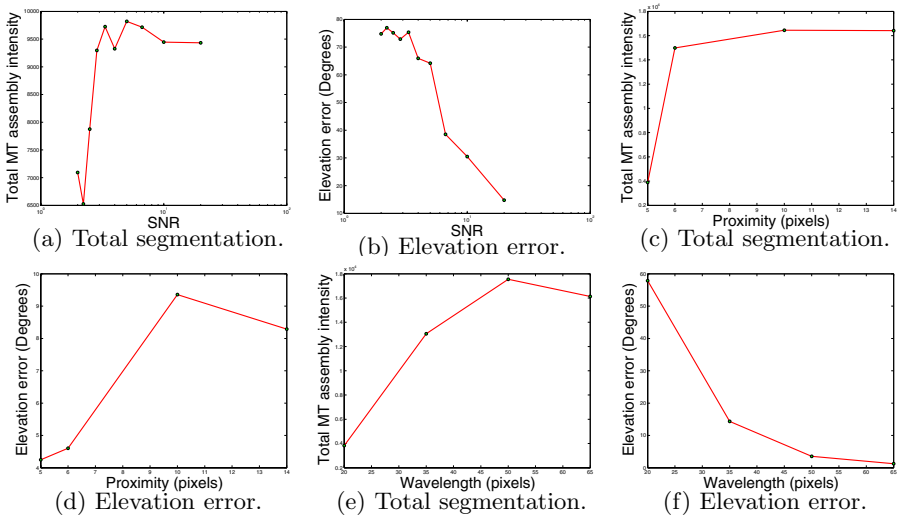


Fig. 4. Evaluation plots as a function of the corresponding phantom set parameter

3–10 pixels in each image. The wavelength takes values $\lambda = 20–65$ pixels across the set. An example is in figures 3 (g-i). The plots in figures 4 (e-f) show that

Table 1. Statistics of microtubules for the real data

Image	Size	d_p (nm)	Elevation (Degrees)		Depth (nm)		Evaluation error (pixels)
			Mean	St. dev.	Mean	St. dev.	
figure 5(a)	179×150	45	0.24°	0.23°	42.7	16.0	1.6
figure 5(d)	205×140	165	1.11°	0.70°	70.0	37.0	1.5
figure 5(g)	125×56	45	0.20°	0.05°	8.5	6.6	1.3

the algorithm fails only for very high curvature. The last set of phantom images tested the sensitivity of the algorithm with respect to microtubule intersections. The resolution of microtubules succeeds for intersection angles $\omega \geq 28^\circ$. This is shown in figures 3 (j) and (k).

The algorithm was used to analyze several images of living epithelial cells acquired with TIRFM. Some examples are shown in the left column of figure 5. The images in figure 5 are in reverse intensity to improve visualization. The extracted microtubules are shown next to them in the same row drawn with different colors. The images in the right column show a viewpoint of the 3D tangents of the microtubules at their tips in red. The data and reconstruction statistics are in table 1.

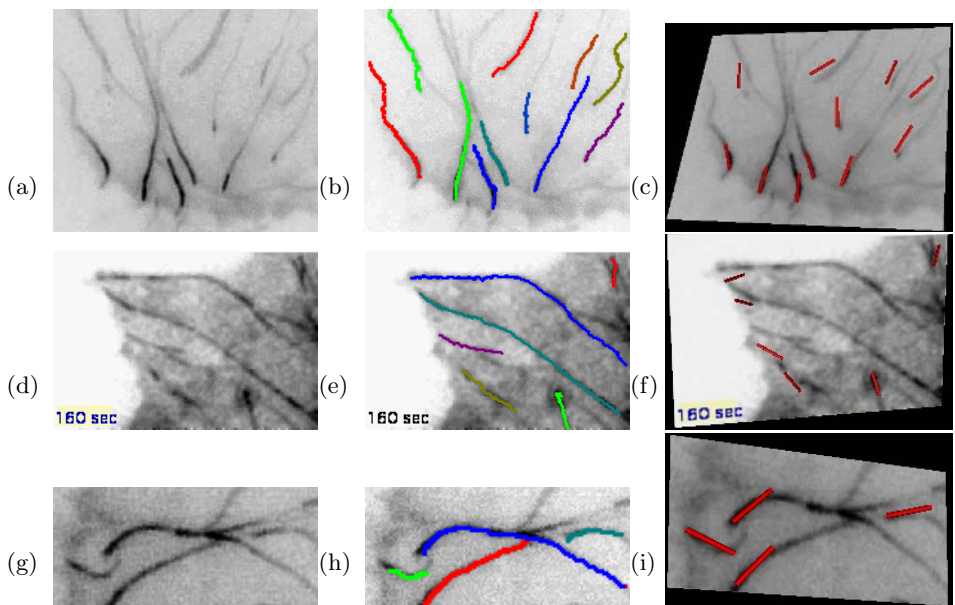


Fig. 5. The data are in the left column. The next column shows the segmented microtubules in color, and the right column shows the 3D tip tangents in red.

The 2D segmentation was evaluated with ten manually traced microtubule points starting from their tips. The error for each specified point was its lowest distance from the segmented centerline of that microtubule. The total segmentation error was the average error over the specified points. The errors for the segmentations shown in figure 5 (b), figure 5 (e), and figure 5 (h) are in table 1. The low error demonstrates the accuracy of the extraction of the microtubule centerline. The experiments were performed with a Xeon CPU. The average total computation time of the algorithm for each phantom image was ≈ 29 sec, and for the real data it was ≈ 3 min 13 sec. The low run time demonstrates the efficiency of the algorithm.

4 Discussion

The interaction of the microtubules with the cell cortex is pivotal for cell physiology. The microtubules are also involved in cell pathological states such as cancer. TIRFM is a convenient method to examine the critical region close to the cell cortex. We have developed a semi-automatic algorithm for high throughput quantitative study of TIRFM images of microtubules. It first extracts microtubules in 2D in terms of consecutive segments. The extraction of the segments is robust to noise and along their centerlines. The sequential extraction of the microtubules allows microtubule intersections, the extraction from a single starting point, as well as robustness to microtubule proximity. Subsequently, the algorithm implements a 3D tip tangent and depth reconstruction. The algorithm has been applied to phantom as well as real data. It was shown to be efficient and accurate.

References

1. Desai, A., Mitchison, T.: Microtubule polymerization dynamics. *Annu. Rev. Cell Dev. Biol.* **13** (1997) 83–117
2. Krylyshkina, O., Anderson, K., Kaverina, I., Upmann, I., Manstein, D., Small, J., Toomre, D.: Nanometer targeting of microtubules to focal adhesions. *The Journal of Cell Biology* **61** (2003) 853–859
3. Waterman-Storer, C., Salmon, W., Salmon, E.: Feedback interactions between cell–cell adherens junctions and cytoskeletal dynamics in newt lung epithelial cells. *Molecular Biology of the Cell* **11** (2000) 2471–2483
4. Hadjidemetriou, S., Duncan, J., Toomre, D., Tuck, D.: Automatic quantification of microtubule dynamics. In: *Proc. of IEEE ISBI*. (2004) 656–659
5. Noordmans, H., Smeulders, A.: High accuracy tracking of 2D/3D curved line-structures by consecutive cross-section matching. *Pattern Recognition Letters* **19** (1998) 97–111
6. Deschamps, T., Cohen, L.: Fast extraction of minimal paths in 3D images and applications to virtual endoscopy. *Med. Image Analysis* **5** (2001) 281–299
7. Parker, C., Wheeler-Kingshott, C., Barker, G.: Estimating distributed anatomical connectivity using fast marching methods and diffusion tensor imaging. *IEEE Trans. in Medical Imaging* **21** (2002) 505–512

8. Danuser, G., Tran, P., Salmon, E.: Tracking differential interference contrast diffraction line images with nanometre sensitivity. *Journal of Microscopy* **198** (1999) 34–53
9. Ponti, A., Valotton, P., Salmon, W., C.M. Waterman-Storer, Danuser, G.: Computational analysis of f-actin turnover in cortical actin meshworks using fluorescent speckle microscopy. *Biophysical Journal* **84** (2003) 3336–3352
10. Axelrod, D.: Total internal reflection fluorescence at biological surfaces, in *Noninvasive techniques in cell biology*. Wiley-Liss, New York, New York, Inc. (1990)
11. Toomre, D., Manstein, D.: Lighting up the cell surface with evanescent wave microscopy. *Trends in Cell Biology* **11** (2001) 298–303
12. Sethian, J., Vladimirov, A.: Ordered upwind methods for static Hamilton-Jacobi equations: Theory and algorithms. *SIAM J. Numer. Anal.* **41** (2003) 325–363
13. Lin, Q.: Enhancement, Extraction, and Visualization of 3D Volume Data. Thesis Licentiate, Dep. of Electrical Engineering, Linköping University (2001)

Ligament Strains Predict Knee Motion After Total Joint Replacement

A Kinematic Analysis of The Sigma Knee

Elvis C.S. Chen¹, Joel L. Lanovaz², and Randy E. Ellis^{1,2,3}

¹ School of Computing

chene@cs.queensu.ca

² Department of Mechanical Engineering

lanovaz@me.queensu.ca

³ Department of Surgery, Queen's University, Kingston, Ontario, Canada

ellis@cs.queensu.ca

Abstract. A passive forward kinematics knee model was used to predict knee motion of a total joint replacement. Given a joint angle, maps of articular surfaces, and patient-specific ligament properties, this model predicted femorotibial contact locations based on the principle of ligament-strain minimization. The model was validated by physical experiments on a commonly implanted knee prosthesis, showing excellent correspondence between the model and actual physical motion. Results suggest that the knee prosthesis studied required an intact posterior cruciate ligament to induce the desirable roll-back motion, and that a single-bundle model of major knee ligaments generated kinematics similar to that of a multi-bundle ligament model. Implications are that a passive model may predict knee kinematics of a given patient, so it may be possible to optimize the implantation of a prosthesis intraoperatively.

1 Introduction

Total knee replacements (TKRs) are currently designed with noncongruent articular surfaces to accommodate human biomechanics and wear properties of the tibial component [1]. It is known that, when no external forces are present, tensile forces stored in knee ligaments move the knee joint to an equilibrium point where ligament strain is minimized [2].

We have developed and tested a validation protocol for a Forward Knee Kinematics (FKK) model of how human knees move after joint-replacement surgery. A contact-determination algorithm was developed to depict *in situ* femorotibial contact. This algorithm was independently validated using pressure-sensitive film that established the actual femorotibial contact under realistic loading conditions produced by a custom knee jig. The FKK model simulated the physics of the knee jig, which produced a set of *in vivo* femorotibial contact (through the contact determination algorithm) that was compared to *in vitro* contacts predicted by the FKK model. The FKK model was validated; various ligament configurations were considered; and their derived knee kinematics are examined.

2 Materials and Methods

Articular surfaces of a size-3 Sigma Knee (Johnson & Johnson) were laser-scanned at a resolution of 0.4mm , resulting in two point clouds of approximately 31,000 and 19,000 for each of the femoral and tibial component, respectively. Joint coordinate frames [3] were assigned to these TKR components. The *absolute*, space-fixed coordinate frame was associated with the tibial component, whereas the *relative*, body-fixed coordinate frame was associated with the femoral component. Without loss of generality, the Z -axes were aligned with the anatomical axes of the lower limbs. The X -axes were perpendicular to the Z -axes lying on the sagittal plane with the anterior direction being positive. The Y -axes were derived as the cross product of the two: $Y = Z \times X$.

These two coordinate frames were related by a homogeneous transformation. If \bar{p} is a 3×1 column vector that measures the coordinate of a point in the tibial system, then its corresponding femoral location \bar{q} can be expressed as:

$$\begin{bmatrix} \bar{q} \\ 1 \end{bmatrix} = \begin{bmatrix} R & \bar{d} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{p} \\ 1 \end{bmatrix} \quad (1)$$

where R is a 3×3 orthogonal rotational matrix and \bar{d} is a 3×1 displacement vector. Because the tibia was assumed to be fixed, R and \bar{d} represented the relative joint angle and position of the femoral component. The rigid-body transformation of a point is a rotation (R) to the mobile coordinate frame followed by a linear displacement (\bar{d}).

2.1 Forward Knee Kinematics

The passive forward kinematics knee model proposed by Chen *et. al.* [2] was implemented with a minor modification. The strain energy for each ligament was calculated as

$$E = \begin{cases} .5 \times K \times (L - \tilde{L})^2 + B \times (L - \tilde{L}), & \text{if } L \geq \tilde{L} \\ 0, & \text{if } L < \tilde{L} \end{cases} \quad (2)$$

where \tilde{L} was the neutral length of the ligament, L was the Euclidean distance between ligament attachment points, and K and B were the spring constants.

2.2 Contact Determination Algorithm

Mathematically, two points are in contact if they coincide in space and have point normals opposite in direction. Let \bar{p} and \bar{q} be contacting points on the femoral and tibial components, respectively, with associated normals \bar{p}_n and \bar{q}_n being

$$\| p - q \| \leq \delta \quad (3)$$

$$-(p_n \cdot q_n) \geq 1 - \epsilon \quad (4)$$

where $\| \bullet \|$ denotes the Euclidean vector norm, and δ and ϵ are distance and angular tolerances, respectively. These tolerance are necessary because the articular surfaces were sampled at a finite resolution. Equation (3) and (4) were the *contact conditions* used to determine points on the contacting surfaces.

To validate the contact determination algorithm, a static contact analysis experiment was performed using a commercial 4-DOF knee implant wear test machine (Force 5, Advanced Mechanical Technologies Inc., USA). The femoral and tibial components were mounted on custom jigs using a polymethylmethacrylate cement that is typically used to affix knee prostheses to human bone. Dynamic Reference Bodies (DRBs) were rigidly fixed to the femoral and tibial components. An accurate optical system (Optotrak 3020, Northern Digital, Canada) was used to record the 6D poses. Registration between the DRBs and the component models were obtained using the Iterative Closest Point [4] algorithm.

For the static trials, the neutral (0° flexion) pose of the components was found using the guidelines given in ISO Standard 14243-3. This standard was used for TKR wear testing and supplies four control waveforms (vertical force, flexion angle, anterior/posterior position and internal/external rotation) for a typical walking cycle. Six poses corresponding to 0, 13, 45, 56, 72, and 88 percent of the standard walking cycle were chosen for testing. These corresponded to start, 1st max vertical force, 2nd max vertical force, max AP position, max FE angle and max IE angle.

For each test pose, a Fuji Prescale pressure sensitive film (Ultra Super Low Pressure Grade), cut to fit the tibial contact surfaces, was fixed to the tibial component using two-sided tape on the anterior region of the component (where femorotibial contact was not possible). A small vertical load was applied to the tibia to bring the components into initial contact with minimal sliding. The 6D pose for each component was collected and was used in the contact-determination algorithm.

After the components were unloaded, a tracked optical probe was used to trace the outline of the Fuji film stains while the film was still in place on the tibial surface. The 3D locations of traced points were superimposed on the contact regions determined by the contact determination algorithm, depicted in Fig. 1.

2.3 Spring Ligament Apparatus

To validate the FKK model, a physical apparatus was constructed to simulate passive knee flexion after TKR. The femoral component was mounted on a rigid frame and rotated so that the femoral long axis was perpendicular to gravitation. Plexiglas plates were rigidly attached to the apparatus frame approximately 3cm from the medial and lateral sides of the femoral component. The tibial component was mounted on the proximal end of a simulated shank (i.e., the distal lower limb segment). The shank consisted of four long threaded rods, arranged to form a $6\text{cm} \times 6\text{cm} \times 30\text{cm}$ parallelepiped shape and held in place by Plexiglas end plates. A 2.2kg mass was fixed at the distal end of the shank, approximating the inertia of the median North American foot.

A DRB was rigidly fixed to the femoral mounting jig and a second DRB was rigidly attached to the shank. The component models were registered to the kinematic system with the same procedure used in the static contact experiments.

Steel tension springs were used to simulate some of the ligament constraints of the knee joint. The ends of each ligament spring were fixed to spherical rod-end bearings. The spherical bearings for each ligament spring were then fixed on one end to Plexiglas plates beside the femur and on the other end to custom ABS jigs attached to the shank segment. The apparatus was designed to allow for one, two, or three springs to

be attached on each of the medial and lateral sides of the joint and at various positions along the proximal/distal axis of the shank. This simulated the ligament constraints of the MCL and LCL.

Two sets of experiments were conducted with this apparatus. The first one involved a total of 6 springs, 3 simulating the MCL and 3 simulating the LCL. The second experiment involved only 2 springs, 1 simulating the MCL with the other simulating the LCL. The mechanical properties of the springs are listed in Table 1. In both experiments, the springs were positioned symmetrically with respect to the medial and lateral sides. The exact physical insertion locations for the springs, relative to the component models, were obtained using the calibrated optical tracking probe.

Table 1. The mechanical properties of steel tension springs used in Eqn. 2

	MCL 1	MCL 2	MCL 3	LCL 1	LCL 2	LCL 3	MCL	LCL
unloaded length (mm)	65.0	65.0	65.0	65.0	65.0	65.0	124.0	125.0
K (N/mm)	.66	.66	.65	.66	.66	.66	1.36	1.37
B (N)	5.90	5.80	6.00	6.80	4.00	6.50	2.57	5.53

The goal of the experimental setup was to simulate passive joint flexion. The simulated joint was extended manually in discrete increments from approximately 100° of flexion to 0° of flexion and then flexed back to 100° of flexion. A total of 171 and 36 poses were obtained with 6 and 2 springs respectively. At each pose, the joint was allowed to rest to a stable configuration with minimal force applied to the shank.

2.4 Patient-Specific Kinematics

It was technically difficult to include a PCL mechanism in our apparatus. Instead, we generated a physiologically plausible knee kinematics by adapting the patient-specific ligament data of Chen *et. al.* [2]. A total of 11 ligament bundles were taken into consideration: 3, 4, and 4 for each of the PCL/MCL/LCL, respectively. The MCL was symmetrical to the LCL.

Four simulations were performed. The first simulation utilized all 11 ligament bundles. The second simulation considered the knee kinematics without the PCL. The 3rd and the 4th simulations were the same as the first two, but with single-bundle ligaments in place of the detailed bundles. The single-bundle ligament was artificially generated by taking the geometrical mean of the ligament insertions in the multi-bundle ligament configuration, using summed spring constants.

3 Results

3.1 Contact Determination Algorithm

Figure 1 depicts a typical experimental result between the Fuji film produced with the Force 5 knee tester and the contact determination algorithm (with $\delta = 0.4mm$ and $\epsilon = 0.004$). The peripheries of the Fuji contacts were digitized and superimposed to the region produced by the algorithm. They show high degree of agreement.

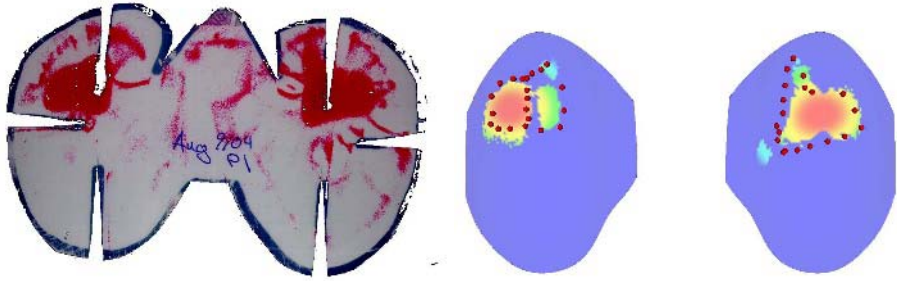


Fig. 1. Contacts determined using Fuji film and the contact determination algorithm

3.2 Apparatus Kinematics

For each recorded joint pose, two types of contact locations were generated. First, the *in situ* contact locations were determined using the contact determination algorithm. The same joint angle was used in the FKK model and, in conjunction with spring information, the *in vitro* contact locations were calculated. Figure 2 depicts a typical result: for the given joint angle, an *energy map* was produced depicting all feasible contact locations (Fig. 2(a)). The subset of the feasible contacts resulting in the minimal ligament strain energy formed the *in vitro contact patch* (Fig. 2(b)). The *instantaneous contact points* were calculated as the point on articular surfaces closest to the centroid of the contact patch, and they constrained the amount of the femoral component displacement (\bar{d} in (1)).

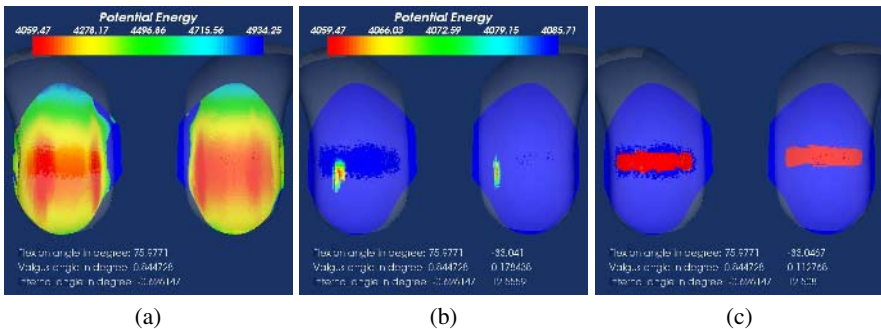


Fig. 2. Contact predicted by the FKK model and the contact determination algorithm

For experiments involving 6 springs, the difference in the displacement vector \bar{d} calculated by the two methods were on average $0.43mm$, with a standard deviation of $0.35mm$. For experiments involving 2 springs, the mean difference was $0.69mm$ with a standard deviation of $0.15mm$.

3.3 Sigma Knee Kinematics

To analyze the kinematics generated by the mechanical springs, contact locations from 0° to 120° flexion were calculated at 1° increments using the FKK model. Figure 3 depicts the contact locations generated using the 6-spring and 2-spring configurations. The placement of contact locations suggest that the kinematics for both spring configuration were basically the same, and that the femoral component spun in place throughout flexion with no obvious anterior-posterior translation. To further demonstrate the difference in kinematics generated, Fig. 4(a) depicts the femoral translation through flexion angle and Fig. 4(b) depicts the Euclidian distance in the displacement vectors \vec{d} for each spring configuration. mean difference was 0.62mm with a standard deviation of 0.36mm , which is negligible.

Four simulations were generated with the physiologically plausible ligament configurations. Contact locations were calculated from 0° to 120° flexion at 1° increments (Fig. 5). The amount of ligament strain stored in knee ligament are depicted in Fig. 5.

In simulations accounting for a PCL mechanism, contact locations at full extension were located at the anterior portion of the tibia; for these, the MCL and LCL were both taut and the PCL was relaxed. As the knee flexed, the contact locations gradually moved posteriorly as the PCL tightened. At full flexion, the MCL and LCL were relaxed but the PCL was taut, pulling the femur posteriorly. In all cases, at the beginning and the end of flexion there were some spinning motion.

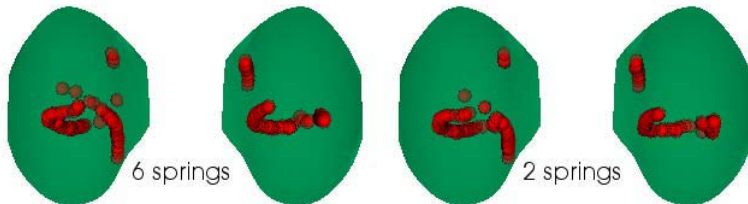


Fig. 3. Contact locations for the 6-springs (a) and 2-springs (b) knee jig from full extension to 120° flexion. In both configurations, the femoral component spun in place through the flexion.

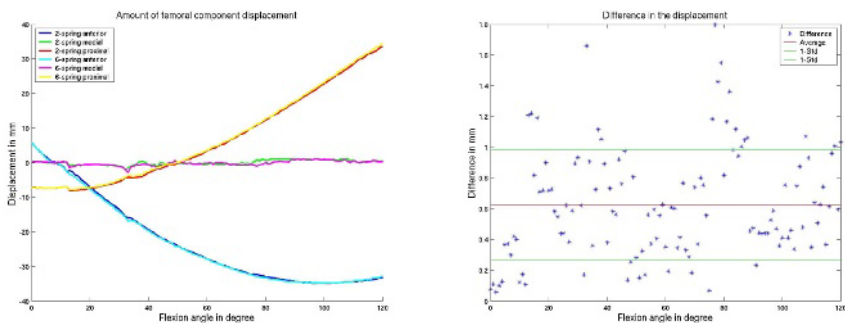


Fig. 4. Plot of the femoral displacement determined with 6-spring and 2-spring knee jig

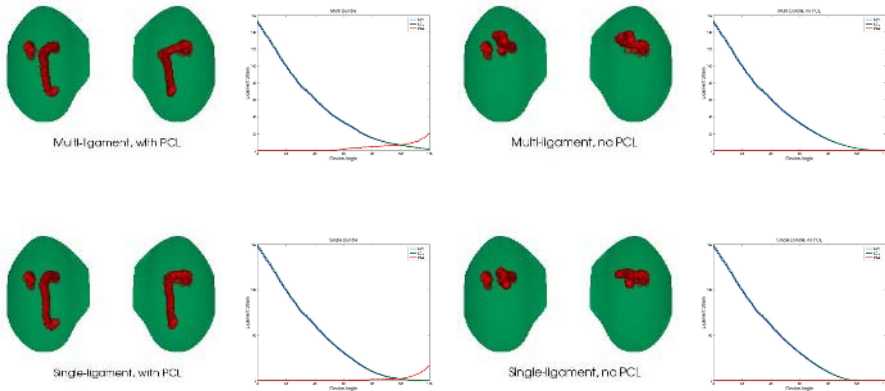


Fig. 5. Contact paths and ligament profile of different ligament configurations: (a) multi-bundle ligament, (b) multi-bundle without PCL, (c) single-bundle ligament, (d) single-bundle without PCL

4 Discussion

The contact determination algorithm was based on (3) and (4). The querying mechanism was implemented with a KD-tree [5] for a speedy retrieval. On a 2GHz PC, the plots depicted in Fig. 1 were generated in less than 1 second, which is acceptable for intraoperative use. This algorithm was validated by a Fuji film study, which is widely accepted as the gold standard for knee kinematics [6].

A custom apparatus was constructed and used to further validate the passive forward kinematics knee model. Two physical experiments involving different ligament configurations were conducted. In each case, the contact locations predicted by the FKK model agreed with the gold-standard *in vitro* contact locations determined with the apparatus with sub-millimeter accuracy.

After the FKK model was validated, various physiologically plausible ligament configurations were used to analyze the kinematics of a specific prosthesis (the Sigma Knee). In simulations without a PCL, the contact locations for all flexion angles were located in the anterior portion of the tibia. This implied that the femur was basically spinning in place, exactly as was observed in the apparatus with mechanical springs. With a PCL present, the contact location gradually moved posteriorly as the knee flexed. In all simulations, there was no obvious difference in the kinematics generated with a multi-bundle ligament model versus a single-bundle ligament model.

5 Conclusion

To determine the *in situ* femorotibial contact location, a fast contact-determination algorithm was developed and validated with the gold-standard pressure-sensitive Fuji film. Physical experiments demonstrated that the Forward Kinematics Knee (FKK) model [2] predicted knee motion with sub-millimeter accuracy.

After the FKK model was validated, various ligament configurations were used to determine their influence on the predicted knee kinematics. Simulation results suggested that:

- The Sigma Knee requires a PCL mechanism to produce a roll-back motion;
- In the absence of a PCL mechanism, the Sigma Knee spins in place; and
- In both simulated and experimental results, multi-fiber and single-fiber ligament configuration produced similar kinematics.

This study is limited by the number of ligaments examined, the configurations of the ligaments, and that only one prosthesis was simulated. Further enquiry into this subject is indicated.

These results have implications for both knee modelers and for surgeons. For knee modelers, the results suggest that a good simple model of knee ligaments suffice to model the Sigma Knee, so it is better to get a good rough guess of knee-ligament geometry rather than to toil for a detailed model. For surgeons, these results suggest that the high surface conformity of the Sigma Knee bearing surfaces are relatively insensitive to implantation geometry, that is, that the minor variations of surgical implantation are not likely to produce major changes in kinematics. Our results suggest that the Sigma Knee, in the hands of an experienced surgeon, may produce motion consistent with a normal healthy knee.

References

1. O'Connor, J.J., Goodfellow, J.W.: The role of meniscal bearing vs. fixed interface in unicondylar and bicondylar arthroplasty. In Goldberg, V.M., ed.: *Controversies of Total Knee Arthroplasty*. Raven Press (1991) 27–49
2. Chen, E., Ellis, R.E., Bryant, J.T., Rudan, J.F.: A computational model of postoperative knee kinematics. *Medical Image Analysis* **5** (2001) 317–330
3. Grood, E.S., Suntay, W.J.: A joint coordinate system for the clinical description of three-dimensional motions: application to the knee. *Journal of Biomechanical Engineering* **105** (1983) 136–144
4. Besl, P.J., McKay, N.D.: A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14** (1992) 239–256
5. Bentley, J.L.: Multidimensional binary search trees used for associative searching. *Communication of the ACM* **18** (1975) 509–517
6. Zdero, R., Fenton, P.V., Rudan, J., Bryant, J.T.: Fuji film and ultrasound measurement of total knee arthroplasty contact areas. *The Journal of Arthroplasty* **16** (2001) 367–375

A Boundary Element-Based Approach to Analysis of LV Deformation

Ping Yan^{1,3}, Ning Lin^{1,3}, Albert J. Sinusas^{2,4}, and James S. Duncan^{1,2,3}

Departments of ¹ Electrical Engineering,
² Diagnostic Radiology,
³ Biomedical Engineering and ⁴ Medicine,
Yale University PO Box 208042, New Haven, CT 06520-8042, USA
ping.yan@yale.edu

Abstract. Quantification of left ventricular (LV) deformation from 3D image sequences (4D data) is important for the assessment of myocardial viability, which can have important clinical implications. To date, feature information from either Magnetic Resonance, computed tomographic or echocardiographic image data has been assembled with the help of different interpolative models to estimate LV deformation. These models typically are designed to be computationally efficient (e.g. regularizing strategies using B-splines) or more physically realistic (e.g. finite element approximations to biomechanical models), but rarely incorporate both notions. In this paper, we combine an approach to the extraction and matching of image-derived point features based on local shape properties with a boundary element model. This overall scheme is intended to be both computationally efficient and physically realistic. In order to illustrate this, we compute strains using our method on canine 4D MR image sequences and compare the results to those found from a B-spline-based method (termed extended free-form deformation (EFFD)) and a method based on finite elements (FEM). All results are compared to displacements found using implanted markers, taken to be a gold standard.

1 Introduction

Quantitative analysis of left ventricular (LV) deformation is known to be a sensitive index of myocardial ischemia and injury. However, while there have been many methods proposed for the estimation of LV deformation [1, 2, 6, 7], most employ some form of modeling to interpolate dense displacement fields from sparse image derived features. These features include shape-based measures [1], MR tags [3], MR phase velocity [4] and echocardiographic features [5]. These models suffer from an inherent tradeoff between the computation time and the complexity of the model. A biomechanical model constructed using FEM, which incorporates the microstructure of the LV, is considered as the model closest to the physical reality of the LV but solving for the parameters embedded in these models is usually time-consuming [2,6]. Recently, simple deformation models such as B-spline-based EFFD [8] have been developed to estimate the LV deformation [7]. This method is computationally efficient but typically doesn't reflect the true physical properties of LV. Hence these

models are typically less accurate. This line of reasoning leads us to the following questions: Is it reasonable to trade off the approximation of physical reality with the computation time? Is there a method that is computationally more efficient than FEM yet still is used to directly incorporate physical parameters?

Over the several past decades, the Boundary Element Method (BEM) has emerged as a versatile and powerful alternative to FEM for the solution of engineering problems [9]. The most important advantage of BEM over the FEM is that it only requires discretization of the surface rather than the volume. Hence it can speed up the solution and it is easier to generate a mesh. The partitioning of only the boundary also makes BEM very suited to applications in the LV motion estimation problem where most feature points are on the boundary. Thus it is our view in this paper that BEM comes close to the physical reality of FEM modeling but is closer to EFFD in computation time.

Our approach, as described in this paper, combines an approach to shape-based feature tracking termed generalized robust point matching (G-RPM) [10] with a biomechanical model constructed using BEM to derive the dense strains from LV image sequences. The attractiveness of G-RPM is its ability to estimate non-rigid correspondences with only one-time *rough* segmentation in the first frame (see [10]).

To our knowledge no prior work has been reported that compares a B-spline-based deformation model with the biomechanical model in this domain. In this paper, we present the comparisons among the B-spline-based EFFD, and biomechanical model-based BEM and FEM.

The organization of the paper is as follows. In Section 2, we first briefly review the G-RPM algorithm and BEM. The computational cost of FEM and BEM are then compared in this section. In Section 3, Canine MRI cardiac datasets are analyzed using our algorithm. We estimate each algorithm’s accuracy by comparing the results to the displacements found using implanted markers. Meanwhile, the comparisons of strain results among the FEM, BEM and EFFD are presented in this section. Finally, Section 4 discusses the results and proposed future research.

2 Methods

2.1 The G-RPM Algorithm

G-RPM is an extension of robust point matching (RPM) [7]. We have previously reported results using this approach with B-spline in [7]. In this work, we embedded shape-based information into RPM and achieved more accurate results than when using point distances alone. Assume we have a data point-set $X = \{x_i, i = 1, 2, \dots, N\}$ and a template point-set $Y = \{y_a, a = 1, 2, \dots, K\}$. We denote the outlier cluster as x_{N+1} and y_{K+1} , and their unique temperature as T_0 . To match set Y onto X , it proposed to minimize the following objective energy function:

$$\begin{aligned}
 E(M, \alpha) = & \sum_{i=1}^{N+1} \sum_{a=1}^{K+1} m_{ai} \left[\|x_i - f(y_a, \alpha)\|^2 + \lambda_A g(\|A_X(x_i) - A_Y(y_a)\|^2) \right] + \lambda \|Lf\|^2 + \\
 & \zeta \sum_{i=1}^N \sum_{a=1}^K m_{ai} + T \sum_{i=1}^N \sum_{a=1}^K m_{ai} \log m_{ai} + T_0 \sum_{a=1}^K m_{a,N+1} \log m_{a,N+1} + T_0 \sum_{i=1}^N m_{K+1,i} \log m_{K+1,i}
 \end{aligned}
 \tag{1}$$

Where f is a non-rigid transformation function with parameters α . T is the annealing temperature, gradually decreasing to zero as the matching iteration begins.

Correspondence matrix M or $m_{ai} \in [0,1]$ is subject to the constraints: $\sum_{i=1}^{N+1} m_{ai} = 1$, for $a \in \{1,2,\dots,K\}$ and $\sum_{a=1}^{K+1} m_{ai} = 1$, for $i \in \{1,2,\dots,N\}$. $g(\cdot)$ is a strictly increasing function and λ_A balances the significance between the distance and new information. $A_X(x_i)$ and $A_Y(y_a)$ are the curvatures of two point-sets, respectively.

Minimization of equation 1 is solved by an alternating update process where two steps are repeated until the algorithm converges. One step is to update the correspondence matrix m_{ai} . The closed form solution can be obtained by differentiating the objective function (1) w.r.t m_{ai} and setting the result to zero:

$$m_{ai} = e^{-\frac{[\|x_i - f(y_a, \alpha)\|^2 + \lambda_A g(\|A_X(x_i) - A_Y(y_a)\|^2)] - \zeta}{T}} \tag{2}$$

A second step to update transformation parameter α involves a least-squared approach given the m_{ai} :

$$\min_{\alpha} E(\alpha) = \min_{\alpha} \sum_{a=1}^K \frac{\|\hat{x}_a - f(y_a, \alpha)\|^2}{\sigma_a^2} + \lambda \|Lf\|^2 \tag{3}$$

Where $\hat{x}_a = \sum_{i=1}^N m_{ai} x_i / \sum_{i=1}^N m_{ai}$ and $1/\sigma_a^2 = \sum_{i=1}^N m_{ai}$. In this algorithm, we used curvature derived from area-based operators [10]. For example, the Gaussian curvature \mathcal{K} from 3D volume data is given by:

$$\mathcal{K} = \frac{L_x^2 L_y L_{zz} + L_y^2 L_x L_{zz} + L_z^2 L_{xx} L_{yy}}{(L_x^2 + L_y^2 + L_z^2)^2} \tag{4}$$

Where $L(x, y, z)$ denotes image function. L_x, L_y and L_z are the first derivatives of L in x, y and z , and L_{xx}, L_{yy} and L_{zz} are the second partials in x, y and z .

2.2 The Boundary Element Method

The boundary element method is a technique for solving partial differential equations by reformulating the original PDE into an integral equation over the boundary of a solid object [9]. Because the integral equations are over the boundary of the object, only the boundary of the object needs to be partitioned. The elastic problem can be expressed by Navier’s equation:

$$\mu \nabla^2 \vec{u} + \frac{\mu}{1-2\nu} \nabla(\nabla \cdot \vec{u}) = 0 \tag{5}$$

Where μ and ν are the shear modulus and Poisson’s ratio of the elastic material. \vec{u} is the displacement vector. The formulation of the boundary integral equations for elastic problems requires the knowledge of the solution of the elastic problems with the same material properties as the body under consideration but corresponding to an infinite domain loaded with a concentrated unit point load. This is called the fundamental solution. The integral equation that relates interior displacements to boundary displacements and tractions (surface force) is known as Somigliana’s identity [9]:

$$u_i^i + \int_{\Gamma} p_{lk}^* u_k d\Gamma = \int_{\Gamma} u_{lk}^* p_k d\Gamma \tag{6}$$

Where u_{lk}^* and p_{lk}^* represent the fundamental solutions of displacement and traction at any point in the k direction when a unit load is applied at ‘i’ in the l direction. \vec{p} is traction vector. Γ is the boundary. When ‘i’ is taken to the boundary, equation 6 transforms into:

$$c_{lk}^i u_l^i + \int_{\Gamma} p_{lk}^* u_k d\Gamma = \int_{\Gamma} u_{lk}^* p_k d\Gamma \tag{7}$$

When ‘i’ is at a point where the boundary is smooth, $c_{lk}^i = \frac{1}{2} \delta_{lk}$. $\delta_{lk} = \begin{cases} 1, l = k \\ 0, l \neq k \end{cases}$

In order to solve the integral equation numerically, the boundary will be discretized into a series of elements. We can start by defining the \vec{u} and \vec{p} which apply over each element ‘j’, i.e. $\vec{u} = \Phi \vec{u}^j$ and $\vec{p} = \Phi \vec{p}^j$. Φ is the interpolation function matrix of shape functions. To study curved elements first we need to define the way in which we can pass from the global Cartesian system $\{x, y, z\}$ to the local system $\{\xi_1, \xi_2, \eta\}$ defined over the element, where ξ_1, ξ_2 are oblique coordinates and η is in the direction of the normal. The Jacobian of transformation is $d\Gamma = |J| d\xi_1 d\xi_2$. Equation 7 can now be written as:

$$c_{lk}^i \vec{u}^i + \sum_{j=1}^N \left\{ \int_{\Gamma_j} p_{lk}^* \Phi |J| d\xi_1 d\xi_2 \right\} \vec{u}^j = \sum_{j=1}^N \left\{ \int_{\Gamma_j} u_{lk}^* \Phi |J| d\xi_1 d\xi_2 \right\} \vec{p}^j \tag{8}$$

It can also be expressed in matrix form as follows:

$$\sum_{j=1}^N H^{ij} \vec{u}^j = \sum_{j=1}^N G^{ij} \vec{p}^j \tag{9}$$

Where $H^{ij} = \hat{H}^{ij}$, if $i \neq j$ and $H^{ij} = \hat{H}^{ij} + c^i$, if $i=j$.

Somigliana’s identity gives the displacement at any internal point in terms of the boundary displacements \vec{u}^j and tractions \vec{p}^j of each element:

$$\vec{u}^i = \sum_{j=1}^N G^{ij} \vec{p}^j - \sum_{j=1}^N H^{ij} \vec{u}^j \tag{10}$$

2.3 Comparison of the Operation Counts

In our algorithm, the H and G matrix only need to be calculated once in analyzing image sequences. Then displacements of any internal point at any frame can be computed by the equation 9 and 10 which don't need much computation time. For a L-sequence problem with N boundary nodes and M internal nodes, the operation counts of BEM and FEM are $O(N^3)+O(M^3)+O(L(NM))$ and $O(L(N+M)^3)$, respectively. Therefore, the BEM is computationally more efficient than the FEM, especially when we analyze image sequences.

2.4 Implementation

Feature Point Extraction: Not all points in images are suitable candidates for tracking. We choose our feature points as follows: 1) to stabilize curvature values, discard the points whose edge strength $\|\nabla L\|$ is less than a threshold; 2) thin the wide ridges resulting from the previous step by nonmaximum suppression [12] and 3) choose the resulting local maxima as the final candidate points to track.

Fundamental Solutions for Elastic Problem: In isotropic elastic materials, the fundamental solutions are known in exact closed form as shown in equation 11 and 12 [9]. In anisotropic analysis, the closed form of a fundamental solution is hard to get and calculation becomes very time-consuming. However, by using the Wilson & Cruse interpolation scheme, computation time is almost the same. (See [13]). For the 3D isotropic elastic problem, the fundamental solutions of displacement u_{lk}^* and traction p_{lk}^* are:

$$u_{lk}^* = \frac{1}{16\pi\mu(1-\nu)|r|} [(3-4\nu)\delta_{lk} + r_{,l}r_{,k}] \tag{11}$$

$$p_{lk}^* = \frac{1}{8\pi(1-\nu)|r|^2} \left[\frac{\partial r}{\partial n} [(1-2\nu)\delta_{lk} + 3r_{,l}r_{,k}] + (1-2\nu)(n_l r_{,k} - n_k r_{,l}) \right] \tag{12}$$

Where r is the distance between the loading point to any point under consideration. Notice that $r_{,l}$ is the derivative of the r in the l direction. n_l and n_k are the direction cosines of the normal with respect to l and k direction. In this paper, we used $\mu = 17857.1$ and $\nu = 0.4$ for the LV myocardium.

Strain tensor definition: Given a strain tensor E (a 3×3 matrix) which is computed in the Cartesian coordinate system of $\{x, y, z\}$. We can transform it to a local coordinate system $\{\xi_1, \xi_2, \eta\}$. First construct the 3×3 rotation matrix R which results in $\{x, y, z\} \rightarrow \{\xi_1, \xi_2, \eta\}$. Using this matrix R we can write the E in the local coordinate system as: $E_{local} = RER'$. In this paper, the local coordinate system is defined by the circumferential, radial, and longitudinal axis at the epicardium.

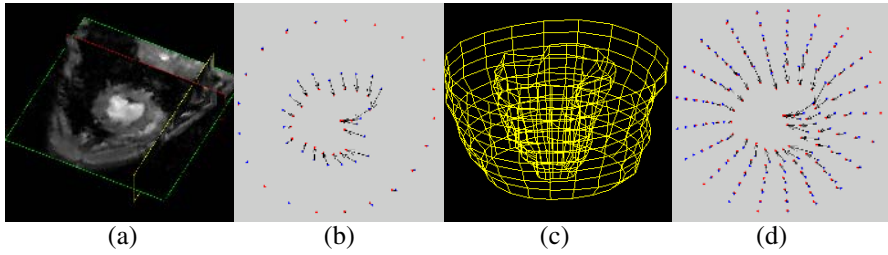


Fig. 1. (a) Original 3D MRI Data in the Orthogonal-slice view, (b) Boundary displacement results for one of its 2D slices obtained using G-RPM, (c) 3D BEM mesh using bilinear element, (d) Dense displacement results for one 2D slices through the 3D results.

3 Experimental Results

4D MR canine cardiac datasets were analyzed using our new algorithm. An example of the results from one of our datasets was shown in figure 1. For each dataset, strains

were calculated between end-diastole (ED) and end-systole (ES). The resulting 3D-derived radial strains (E_{rr}) and Circumferential strains (E_{cc}) of one dataset at frame 1, 3, 4 and 6 were illustrated in figure 2. Frame 1 was at ED and Frame 6 was at ES. Note the normal behavior in the LV, showing positive radial strain (thickening) and negative circumferential strain (shortening) as we move from ED to ES.

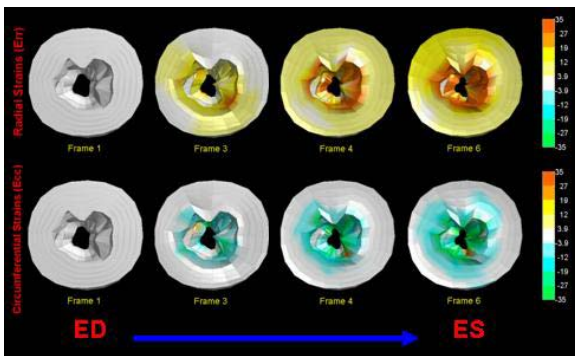


Fig. 2. E_{rr} (Top) and E_{cc} (Bottom) between ED (frame 1) to ES (frame 6) of dataset 1

3.1 Comparisons with Implanted Markers

To further quantitatively validate resulted motion trajectories, we used four canine MRI datasets with implanted markers for point-by-point displacement comparison (see [6] for more details on the marker implantation and localization.). The mean displacement errors of FEM, BEM and EFFD methods were illustrated in the figure 3. It can be seen that FEM-based models show the smallest error while the EFFD approach shows the largest.

3.2 Comparisons Between FEM and BEM, FEM and EFFD

In order to see the difference between the B-Spline based deformation model and the biomechanical model, we compared strain results obtained using FEM, BEM and

EFFD. The strains were calculated between ED to ES using BEM, FEM and EFFD, respectively. The comparisons of strain results between FEM and BEM, FEM and EFFD for 4 canine image sequences are illustrated in the figure 3. The mean differences of E_{cc} , E_{rr} and E_{ll} between FEM and BEM were less than 1.3%, 4.05% and 4.24%, while the mean differences of E_{cc} , E_{rr} and E_{ll} between FEM and EFFD were about 3 times higher. These results show that the BEM came closer to the FEM model which we presume to be more physically realistic than the EFFD. EFFD is a powerful tool for modeling deformable objects with arbitrary geometries but it may not be able to fully approximate the physical properties of LV.

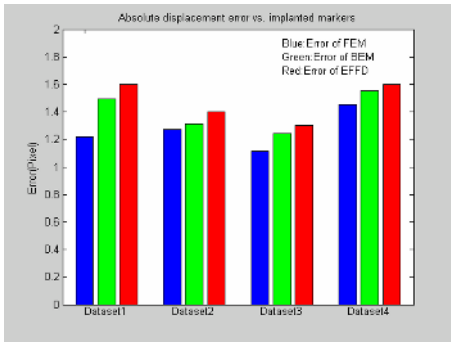


Fig. 3. Absolute displacement error vs. implanted markers. The motion is estimated between ED to ES and compared by the FEM, BEM and EFFD. **Blue:** Error of FEM; **Green:** Error of BEM; **Red:** Error of EFFD

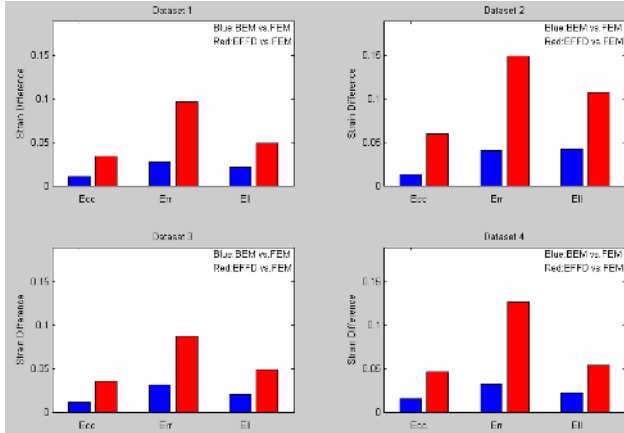


Fig. 4. Comparisons of strains between FEM and BEM, FEM and EFFD: **Blue:** Mean difference between the strain results obtained by FEM and BEM. **Red:** Mean difference between the strain results obtained by FEM and EFFD.

4 Conclusions and Future Work

We conclude that the BEM-based algorithm presented in this paper is both computationally efficient and physically realistic for the following reasons: 1) Only *rough* boundary segmentation is required in the first frame; 2) Only the discretization of surface is required; 3) The fundamental solution matrix only needs to be calculated once in analyzing image sequences; and 4) The deformation is estimated by a biomechanical model of the myocardium. In this paper, by comparing displacement

results using implanted markers, we identified that the FEM has the least errors, while the BEM is the second and EFFD has the largest error. We also found that the difference of strain results using EFFD has larger differences from a FEM approach than the newer BEM-based approach. Further extensions to this work could be using a more complicated biomechanical model, which incorporates the fiber directions of LV.

References

1. P. Shi, A. J. Sinusas, R. T. Constable, E. Ritman, and J. S. Duncan, "Point-tracked quantitative analysis of left ventricular motion from 3D image sequences". *IEEE Transactions on Medical Imaging*, 19(1): 26–50, 2000.
2. J. M. Guccione and A. D. McCulloch. "Finite element modeling of ventricular mechanics", P. J. Hunter, A. McCulloch, and P. Nielsen, editors, *Theory of Heart*, pages 122–144. Springer-Verlag, Berlin, 1991.
3. E. Haber, D. N. Metaxas, and L. Axel. "Motion analysis of the right ventricle from MRI images". In *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, Cambridge, MA, October 1998.
4. F. G. Meyer, R. T. Constable, A. J. Sinusas, and J. S. Duncan. "Tracking myocardial deformation using phase contrast MR velocity fields: A stochastic approach.", *IEEE Transactions on Medical Imaging*, Vol. 15, No. 4, August 1996.
5. K. Kaluzynski, X.Chen, S. Emelianov, A. Skorovoda and M.O'Donnell, "Strain Rate Imaging Using Two-Dimensional Speckle tracking," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, volume 48, 2001: 1111-1123.
6. X.Papademetris, A.J.Sinusas, D. P.Dione, R.T.Constable and J.S.Duncan, "Estimation of 3D Left Ventricular Deformation from Medical Images Using Biomechanical Models," *IEEE Transactions on Medical Imaging*, volume 21, no. 7, July, 2002: 786-799
7. N. Lin and J.S.Duncan, "Generalized Robust Point Matching Using an Extended Free-Form Deformation Model: Application to Cardiac Images," *IEEE International Symposium on Biomedical Imaging (ISBI)*, Arlington, VA, April, 2004
8. S. Coquillart, "Extended free-form deformation: A sculpturing tool for 3D geometric modeling," *Computer Graphics*, vol. 24, no. 4, pp. 187–196, 1990.
9. C.A.Brebbia and J.Dominguez, "Boundary Elements An Introductory Course", 2nd ed, *Computational Mechanics Publications*, McGraw-Hill Book Company
10. N. Lin, X. Papademetris, A. Sinusas, and J.S. Duncan, "Analysis of left ventricular motion using a general robust point matching algorithm," in *Medical Image Computing and Computer-Assisted Intervention*, _Montreal, Canada, November 15-18, 2003, pp. 556–563, (LNCS 2878) Springer Verlag.
11. B.M. Ter, Haar Romeny, L.M.J.Florack,A.H.Salden, and M.A.Viergever, "Higher order differential structure of images", *Image and Vision Computing*, 12(16):317-325,1994
12. E.Trucco and A.Verri., "Introductory Techniques for 3D Computer Vision.", *Prentice Hall*, 1998
13. Wilson, R.B. and Cruse, T.A. (1978), "Efficient implementation of anisotropic three dimensional boundary-integral equation stress analysis", *International Journal for Numerical Methods in Engineering*, 12, 1383-1397

Reconstruction of Cerebrospinal Fluid Flow in the Third Ventricle Based on MRI Data

Vartan Kurtcuoglu¹, Michaela Soellinger², Paul Summers³, Kevin Boomsma¹, Dimos Poulikakos¹, Peter Boesiger², and Yiannis Ventikos⁴

¹ Laboratory of Thermodynamics in Emerging Technologies,
Department of Mechanical and Process Engineering, ETH Zurich, Switzerland
Vartan.Kurtcuoglu@ethz.ch

² Institute for Biomedical Engineering,

University of Zurich and ETH Zurich, Switzerland

³ Department of Neuroradiology, Nuffield Department of Surgery,
University of Oxford, UK

⁴ Fluidics and Biocomplexity Group and Institute of Biomedical Engineering,
Department of Engineering Science, University of Oxford, UK

Abstract. A finite-volume model of the cerebrospinal fluid (CSF) system encompassing the third ventricle and the aqueduct of Sylvius was used to reconstruct CSF velocity and pressure fields based on MRI data. The flow domain geometry was obtained through segmentation of MRI brain anatomy scans. The movement of the domain walls was interpolated from brain motion MRI scans. A constant pressure boundary condition (BC) was specified at the foramina of Monro. A transient velocity BC reconstructed from velocimetric MRI scans was employed at the inferior end of the aqueduct of Sylvius. It could be shown that a combination of MRI scans and computational fluid dynamics (CFD) simulation can be used to reconstruct the flow field in the third ventricle. Pre-interventional knowledge of patient-specific CSF flow has the potential to improve neurosurgical interventions such as shunt placement in case of hydrocephalus.

1 Introduction

The cerebrospinal fluid (CSF) is contained within and surrounds the brain and spinal cord [1]. It suspends the brain through its buoyancy force and protects it from impact on the cranial vault walls in cases of sudden head motion. The CSF further serves as an intermediary between blood and nervous tissue, providing the latter with nutrients and removing waste products.

CSF is produced in the choroid plexuses of the brain and is drained mainly through the superior sagittal sinus. A periodic pulsatile motion governed by the cardiac cycle is superimposed upon the steady flow caused by the CSF production. Within the skull, the cerebrospinal fluid is contained in the ventricles and the subarachnoid space. The ventricles are four cavities interconnected by pathways (see Fig. 1).

In contrast to the cardiovascular system, computational representations of the cerebrospinal fluid space are still in their infancy, despite the first attempts at such models being made as early as 1969 [2]. Since 1999, a series of models based on CFD

and numerical structural mechanics have been published, all of which rely on either or both extensively simplified domain geometry and boundary conditions [3-6]. The finite-volume model presented here takes into account the wall motion of the third ventricle and aqueduct of Sylvius in feet-head direction based on MRI measurements. The domain geometry is obtained using anatomic MRI scans. The boundary condition at the inferior end of the aqueduct of Sylvius is implemented by reconstruction of a Womersley velocity profile based on MRI velocimetry. All of these features are novel.

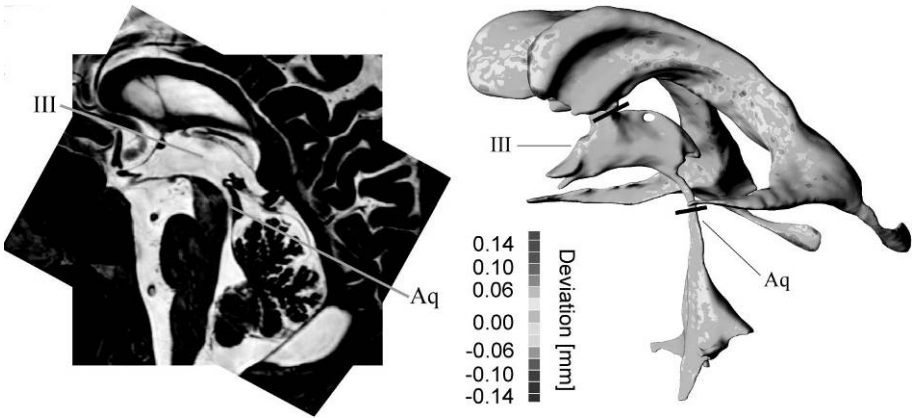


Fig. 1. Left: Composite of two registered slices of MRI anatomy scans of the ventricular system. Right: Rendering of the segmented ventricular system after conversion to NURBS surfaces. The color coding indicates the deviation between the original voxel-based segmentation and the NURBS representation. The bars indicate the boundaries of the system treated in this paper, encompassing the third ventricle (III) and aqueduct of Sylvius (Aq).

2 Methods

2.1 Acquisition and Processing of Anatomy Data

Structural MR imaging to define the ventricle and aqueduct boundaries consisted of a strongly T2-weighted, 3D, turbo spin-echo sequence [7] performed on a 27 year old healthy male volunteer, on a high-field clinical imaging system (Achieva 3T, Philips Medical Systems, Best, The Netherlands). The imaging parameters were: repetition time (TR) = 2325 ms, echo time (TE) = 160 ms, number of signal averages = 1 and an echo train length of 36 echoes. The reconstructed spatial resolution was 0.22 x 0.22 x 0.4 mm.

To accommodate for the slight anisotropy in the spatial resolution of imaging data, the sequence was performed three times, each acquisition in a different cardinal orientation. These three data sets were aligned by minimizing the Euclidean distance between similar gray values. The registration was further enhanced using a quasi-Newton optimizer [8] and the datasets were merged using Lanczos interpolation [9]. The third ventricle and the aqueduct of Sylvius were manually segmented and the

obtained voxel-based 3D representation was converted to NURBS surfaces (non-uniform rational B-spline). This approach ensured adequate smoothing of the domain surface which is necessary for the generation of a high-quality computational grid.

2.2 Acquisition and Processing of Brain Motion Data

For measurement of the brain tissue displacement, the aforementioned volunteer was scanned with a complementary spatial magnetization modulation (CSPAMM) [10] prepared sequence in a mid-sagittal slice using a tag distance of 2 mm. The scans were performed on a whole-body MRI system (Intera 1.5T, Philips Medical Systems, Best, The Netherlands). Separate tag line preparation was performed in two orthogonal in-plane directions (cranio-caudal and anterior-posterior). The number of cardiac phases was chosen such that at least 120% of the R-R interval could be covered to allow indirect validation of the tracking error. Reproducibility of the measurements was assessed by two subsequent measurements on the same volunteer. After tag preparation, the data were acquired with a turbo field echo [7] sequence with a turbo factor of 3 and TR/TE: 8.13/3.8 ms. The spatial resolution was approximately 1 mm and the temporal resolution 60 ms. Cranio-caudal displacement was finally quantified using harmonic phase (HARP) [11] post-processing with a shutter width of 24 pixels which yielded a HARP tracking resolution of about 5.3 mm [12]. Subject to the limits of resolution of the data acquired, displacements in the anterior-posterior direction were too small to be resolved. Using least-squares error minimization, a fourth order Fourier series was fit to the time-dependent displacement data of each measured brain location. The displacement of the third ventricle and aqueduct boundaries were approximated from the measured locations using thin-plate splines.

2.3 Acquisition and Processing of CSF Flow Velocity Data

A standard phase contrast velocity mapping sequence [7] was also used on the same volunteer. The measured slices were planned perpendicular to the aqueduct of Sylvius and carried out in its caudal aspect where the diameter is larger than in the cranial section. The data were acquired with a turbo field echo (turbo factor = 3) read-out and the spatial scan resolution was set to 288 x 216 (reconstruction to 512 x 512, covering a field of view of 250 x 211 x 5 mm). As maximum velocity values in the range of 5 cm/s [13] were expected, the encoding velocity was set to 7 cm/s in order to avoid phase wraps. A total of 40 cardiac phases were generated with a temporal resolution of 26 ms. Three signal averages were used to improve the signal-to-noise ratio. As both the spatial and temporal resolution of the velocimetric MR data are lower than what is needed for accurate CFD calculations, the algebraic solution of the flow field at the inferior end of the aqueduct was reconstructed from the measured data using the theory of pulsatile flow in a pipe [14].

For this purpose, the aqueduct was regarded as a circular cylinder and the flow at its outlet section was assumed to be fully developed. The former is a plausible simplification, as the shape of the aqueduct is indeed approximately cylindrical, albeit with variable radius. The plausibility of the latter assumption can be verified as follows: The maximum Reynolds number (Re) and minimum Womersley parameter (α) for the flow in the aqueduct can be calculated from the velocimetric MR data,

$$\text{Re}_{\max} = \frac{u_{\max} D}{\nu}, \quad \alpha_{\min} = \frac{D}{2} \sqrt{\frac{2\pi f_{\min}}{\nu}}, \quad (1)$$

where u_{\max} is the highest axial velocity – both locally and temporally – at the inferior end of the aqueduct of Sylvius, D is the diameter of the aforementioned cylinder, ν is kinematic viscosity and f_{\min} is the cardiac cycle base frequency. In the case at hand, Re_{\max} is 70.9 and α_{\min} is 2.9, which results in a hydrodynamic entrance length acceptably shorter than the length of the aqueduct [15, 16]. Consequently, the flow field at the inferior end of the aqueduct can be regarded as being fully developed. Assuming a harmonic driving pressure gradient, the velocity field can be calculated as

$$u(r, t) = \frac{i k R^2}{\mu \alpha^2} \left(1 - J_0 \left(A \frac{r}{R} \right) \left(J_0(A) \right)^{-1} \right) e^{i \omega t}, \quad A = \left(\frac{i-1}{\sqrt{2}} \right) \alpha, \quad (2)$$

[17], where u is the axial velocity, r is the radial location in the cylinder cross-section, t is time, i is $\sqrt{-1}$, k is the amplitude of the driving pressure gradient, R is the pipe radius, μ is dynamic viscosity, α is the Womersley parameter, J_0 is the Bessel function of first kind and order zero and ω is the angular frequency of the pressure gradient. By integrating Eq. (2), the mass flow rate is obtained,

$$\dot{m}(t) = \frac{k B}{\alpha^2} e^{i \omega t}, \quad B = \frac{i \pi R^4}{\nu} \left(1 - \frac{2 J_1(A)}{A J_0(A)} \right), \quad (3)$$

where J_1 is the Bessel function of first kind and order. As the pressure gradient driving the flow in the aqueduct of Sylvius is periodic, it can be described using a Fourier series, so that from Eqs. (1, 3) and Poisseuille's law [18] the mass flow rate can be expressed as

$$\dot{m}(t) = \frac{-k_0 \pi R^4}{8 \nu} + \frac{1}{\alpha^2} \sum_{n=1}^{\infty} \left(k_{2n-1} \text{Re} \left(B e^{i \omega t} \right) + k_{2n} \text{Im} \left(B e^{i \omega t} \right) \right), \quad (4)$$

where Re and Im designate real and imaginary parts, respectively. This mass flow rate has to be equal to the one obtained by integrating the velocity field measured with velocimetric MR. Using weighted least-squares error minimization, the coefficients k were calculated and, hence, the velocity profile at the inferior end of the aqueduct was determined.

2.4 Reconstruction of the Flow Field

For the flow field reconstruction, the CSF was regarded as an incompressible Newtonian fluid with the same material properties as water at 37° Celsius [6, 19]. Three sets of non-uniform, unstructured computational grids, respectively consisting of about 257,000, 558,000 and 811,000 tetrahedral elements were used to carry out the finite-volume calculations with the CFD code Fluent (Fluent Inc., Lebanon, NH) using an implicit Algebraic Multigrid Scheme with PISO pressure correction [20].

Grid independence tests showed that the intermediate grid was fine enough to resolve all of the flow features accurately. Time step independence tests carried out with time step sizes of $T/100$, $T/1,000$ and $T/10,000$, where T is the length of one

period, showed that a time step size of $T/1,000$ was necessary to adequately model the flow. A period independence study demonstrated that three periods were necessary for the transients originating from the zero velocity and zero pressure initial condition to abate.

3 Results

Figure 3 shows contour lines of velocities perpendicular to the corresponding section planes, whose positions are shown in Figure 2, right side. The states at 0%, 25%, 50% and at 75% of the duration of the cycle are depicted. Initially, the cerebrospinal fluid flows in cranial direction (see Fig. 2, left). A jet directed towards the foramina of Monro exits the aqueduct of Sylvius with a peak velocity of slightly less than 12 cm/s. Before flow reversal occurs at around 19% of the cycle, an area of recirculation develops above the jet, with its center at about the height of the recessus pinealis. After flow reversal, the recirculation splits the path of CSF flowing in caudal direction from the foramina of Monro into two: one part of the CSF is deflected in cranial direction, flows past the recirculation along the top wall of the ventricle and then in caudal direction towards the aqueduct, without entering the recessus suprapinealis. The second CSF stream is only slightly disturbed and flows caudal of the area of recirculation almost straight into the aqueduct. Before the second flow reversal occurs at about 75%, this first recirculation has already dissipated and a second one is formed approximately half way between and caudal to the foramina of Monro and the adhesio interthalamica. By the time the cycle is 95% through, this second, bigger, recirculation has also dissipated.

The flow velocities in the treated domain are in general below 1 cm/s. Exceptions are the flow in the aqueduct with peak values of approximately 15 cm/s at the narrowest cross-sections, the jet emerging from the aqueduct and the foramina of Monro, where peak velocity values of slightly above 2 cm/s are reached. Maximum absolute velocities below 0.1 cm/s are observed in all of the recessi.

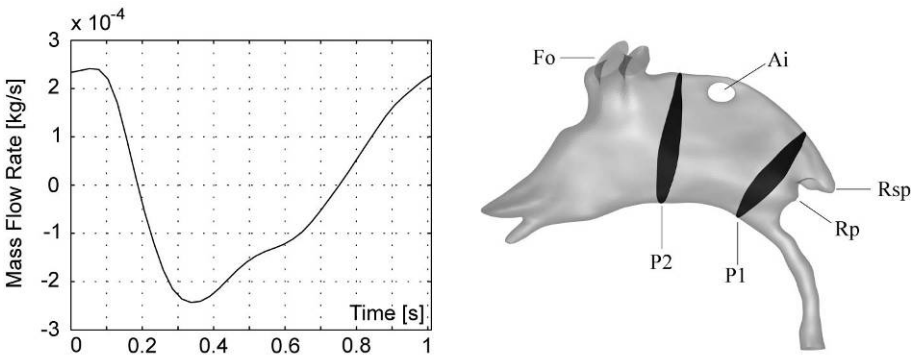


Fig. 2. Left: Mass flow rate through the inferior end of the aqueduct of Sylvius. Positive values indicate cranial direction. Right: Flow domain on which the finite-volume calculations were performed. P1 and P2 are section planes referred to in Fig. 3. Fo: foramina of Monro, Rp: recessus pinealis, Rsp: recessus suprapinealis, Ai: adhesio interthalamica.

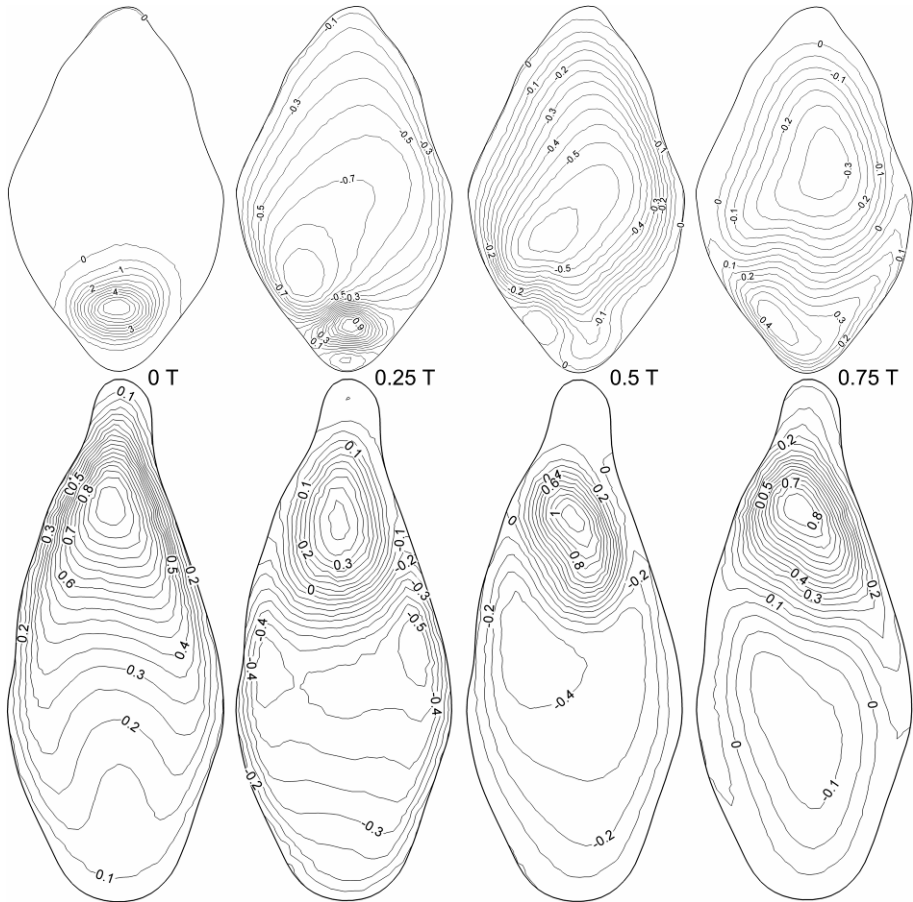


Fig. 3. Top row: Iso-contours of normal velocities in cm/s in the section plane P1 (see Fig. 2, right) at 0%, 25%, 50% and 75% of the period T (Fig. 2, left). Positive values indicate cranial direction. Bottom row: Normal velocities in the plane P2.

4 Discussion

The basis of the simulation of cerebrospinal fluid flow in the third ventricle and the aqueduct of Sylvius as presented above are MRI scans of an individual. Therefore, the accuracy of the simulation is first and foremost governed by the accuracy of the MRI data acquisition. Due to the relatively high flow velocities occurring in the aqueduct (in the order of several cm/s) and the aqueduct's nearly cylindrical shape, the corresponding boundary condition can be reconstructed fairly accurately with a relative error well below 10%. Anatomic MRI scans cannot resolve the detailed surface structures of the choroid plexus, which are in the order of micrometers. We believe that neglecting these structures in the third ventricle will have limited influence on the flow field, as they are located predominantly at the ventricle's top wall where the flow velocities are low.

The motion of the brain tissue is responsible for the pulsatile flow of the cerebrospinal fluid. Hence, it is important to take it into account when reconstructing the CSF flow. MRI tagging can be used to capture the brain motion in feet-head direction (maximum amplitude 0.25 mm). However, it will not capture the motion in the remaining directions which generally portray amplitudes less than 0.15 mm. The velocity field specified at the inferior end of the aqueduct of Sylvius makes up for the lack of that data to a large extent. For clinical use, however, more advanced techniques will have to be used in order to take into account brain motion in every direction [21].

5 Conclusions

The resolution of current MRI scanners is not high enough to capture the fine details of the cerebrospinal fluid flow in the brain ventricular system. The use of idealized CFD simulations, on the other hand, will not capture patient-specific flow features that may be crucial to the outcome of the intervention. By relying on MRI measurements to capture the patient's anatomic structures and to provide boundary conditions for the computer simulations, these shortcomings can be eliminated. Precise knowledge of cerebrospinal fluid flow would allow for better planning of surgical interventions in those cases where the CSF space is involved. The prime example is shunt placement for the treatment of hydrocephalus and other conditions where CSF drainage or control is important.

We have shown that it is possible to combine anatomic, velocimetric and tagged MRI scans with CFD simulations to reconstruct the CSF flow field in the third brain ventricle and the aqueduct of Sylvius. The methods described herein can be applied without significant modifications to the entire ventricular system.

Acknowledgements. The financial support of the ETH Zurich Research Commission has made this research possible and is kindly acknowledged.

References

1. Davson, H., Segal, M.B.: *Physiology of the CSF and Blood-Brain Barriers*. CRC Press, Boca Raton (1996)
2. Agarwal, G.C., Berman, B.M., Stark, L.: A Lumped Parameter Model of the Cerebrospinal Fluid System. *IEEE Transactions on Biomedical Engineering* 16 (1969) 45–53
3. Jacobson, E.E., Fletcher, D.F., Morgan, M.K. et al.: Computer Modelling of the Cerebrospinal Fluid Flow Dynamics of Aqueduct Stenosis. *Medical & Biological Engineering & Computing* 37 (1999) 59–63
4. Loth, F., Yardimci, M.A., Alperin, N.: Hydrodynamic Modeling of Cerebrospinal Fluid Motion within the Spinal Cavity. *ASME Journal of Biomechanical Engineering* 123 (2001) 71–79
5. Fin, L., Grebe, R.: Three Dimensional Modeling of the Cerebrospinal Fluid Dynamics and Brain Interactions in the Aqueduct of Sylvius. *Computer Methods in Biomechanics and Biomedical Engineering* 6 (2003) 163–70

6. Kurtcuoglu, V., Poulidakos, D., Ventikos, Y.: Computational Modeling of the Mechanical Behavior of the Cerebrospinal Fluid System. *ASME Journal of Biomechanical Engineering* 127 (2005) 264–269
7. Haacke, E.M., Brown, R.W., Thompson, M.R. et al.: *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. John Wiley & Sons, New York (1999)
8. Martinez, J.: Practical Quasi-Newton Methods for Solving Nonlinear Systems. *Journal of Computational and Applied Mathematics* 124 (2000) 97–121
9. Lanczos, C.: *Applied Analysis*. Pitman, London (1957)
10. Fischer, S.E., McKinnon, G.C., Maier, S.E. et al.: Improved Myocardial Tagging Contrast. *Magnetic Resonance in Medicine* 30 (1993) 191–200
11. Osman, N.F., Kerwin, W.S., McVeigh, E.R. et al.: Cardiac Motion Tracking using CINE Harmonic Phase (HARP) Magnetic Resonance Imaging. *Magnetic Resonance in Medicine* 42 (1999) 1048–60
12. Soellinger, M., Ryf, S., Kozerke, S. et al.: Brain Motion Measurements using CSPAMM. In 21th Annual Scientific Meeting of the ESMRBM, Copenhagen (2004)
13. Stahlberg, F., Mogelvang, J., Thomsen, C. et al.: A method for MR Quantification of Flow Velocities in Blood and CSF using Interleaved Gradient-Echo Pulse Sequences. *Magnetic Resonance Imaging* 7 (1989) 655–67
14. Womersley, J.R.: Oscillatory Motion of a Viscous Liquid in a Thin-Walled Elastic Tube. 1. The Linear Approximation for Long Waves. *Philosophical Magazine* 46 (1955) 199–221
15. Friedmann, M., Gillis, J., Liron, N.: Laminar Flow in a Pipe at Low and Moderate Reynolds Numbers. *Applied Scientific Research* 19 (1968) 426–438
16. He, X.Y., Ku, D.N.: Unsteady Entrance Flow Development in a Straight Tube. *ASME Journal of Biomechanical Engineering* 116 (1994) 355–360
17. Zamir, M.: *The Physics of Pulsatile Flow*. Springer Verlag, New York (2000)
18. Wilcox, D.C.: *Basic Fluid Mechanics*. DCW Industries, La Cañada (1999)
19. Bloomfield, I.G., Johnston, I.H., Bilston, L.E.: Effects of Proteins, Blood Cells and Glucose on the Viscosity of Cerebrospinal Fluid. *Pediatric Neurosurgery* 28 (1998) 246–51
20. Issa, R.I.: Solution of the Implicitly Discretised Fluid Flow Equations by Operator Splitting. *Journal of Computational Physics* 62 (1985) 40–65
21. Soellinger, M., Rutz, A., Kozerke, S. et al.: Direct Quantification of Brain Motion. In 43rd Annual Meeting of the ASNR, Toronto (2005)

Schwarz Meets Schwann: Design and Fabrication of Biomorphic Tissue Engineering Scaffolds

Srinivasan Rajagopalan and Richard A. Robb

Biomedical Engineering Resource, Mayo Clinic College of Medicine, Rochester, USA
{srini, robb.richard}@mayo.edu

Abstract. Tissue engineering is a discipline at the leading edge of the field of computer assisted intervention. This multidisciplinary engineering science is based on the notion of design and fabrication of scaffolds- porous, three-dimensional "trellis-like" biomimetic structures that, on implantation, provide a viable environment to recuperate and regenerate damaged cells. Existing CAD-based approaches produce porous labyrinths with *contra-naturam* straight edges. The biomorphic geometry that mimics the *secundam-naturam* substrate would be one that is continuous through all space, partitioned into two not-necessarily-equal sub-spaces by a non-intersecting, two-sided surface. *Minimal surfaces* are ideal to describe such a space. We present results on the premier attempt in computer controlled fabrication and mechanical characterization of Triply Periodic Minimal Surfaces [TPMS]. This initiative is a significant step to link Schwann's 1838 cell theory with Schwarz's discovery of TPMS in 1865 to fabricate the previously elusive optimal biomorphic tissue analogs.

1 Introduction

Millions of surgical procedures requiring organ (tissue) substitutes to repair or replace malfunctioning organs (tissues) are performed worldwide every year. The ever-widening gap between the demand and supply of transplant organs (tissues) has resulted in natural and biomimetic solutions. Autografting, allografting, and tissue engineering are the classical techniques currently pursued to carry out organ transplantation [1]. Tissue engineering involves regenerating damaged or malfunctioning organs from the recipient's own cells. The cells are harvested and embedded on a natural or synthetic carrier material, *scaffold*, that is ultimately implanted at the desired site. The seeded cells cling, crawl and proliferate within the scaffold. On implantation and integration, blood vessels attach to the new tissue, the scaffold dissolves, and the newly grown tissue eventually blends in with its surroundings. The scaffolds suitable for tissue engineering, apart from being biocompatible and biodegradable, should have highly interconnected pores; have sufficient mechanical properties; and, provide a suitable substrate for cell attachment, migration and proliferation [2].

Tissue engineering scaffolds can be manufactured reproducibly using rapid prototyping approaches like Solid Freeform Fabrication (SFF)- computerized fabrication techniques for rapidly producing complex three-dimensional objects using data generated by CAD systems, medical imaging modalities and digitizers. This

technique offers unique ways to precisely control substrate architecture for biomimetic structures varying in design and material composition, thereby enhancing control over mechanical properties, biological effects and degradation kinetics of the scaffold. However, the majority of existing approaches under-utilize SFF by producing CAD-based scaffolds with straight edges and sharp turns or those derived from Boolean intersections of geometric primitives such as spheres and cylinders. Neither of these partitions provides a biomorphic environment suitable for cell attachment, migration and proliferation [3]. The aggregates of cells, *foam/extracellular matrix*, typically have cells separated by curved partitions. The biomorphic geometry that best mimics this structural configuration would be surfaces that are continuous through space and divided into two (pore and non-pore) not-necessarily-equal sub-spaces by a non-intersecting two-sided surface. Minimal surfaces are ideal to describe such a space [4]. This paper presents a first practical application of Triply Periodic Minimal Surfaces (TPMS) for the construction of tissue engineering scaffolds. The optimal stress distributions determined from finite element simulations, and mechanical testing, and initial cell viability studies conducted on these scaffolds provide compelling evidence to use them as orthopedic tissue analogs.

2 Triply Periodic Minimal Surfaces

A surface $f(x,y,z) = 0$ is a minimal surface if and only if the mean curvature (H) at every point of the surface is zero. ie.

$$H=0=\frac{(1+f_y^2)f_{xx}-2f_xf_yf_{xy}+(1+f_x^2)f_{yy}}{2(1+f_xf_x+f_yf_y)^{3/2}} \quad (1)$$

Any small patch cut from such a surface will have the least area of all the surface patches with the same boundary [5]. TPMS are minimal surfaces periodic in three independent directions, extending infinitely and, in the absence of self-intersections, partitioning the space into two labyrinths. Figure 1 shows the tessellated versions of the TPMS Primitive (P) surface and the Diamond (D) surface discovered by Schwarz [6] and the Gyroid (G) surface discovered by Schoen [7]. The natural manifestation of two, continuous, inter-penetrating, non-intersecting networks in these surfaces has been adapted in natural and man-made environments to concisely describe many seemingly unrelated structures [8]. Though this architecture is a natural fit for the preferred porous and solid subspace topology for tissue engineering scaffolds, to the best of our knowledge, it has not been previously investigated and exploited.

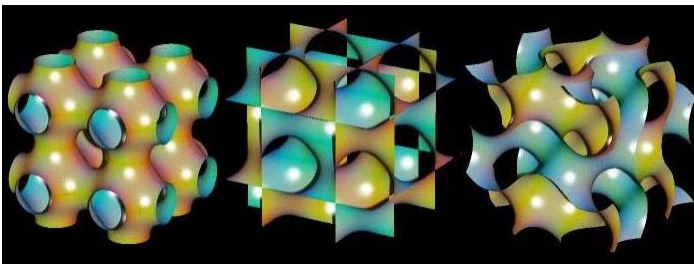


Fig. 1. 3-dimensional tessellations of Schwarz's Primitive (left), and Diamond (centre) and Schoen's Gyroid (right) Triply Periodic Minimal Surfaces

3 Geometric Modeling of TPMS

There are a number of approaches to generate discrete minimal surfaces. Coppin's approach [9] is based on the construction of a quasimolecular model that approximates the molecular structure of a soap film which is a prominent natural manifestation of minimal surfaces. Mackay's construction [10] uses the concept of nodal and Fermi surfaces within a Finite Element formulation. Brakke developed *Surface Evolver* [11]– a general purpose application that minimizes the energy of a polygonal mesh subject to constraints. Using this approach it is possible to evolve the straight edged, sharply turning cubic labyrinthine network conventionally used in scaffold fabrication into a smooth, curved, non-intersecting, bicontinuous partition. The energy (based on surface tension, gravity and other forms) minimization is performed by evolving the ordered polygonal mesh surface down the energy gradient subject to geometrical constraints on vertex positions and other constraints on integrated quantities such as body volumes. The fully evolved minimal surface can be converted into geometric models that can be realized physically with existing rapid prototyping devices.

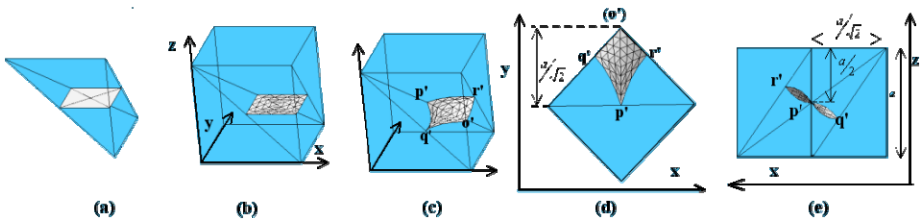


Fig. 2. Evolution of triangulated plane into a P-surface patch

To generate the fundamental patch of the Schwarz P-surface, quadri-rectangular tetrahedron is used as the fundamental region that confines the initial plane and its subsequent evolved surface. Figure 2 shows the evolution of the plane into a P-surface patch. The plane embedded orthogonally with respect to the faces of the fundamental region (panel a) is refined by subdividing each facet to create a finer triangulation (panel b). The triangulated patch is iteratively evolved to form the P-surface patch (panel c). Panels d and e show the projection of the patch onto the xy and xz planes, respectively. The cube shown in Panels b and c is the bounding cell with edge length a .

To form the unit P-cell, the evolved surface patch is reflected around the fundamental region using the line and plane symmetry periodic properties. If part of the boundary of a minimal surface is a straight line, the reflection of the surface across the line, when added to the original surface, forms another minimal surface. If a minimal surface meets a plane at right angles, the mirror image of the surface about the plane, when added to the original surface, also forms a minimal surface. Figure 3 shows the formation of the unit P-cell. Tessellation of the embedded plane (panel a) using the above mentioned properties results in a hollow unit cubic cell (panel b). Tessellation of this cell produces the cubic labyrinthine network (panel c) reminiscent of the existing CAD-based scaffolds. Evolution of (a) leads to the fundamental P-

patch (panel e) which when replicated using the same tessellation rules produces the unit P-cell (panel f). Tessellation of the unit cells produce the bicontinuous network (panels c,d,g,h). Note that the central hole in both the cubic and P-cell tessellations does not contribute to a pore since the model is not defined therein and hence is treated as belonging to the solid subspace.

Since the P-cell equally partitions the unit cube, the volume should be a half unit; for a continuous P-surface the surface area is 2.3451 [12]. Evolution based on finer triangulation of the initial plane provides better approximation of the discrete surface. However, the complexity of machining increases with the number of polygons. While the volume was preserved across all the triangulations, the optimal balance between polygon count (3072 triangles), surface smoothness and surface area (2.352) was achieved at $2x$ refinement.

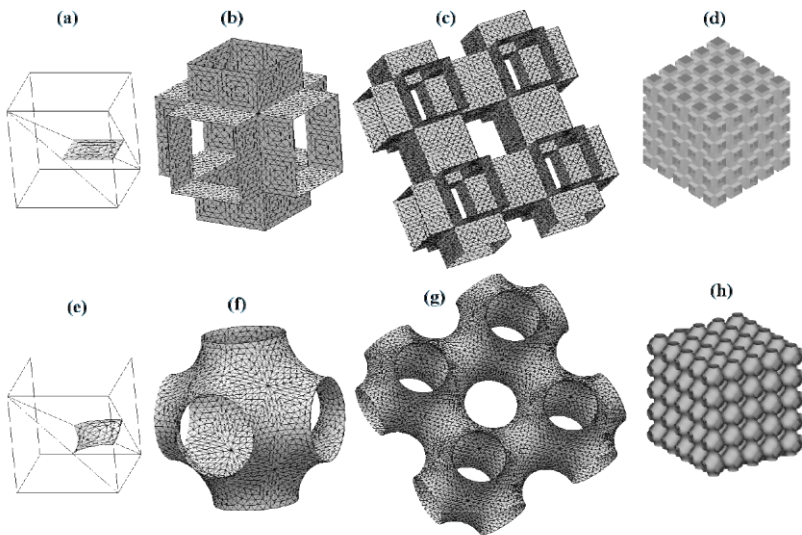


Fig. 3. Using the symmetric periodic properties of minimal surface patches to form the cubic and TPMS unit cells and their tessellations to form scaffolds

4 Scaled-Truncated TPMS

The bicontinuous, equi-partitioned triply periodic minimal surfaces have a maximum void area (porosity) of fifty percent. To minimize the amount of material to be washed out on degradation, and to provide sufficient void space for the cells to proliferate, the porosity should be higher. Since a scaled and clipped minimal surface also results in a minimal surface, unit P-cells can be scaled and truncated with a unit cube. The resultant scaled-truncated TPMS (ST-TPMS) would have an increased porosity. Figure 4 shows the effect of scaling the P- and cubic unit cells by 1.25 and truncating the resultant cells by a unit cube. Figure 5 shows the effect of scaling and truncation on the porosity. As is evident from the graph, for the same scale factor, the porosity of the P-cells is higher than that of the cubic cells. At a scaling factor of 1.25, the strut

width of 200 microns is the minimum feature size that a majority of rapid prototyping devices support. At this level, while the cubic unit cell has a porosity of 68%, the scaled-truncated P-cell has a porosity of 73%. However, the scaling-truncation approach reduces the contact angle between the unit cells. Figure 6 shows the effect of the scaling factor on the contact angle. Though the contact angle decreases, it is still better than the sharp ninety degree turns within the cubic unit cells across all scale factors.

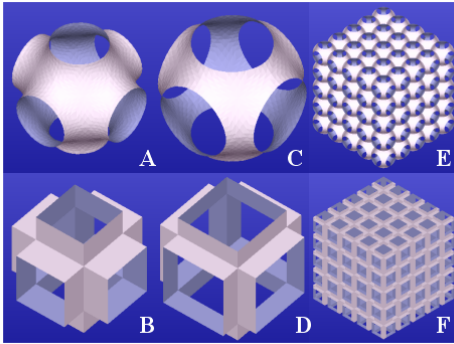


Fig. 4. Scaling and truncation of P- (A) and cubic (B) unit cells to form unit cells (C, D) and their tessellations (E, F) with increased porosity. Unit cells in C and D are obtained at a scale factor of 1.25

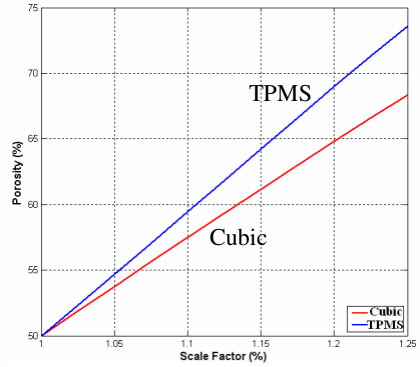


Fig. 5. Effect of scaling and truncation on the porosity of cubic and P- unit cells

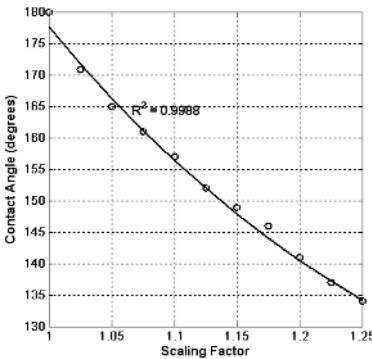


Fig. 6. Effect of scaling on the contact angle between the unit P-cells

5 Fabrication of P-Scaffolds

The unit cells obtained by evolving and tessellating the plane triangulated at 2x refinement can be directly converted into the .STL stereolithographic format suitable for fabrication with rapid prototyping devices. Multiple copies of the 5mm³ scaffolds with strut width and pore diameter of 500 microns were fabricated using the Patternmaster rapid prototyping machine (Solidscap, Merrimack, NH) which uses thermoplastic and

wax to print the solid and pore spaces, respectively. The thermoplastic phase was dissolved in acetone and a biodegradable poly (propylene fumarate) (PPF) polymerizing macromer was injected into the wax mold under a vacuum of 100 mmHg. Finally, the wax was melted to form the biodegradable PPF scaffold.

6 Mechanical Characterization of P-Unit Cells and Scaffolds

Mechanical Characterization of the unit cells was performed with bulk compressive simulation using ABAQUS (an FEM package). Force of 0.3031N was applied to each of the 32 face nodes in the six faces. The material properties (material: PMMA; Young's modulus: 2.4 GPa; Poisson ratio: 0.375) were set identically for both the cubic and P- unit cells. Figure 7 shows the resultant Von Mises stress and Principal strain distribution for the unit cells. Figure 8 shows the same for the unit cells obtained by scaling the original cells by 1.25 and truncating the resultant with a unit cube. Note that at this scale factor, the porosities of P-cell and cubic unit cells are 73% and 68% respectively. It is clearly seen that the stress and strain concentration along sharp edges of the cubic unit cell have been significantly reduced in the unit P-cell. The simulation also reveals the optimal stress distribution and significant strain reduction on the TPMS unit cells.

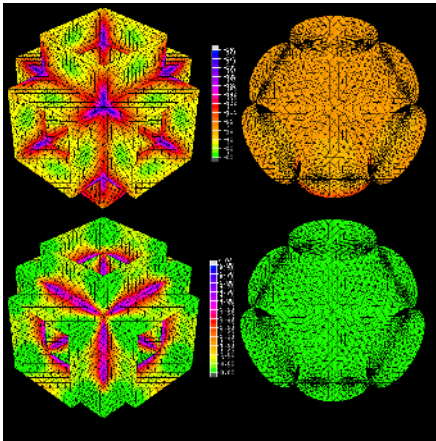


Fig. 7. Von Mises stress (top) and Principal Strain (bottom) maps under bulk compression for Cubic (left) and TPMS (right) unit cells (scale factor 1.0) with identical loading conditions and material properties

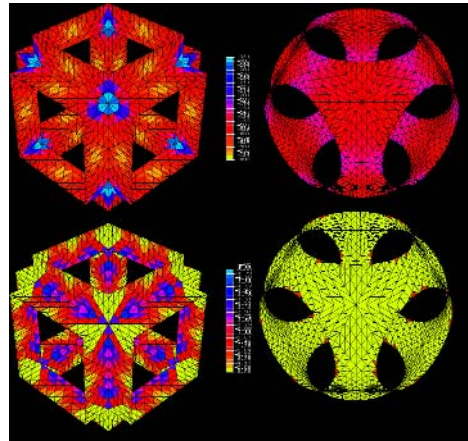


Fig. 8. Von Mises stress (top) and Principal Strain (bottom) maps under bulk compression for Cubic (left) and TPMS (right) unit cells (scale factor 1.25) with identical loading conditions and material properties

To further validate the simulations and to characterize the whole scaffold, uniaxial compression of the scaffolds and unit cells was performed on a Dynamic Mechanical Analyzer (DMA: TA Inst., New Castle, DE). Fourteen specimens of both the scaffolds were used for the mechanical testing. The specimens were uniaxially

compressed with parallel plates by applying a ramp force of 4N/min for 4.5 minutes. Figure 9 compares the linear modulus of the cubic and P-cell based unit cells and scaffolds. Figure 10 shows the effect of scaling on the linear modulus of the scaffolds. Both the results correlate with the findings from the finite element simulations.

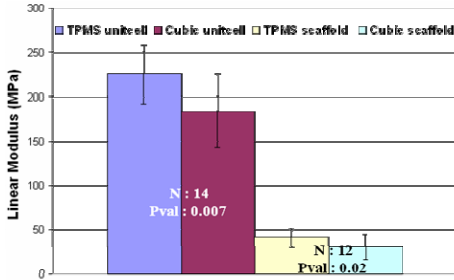


Fig. 9. Linear modulus of TPMS and cubic scaffolds and unit cells based on uniaxial compression on DMA under identical loading conditions (4 N/min. ramp force for 4.5 minutes)

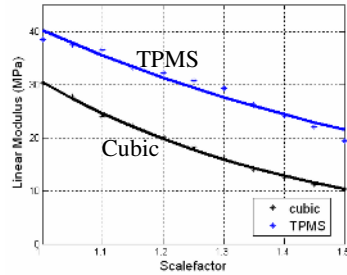


Fig. 10. Effect of unit cell scaling on Linear Modulus based on DMA testing of TPMS and cubic ST scaffolds

7 Cell Seeding on TPMS Scaffolds

It is well known that organelles in plant and mammalian cells assume TPMS forms during normal differentiation [13]. To test the suitability of TPMS morphology in tissue engineering, ATDC cells (mouse chondrogenic cell line) were cultured on the scaffolds and examined for viability after 1 day using Viability/ Cytotoxicity kit. Living cells were stained green and dead cells were red when visualized with confocal scanning microscope. A high viability (>95%) was observed for the adherent cells to the surface of scaffolds. Figure 11 shows the distribution of cells on the surface of the scaffold. The cells are active and well distributed through the edges of the pores. The extended cell processes along the scaffold surface reveals the suitability of the substrate.

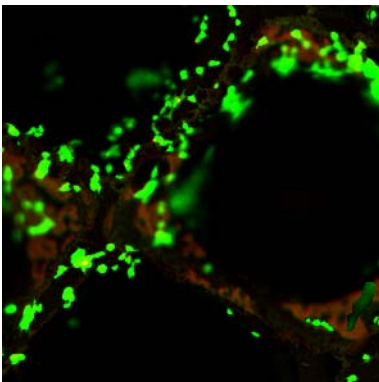


Fig. 11. Suitability of TPMS scaffolds as viable substrate for cell growth demonstrated with confocal scanning microscopy. The live cells are stained in green and the dead cells (representative shown by arrow) in red. As is evident, cell viability is high.

8 Conclusions

Though the natural manifestations of triply periodic minimal surface forms have been observed in living and non-living forms, and the computational framework to create such forms have been in place for decades, to the best of our knowledge, no attempt has been made so far to fabricate them using rapid prototyping devices let alone investigate their potential use as tissue surrogates. This paper describes a first attempt to physically realize and mechanically characterize the Schwarz primitive triply periodic minimal surfaces. The preliminary outcomes of our investigation provide the previously elusive justification for the *TPMS-philicity* of plant and mammalian organelles. This viable morphology, when replicated at macro (tissue) level may also have profound influence on cell migration and tissue growth and may provide an optimal biomorphic tissue analog.

References

1. Yaszemski MJ, Payne RJ, Hayes WC, Langer R, and Mikos AG, "Evolution of bone transplantation: molecular, cellular and tissue strategies to engineer human bone", *Biomaterials*, **17**(1996) 175-185.
2. Huttmacher DW, Sittinger M, and Risbud MV, "Scaffold-based tissue engineering: rationale for computer-aided design and solid free-form fabrication systems", *Trends in Biotechnology*, **22**(7) (2004) 354-362.
3. Spalazzi JP, Dionisio KL, Jiang J, and Lu KK, "Osteoblast and Chondrocyte interactions during coculture on scaffolds", *IEEE Eng. Med & Biol.*, **22**(5) (2003) 27-34.
4. Lord EA, and Mackay AL, "Periodic minimal surfaces of cubic symmetry", *Current Science*, **85**(3) (2003) 346-362.
5. Meusnier JBM, *Mem. Mathem. Phys. Acad. Sci. Paris, pres. par. div. Savans*, **10** (1785) 477-510.
6. Schwarz H, *Gesammelte Mathematische Abhandlungen*, Vols 1, 2 (1890), Springer, Berlin.
7. Schoen AH, "Infinite periodic minimal surfaces without self-intersections" *NASA Technical report TN D-5541*, (1970), Washington DC.
8. Hyde S, Andersson S, Larsson Z, Blum T, Landh S, Lidin BW, and Ninham BW, *The Language of Shape. The Role of curvature in Condensed Matter: Physics, Chemistry and Biology*, (1997), Elsevier, Amsterdam.
9. Pinkall U, and Polthier K, "Computing Discrete Minimal Surfaces and their conjugates", *Experimental Mathematics*, **2** (1993) 15-36.
10. Mackay AL, "Periodic minimal surfaces from finite element methods", *Chem. Phys. Lett.* **221** (1994) 317-321.
11. Brakke KA, "The Surface Evolver", *Experimental Mathematics*, **1**(2) (1992) 141-165.
12. Mackay AL, "Periodic minimal surfaces", *Nature*, **314** (1985) 604-606.
13. Landh T. *FEBS Lett*, **369**:13-13, 1995.

Automatic Detection and Segmentation of Robot-Assisted Surgical Motions

Henry C. Lin¹, Izhak Shafran², Todd E. Murphy³, Allison M. Okamura³,
David D. Yuh⁴, and Gregory D. Hager¹

¹ Department of Computer Science,

² Department of Electrical and Computer Engineering,

³ Department of Mechanical Engineering,

The Johns Hopkins University, Baltimore, MD, USA

⁴ Division of Cardiac Surgery,

Johns Hopkins Medical Institutions, Baltimore, MD, USA

hcl@cs.jhu.edu

Abstract. Robotic surgical systems such as Intuitive Surgical’s da Vinci system provide a rich source of motion and video data from surgical procedures. In principle, this data can be used to evaluate surgical skill, provide surgical training feedback, or document essential aspects of a procedure. If processed online, the data can be used to provide context-specific information or motion enhancements to the surgeon. However, in every case, the key step is to relate recorded motion data to a model of the procedure being performed. This paper examines our progress at developing techniques for “parsing” raw motion data from a surgical task into a labelled sequence of surgical gestures. Our current techniques have achieved >90% fully automated recognition rates on 15 datasets.

1 Introduction

Surgical training and evaluation has traditionally been an interactive and slow process in which interns and junior residents perform operations under the supervision of a faculty surgeon. This method of training lacks any objective means of quantifying and assessing surgical skills [1–4]. Economic pressures to reduce the cost of training surgeons and national limitations on resident work hours have created a need for efficient methods to supplement traditional training paradigms. While surgical simulators aim to provide such training, they have limited impact as a training tool since they are generally operation specific and cannot be broadly applied [5–8].

Robot-assisted minimally invasive surgical systems, such as Intuitive Surgical’s da Vinci, introduce new challenges to this paradigm due to its steep learning curve. However, their ability to record quantitative motion and video data opens up the possibility of creating descriptive, mathematical models to recognize and analyze surgical training and performance. These models can then be used to help evaluate and train surgeons, produce quantitative measures of surgical proficiency, automatically annotate surgical recordings, and provide data for a variety of other applications in medical informatics.

Recently, several approaches to surgical skill evaluation have had success. In the area of high-level surgical modeling, Rosen et al. [9–11] have shown that statistical models derived from recorded force and motion data can be used to classify surgical skill level (novice or expert) with classification accuracy approaching 90%. However, these results rely on a manual interpretation of recorded video data by an expert physician. In the area of low-level surgical data analysis, the MIST-VR laparoscopic trainer has become a widely used system [12]. These systems perform low-level analysis of the positions, forces, and times recorded during training on simulator systems to assess surgical skill [13–15]. Similar techniques are in a system developed by Darzi et al., the Imperial College Surgical Assessment Device (ICSAD) [16]. ICSAD tracks electromagnetic markers on a trainee’s hands and uses the motion data to provide information about the number and speed of hand movements, the distance traveled by the hands, and the overall task time. ICSAD has been validated and used extensively in numerous studies, e.g. [17, 18]. Verner et al. [19] collected da Vinci motion data during performance of a training task by several surgeons. Their analysis also examined tool tip path length, velocities, and time required to complete the task.

It is important to note that ICSAD, MIST-VR, and most other systems mentioned above simply count the number of hand motions, using hand velocity as the segmentation criteria, and do not attempt to identify surgical gestures. In this paper we have developed automatic techniques for not only detecting surgical gestures but also segmenting them. This would allow for the development of automatic methods to evaluate overall proficiency and specific skills.

2 Modeling Robot-Assisted Surgical Motion

Evaluating surgical skill is a complex task, even for a trained faculty surgeon. As a first step, we investigate the problem of recognizing simple elementary motions that occur in a simplified task. Robot motion analysis of users with varying da Vinci experience were studied. Automatic recognition of elementary motion requires complex machine learning algorithms, and, potentially, a large number of parameters. To guide the choice of techniques and to gain useful insight into the problem, we divided the task into functional modules, illustrated in Fig. 2, and akin to other pattern recognition tasks such as automatic speech recognition. In this section, we will describe the data used for this study, the paradigm for training and testing, and a solution for the motion recognition problem.



Fig. 1. A video frame of the suture task used for this study

2.1 Corpus for the Experiments

The da Vinci API data consists of 78 motion variables acquired at 10 Hz during operation. Of these, 25 track each of the master console manipulators, and 14

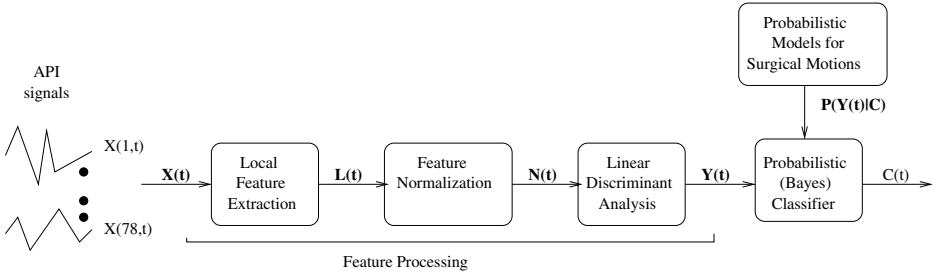


Fig. 2. Functional block diagram of the system used to recognize elementary surgical motions in this study

track each of the patient-side manipulators. We selected the suturing task (Fig. 1) as the model in which our motion vocabulary, $m(s)$, would be defined.

The eight elementary suturing gestures are:

1. reach for needle (gripper open)
2. position needle (holding needle)
3. insert needle/push needle through tissue
4. move to middle with needle (left hand)
5. move to middle with needle (right hand)
6. pull suture with left hand
7. pull suture with right hand
8. orient needle with two hands

2.2 Recognizing Surgical Motion

The task of recognizing elementary surgical motions can be viewed as a mapping of temporal signals to a sequence of category labels. The category labels belong to a finite set C , while the temporal signals are real valued stochastic variables, $\mathbf{X}(k)$, tapped from the master and patient-side units. Thus, the task is to map:

$$\mathcal{F} : \mathbf{X}(1 : k) \mapsto C(1 : n)$$

Our work adopts a statistical framework, where the function \mathcal{F} is learned from the data. The task of learning \mathcal{F} can be simplified by projecting $\mathbf{X}(k)$ into a feature space where the categories are well separated. The sequence of operations is illustrated by the functional block diagram in Fig. 2.

2.3 Feature Processing

The goal of feature processing is to remove redundancy in the input features while retaining the information essential for recognizing the motions with high accuracy. As noted earlier, the input feature vectors consist of 78 position and velocity measurements from the da Vinci manipulators. Feature processing reduces the dimension from 78 to less than 6 features without any loss of performance. In this work, we have found the following feature processing steps to be effective.

1. **Local Temporal Features:** Surgical motion seldom changes from one gesture to another abruptly. Thus information from adjacent input samples can be useful in improving the accuracy and robustness of recognizing a surgical motion. As in automatic speech recognition, this information can be incorporated directly by concatenating the feature vector $\mathbf{X}(k_t)$ at time t with those from neighboring samples, $t - m$ to $t + m$, to make it vector of dimension $(2m + 1)|\mathbf{X}(k_t)|$.

$$\mathbf{L}(k_t) = [\mathbf{X}(k_{t-m})|\mathbf{X}(k_{t-m+1})|\dots|\mathbf{X}(k_t)|\dots|\mathbf{X}(k_{t+m-1})|\mathbf{X}(k_{t+m})]$$

In addition, derived features such as speed and acceleration were included as a part of each local temporal feature.

2. **Feature Normalization:** Since the units of measurements for position and velocity are different, the range of values that they take are significantly different. This difference in dynamic range often hurts the performance of a classifier or a recognition system. So, the mean and variance of each dimension is normalized by applying a simple transformation,

$$\mathbf{N}_i(k) = \frac{1}{\sigma_i^2}(\mathbf{L}_i(k) - \mu_i),$$

where $\mu_i = \frac{1}{N}\mathbf{L}_i(k)$ and $\sigma_i^2 = \frac{1}{N}(\mathbf{L}_i(k) - \mu_i)^2$.

3. **Linear Discriminant Analysis:** When the features corresponding to different surgical motions are well separated, the accuracy of the recognizer can be considerably improved. One such transformation is the linear discriminant analysis [20].

$$\mathbf{Y}(k) = \mathbf{W}N(k)$$

The linear transformation matrix W is estimated by maximizing the Fisher discriminant, which is the ratio of distance between the classes and the average variance of a class. The transformation that maximizes the ratio projects the features into a space where the classes are compact but away from each other.

2.4 Bayes Classifier

The discriminant function, \mathcal{F} , could be of several forms. When all errors are given equal weight, it can be shown that the optimal discriminant function is given by Bayes decision rule.

$$\begin{aligned} \hat{C}(1:n) &= \arg \max_{C(1:n)} P(C(1:n)|\mathbf{Y}(1:k)) \\ &= \arg \max_{C(1:n)} P(\mathbf{Y}(1:k)|C(1:n))P(C(1:n)) \end{aligned}$$

In other words, the optimal decision is to pick the sequence whose posterior probability, $P(C(1:n)|\mathbf{Y}(1:k))$, is maximum. Using Bayes chain rule, this can

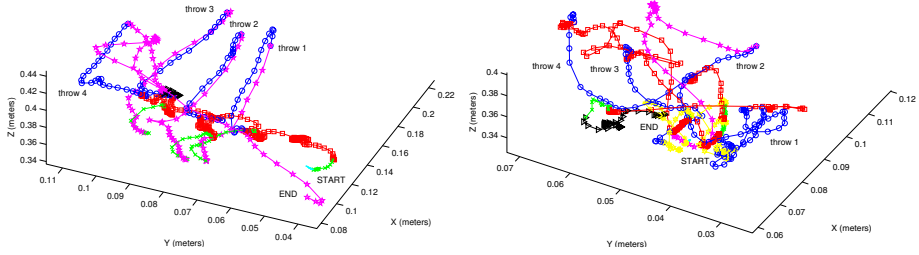


Fig. 3. A plot of the Cartesian positions of the da Vinci left master manipulator, identified by surgical gesture, during performance of a 4-throw suturing task. The left plot is that of an expert surgeon while the right is of a less experienced surgeon.

be rewritten as the product of prior probability of the class sequence, $P(C(1 : n))$, and the generative probability for the class sequence, $P(\mathbf{Y}(1 : k)|C(1 : n))$.

As a first step, we make the simplifying assumption that each time frame in the input sequence is independently generated. That is, $P(C(1 : k)|\mathbf{Y}(1 : k)) = \prod_{i=1}^k P(C(i)|\mathbf{Y}(i))$. Thus, the decision is made at each frame independent of its context.

2.5 Cross-Validation Paradigm

The data used for this study contains 15 expert trials and 12 intermediate trials of performing a suturing task, consisting of 6 to 8 different elementary surgical motions. To improve the statistical significance of the results, we performed a 15-fold cross validation on the expert data. That is, the machine learning algorithm was evaluated by performing 15 different tests. In each test, two trials were held out for testing and the statistical models were trained on the rest of the data. The average across 15 such tests were used to measure the performance of various settings of the parameters.

3 Results

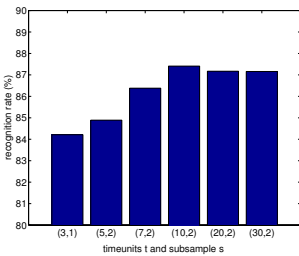


Fig. 4. Results of varying the temporal length t and sampling granularity s

To guide the choice of parameters, our initial experiments were performed on the data collected from 15 trials by an expert da Vinci surgeon, performing a suturing task involving 4 throws (Fig. 1) in each trial. Subsequently, we applied the recognition and segmentation techniques on 12 trials of a surgeon with limited da Vinci experience (intermediate) and compared the results.

After preliminary observation of the data, a few preprocessing steps were carried out before modeling the surgical motions. Of the eight motions defined in Sec. 2.1, the expert surgeon did not utilize motion 5 and 7, so they were not modeled. Each

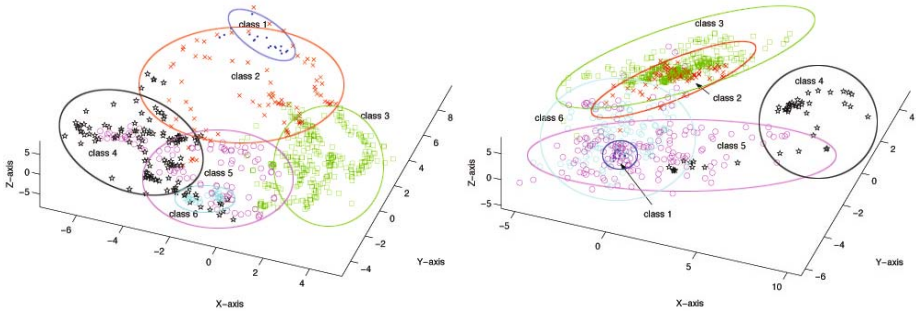


Fig. 5. The result of LDA reduction with $m=6$ and $d=3$. The expert surgeon’s motions (left) separate more distinctly than the less experienced surgeon’s (right).

dimension of the feature vector from the expert surgeon contained about 600 samples. For example, Fig. 3 illustrates Cartesian positions of the left master during one of the trials.

3.1 Local Temporal Features

The length and sampling rate of the temporal feature “stacking” was varied to determine the optimal length and granularity of motion to consider. Our results showed, as one would expect, that too little temporal length results in a disappearance of any advantage, whereas too large of a temporal length increased the chance of blurring the transition between neighboring motions. Fig. 4 shows the results of varying the temporal length (t) and sampling granularity (s). Due to its high recognition rates, we use $t=10$ and $s=2$ for the rest of our experiments.

3.2 Linear Discriminant Analysis

Fig. 5 shows the reliability of LDA in separating motion data into 6 distinct regions in a 3-dimensional projection space. An intermediate surgeon’s motions tend to not separate as well, indicating less consistent motions.

These initial experiments validated the hypothesis that LDA could be used to simplify the original data into a simpler, low-dimensional data set. A second set of experiments examined the effect of varying the number of motion classes, $C(1:\{4,5,6\})$, and the dimensionality of the projection, $d = \{3,4,5\}$. The cross-validation paradigm described in Sec. 2.5 was applied in all experiments to compute a recognition rate. Table. 1 shows the recognition rates of the Bayes classifier after the LDA reduction with varying C and d values.

Having fine tuned the classifier for surgical motion, we then applied the algorithm to produce segmentations. Fig. 6 shows the comparison of segmentation generated by the algorithm and by a human for a randomly chosen trial of the expert surgeon. In spite of the fact that the model only incorporates weak temporal constraints through the local temporal features described in Sec. 2.3, the

Table 1. The results of grouping the motion categories and varying the dimension of the projected space. In the second column, the number of unique integers indicates the number of motion categories, and the position of the integer indicates which motions belong to that category.

<i>n</i>	<i>class_membership</i>	<i>LDA_dimensions</i>	% correct
1	12345566	3	91.26
2	12345566	4	91.46
3	12345566	5	91.14
4	12345555	3	91.06
5	12345555	4	91.34
6	11234455	3	92.09
7	11234455	4	91.92
8	12234444	3	91.88

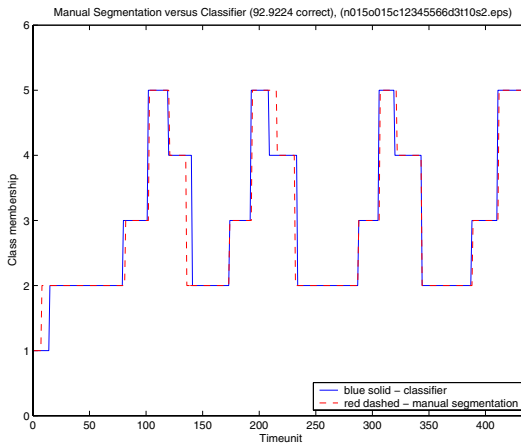


Fig. 6. Comparison of automatic segmentation of robot-assisted surgical motion with manual segmentations. Note, most errors occur at the transitions.

segmentation produces surprisingly good results. In most trials, the errors are largely at the transition, as shown in Fig. 6. While using the robotic system, transitions from one motion to the next are often performed without any pause, and as a result it is difficult even for a human to mark a sharp transition boundary. Consequently, we removed a 0.5 second window at each boundary, so as to avoid confidence issues in the manual segmentation. The 0.5 second window is statistically insignificant because an average surgical motion lasts over 3 seconds.

4 Discussion

We have shown that linear discriminant analysis is a robust tool for reducing and separating surgical motions into a space more conducive to gesture recogni-

tion. In our highest rated test, we reduced 78 feature vectors into 3 dimensions with 6 classes and still achieved nearly 90% in recognition. With refinement and the combination of other statistical methods, such as Hidden Markov Models (HMMs), we believe mid-90s recognition rates are possible. We have also suggested how this framework can support objective evaluation of surgical skill levels by varying different parameters in our mathematical model. Our experiments have shown that the motions of an expert surgeon are very efficient and thus can be used as a skill evaluator or training model. In ongoing work, we have begun combining the training of expert and intermediate surgeon data to create one model that can distinguish varying skill levels.

Acknowledgements

This research was supported by NSF and the Division of Cardiac Surgery at the Johns Hopkins Medical Institutions. The authors thank Dr. Randy Brown, Sue Eller, and the staff of Minimally Invasive Surgical Training Center at Johns Hopkins Medical Institutions for access to the da Vinci, and Intuitive Surgical, Inc. for use of the da Vinci API.

References

1. King, R.T.: New keyhole heart surgery arrived with fanfare, but was it premature? *Wall Street Journal* (1999) 1
2. Haddad, M., et al.: Assessment of surgical residents' competence based on postoperative complications. *Int Surg* **72** (1987) 230–232
3. Darzi, A., Smith, S., Taffinder, N.: Assessing operative skill needs to become more objective. *British Medical Journal* **318** (1999) 887–888
4. Barnes, R.W.: But can s/he operate?: Teaching and learning surgical skills. *Current Surgery* **51(4)** (1994) 256–258
5. Wilhelm, D., et al.: Assessment of basic endoscopic performance using a virtual reality simulator. *J Am Coll Surg* **195** (2002) 675–681
6. Cowan, C., Lazenby, H., Martin, A., et al.: National health care expenditures, 1998. *Health Care Finance Rev* **21** (1999) 165–210
7. Acosta, E., Temkin, B.: Dynamic generation of surgery specific simulators - a feasibility study. *StudHealth Technology Inform* **111** (2005) 1–7
8. R, M., CP, D., SS, M.: The anterior abdominal wall in laparoscopic procedures and limitations of laparoscopic simulators. *Surg Endosc* **10(4)** (1996) 411–413
9. Rosen, J., et al.: Markov modeling of minimally invasive surgery based on tool/tissue interaction and force/torque signatures for evaluating surgical skills. In: *IEEE Trans Biomed Eng.* Volume 48(5). (2001) 579–591
10. Rosen, J., et al.: Task decomposition of laparoscopic surgery for objective evaluation of surgical residents' learning curve using hidden Markov model. In: *Computer Aided Surgery.* Volume 7(1). (2002) 49–61
11. Richards, C., Rosen, J., Hannaford, B., Pellegrini, C., Sinanan, M.: Skills evaluation in minimally invasive surgery using force/torque signatures. In: *Surgical Endoscopy.* Volume 14. (2000) 791–798

12. Gallagher, A.G., Satava, R.M.: Virtual reality as a metric for the assessment of laparoscopic psychomotor skills. In: *Surg. Endoscopy*. Volume 16(2). (2002) 1746–1752
13. Cotin, S., et al.: Metrics for laparoscopic skills trainers: the weakest link. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Volume 2488. (2002) 35–43
14. O'Toole, R.V., Playter, R.R., Krummel, T.M., Blank, W.C., Cornelius, N.H., Roberts, W.R., Bell, W.J., Raibert, M.R.: Measuring and developing suturing technique with a virtual reality surgical simulator. *Journal of the American College of Surgeons* **189**(1) (1999) 114–127
15. Yamauchi, Y., et al.: Surgical skill evaluation by force data for endoscopic sinus surgery training system. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Volume 2488. (2002)
16. Darzi, A., Mackay, S.: Skills assessment of surgeons. *Surg.* **131**(2) (2002) 121–124
17. Datta, V., et al.: The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in laboratory-based model. *Journal of the American College of Surgery* **193** (2001) 479–485
18. Datta, V., et al.: Relationship between skill and outcome in the laboratory-based model. *Surgery* **131**(3) (2001) 318–323
19. Verner, L., Oleynikov, D., Holtman, S., Haider, H., Zhukov, L.: Measurements of the level of expertise using flight path analysis from da vinci robotic surgical system. In: *Medicine Meets Virtual Reality II*. Volume 94. (2003)
20. Fisher, R.: The use of multiple measurements in taxonomic problems. *Annals of Eugenics* **7** (1936) 179–188

DaVinci Canvas: A Telerobotic Surgical System with Integrated, Robot-Assisted, Laparoscopic Ultrasound Capability

Joshua Leven¹, Darius Burschka¹, Rajesh Kumar², Gary Zhang²,
Steve Blumenkranz², Xiangtian (Donald) Dai¹, Mike Awad¹, Gregory D. Hager¹,
Mike Marohn¹, Mike Choti¹, Chris Hasser², and Russell H. Taylor¹

¹ The Johns Hopkins University, Baltimore, Maryland 21218 USA
rht@jhu.edu

² Intuitive Surgical, Inc., Sunnyvale, CA 94086
Chris.Hasser@intusurg.com

Abstract. We present daVinci Canvas: a telerobotic surgical system with integrated robot-assisted laparoscopic ultrasound capability. DaVinci Canvas consists of the integration of a rigid laparoscopic ultrasound probe with the daVinci robot, video tracking of ultrasound probe motions, endoscope and ultrasound calibration and registration, autonomous robot motions, and the display of registered 2D and 3D ultrasound images. Although we used laparoscopic liver cancer surgery as a focusing application, our broader aim was the development of a versatile system that would be useful for many procedures.

1 Introduction

Open surgery serves as the gold standard treatment for many disorders, despite the increased morbidity associated with laparotomy. Minimally invasive techniques have developed to reduce this morbidity and improve the quality of life following surgery. In exchange for this lower morbidity, clinicians and patients have, at times, accepted lower efficacy, and surgeons have tolerated more difficult working conditions. The daVinci® surgical system gives surgeons dexterous laparoscopic tools, intuitive motion, and a high-fidelity 3D vision system, allowing them to achieve outcomes comparable or better to that of open surgery, with the low morbidity of minimally invasive surgery (e.g., [1]). However, the “classical telesurgery” approach of the daVinci essentially is limited to replicating 20th century open surgery in an MIS environment.

Primary liver cancer is one of the most common malignancies, accounting for more than 1 million cases per year [2]. Hepatic metastases, or secondary liver cancer, is also a common disease facing the clinician. IOUS-directed procedures such as liver biopsy, tumor ablation, and hepatic resection today often require trade-offs between efficacy and morbidity. While liver resection is most often performed using open laparotomy, percutaneous approaches are often used to perform needle biopsy or ablation. This technique has the advantage of potentially lower morbidity, at least compared to open surgery, and it can employ various imaging modalities, including US, CT, and MRI. However, there are several significant advantages to the performance of ablation during laparotomy or laparoscopy. First, surgical procedures

provide enhanced staging, as laparotomy and laparoscopy afford the opportunity to identify both hepatic and extra-hepatic metastases not visualized on preoperative imaging. Intraoperative ultrasonography (IOUS) capitalizes on the ability to place the ultrasound probe directly on the liver surface. Second, laparotomy or laparoscopy affords improved access to tumors in difficult locations. Third, in cases of multiple tumors, biopsy and ablation can be combined with resection when using an operative approach. Finally, clinical studies demonstrate that open surgical ablation likely has improved outcomes compared to percutaneous approaches [3].

IOUS of the liver [4] is widely used for staging, ablation, and planning for resection. It is the most accurate method for detecting liver metastases, with accuracy rates above 90 percent [27]. Despite theoretical advantages, intraoperative laparoscopic ultrasound (LUS) is not widely practiced. In situations in which LUS has been utilized, the results are not comparable to that of open operative ultrasonography. Limitations to this technique among general and hepatobiliary surgeons relate to the its technical difficulty, to the imprecise methods of manual free-hand probe manipulation, limited degree of positioning, and 2D imaging [5].

This paper reports our initial steps in integrating robot-assisted laparoscopic ultrasound (RLUS) into the daVinci surgical robot system. Our general goal is to enable surgeons to perform minimally invasive liver interventions with the efficacy of open surgery. Although we have used liver surgery as a focusing application, our broader aim is to develop a generally useful LUS capability for many surgical fields, including general, cardiac, gynecologic, and urologic surgery.

2 Background

Several groups have active programs in robotically-assisted ultrasonography. Fenster, *et al.* have a long record of using tracked and robotically-manipulated 2D US probes to produce 3D US images, as well as recent work on robotically-assisted TRUS-guided prostate brachytherapy [6]. Several groups have reported systems using ultrasound-based targeting to assist in other robotically-assisted percutaneous procedures [7,8] and other groups (e.g., [9-12]) have developed robotically-manipulated extracorporeal ultrasound systems. None of these systems involve LUS or integrate US into an interventional procedure. There has been one experimental system for remote LUS probe manipulation [13] as part of a 1998 EU telemedicine initiative. We have not been able to find any subsequent publication of this work.

Many groups have explored the use of navigational tracking devices with transcatheter 2D and 3D ultrasound, including work targeted at the liver (e.g., [14, 15]) There have been recent efforts to apply navigation techniques to LUS (e.g., [16-18]) and work on “augmented reality” for US using head mounted displays [19]. There has also been a body of work on human-machine cooperative systems implementing “virtual fixtures” or the equivalent for telesurgery (e.g., [20-22]).

3 System Overview

A block diagram of our system appears in Fig 1-A. It consists of a standard daVinci surgical robot equipped with a special LUS tool (Fig 1-B,C), consisting of a 10 mm

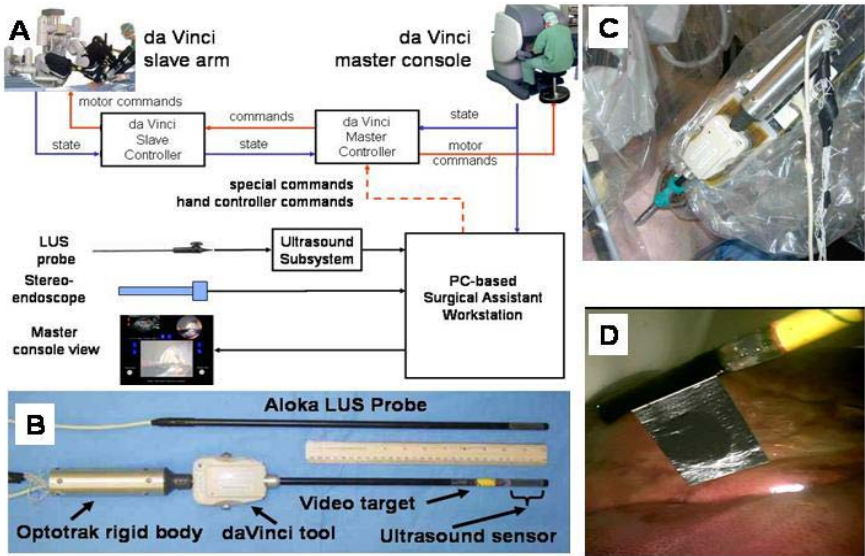


Fig. 1. (A) Block diagram; (B) LUS tool with Optotrak tracking body; (C) LUS tool held by daVinci in pig cadaver; (D) view through surgeon console of LUS probe with video tracking target and video overlay of LUS image

diameter Aloka LUS probe held by a specially constructed daVinci interface unit that is capable of rotating the LUS probe about its axis. We have added a distinctive pattern to the shaft of the probe near the LUS sensor (Fig 2-A) that can easily be located by computer vision techniques. The back end of the LUS probe is equipped with a specially constructed Optotrak[®] tracking body for use in system calibration and validation experiments. To this, we added a PC-based “surgical assistant” workstation which communicated with the daVinci master console through serial links and research application programming interface (API) provided under special agreement by Intuitive Surgical, Inc. Both video channels from the daVinci’s stereoendoscope were fed through video frame grabber cards attached to the workstation, permitting us to perform real time image processing. LUS images were captured from the Aloka video console through a third frame grabber. State information from the daVinci was

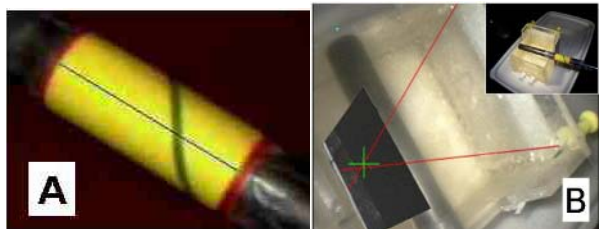


Fig. 2. (A) Example of the ultrasound probe with tracking marker and detected shaft axis (blue line -- graphic overlay). (B) Overlay Accuracy Phantom being used with the LUS probe. (Inset) Probe sitting in the phantom. (Main image) The extrapolated location of the intersection point in camera view (red), and the ultrasound crosswire segmentation (green).

transmitted to the workstation through Ethernet and a limited repertoire of motion commands were sent to the daVinci through a serial interface.

Video tool tracking: To demonstrate the feasibility of visual tool tracking, the ultrasound probe was marked with a yellow sticker with green and red stripes as shown in Fig 2-A, and a segmentation algorithm located the center of the sticker in 2D image space. The yellow sticker helped locate the shaft centerline, the green spiral line helped compute the twist angle of the shaft while the red stripes helped the extraction of the end points of this centerline.

Once the centerlines are detected in both stereo cameras, we compute the 3D axis of the shaft from the intersection of two planes formed by the shaft centerline and the optical origin in the left and right cameras respectively.

Calibration and registration: DaVinci Canvas requires several calibration and registration steps, all implemented by leveraging prior work in our laboratory.

Laparoscopic Ultrasound Calibration - Ultrasound image coordinates were calibrated to the Optotrak rigid body using an $AX=XB$ formulation [23]. The ultrasound probe was placed in three known orientations in a specially constructed calibration phantom {Viswanathan, 2004 #1599}. The three poses allow three relative transformations based on Optotrak readings (A) and three relative transformations based on the ultrasound images (B) for the $AX=XB$ registration.

Stereo Endoscope Calibration and Registration - In order to determine the intrinsic and extrinsic parameters of the stereo endoscope, we used our checkerboard phantom with the multi-plane formulation provided by the Caltech Image Calibration Toolkit. We added Optotrak markers to a typical checkerboard video calibration phantom and digitized each corner of the checkerboard using a calibrated Optotrak pointer. These corner coordinates were then used with the camera calibration to perform a point-cloud to point-cloud registration between the endoscope rigid body and camera frame.

DaVinci Robot API Registration - The daVinci robot API uses the robot kinematics to report a coordinate frame for the endoscope and ultrasound tool tip. However, due to inaccuracies in the setup joint encoders, both of these coordinate frames were offset from their correct values. Thus, it was necessary to register the offsets between the real camera frame and the camera frame calculated from the kinematics as well as between the real and kinematic ultrasound probe frames. Then, the kinematics could be used in place of the Optotrak readings to determine ultrasound image overlay placement.

We assumed a constant transformation between the kinematic tool tip and the LUS probe Optotrak rigid body. This is reasonable so long as the position of the camera does not overly change. Using an $AX=XB$ formulation, the probe was moved to several positions, and the

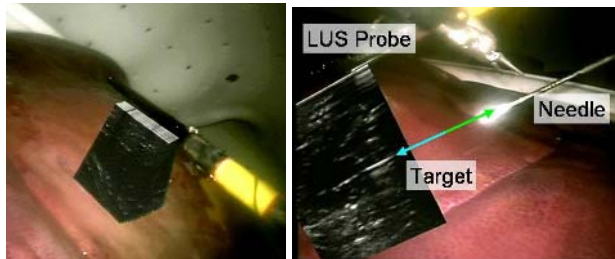


Fig. 3: (left) 3D LUS image obtained by “rocking”; (right) 2D LUS of lesion during targeting task (graphics added manually)

static offset between tool tip and rigid body was registered. Knowing this offset, the camera offset could be calculated directly.

Ultrasound Image Display: DaVinci Canvas includes the ability for the user to view the ultrasound image as a picture-in-picture (PIP) insert, or as an overlay that appears to fan out from the ultrasound transducer itself, providing a visualization similar to the “ultrasound flashlight” concept of Stetten *et al.* [25](Fig 1-D). Our system includes a simple 3D ultrasound overlay volume rendering implementation based on [26] (Fig. 3 - Left), and was evaluated by performing a needle insertion task.

Autonomous Motions: Robotic ultrasound has the potential to reduce variability in the images produced, compared to freehand scanning, and can reduce operator workload and difficulty. Behaviors as simple as rocking the probe back and forth can maintain an updated 3D image without operator intervention. In daVinci Canvas, surgeons can rock the probe to one extreme, squeeze the daVinci grip, rock to the other extreme, and release the grip to easily specify automated rocking. Surgeons could also use this same click and drag paradigm to initiate repeating trajectories in Cartesian space.

4 System Accuracy

We evaluated the accuracy of the system under various calibration conditions and determined an error of 2.83 ± 0.83 mm using Optotrak exclusively, and 2.16 ± 1.43 mm using registered daVinci kinematics. We defined accuracy error in this case as the 3D displacement between where a feature appears in the ultrasound image overlay, and where that same feature appears in the endoscope video.

Accuracy assessment was performed using our overlay accuracy phantom (Fig 2-B). The phantom is essentially a crosswire calibration phantom with a significant fraction of the two wires visible above the water line. The crosswire point was localized using the ultrasound overlay and compared to the intersection point of the two wires visible above the water. The extracted points were used to determine two 3D positions in the camera frame – one due to the video, and one due to the overlay. We considered the distance between these two points to be the accuracy error.

5 Surgeon Evaluation

The clinical evaluation team included two attending general surgeons with significant robotic surgery and ultrasound experience, three surgery residents, and a physician with extensive robotic surgery and customer interface experience. The clinical evaluation team assessed ease of use, reliability, clinical strengths, and limitations of the robot assisted laparoscopic ultrasound probe.

Staging Task (Lesion Finding): Our objective in the staging task was to compare the efficacy of freehand laparoscopic ultrasound with robotic laparoscopic ultrasound in locating simulated liver lesions. Lesions were created using 2% agarose in water, and placed in ex-vivo porcine liver. Four livers were used for the staging experiment. The number, size, and location of lesions within each liver were randomized. Six surgeons

found the lesions with freehand LUS, daVinci LUS without automated motions, and daVinci LUS with automated motions.

Results: As expected during freehand ultrasound, advanced surgeons performed better than intermediate and novice surgeons in nearly all categories studied. Likewise, intermediates performed better than novices in the number of lesions correctly identified and identifying correctly the location and size of lesions. Best performance was observed from advanced surgeons using freehand ultrasound. Advanced, intermediate, and novice surgeons performed similarly when using robotic assistance. The longest time to identify each lesion was observed when using robotic assistance with automation.

Discussion: While the robotic techniques did not perform at the level of advanced surgeons using the freehand technique, the robot served as the “equalizer,” allowing novices to perform at the level of more experienced surgical ultrasonographers or better. The rigidity of the ultrasound probe significantly limited motion within the simulated abdomen. Almost all surgeons felt it would be important to have a flexible ultrasound probe that would return degrees of freedom characteristic of the daVinci end effectors. Another important feature universally desired by surgeons was a marking tool to determine if it was a new lesion or one previously encountered. Clinically, this was felt to be important in reducing the number of missed lesions.

Of all the robotic ultrasound features available to the surgeon, overlay and PIP were found to be the most useful. These allowed rapid identification and subsequent characterization of lesions. Automation features were less useful, although these features may be useful with daVinci models having 3 instrument manipulation arms in addition to the endoscope arm. In this case, the surgeon could use two arms to perform the procedure while the third instrument holds the LUS probe and initiates “rocking” to obtain 3D images or 2D sweeps over a lesion or critical structure.

This pilot experiment served as a learning exercise, with a rough prototype and small sample size not intended to show quantitative superiority of robotic ultrasound over freehand ultrasound. The key result was qualitative: Surgeons felt that with an articulated probe and the ability to annotate lesions, the system would have a competitive advantage over freehand ultrasound, particularly when used as an aid to needle targeting or resection tasks. They also felt that it would lower the barriers to the use of laparoscopic ultrasound by making it easier to learn.

Lesion Targeting Task: We performed some preliminary investigation into the use of registered ultrasound overlays in guiding needles to target lesions. Livers were prepared as in the staging task. A laparoscopic surgeon skilled in ultrasound techniques performed the targeting task using either freehand ultrasound and needle or using daVinci assistance with a long rigid needle in the left grasper and the ultrasound probe in the right arm. The long needle was passed through the simulated torso and its pivot point was fixed at the “skin.” Three targeting tasks were performed, one laparoscopic freehand (long needle only) and twice robotically.

Results and Discussion: Freehand placement accuracy was 5.4 mm and robot-assisted accuracy was comparable (4.5 mm & 7.5 mm). For this anecdotal (N=1) experiment, the surgeon comments were more informative. The surgeon suggested it would be helpful to have a flexible daVinci ultrasound probe, and an extrapolation of the needle trajectory displayed in the daVinci console.

6 Conclusions and Future Work

This work was phase one of a larger effort directed toward a complete integrated robotic system for LUS-assisted hepatic surgery and (more generally) toward development of a new generation of information rich surgical assistant systems. Our initial results are very encouraging. The ability of the ultrasound dynamic image to be displayed in 'real time' at the point of scanning is a powerful tool, enhancing ease of use and surgeon situational awareness.

Clinically acceptable accuracies can be achieved using an Optotrak or directly using daVinci kinematics. The rigid ultrasound probe (versus flexible) was a limitation in scanning selected sectors of the operative field. We had hoped that a volume display would improve the speed and accuracy of such a procedure. However, in its current state, it was found to be less helpful and even distracting when compared to the 2D ultrasound overlay. The registered 2D overlay exceeded our expectations by proving capable of assisting needle targeting procedures.

Future plans include development of more robust, capable tools & image processing algorithms, further evaluation *in vivo*, and eventual clinical development

Acknowledgments

At the risk of some injustice to the many colleagues who assisted in various ways, we would like to extend special thanks to Emad Boctor for ultrasound expertise; to Iulian Iorchidata for help with test apparatus; and to Randy Brown and Sue Eller of JHU's MIS Training Center for help with JHU animal trials. This work was funded in part by NIH grant R41 RR019159, in part by the NSF under agreement EEC9731478, and in part by JHU internal funds. We thank Aloka for the loan of LUS equipment.

References

1. T. Ahlering, D. Woo, E. L., and others, "Robot-assisted versus open radical prostatectomy: a comparison of one surgeon's outcomes.," *Urology*, vol. 63, pp. 819-822, 2004.
2. E. Nakakura and M. Choti, "Hepatocellular Carcinoma: Current Management Recommendations," *Advances on Onc*, vol. 16(2), pp. 12-18, 2000.
3. T. Wood and others, "Radiofrequency ablation of 231 unresectable hepatic tumors: indications, limitations, and complications," *Ann. Surg. Oncol.*, vol. 7, pp. 593-600, 2000.
4. R. A. Kane, "Intraoperative Ultrasonography, History, Current State of the Art, and Future Directions," *J Ultrasound Med*, vol. 23, pp. 1407-1420, 2004.
5. B. Rau, M. Hunerbein, and P. Schlag, "Is There Additional Information From Laparoscopic Ultrasound in Tumor Staging?," *Dig Surg*, vol. 19(6), pp. 479-83, 2002.
6. Z. Wei, G. Wan, L. Gardi, and others, "Robot-assisted 3D-TRUS guided prostate brachytherapy: system integration and validation," *Medical Physics*, vol. 31(3), pp. 539-548, 2004.
7. J. Stoll, P. Dupont, and R. Howe, "Ultrasound-based servoing of manipulators for telesurgery," in *SPIE Telesurgery and Telepresence Technologies VIII*, vol. 4570. Newton, Mass.: SPIE, 2001.
8. E. Boctor and others, "Dual-Armed Robotic System for Intraoperative Ultrasound Guided Hepatic Ablative Therapy: A Prospective Study," in *IEEE 2004 International Conference on Robotics and Automation*. New Orleans, LA, 2004, pp. 2517-2522.

9. E. Degoulange, L. Urbain, P. Caron, S. Boudet, J. Garipey, L. Megnien, F. Perrot, and E. Dombre, "HIPPOCRATE: an intrinsically safe robot for medical applications," in *IEEE/RSH International Conference on Intelligent Robots and Systems*. Victoria, 1998, pp. 959-964.
10. P. Abolmaesumi, S. E. Salcudean, W. H. Zhu, M. R. Siroospour, and S. P. DiMaio, "Image-Guided Control of a Robot for Medical Ultrasound," *IEEE Transactions on Robotics and Automation*, vol. 18(1), pp. 11-23, 2002.
11. M. Mitsuishi, S. i. Warisawa, T. Tsuda, T. Higuchi, N. Koizumi, H. Hashizume, and K. Fujiwara, "Remote Ultrasound Diagnostic System," in *Proc. IEEE Conf. on Robotics and Automation*. Seoul, 2001, pp. 1567-1574.
12. A. Vilchis, J. Troccaz, P. Cinquin, K. Masuda, and F. Pellissier, "A New Robot Architecture for Tele-Echography," *IEEE Trans. Robotics & Automation*, vol. 19(5), pp. 922-926.
13. D. d. Cunha, P. Gravez, C. Leroy, E. Maillard, J. Jouan, P. Varley, M. Jones, M. Halliwell, D. Hawkes, P. N. T. Wells, and L. Angelini, "The MIDSTEP System for Ultrasound guided Remote Telesurgery .," in *IEEE EMBS*, vol. 20/3: IEEE, 1998, pp. 1266-1269.
14. J. Stefansic, A. Herline, Y. Shyr, W. Chapman, J. Fitzpatrick, B. Dawant, and R. J. Galloway, "Registration of Physical Space to Laparoscopic image space for use in minimally invasive hepatic surgery," *IEEE Trans Med Imaging*, vol. 19(10), pp. 1012-1023, 2000.
15. T. Lange, S. Eulenstein, M. Hunerbein, H. Lamecker, and P.-M. Schlag, "Augmenting Intraoperative 3D Ultrasound with Preoperative Models for Navigation in Liver Surgery," in *MICCAI 2004*, vol. LNCS 3217. St. Malo: Springer, 2004, pp. 534-541.
16. D. Wilhelm, H. Feussner, A. Schneider, and J. Harms, "Electromagnetically navigated laparoscopic ultrasound," *Surg. Technol. Int.*, vol. 11, pp. 50-54, 2003.
17. J. Ellsmere, J. Stoll, D. Rattner, and D. Brooks, "A navigation system for augmenting laparoscopic ultrasound," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003*: Springer, 2003, pp. 184-191.
18. P. M. Novotny, J. W. Cannon, and R. D. Howe, "Tool Localization in 3D Ultrasound Images," in *MICCAI 2003*, vol. LNCS 2879. Montreal: Springer, 2003, pp. 969-970.
19. H. Fuchs, M. A. Livingston, R. Raskar, D. n. Colucci, K. Keller, A. State, J. R. Crawford, P. Rademacher, S. H. Drake, and A. A. Meyer, "Augmented Reality Visualization for Laparoscopic Surgery," in *MICCAI '98*. Boston: Springer, 1998.
20. B. Davies, S. Harris, M. Jakopec, and J. Cobb, "A novel hands-on robot for knee replacement surgery," in *Computer Assisted Orthopaedic Surgery USA (CAOS USA)*, A. DiGioia and B. Jaramaz, Eds. Pittsburgh: UPMC Medical Center, 1999, pp. 70-74.
21. A. Bettini, S. Lang, A. Okamura, and G. Hager, "Vision Assisted Control for Manipulation Using Virtual Fixtures: Experiments at Macro and Micro Scales," *IEEE International Conference on Robotics and Automation*, 2002.
22. M. Li and R. H. Taylor, "Performance of Teleoperated and cooperatively controlled surgical robots with automatically generated spatial virtual fixtures.," in *IEEE International Conference on Robotics and Automation*, 2005, pp. accepted.
23. E. Boctor, A. Viswanathan, M. Choti, R. Taylor, G. Fichtinger, and G. Hager, "A Novel Closed Form Solution For Ultrasound Calibration," in *International Symposium on Biomedical Imaging*. Arlington, VA, 2004, pp. 527-530.
24. A. Viswanathan, E. M. Boctor, R. H. Taylor, G. D. Hager, and G. Fichtinger, "Immediate Ultrasound Calibration with Three Poses and Minimal Image Processing," in *MICCAI*, vol. 2, C. B. D. R. H. P. Hellier, Ed. Saint-Malo, France, 2004, pp. 446-454.
25. G. Stetten and V. Chib, "Overlaying Ultrasound Images on Direct Vision," *Journal of Ultrasound in Medicine*, vol. 20(3), pp. 235-240, 2001.
26. W. F. Garrett, H. Fuchs, M. C. Whitton, and A. State, "Real-Time Incremental Visualization of Dynamic Ultrasound Volumes Using Parallel BSP Trees," in *Visualization '96*, 1996.
27. J. Zacherl, C. Scheuba, M. Imhof, M. Zacherl, F. Langle, P. Pokieser, F. Wrba, E. Wenzl, F. Muhlbacher, R. Jakesz, R. Steininger. Current value of intraoperative sonography during surgery for hepatic neoplasms. *World J Surg*. 2002 May;26(5):550-4.

Design and Control of In-Vivo Magnetic Microrobots

K. Berk Yesin, Philipp Exner, Karl Vollmers, and Bradley J. Nelson

Institute of Robotics and Intelligent Systems,
Swiss Federal Institute of Technology (ETH) Zurich, Switzerland
{byesin, kvollmers, bnelson}@ethz.ch, exnerp@student.ethz.ch

Abstract. This paper investigates fundamental design, modeling and control issues related to untethered biomedical microrobots guided inside the human body through external magnetic fields. Immediate application areas for these microrobots include cardiovascular, intraocular and inner-ear diagnosis and surgery. A prototype microrobot and steering system are introduced. Experimental results on fluid drag and magnetization properties of the robots are presented along with an analysis of required magnetic fields for application inside blood vessels and vitreous humor.

1 Introduction

The state-of-the-art in MEMS technology is progressing from individual, chip-level microsensor and microactuator devices to complete integrated microrobot systems. These types of systems will impact future minimally invasive surgical techniques by providing sub-mm untethered biomedical microrobots capable of performing a number of new procedures. The benefit will be even less injury to the patient resulting in correspondingly faster recovery times. Proven MEMS technologies such as micro-needles, micro-pumps and force and chemical sensors will be carried on-board for a variety of surgical and diagnostic tasks. Possible areas of application for these microrobots include cardiovascular, intraocular and inner-ear diagnosis and surgery.

Although electronic and mechanical systems have been miniaturized by VLSI and MEMS technologies, no counterpart to these exists for electro-chemical energy storage. Currently, the only viable option for the actuation and steering of such a microrobot is external energy transfer. The use of magnetic fields generated ex-vivo for energy transfer and actuation can provide a solution to this problem. In this paper a prototype microrobot system that was designed to investigate this idea is introduced. The nature of magnetic and viscous drag forces are discussed and experimental results on the magnetic and drag properties of the microrobots are presented. With these results, the necessary magnetic field strengths for application inside different body-fluids are analyzed. Finally, a strategy for magnetic steering is proposed and a prototype steering system is presented.

2 Microrobot with Hybrid MEMS Design

The microrobot prototype is a three dimensional structure built by microassembly of individual parts shown in Figure 1(a). Currently four different types of robots are being assembled from a combination of these parts and bonded with UV activated glue (Figure 1(c)). The "winged-ellipsoid" shape has an axis of symmetry along the long axis of the ellipsoid. An external magnetic field acts to align and pull the robot along this axis (i.e. magnetic torque and force) due to the shape anisotropy effect, much like a needle always becoming magnetized along its long axis. On the other hand, the winged shape acts to reduce the side-ways drift of the microrobot by increasing the fluid drag along the axes perpendicular to the long axis.

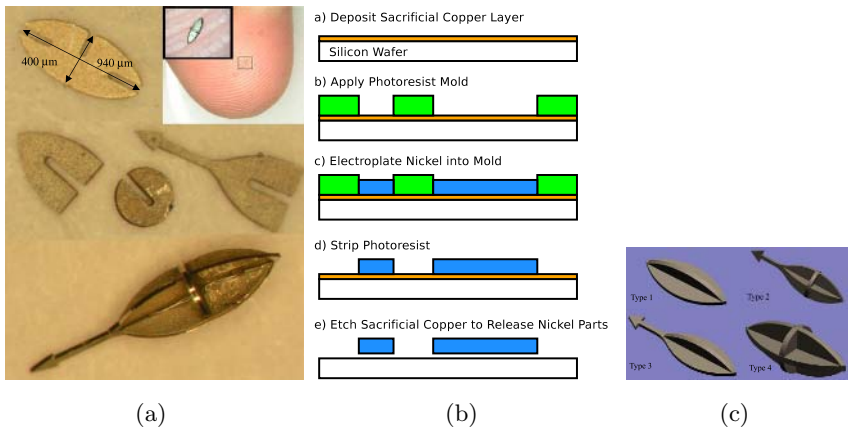


Fig. 1. a) Nickel microrobot components and an assembled microrobot. The inset shows another microrobot on a finger tip. b) Manufacturing steps for the microrobots. c) Different configurations of microrobots.

The significant advantage of the hybrid design is that the individual parts of the assembly can be produced with standard MEMS manufacturing processes which create planar geometries. This way, different sub-systems of the robot can be manufactured using the most suitable process for the purpose. Figure 1(b) shows the steps of the micromanufacturing process for the nickel parts. Initially the wafers are cleaned and metallized with a titanium adhesion layer and sacrificial copper seed for subsequent electroplating. The thickness of the seed layer is 300 nm. After metallization the wafer is covered with a 80 μm thick layer of negative photoresist that is patterned to form the microrobot parts. The nickel parts are then deposited inside a bench top plating system at the rate of 1 $\mu\text{m}/\text{min}$ to a thickness of 50 μm . Following the plating step, the photoresist is stripped and the copper seed layer is etched in a solution of ammonium persulfate which does not attack nickel at low concentrations.

3 Wireless Actuation Through Ex-Vivo Magnetic Fields

3.1 Nature of Magnetic Forces

The primary vectors that define the magnetostatic field in magnetized matter are (external) magnetic field strength, \mathbf{H} (A/m), magnetization of the matter \mathbf{M} (A/m) and magnetic flux density \mathbf{B} (Tesla). The relationship between these vectors is

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}) \quad (1)$$

where μ_0 is the magnetic permeability of free space defined as $4\pi \times 10^{-7}$ Tm/A. For the idealized case of linear, isotropic and homogeneous media the following relationships simplify (1) as

$$\mathbf{M} = \chi\mathbf{H} \quad (2)$$

$$\mathbf{B} = \mu_0(1 + \chi)\mathbf{H} = \mu_0\mu_r\mathbf{H} \quad (3)$$

where χ and μ_r are the susceptibility and relative permeability of the media, respectively. In general, these values are not constant but change with magnetization, approaching zero as the magnetization reaches a material dependent limit called the *saturation magnetization* M_s . Within the saturation limits, the permeability can be thought as an amplification factor that creates a net magnetic field inside the matter through an external field.

The magnetic force and torque that are exerted on an object with uniform magnetization \mathbf{M} in a magnetic field with flux density \mathbf{B} are defined as

$$\mathbf{F}_m = V_m(\mathbf{M} \cdot \nabla)\mathbf{B} \quad (4)$$

$$\boldsymbol{\tau}_m = V_m\mathbf{M} \times \mathbf{B} \quad (5)$$

where V_m is the volume of the magnetized object. Notice that the magnetic torque is dependent on \mathbf{B} whereas the magnetic force is dependent on the gradient of \mathbf{B} . Equations (4) and (5) indicate that magnetic forces are volumetric. Therefore, the required fields and field gradients to exert a certain torque and force on a magnetized object increase rapidly as the object gets smaller.

3.2 Magnetic Properties of the Microrobot

The relationship between the external magnetic field and the resulting magnetization of the robot is dependent on geometry and material properties. However, the magnetic properties of the material are also significantly affected by the manufacturing process, which is difficult to model. Therefore, it was necessary to perform experimental evaluation of the magnetization curve for the microrobots in order to obtain more reliable values. The microrobots were tested using a coercivity spectrometer device. The results of these tests covered the full range of magnetization as shown in Figure 2. The robots reached a saturation magnetization of 5 to 8×10^5 A/m with a 0.2 T external field.

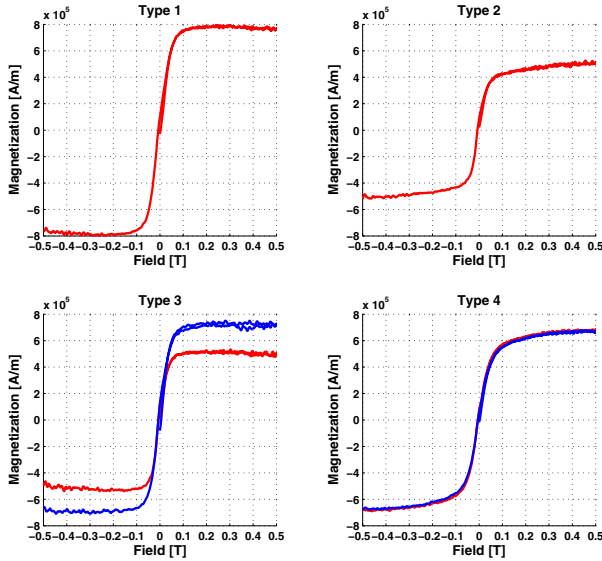


Fig. 2. Magnetization curves for the different types of microrobots (Figure 1(c)). Types 3 and 4 were tested with two different robots.

3.3 Fluid Drag Forces on the Microrobot

Other than the magnetic forces analyzed above, gravitation/bouyancy and viscous fluid drag forces are also acting on the microrobot. The drag force on a spherical object in laminar flow regime (i.e. $Re < 1.0$) can be expressed as

$$F_d = 3\pi D\mu_f\nu \tag{6}$$

where D is the diameter of the sphere, μ_f is the dynamic viscosity of the fluid and ν is the relative velocity of the robot with respect to the fluid media. Equation (6), also known as Stoke’s Drag, is frequently used for calculating the drag force on magnetic beads and other small particles, approximating their shape as a sphere [1]. In addition to the drag force, a net buoyancy force will be acting on the robot

$$F_b = V_r(\rho_r - \rho_f)g \tag{7}$$

where V_r and ρ_r are the density and volume of the robot respectively and g is the gravitational acceleration. The most important outcome of (4), (5), (6) and (7) is that, whereas magnetic and buoyancy forces are volumetric, the fluid drag forces are linearly dependent on the size. For this reason, as the size of the robot gets smaller the required field gradient to move at a particular speed rapidly increases.

3.4 Viscous Drag Force Experiments

We performed experiments in order to quantify the drag forces on the winged-ellipsoid shape of the microrobots. The experiments were conducted by releasing

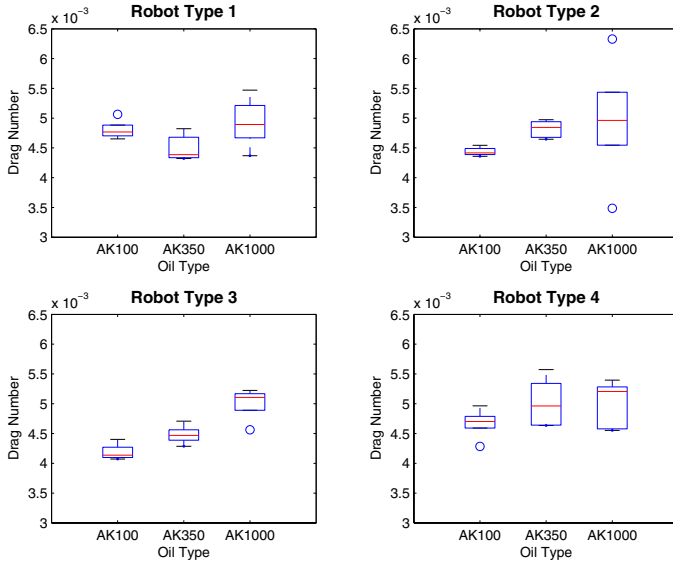


Fig. 3. Drag numbers calculated for each robot type for oils of three different viscosities

the microrobots inside a chamber filled with a fluid of known viscosity and observing them with a microscope-camera in order to determine their terminal velocity under the net effect of fluid drag, gravity and buoyancy forces. The experiments were performed using three different silicon oils AK100, AK350 and AK1000, each having room temperature dynamic viscosities of 100, 350 and 1000 cP (centiPoise), respectively. The actual viscosity values were corrected for the measured temperature of the chamber. The experimental setup was verified using steel ball-bearing spheres of known size and weight. The measurements for the steel spheres were within 5% of the calculated values.

The results of experiments indicated a linear dependency of the drag force on velocity, similar to the case of a sphere. A drag number, D_N was calculated as $D_N = F_d/\mu_f\nu$ for each of the four types of the robot. The resulting data are shown in Figure 3. The boxplots present the distribution of the repeated measurements. The height of the box indicates the interquartile range between 25% and 75% percentiles of the data, while the line inside the box shows the median. The drag number measurements show a variation of about 15%, clustered around a value of 4.5×10^{-3} m. If this value is converted to an effective sphere diameter through the relation $D_N = 3\pi D$, we find a diameter of $477 \mu\text{m}$ which is close to the cross-section diameter ($400 \mu\text{m}$) of the ellipsoid shape.

3.5 Required Magnetic Fields in Body Fluids

With the analysis of magnetic and viscous drag properties of the microrobots, the necessary magnetic field strengths for medical applications can be estimated. For

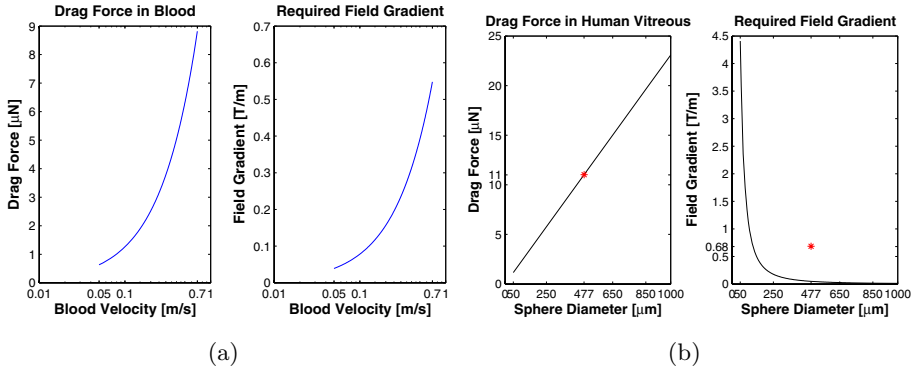


Fig. 4. a) Drag forces on the microrobot in human blood and the required field gradient to resist the fluid drag. b) Drag forces on a spherical magnet moving at 0.5 mm/s in the human vitreous and the required field gradient for various sphere diameters. Magnetization is $M = 5 \times 10^5$ A/m and viscosity is 4900 cP. The values for the prototype microrobot are shown with *.

this analysis the microrobots will be assumed to be magnetized to saturation at 5×10^5 A/m. The volume of a Type 3 robot will be considered ($3.22 \times 10^{-11} \text{m}^3$). For fluid drag, the average drag number of 4.5×10^{-3} m is used as an approximation.

For cardiovascular applications, the fluid media that the microrobot moves in is blood. The flow velocity of blood in the human body ranges between 0.05 m/s in the capillaries to 0.7 m/s at the exit of the aorta with a viscosity of about 2.8 cP [2]. Figure 4(a) shows the drag forces on the microrobot and the required field gradient to balance the robot against the flow.

The human eye is another challenging workspace for a biomedical microrobot. The vitreous humor fills the posterior cavity of the eye between the lens and the retina and has the properties of a viscoelastic liquid with high viscosity [3]. A complex model of the vitreous with spring and dashpot elements representing the elastic and viscous characteristics of the material was developed [4][5]. The viscosity of the Maxwell-dashpot element of this model which represents the irreversible flow of the material under constant stress is the main source of the viscous drag forces on the robot. The mean values of viscosity were 1398 cP at the anterior, 2179 cP at the central and 4862 cP at the posterior regions. Figure 4(b) shows the drag forces on a spherical shaped robot with a diameter between 50 to 1000 μm and magnetization 5×10^5 A/m moving inside the vitreous at a speed of 0.5 mm/sec. The viscosity at the posterior region is used. In addition, the necessary field gradient to balance the drag forces is also shown. The location of the prototype microrobot with the winged-ellipsoid shape is also indicated.

4 Magnetic Steering System

Equations (4) and (5) suggest that controlled external magnetic fields can be used to induce torques and forces on a magnetized object and control its orien-

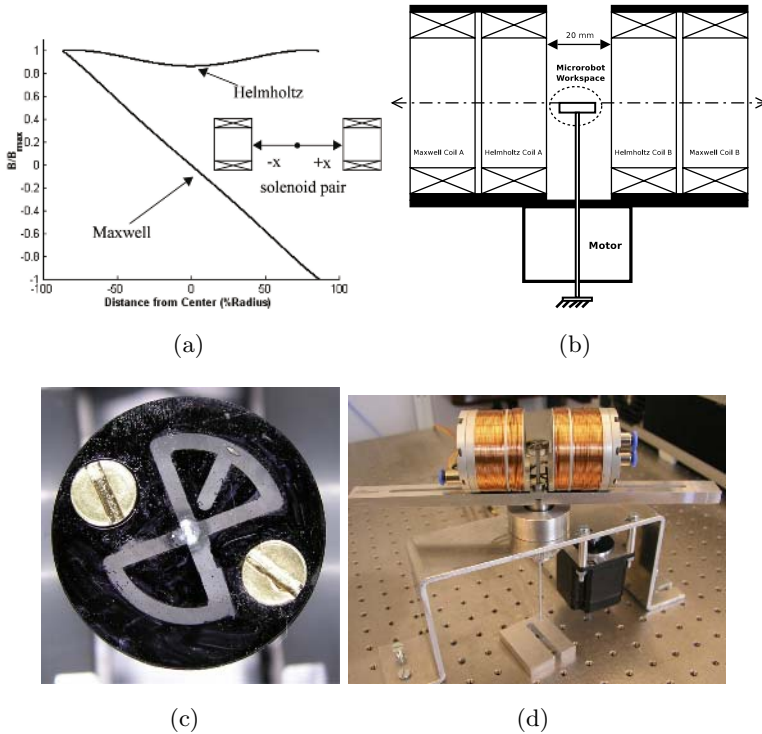


Fig. 5. a) Normalized magnetic field along the central axes of Helmholtz and Maxwell coils. b) Steering system with co-axial Maxwell and Helmholtz coils. A maze structure is fixed at the center of the workspace between the rotating coils. c) Microrobot steering maze. The channels are $1000\ \mu\text{m}$ wide and $300\ \mu\text{m}$ deep. The microrobot is also visible at the upper part. d) Magnetic steering system.

tation and position. An important issue related to control is the highly nonlinear nature of the field and gradients that are created by electromagnet coils. One way of reducing the effect of such nonlinearities is to create uniform magnetic fields and field gradients using various coil configurations [6]. For example, the Helmholtz coil configuration consists of two identical coils that are placed on the same axis and separated by a distance equal to the radius of the coils. This arrangement generates a uniform field close to the center of the coil pair when current passes in the same direction in both coils. A variation of this configuration with opposing direction of currents is the Maxwell coil which generates a uniform gradient. Figure 5(a) shows the plot of the fields of Helmholtz and Maxwell coils. Both of these coils are commonly used in MRI systems. Although, the uniform fields of commercial MRI machines are quite high (0.5-2.0 T), the gradient fields they can generate continuously are in the 0.01 to 0.05 T/m range.

The magnetic steering principle was investigated using a coil system as illustrated in Figure 5(b). The magnetic torque and force on the microrobot were controlled independently through the Maxwell and Helmholtz fields, respectively.

This way the orientation and direction of motion of the microrobot could be commanded on a two dimensional plane. Steering tests were performed in a water filled, maze-like structure etched into a plastic substrate (Figure 5(c)). The maze was fixed at the center of the rotating coils, and a microscope camera system was used to obtain a top view (Fig. 5(c)). The orientation of the coils were controlled through a PC interface. Numerous trials with the system confirmed that the independent orientation/thrust control principle was successful. The microrobot was observed to follow the orientation of the coils without translating unless the Maxwell coils are energized.

5 Conclusions

This paper investigated the concept of steering sub-mm sized untethered microrobots inside body fluids with external magnetic fields. A prototype microrobot was introduced. Estimations on the necessary field gradient to resist fluid drag forces in different body fluids showed that a minimum gradient of about 0.7 T/m would be necessary. A magnetic steering principle based on independent control of orientation and thrust with constant field and field-gradient generating coils was successfully tested. These results confirm that magnetic actuation is a suitable mode of energy transfer for untethered biomedical microrobots.

References

1. D.L. Holligan, G.T. Gilles, J.P. Dailey, "Magnetic Guidance of Ferrofluidic Nanoparticles in an In Vitro Model of Intraocular Retinal Repair", *IOP Nanotechnology*, vol. 14, pp. 661–666, 2003.
2. A. Despopoulos, S. Silbernagl, *Color Atlas of Physiology*, New York: Thieme Medical Publishers, pp. 156, 1991.
3. T.V. Chirilla, Y. Young, "The Vitreous Humor", *Handbook of Biomaterial Properties*, J. Black and G. Hastings et al., Chapman & Hall, London, 1998.
4. B. Lee, M. Litt, G. Buchsbaum, "Rheology of the Vitreous Body. Part I: Viscoelasticity of Human Vitreous", *Biorheology*, vol. 29, pp. 521–533, 1992.
5. B. Lee, *Comparative Rheological Studies of the Vitreous Body of the Eye*, Ph.D. Thesis, University of Pennsylvania, 1992.
6. J. Jin, *Electromagnetic Analysis and Design in Magnetic Resonance Imaging*, Florida: CRC Press, 1999.

3D Needle-Tissue Interaction Simulation for Prostate Brachytherapy

Orcun Goksel¹, Septimiu E. Salcudean¹, Simon P. DiMaio²,
Robert Rohling¹, and James Morris³

¹ Department of Electrical and Computer Engineering,
University of British Columbia, Vancouver, Canada
{orcung, tims, rohling}@ece.ubc.ca

² Surgical Planning Laboratory, Department of Radiology,
Brigham and Women's Hospital, Boston, USA

³ Vancouver Center, British Columbia Cancer Agency, Vancouver, Canada

Abstract. This paper presents a needle-tissue interaction model that is a 3D extension of a prior work based on the finite element method. The model is also adapted to accommodate arbitrary meshes so that the anatomy can effectively be meshed using third-party algorithms. Using this model a prostate brachytherapy simulator is designed to help medical residents acquire needle steering skills. This simulation uses a prostate mesh generated from clinical data segmented as contours on parallel slices. Node repositioning and addition, which are methods for achieving needle-tissue coupling, are discussed. In order to achieve real-time haptic rates, computational approaches to these methods are compared. Specifically, the benefit of using the Woodbury formula (matrix inversion lemma) is studied. Our simulation of needle insertion into a prostate is shown to run faster than 1 kHz.

1 Introduction

Prostate cancer is the most common cancer in the US with 232,090 new cases and 30,350 deaths estimated in 2005. Brachytherapy is the single most commonly used curative treatment for prostate cancer in North America. It is a conformal radiation therapy method that involves the percutaneous insertion of radioactive sources directly in or adjacent to malignant tissues. A conceptual overview of the procedure can be seen in Fig. 1.

Despite the low risk of brachytherapy, seed placement errors are still common. In [1], an experienced physician implanting seeds in 20 patients achieved an average placement error of 0.63 cm, which is a substantial error of 21% of average prostate diameter (3 cm). Errors decrease the effectiveness of the treatment and may also lead to subsequent complications such as urinary incontinence and impotence. Some procedural challenges of brachytherapy are needle steering and visualizing the 3D volume from 2D medical images. The required skill set is acquired by medical residents using mannequins, animals, or cadavers, which are either not very realistic or may involve ethical issues. Residents also practice

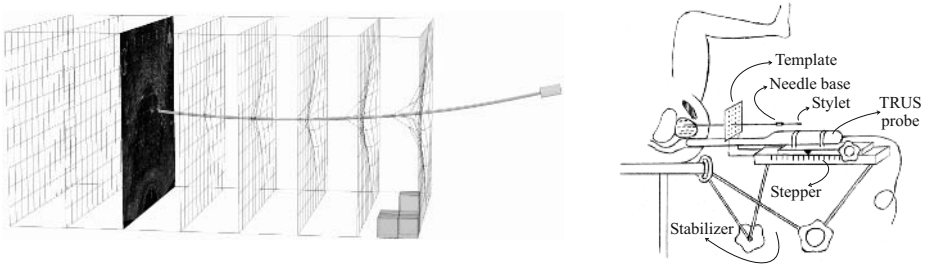


Fig. 1. Representation of 3D prostate volume deformation with a needle inserted (left) and conceptual overview of the procedure (right)

on patients which may lead to possibly irreversible consequences. The identification of this need for better training schemes led us to propose a prostate brachytherapy simulator.

The simulation of needle insertion differs from the simulation of other medical tool-tissue interactions in several ways: the needle does not manipulate only the organ surface; friction is a significant force during its interaction; and it is flexible unlike many rigid surgical tools. These issues were addressed in the 2D case with a finite element model (FEM) using the condensation approach, a stick-slip friction model, and local frame rotations with low-rank matrix updates, respectively [2]. Fluid pockets were then introduced onto this system [3].

Online mesh modification is a common challenge of real-time implementations (e.g., incision in [4]). To retain tissue mesh nodes on the needle, element subdivision and node relocation techniques were implemented in [2, 4, 5]. Specifically, the 2D haptic implementation in [2] simply constrains the closest tissue node onto the needle without any processing of the nominal mesh which locally deforms the tissue unrealistically. This work was also the first to model a decoupled tip cutting force and a velocity-dependent shaft friction force.

There are models [6] and graphical interfaces [7] designed for prostate brachytherapy. However, the only available physically-based brachytherapy model [5] is designed for offline planning simulation and it runs in 2D neglecting spatial deformations such as transversal rotation. In contrast, our proposed model runs at haptic rates while simulating needle insertion in 3D. Extension of algorithms from 2D to 3D is not a trivial task. It also increases the computational complexity significantly. Control of such a 3D simulation at haptic rates require additional computational treatment on the algorithms, some of which are presented in this paper. An application on a clinical case such as brachytherapy needs to also address issues related to optimal discretization of the tissue for the model.

2 Methods

In the simulation, while the needle is inserted, it interacts with the tissue through tissue mesh nodes that are confined onto the needle shaft. These nodes are called

contact nodes. When the needle base is manipulated by the user, the needle shaft poses a displacement boundary condition for the tissue mesh via contact nodes. The reaction forces of the tissue are then applied back on the needle model to solve for needle bending and to find feedback forces onto the user's hand. The following steps were taken to implement this system in 3D.

2.1 Mesh Generation

Considering FEM is an approximation technique for elastic deformation, its approximation error largely depends on the element discretization geometry. This can be achieved simply by using more elements. However, the speed of many mesh-processing algorithms is significantly affected by the number of nodes. These pose the two basic constraints: a high quality mesh with few nodes. A case with multiple sub-regions with different physical properties, in addition, necessitates a *conforming mesh* having element surfaces as boundaries between separate anatomical structures. Considering that the segmentation by the physician is simply in the form of parallel contours, generating a prostate region mesh obeying the above criteria requires additional steps, since the conforming meshing algorithms work on a boundary surface definition. Thus, the following steps are taken on the segmentation of the parallel images from a prostate volume study: first the extraneous nodes are removed, then Nuages [8] is used to obtain the surface definition, and finally the interior (with finer elements) and the exterior (with coarser elements) of the prostate are meshed (see Fig. 2) using the academic version of GiD [9] which utilizes the advancing fronts technique.

The needle bending is modeled using FEM with quadratic strain. Its deflection is solved using the iterative approach of [2]. The deflection response of this model to an applied force was observed to be substantially mesh topology dependent in 3D. Therefore, the symmetric needle segment in Fig. 2 was designed. A needle model constructed from segments similar to the one shown demonstrates axial symmetry for a change of lateral force direction. On the tissue interaction side, this mesh model is assumed to be a set of line segments passing through its center. The needle-tissue coupling is damped in order to ensure stability.

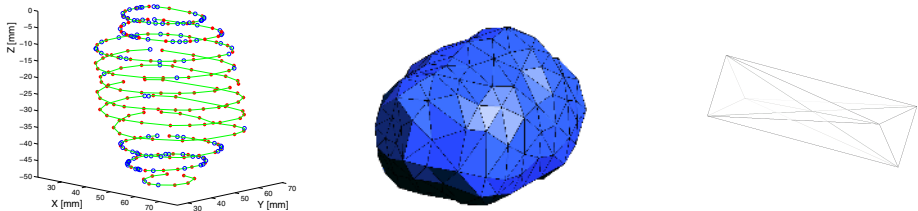


Fig. 2. Prostate segmentation data where the lines denote contours, stars are culled nodes and the circles are the removed ones (left); prostate surface definition (center); and our symmetric design of a needle segment (right)

2.2 Implementation

As opposed to the surrounding relatively soft tissue, a stiffer region connects the prostate to the pubic arch from its anterior inferior end. This creates a torquing movement of the prostate as the needles are inserted. To simulate this effect, the prostate is spatially fixed around this region by setting fundamental displacement constraints. Then the tissue stiffness matrix \mathbf{K} is inverted offline before the simulation starts. A linear quasi-static tissue model is used as an approximation to the deformation characteristics of soft tissue. An overview of the system is given in Fig. 3.

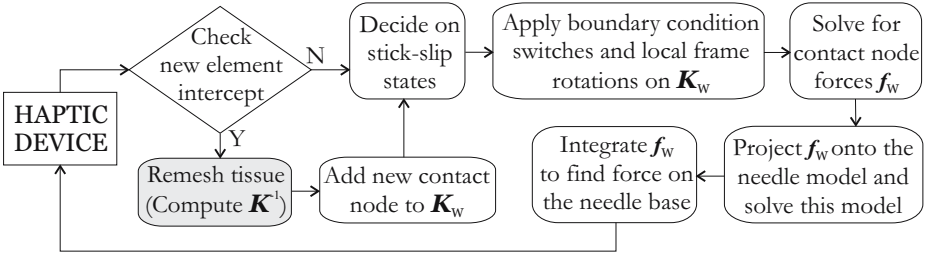


Fig. 3. Flowchart of an iteration during the needle insertion into the tissue

Considering that the needle boundary conditions are applied only on the contact nodes and the feedback forces for the needle are only required on those nodes, a $3w \times 3w$ condensed system can be employed with the equation:

$$\mathbf{x} = \mathbf{K}_W \mathbf{y} \tag{1}$$

where \mathbf{K}_W is the condensed stiffness matrix and w is the number of contact nodes. This condensation approach and the stick-slip friction model applied on it are extended to 3D from the work of [2]. A switch in the stick-slip state of a node is achieved by changing its axial boundary constraint. A force-displacement boundary constraint switch (BCS) on row/column i is in the form of the following low-rank update:

$$\mathbf{K}_W' = \mathbf{K}_W - \frac{1}{k(i, i)} \mathbf{c}_i \mathbf{r}_i \tag{2}$$

where \mathbf{c}_i and \mathbf{r}_i are the corresponding i^{th} column and i^{th} row of \mathbf{K}_W , and $k(i, i)$ is the pivot i . Note that this requires $\mathcal{O}(w^2)$ operations. With the changing orientations of the needle segments caused by either user manipulation or needle shaft bending, \mathbf{K}_W needs to be updated to accommodate the new local frame orientations. This local frame rotation of a contact node differs depending on its stick-slip state. A treatment in 3D is developed based on the 2D case in [2].

The tissue reaction forces of the contact nodes are linearly distributed onto the closest needle mesh nodes to be applied onto the needle model and also to determine force/torque feedback to the user’s hand.

2.3 Tissue Remeshing

Although it is possible to maintain a non-conforming needle-tissue coupling using force/displacement interpolations, in our insertion simulation, preserving this conformity was preferred due to speed considerations. It is achieved by placing a tissue node on the needle tip whenever it encounters a new tissue element. Generation of this node without distorting the tissue requires remeshing (the shaded step in Fig. 3). Two remeshing techniques, node repositioning and node addition, are shown in Fig. 4. Repositioning is equivalent to removing the surrounding elements of P and introducing the elements surrounding the new node P' . In contrast, addition in 2D is removing only two elements and introducing four smaller ones (see Fig. 4(d)).

The 3D simulation uses the inverse of $3n \times 3n$ tissue stiffness matrix \mathbf{K} where n is the number of tissue mesh nodes. Considering the dimensions of \mathbf{K}^{-1} , a real-time full-rank inversion is impractical. Thus, two methods of fast \mathbf{K}^{-1} manipulation are implemented and compared. The first one employs boundary condition switches whereas the second one is a one-step algebraic solution. Nonetheless, both approaches require the differential $3m \times 3m$ stiffness matrix $\Delta\mathbf{K}$ which is the desired change in \mathbf{K} to obtain the remeshed stiffness matrix \mathbf{K}' where m is the total number of nodes involved. $\Delta\mathbf{K}$ can easily be found by subtracting the stiffness matrices of the elements to be removed from the ones to be added. Note that for the 2D cases in figures 4(c) and (d), m is 7 and 5, respectively.

The first method performs boundary condition switches on \mathbf{K}^{-1} for all three axes of each node involved using (2). It then adds $\Delta\mathbf{K}$ and switches the boundary conditions back. This results in $\approx 2(3m)(3n)^2$ operations. Throughout this paper, *operations* refer to floating-point multiplications and divisions only. On the other hand, the second method employs the Woodbury formula as follows:

$$(\mathbf{K} + V\Delta\mathbf{K}V^T)^{-1} = \mathbf{K}^{-1} - \left[\mathbf{K}^{-1}V\Delta\mathbf{K} (I + V^T\mathbf{K}^{-1}V\Delta\mathbf{K})^{-1} V^T\mathbf{K}^{-1} \right] \quad (3)$$

where V is a $3n \times 3m$ zero matrix with $3m$ corresponding rows taken from identity matrix $I_{3m \times 3m}$. A fast implementation of (3) with the low-rank inverse computed using Gaussian elimination costs $\approx 3m(3n)^2 + (3m)^2 3n + 2(3m)^3$ operations.

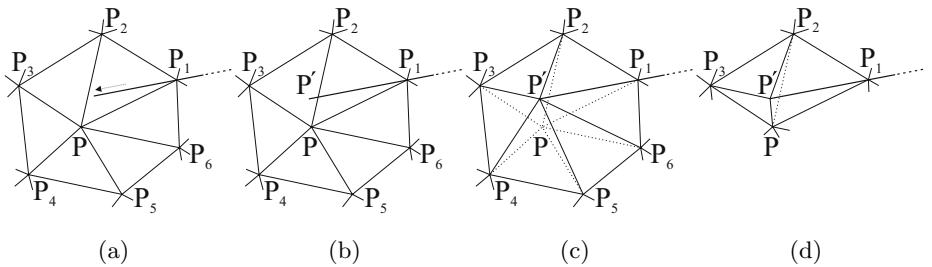


Fig. 4. The simulation steps (a) before and (b) after a new element penetration; and (c)node repositioning and (d) node addition methods for preserving mesh conformity

Although n is the dominant factor, the speed of both approaches (BCS and Woodbury) is also dependent on the dimension of ΔK . For node repositioning, this dimension is determined by node valence which causes a relatively long and unpredictable computation time. Furthermore, repositioning an organ boundary node may pose the issue of distorting the anatomy. In contrast, 3D node addition runs at constant time with a guaranteed ΔK size of 18×18 . The extra rows/columns needed in \mathbf{K}^{-1} by the Woodbury formula for the addition of P' are initially set to zero except for their diagonal elements having very large values to represent infinity. The condition number of the low-rank matrix inverted in (3) is also considered while determining these large numbers. Everytime the needle is completely out of the tissue, \mathbf{K}^{-1} is reset to its initial dimensions and values in order not to crowd the model with nodes after consecutive insertions.

3 Results and Discussion

The two computational approaches, BCS and Woodbury, were implemented on a 2D tissue (see Fig. 5) in Matlab[®] using a P4 3.2 GHz system with 1 MB L2 cache. Although both techniques theoretically have the same order of complexity (for $n \gg m$), it is observed that the Woodbury formula update runs significantly faster for large \mathbf{K}^{-1} (see Fig. 5). This is due to the efficient cache use of the Woodbury algorithm as described below. Each BCS needs to process the *entire* $3n \times 3n$ matrix before the next BCS can start. This causes a *cache miss* for each of the $9n^2$ numbers at every BCS operation that happens $6m$ times in total. Each miss entails a RAM access which is much slower than a cache access. On the other hand, the full-rank multiplication in the Woodbury formula allows for an implementation letting its $6m$ operations be processed in an inner loop. So, as long as the matrix V fits in cache memory, much fewer cache misses occur resulting in significant speed gain as shown in Fig. 5. Here, Matlab's *inv* function running time for rank- $3n$ matrix inversion which has $\mathcal{O}(n^3)$ operations is also presented for comparison.

The tissue mesh that is used in the simulation consists of 570 nodes and 2801 tetrahedra. The shortest edge is 3.06 mm and the longest one is 24.04 mm. The minimum and the maximum dihedral angles are 9.94° and 159.49° , respectively.

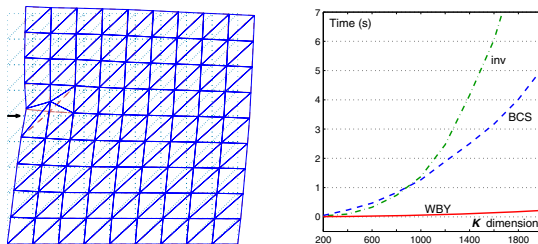


Fig. 5. 2D tissue with a \mathbf{K} of dimension 200 (left) and the comparison of mesh modification techniques while increasing the dimension of \mathbf{K} (right)

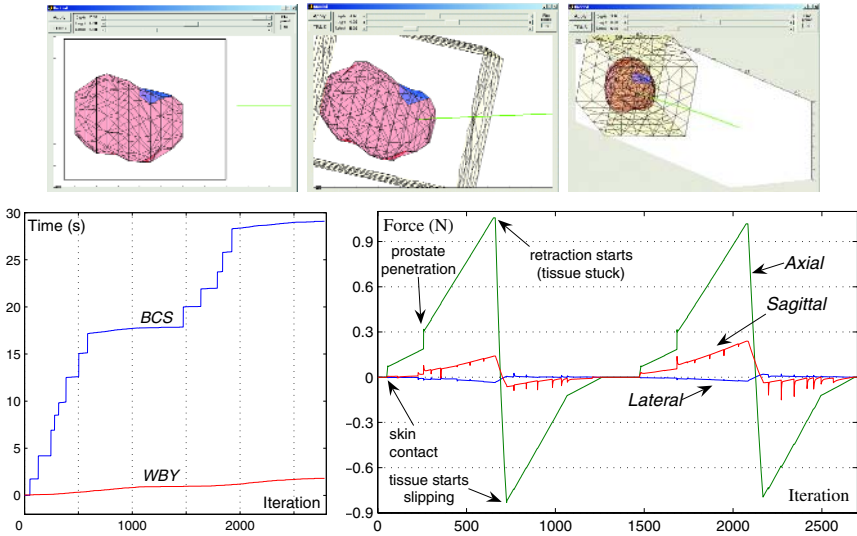


Fig. 6. Interface to the prostate brachytherapy needle insertion simulation (top), cumulative running time since simulation start for two separate instances of needle insertion (lower left), and force on the needle base for Woodbury algorithm (lower right)

Furthermore, only a couple elements have these border-line values showing that the geometry of elements are acceptable for a reasonable FEM approximation.

The 3D needle insertion simulation was implemented in C for the ease of future use with VxWorks[®] which is a real-time operating system used in controlling haptic devices. A Matlab interface, where the needle can be manipulated and the tissue deformation can be observed, was built for testing and debugging purposes (see Fig. 6). The three axis components of the needle base force profile during two insertions having different trajectories can be seen in Fig. 6.

Our first implementation of this simulation used BCS for node repositioning where some iterations took as long as 2.75s. For an haptic implementation any iteration has to complete in less than 2ms. That is why the simulation is reimplemented using the Woodbury formula which led us to less than 1ms iterations. For the two consecutive insertions that the force profile is given above, the cumulative time elapsed using each method is plotted in Fig. 6. The big steps of BCS occur when node repositioning is performed. Since no repositioning is needed during retraction, the cumulative time increases only slightly due to other simulation computations between iterations 650 and 1300.

4 Conclusions and Future Work

A prostate brachytherapy simulator design has been presented using a prostate mesh with conforming anatomical boundaries on surfaces generated from contours on parallel slices. The proposed system successfully couples a conform-

ing quasi-static linear FEM of the tissue with a flexible needle. It is the first physically-based 3D interaction model of a flexible needle insertion into a soft deformable body. The first physically-based 3D simulator of a brachytherapy procedure is programmed with a Matlab interface. The real-time portion of the code is ready to be implemented on a haptic setup. The proposed simulator can also be used for the prediction of prostate motion in brachytherapy planning.

Next, other anatomical structures will be segmented and implemented in the simulator. Particularly, the pubic arch is needed as a fundamental displacement boundary condition. The bladder is also a significant adjacent anatomy to be modeled. Given more anatomical information incorporated into the model, validation of the system will become the next step. TRUS data of the entire procedures were already recorded for this purpose. The prostate shift caused by the needle insertion can easily be seen on those. This observed deformation will be compared with our model for validation when we incorporate at least the pubic arch into our model. Ultimately, the simulation is going to be implemented on a suitable haptic interface. Considering brachytherapy planning, real-time acquisition of the tissue elasticity parameters using elastography which is another future research subject would make accurate intra-op plan adjustments possible when combined with the simulator presented in this paper.

References

1. Taschereau, R., Pouliot, J., Roy, J., Tremblay, D.: Seed misplacement and stabilizing needles in transperineal permanent prostate implants. *Radiotherapy and Oncology* **55** (2000) 59–63
2. DiMaio, S.P., Salcudean, S.E.: Needle insertion modeling and simulation. *IEEE Transactions on Robotics and Automation* **19** (2003) 864–875
3. Gosline, A.H., Salcudean, S.E., Yan, J.: Haptic simulation of linear elastic media with fluid pockets. In: *Proceedings of the 12th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS)*. (2004) 266–271
4. Nienhuys, H.: Cutting in deformable objects. PhD thesis, Utrecht Univ. (2003)
5. Alterovitz, R., Goldberg, K., Pouliot, J., Taschereau, R., Hsu, I.C.: Sensorless planning for medical needle insertion procedures. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems, Las Vegas, NV* (2003)
6. Wang, X., Fenster, A.: A haptic-enhanced 3D real-time interactive needle insertion simulation for prostate brachytherapy. In: *Proceedings of SPIE Medical Imaging 2004: Vis., Image-Guided Procedures, and Display*. Volume 5367. (2004) 781–789
7. Kimura, A.N., Camp, J., Robb, R., Davis, B.: A prostate brachytherapy training rehearsal system – simulation of deformable needle insertion. In: *Proceedings of MICCAI*. (2002) 264–271
8. Geiger, B.: Three-dimensional modeling of human organs and its application to diagnosis and surgical planning. PhD thesis, École Nationale Supérieure des Mines de Paris, France (1993)
9. Löhner, R.: Progress in grid generation via the advancing front technique. *Engineering with Computers* **12** (1996) 186–210

Development and Application of Functional Databases for Planning Deep-Brain Neurosurgical Procedures

Ting Guo^{1,3}, Kirk W. Finnis⁴, Andrew G. Parrent², and Terry M. Peters^{1,3}

¹ Robarts Research Institute and University of Western Ontario

² The London Health Sciences Centre, Department of Neurosurgery
London, Ontario, Canada N6A 5K8

³ Biomedical Engineering Graduate Program, University of Western Ontario,
London, Ontario, Canada N6A 5B9

⁴ Atamai Inc, London, Ontario, Canada N6B 2R4
{tguo, tpeters}@imaging.robarts.ca

Abstract. This work presents the development and application of a visualization and navigation system for planning deep-brain neurosurgeries. This system, which incorporates a digitized and segmented brain atlas, an electrophysiological database, and collections of final surgical targets of previous patients, provides assistance for non-rigid registration, navigation, and reconstruction of clinical image data. The fusion of standardized anatomical and functional data, once registered to individual patient images, facilitates the delineation of surgical targets. Our preliminary studies compared the target locations identified by a non-expert using this system with those located by an experienced neurosurgeon using regular technique on 8 patients who had undergone subthalamic nucleus (STN) deep-brain stimulations (DBS). The average displacement between the surgical target locations in both groups was $0.58\text{mm}\pm 0.49\text{mm}$, $0.70\text{mm}\pm 0.37\text{mm}$, and $0.69\text{mm}\pm 0.34\text{mm}$ in x, y, and z directions respectively, indicating the capability of accurate surgical target initiation of our system, which has also shown promise in planning and guidance for other stereotactic deep-brain neurosurgical procedures.

1 Introduction

Surgical treatments for movement disorders, such as Parkinson's disease and essential tremor, are performed by either creating a small lesion, or placing electrical stimulators at precise locations deep within the brain, using minimally-invasive stereotactic techniques. However localizing the surgical target is challenging due to the incompleteness of the information provided by the regular pre-operative medical images, where neither the motor nuclei of the thalamus, the internal segment of the globus pallidus (GPi), nor the subthalamic nucleus (STN) (the targets for the surgical treatments of Parkinson's disease) can be visualized directly. In clinical practice, additional information from anatomic atlases [1,2] is needed to enhance the accuracy and precision of targeting. Moreover, computerized atlases [3-5] have been implemented to overcome the inherent disadvantages of the printed versions. Digitized atlases can be aligned and fused with individual pre-operative brain images

to facilitate the identification of the surgical targets. Nevertheless existing anatomical atlases should be employed conservatively for planning stereotactic procedures because of their limitations, such as lack of morphometric information, poor generalization, and insufficient statistical representation of the population.

In addition to the anatomical information derived from pre-operative images and brain atlases, functional information obtained from intra-operative electrophysiological measurements is also required to refine the optimal surgical targets, characterize tissue function, and map somatotopy. Accurate pre-operative surgical target planning can reduce the need for invasive exploration and decrease procedure-related complications. In this case, additional standardized electrophysiological information available prior to surgery may assist the surgical target determination. Previously, researchers have developed electrophysiological atlases [6,8,9] and databases [7] containing data from a series of patients, standardized and normalized to a standard brain template to establish the relationship between functional brain organization and anatomic structures, and to estimate the surgical targets [6-9].

This paper expands that of Finnis *et al.* [7] by adding data of 43 new patients (for a total of 131) and incorporates the database into a comprehensive neurosurgical system. Our work focuses on the integration of a 3D visualization and navigation system for stereotactic functional deep-brain neurosurgery planning and guidance in order to improve target localization accuracy and to minimize electrophysiological exploration and patient trauma. We describe preliminary studies evaluating the effectiveness of this system in surgical targeting for STN deep-brain stimulation (DBS) procedures. This system integrates the electrophysiological database, digitized 3D brain atlases [1], segmented deep-brain nuclei, and surgical targets from previous procedures, along with representations of surgical instruments, into the visualization and navigation system. All the standardized functional and anatomical data in this system, once non-rigidly mapped to a patient brain space, can be an important adjunct for pre-operative surgical target planning and intra-operative surgical guidance.

2 Materials and Methods

2.1 Image Registration

Two registration steps are necessary to implement this procedure. The first rigid-body registration step establishes the transformation between a patient image-space and the stereotactic frame, whereas the second performs the non-rigid mapping of the functional data from each patient brain space to the standard database, and vice versa.

Frame-to-Image: An automatic fiducial localization algorithm registers the image volume to the frame in 1.5s by extracting the ‘Z-bar’ fiducial points from the images with a fiducial localization error of approximate 0.5mm for both MR and CT.

Data-to-Database & Database-to-Patient: The rapid 3D non-rigid registration algorithm [10] has been adopted to accommodate the intersubject variability between each patient anatomical brain image and the standard brain template. This algorithm employs a completely unsupervised multi-resolution approach that executes 8 to 12 minutes on a dual PIII 933 MHz machine with an average registration error of 1.04 ± 0.65 mm [10].

2.2 Functional Database Construction

Subjects: 131 patients who had undergone a total of 161 surgeries for symptomatic treatment of Parkinson's disease, chronic pain, and essential tremor at London Health Sciences Centre (LHSC), London, Ontario, Canada, have been recruited, for the functional database construction.

MRI data: The pre-operative MRI images of the patients were acquired with a 1.5T GE Signa scanner using a 3D SPGR sequence (TR/TE 8.9/1.9ms, flip angle 20°, NEX 2, voxel size 1.17mm×1.17mm×1mm, in-slice resolution 256×256).

Standard brain template: The CJH-27 dataset [11] was adopted as the standard brain template (the common coordinate system) of the functional database. CJH-27 consists of 27 registered T₁-weighted MRI scans (20 × 1mm³: TR/TE 18/10ms, flip angle 30°, NEX 1; 7 × 0.78mm³: TR/TE 20/12ms, flip angle 40°, NEX 1) of the same healthy individual averaged into a single volume.

Functional data collection: First of all, the pre-operative brain image of each patient was non-rigidly registered to the CJH-27 template, to establish the 3D transformation and the deformation grid. Then the micro-recording, electrical-stimulation data were coded using a comprehensive coding scheme and intra-operatively entered into single patient brain image space. Finally, the functional data in each patient image space were non-rigidly mapped to the standard brain coordinate according to each 3D non-rigid transformation. The functional data in the population-based database can be applied to the individual brain image using the inverse of the original non-rigid patient-to-database transform.

2.3 Visualization and Navigation System Integration

Electrophysiological Database: Our current electrophysiological database is an expanding version of that reported previously [7]. The functional data in this study were obtained during the procedures performed on the patients, and those relating to a particular firing pattern, a specific body reaction, and certain anatomical regions, can be retrieved and displayed as clusters of spheres in 3D space or density maps on three intersecting orthogonal 2D image planes.

Digitized Brain Atlas and Segmented Deep-brain Nuclei: We non-rigidly mapped a digitized version of Schaltenbrand-Wahren atlas [1] to the standard brain. Deep-brain nuclei were segmented based on the anatomical representation in this atlas, and represented as either 3D objects or triangulated meshes. The centroid of each segmented deep-brain nucleus is shown as a sphere.

Collections of final surgical targets of previous patients: The location of the final target of each surgical procedure was non-rigidly registered to the standard brain coordinate and saved to each categorized database depending on the characteristic of the surgery. Currently we have eight databases containing data from 50 thalamotomy (34 left and 16 right), 59 pallidotomy (30 left and 29 right), 22 thalamus DBS (12 left and 10 right), and 30 STN DBS (18 left and 12 right) procedures. The collection of final targets can be non-rigidly mapped to the pre-operative MR images of individual

patients. The center of mass (COM) and the statistical map of a cluster of target locations are used to estimate the initial surgical target of each individual patient.

Representation of surgical instruments: Up to five multiple virtual probes can be manipulated simultaneously or independently to simulate the real surgical procedures. When implemented with the Frame Finder algorithm, this system also provides the simultaneous display of the tip positions of these trajectories in both image space and stereotactic frame space.

Visualization and Navigation Platform: Within this system, coded functional data, plotted directly onto the patient’s pre-operative image along a virtual trajectory corresponding to the position and orientation of the physical probe, can be non-rigidly transformed to the standard brain template automatically.

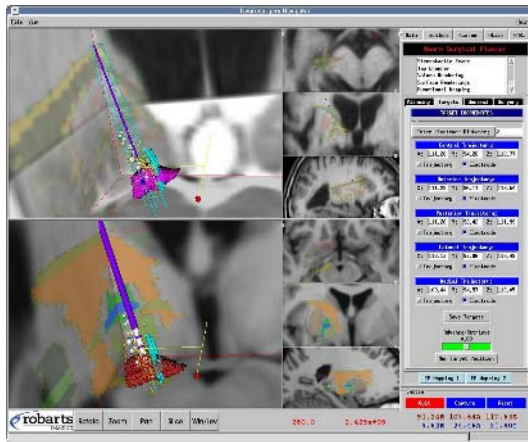


Fig. 1. The primary graphical user interface of the system displays the 3D image volume and 2D slices of a patient (upper) and those of the standard brain template (lower). The digitized atlas is registered and fused with each image. A T2-weighted image is also shown fused with the patient image. The control panel shows the tip locations of the five probes.

3 Clinical Application

Usually, the initial pre-operative surgical target is selected by the neurosurgeon using a geometric technique based on measurements relative to the AC-PC positions. For each of the eight STN DBS cases (5 left and 3 right procedures), an experienced stereotactic neurosurgeon performed the pre-surgical planning and targeting using this standard approach, carried out the surgical procedure according to his plan, and refined the surgical target through electrophysiological recording and stimulation. At the same time, one of the authors (TG), a non neurosurgeon familiar with deep brain anatomy, estimated the surgical target location and trajectory orientation for the placement of a DBS electrode independently of the surgeon using the neurosurgical visualization and navigation system integrated with the customized functional and anatomical data. For each case, the pre-operative MR image of each patient was loaded into our system and

non-rigidly registered to the standard brain template. Then the 3D transformation and non-rigid displacement grid file generated by registration were applied to map the data in the electrophysiological database, the collection of previous surgical targets, the digitized Schaltenbrand atlas, as well as segmented deep-brain nuclei to the patient brain image. In this preliminary study, the effectiveness of our system for surgical targeting was evaluated by comparing the target locations estimated by the non-expert with those identified by the neurosurgeon.

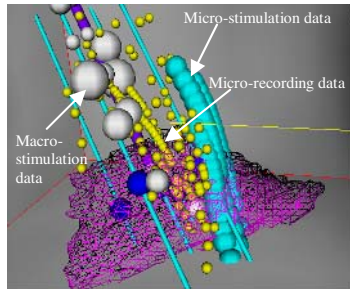


Fig. 2. The magnified version of Fig. 1; Purple lines (central): Central electrodes; Cyan lines (parallel to the central line): surgical trajectories; Yellow spheres (small): micro-recording data; Cyan spheres (medium): micro-stimulation data; White spheres (large): macro-stimulation data; Mesh object: sub-thalamic nucleus (STN); the segmented STN and electro-physiological data are non-rigidly registered from the standard brain space to the patient brain image

4 Results

4.1 Clinical Validation of the Non-rigid Registration Algorithm

Accurate registration plays a critical role in localizing surgical targets with the references of standardized anatomical and functional information. Registering the collection of previous final surgical targets from the standard database to an individual patient image yields a probabilistic estimation of the target location for the patient. To assess our registration algorithm clinically, a cluster of 18 left STN DBS surgical targets contained in the categorized database was non-rigidly transformed to the images of 5 patients who had received similar surgical procedures. That of the 12 right STN DBS targets in another database was mapped to the images of 3 patients undergoing right STN DBS. Table 1 shows the comparison between the center of mass or the most significant position on the probability map of database-initialized locations and the actual surgical targets of the 8 patients. The results demonstrate that the registration algorithm performs well within the homogeneous regions in the deep-brain. Although the average distance between the registered centroid of the collection of previous surgical targets and the target locations of the new patients is $2.19 \pm 0.72 \text{mm}$, this technique nevertheless provides a suitable initial estimate of the pre-operative surgical target, which may be further refined with additional functional and anatomical information available on the neurosurgical system.

Table 1. Absolute differences between the database-initialized and the real surgical targets

<i>Difference</i>	<i>x</i>	<i>y</i>	<i>z</i>	<i>d(x,y,z)</i>
<i>Avg. (mm)</i>	1.22	1.28	1.08	2.19
<i>Max (mm)</i>	2.09	2.67	2.3	3.6
<i>Min (mm)</i>	0.73	0.83	0.33	1.55
<i>Sd (mm)</i>	0.48	0.59	0.73	0.72

4.2 Application of the Segmented Deep-Brain Nuclei

If the registration is accurate, the atlas-based segmented deep-brain nuclei transformed to a patient brain space should have a high percentage of overlap with the patient’s own nuclei. The accuracy of our registration algorithm suggests that the anatomical information provided by a specific segmented nucleus registered to a patient (Fig. 3) should indicate the spatial location of the optimal surgical target within the nucleus. At our institution, the dorsolateral portion of STN is regarded as the most effective stimulation site in STN DBS. Acknowledging the centroid position of the segmented STN and its spatial relationship with respect to the theoretical surgical destination, the neurosurgeon may more confidently localize the surgical target. We compared the centroid of the segmented STN and the real target location for each of the eight patients. Table 2 reports the absolute differences between them. 75 percent of the actual surgical targets are located dorsolateral to the corresponding centroids in our study.

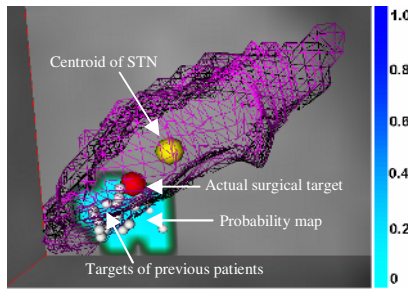


Fig. 3. Mesh object: STN; Yellow sphere: the centroid of STN; White spheres: surgical targets of previous patients; Colour-coded map: the probability map of a collection of left STN DBS targets; Red sphere: the actual surgical target

Table 2. Absolute differences between the centroids of the segmented STN and the real targets

<i>Difference</i>	<i>x</i>	<i>y</i>	<i>z</i>	<i>d(x,y,z)</i>
<i>Avg. (mm)</i>	1.41	1.99	0.89	2.84
<i>Max (mm)</i>	2.24	3.23	1.41	3.52
<i>Min (mm)</i>	0.32	0.01	0.23	1.24
<i>Sd (mm)</i>	0.73	1.15	0.58	0.85

4.3 Effectiveness in Surgical Targeting

We evaluated the effectiveness of our neurosurgical visualization and navigation system in surgical targeting on eight STN DBS patients. The average distance between the non-expert-planned surgical targets and the expert-localized ones was $0.58\pm 0.49\text{mm}$, $0.70\pm 0.37\text{mm}$, and $0.69\pm 0.34\text{mm}$ in x, y, and z directions respectively. In addition, the surgical sites determined with the combined information from both electrophysiological database and the anatomical resources were closer to the final surgical targets chosen by the neurosurgeon than those defined either by regular image-based techniques, or by the mapping of a cluster of surgical targets on the pre-operative images. Practically, these residuals indicate that an initial estimation of target location, made on the basis of this software, is typically within 1.7mm of the real surgical target. Therefore, the final surgical target can be reached by slightly refining the initiation of the target position estimated using this system with greatly reduced electrophysiological exploration.

Table 3. Absolute differences between the system estimated and the real surgical targets

<i>Difference</i>	<i>x</i>	<i>y</i>	<i>z</i>	<i>d(x,y,z)</i>
<i>Avg. (mm)</i>	0.58	0.70	0.69	1.67
<i>Max (mm)</i>	1.28	1.35	1.27	2.75
<i>Min (mm)</i>	0.06	0.30	0.2	0.65
<i>Sd (mm)</i>	0.49	0.37	0.34	0.70

5 Discussion

This visualization and navigation system has been used both pre- and intra-operatively for planning the surgical trajectories, plotting and analyzing functional data in patient pre-operative image space during eight separate surgeries. Even without considering the possibility of brain movement/shift, this system has been proved helpful in facilitating the identification of surgical loci for STN DBS. The visualization capabilities, designed for presentation of all the relevant functional and anatomical data along with multiple virtual surgical instruments, have made simulating real surgical procedures feasible. In conjunction with the function of automatic calculation of frame-to-MRI transforms, the intra-operatively acquired data can be saved in a text file whose header contains the code describing the physical information of the patient and specifications of the probes used during the procedure. Meanwhile, the homologous data in the standard brain space can be stored in the functional database. This initial pilot study, involving only one non-expert and one neurosurgeon, yielded promising results that need to be further validated on more subjects and a study of inter and intra “non-expert” variability for the application of our system to deep-brain neurosurgical procedures. While only a single neurosurgeon was involved in this work to date, we hope to address this problem through a multi-centre study in the future. Despite the difficulty and complexity of the accurate STN segmentation on the patient image files, comparing overlapping ratio between the registered segmented STN of each patient and that of the standard brain atlas could

also provide valuable measurements for the validation of the registration algorithm. Although the system has reached a stage of development where prediction of surgical targets is possible, further clinical evaluation is required for thorough validation and application of this system in stereotactic deep-brain neurosurgical procedures.

Acknowledgements. The authors acknowledge the financial support from the Canadian Institute of Health Research (CIHR), the Ontario Research & Development Challenge Fund (ORDCF), the Canada Foundation for Innovation (CFI), and the Ontario Innovation Trust (OIT).

References

1. G Schaltenbrand and W Wahren, *Atlas for Stereotaxy of the Human Brain*. Stuttgart, Germany: Thieme, 1977.
2. J Talairach and P Tournoux, *Co-planar stereotaxic atlas of the human brain*. Stuttgart, Germany: Thieme, 1988
3. WL Nowinski, TT Yeo, and GL Yang, "Atlas-based system for functional neurosurgery," in *Proc. SPIE Medical Imaging*. Newport Beach, USA: SPIE, 1997, vol. 3031, pp. 92–103, Image Display.
4. T Schiemann, KH Hohne, C Koch, A Pommert, M Riemer, R Schubert, and U Tiede, "Interpretation of tomographic images using automatic atlas lookup," in *Proc. Vis. Biomed. Comput*. Rochester, MN. SPIE, 1994, vol. 2359, pp. 457-465.
5. KA Ganser, H Dickhaus, R Metzner, and CR Wirtz, "A deformable digital brain atlas system according to Talairach and Tournoux," *Med. Imag. Analy.*, vol. 8, pp. 3-22, 2004.
6. RR Tasker, LW Organ, and PA Hawrylyshyn, *The Thalamus and Midbrain of Man*. Springfield, IL: Charles C Thomas, 1982.
7. KW Finnis, YP Starreveld, AG Parrent, AF Sadikot, and TM Peters, "Three-Dimensional Database of Subcortical Electrophysiology for Image-Guided Stereotactic Functional Neurosurgery," *IEEE Trans. Med. Imag.*, vol. 22, pp. 93–104, Jan. 2003.
8. PF D'Haese, E Cetinkaya, C Kao, JM Fitzpatrick, PE Konrad, and BM Dawant, "Toward the creation of an electrophysiological atlas for the pre-operative planning and intra-operative guidance of deep-brain stimulators (DBS) implantation," in *Proc MICCAI 2004*, C Barillot, DR Haynor, and P Hellier (Eds.), LNCS 3216, pp. 729-736, 2004.
9. WL Nowinski, D Belov, and AL Benabid, "An algorithm for rapid calculation of a probabilistic functional atlas of subcortical structures from electrophysiological data collected during functional neurosurgery procedures," *Neuroimage*, vol. 18, pp. 143-155, 2003.
10. T Guo, YP Starreveld, and TM Peters, "Evaluation and validation methods for intersubject non-rigid 3D image registration of the human brain," in *Proc. SPIE Medical Imaging*. Bellingham, WA: SPIE, 2005, vol. 5744, pp. 594-603, Visualization, Image-Guided Procedures, and Display.
11. CJ Holmes, R Hoge, DL Collins, R Woods, AW Toga, and AC Evans, "Enhancement of MR images using registration for signal averaging," *J. Comput. Assist. Tomogr.*, vol. 22, no. 2, pp. 324–333, 1998.

Gaze-Contingent Soft Tissue Deformation Tracking for Minimally Invasive Robotic Surgery

George P. Mylonas, Danail Stoyanov, Fani Deligianni,
Ara Darzi, and Guang-Zhong Yang

Royal Society/Wolfson Medical Image Computing Laboratory,
Imperial College London, London, United Kingdom
{george.mylonas, danail.stoyanov, fani.deligianni, a.darzi,
g.z.yang}@imperial.ac.uk

Abstract. The introduction of surgical robots in Minimally Invasive Surgery (MIS) has allowed enhanced manual dexterity through the use of microprocessor controlled mechanical wrists. Although fully autonomous robots are attractive, both ethical and legal barriers can prohibit their practical use in surgery. The purpose of this paper is to demonstrate that it is possible to use real-time binocular eye tracking for empowering robots with human vision by using knowledge acquired *in situ*. By utilizing the close relationship between the horizontal disparity and the depth perception varying with the viewing distance, it is possible to use ocular vergence for recovering 3D motion and deformation of the soft tissue during MIS procedures. Both phantom and in vivo experiments were carried out to assess the potential frequency limit of the system and its intrinsic depth recovery accuracy. The potential applications of the technique include motion stabilization and intra-operative planning in the presence of large tissue deformation.

1 Introduction

The field of Minimally Invasive Robotic Surgery (MIRS) is increasingly attracting considerable attention from the computer-assisted surgical research community. In robotic surgery, dexterity is enhanced by microprocessor controlled mechanical wrists, which allow motion scaling for reducing gross hand movements and the performance of micro-scale tasks that are otherwise not possible. The continuing evolution of the technology, including force feedback and virtual immobilization through real-time motion adaptation, will permit more complex procedures such as beating heart surgery to be carried out under a static frame-of-reference. The use of robotically assisted minimally invasive surgery provides an ideal environment for integrating patient specific pre-operative data for performing image guided surgery and active constraint control, which can all be conducted without the need of the surgeon to remove his/her eyes from the operating field of view. Under the dichotomy of autonomous and manipulator technologies in robotics, intelligence of the robot is typically pre-acquired through high-level abstraction and environment modelling. For systems that require robotic vision, this is known to create major difficulties. Additionally, the ethical and legal barriers imposed on interventional surgical robots

give rise to the need of a tightly integrated perceptual docking between the operator and the robot.

The purpose of this paper is to demonstrate that eye gaze derived from binocular eye-tracking can be effectively used to recover 3D motion and deformation of the soft tissue during MIRS. Compared to the use of other input channels, eye gaze is the only input modality that implicitly carries information about the focus of the user's attention at a specific point in time. This allows seamless *in vivo* registration of the motion and deformation fields within the area that is directly under foveation. This research extends our existing work in eyetracking and saccadic eye movement analysis, for gaze-contingent compensation of complex motion of the soft tissue.

2 Methods

2.1 Ocular Vergence Measurement for Gaze-Contingent 3D Deformation Recovery

One of the strongest depth cues available to human is the horizontal disparity that exists between the two retinal images. There is a close relationship between the horizontal disparity and the depth perception, varying with the viewing distance. More specifically, as the fixation point moves away from the observer, the horizontal disparity between the two retinal images is diminished and vice-versa. In order to extract quantitative information regarding the depth of the fixation point, ocular vergence needs to be measured, thus providing a veridical interpretation of stereoscopic depth [1]. One technique of achieving that is video-oculography. This is a non-intrusive, video based, approach used to measure the corneal reflection from a fixed IR light source in relation to the centre of the pupil [2]. The two centres define a vector, which can be mapped to a unique eye gaze direction. The combined tracking of both eyes provides the binocular vergence measure, which in turn determines the fixation point. In order to establish the relationship between pupil-glint vectors and points in 3D space, calibration is required prior to any eye tracking session. Different methods of binocular eye-tracking calibration exist. In [3] and [4], two of these calibration methods are described. Our eye-tracking calibration method however, takes also into consideration the intrinsic and extrinsic robot camera characteristics, including their relative pose.

2.2 Experimental Design and System Setup

In order to assess the basic characteristics of using ocular vergence for 3D depth recovery, a replication of the DaVinci (Intuitive Inc, CA) robotic surgical environment was created. The system consists of a stereoscopic console and an industrial robot equipped with a stereo-camera rig. The stereo console allows the user to look at 3D video captured by the two cameras held by the robot. Inside the stereoscope, a video based binocular eye tracking system was installed [4]. With binocular eyetracking, it is possible to determine the 3D position of the fixation point of the user while observing the stereo images. The robot used in this experiment is a Stäubli RX60 robotic arm with six degrees of freedom (DOF) and a repeatability accuracy of $\pm 0.02\text{mm}$ at high speed and acceleration. For accurate positioning of the

camera tips, a Polaris (Northern Digital Inc, Ontario) 6-DOF tracker was used. The Polaris tracker is able to simultaneously track a number of passive, active, wired and wireless IR tools in real time. Data interfacing is achieved through RS-232/RS-422 and the provided tracking accuracy is 0.35 mm RMS at a sampling rate of 60Hz.

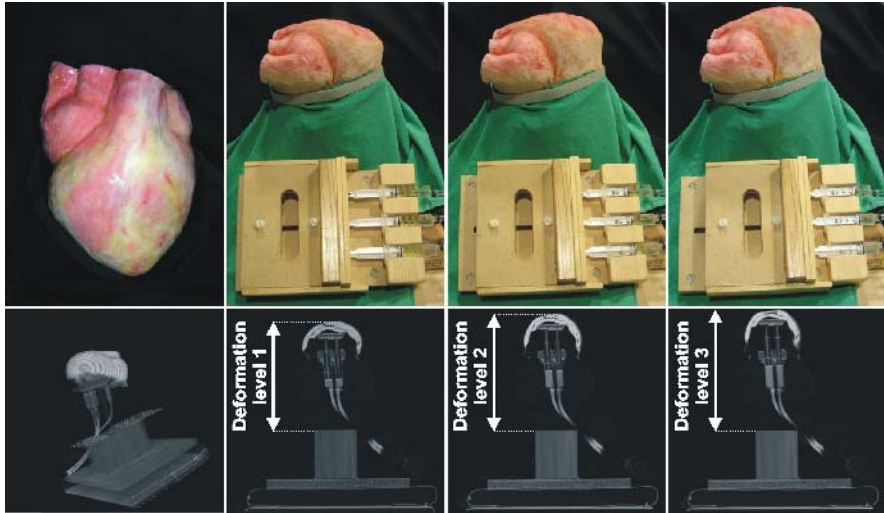


Fig. 1. *Top:* The phantom heart at different deformation levels. The oil filled pistons (only 3 shown here) allow for reproducible deformation control through numbered injection levels. *Bottom:* The reconstructed phantom heart from a series of CT slices. On the right, CT slices at exactly the same position of the phantom and at three deformation levels.

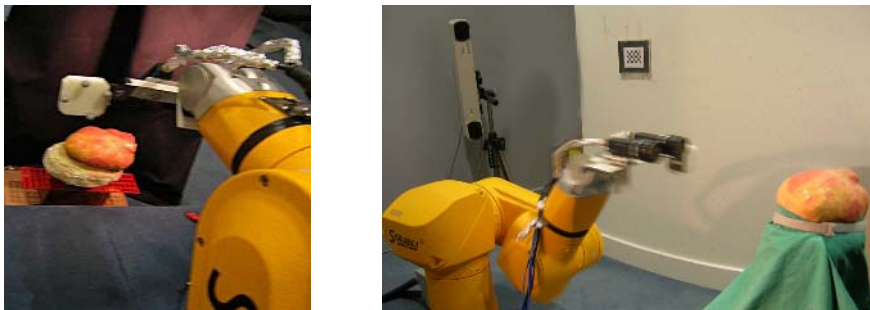


Fig. 2. On the left, the robot with the mounted optical-tracker retroreflectors and the stereo-camera rig. On the right, the polaris optical tracker is visible on the background.

For assessing the accuracy of the proposed 3D depth recovery framework through ocular vergence, a phantom heart model was created by using thixotropic silicone mould rubber and pre-vulcanised natural rubber latex with rubber mask grease paint to achieve a specular appearance and high visual fidelity. The phantom is deformable by means of a system of 4 oil filled pistons with controllable injection levels. In this

way, the amount of deformation can be accurately controlled and reproduced at any time as this can be seen in Fig. 1. The same figure also depicts the reconstructed phantom heart from a series of CT slices at different deformation levels. In Fig. 2, the entire setup can be seen. For determining the exact position of the epicardial surface, an Aurora (Northern Digital Inc, Ontario) 5-DOF electromagnetic catheter-tip tracker was used. This device has an accuracy of 0.9-1.3mm RMS, depending on the distance of the tool from the magnetic field generator. The maximum sampling rate is 45Hz for up to 5 tracked tools, which drops down to 22Hz when more than 6 tools are tracked.

2.3 Autoregressive Moving Average Modelling (ARMA) for Visual System Identification subjected to Oscillatory Stimulus

In order to study the oculomotor response, it is necessary to obtain a model that can closely describe the system. To this end, the first part of our experiment is to record a set of data for deriving the input response of the eye-tracking system. By collecting the stimulus and the response data sets, the Steiglitz-McBride method was used [5] to derive the rational transfer function of the proposed ocular vergence 3D depth recovery system. This method is commonly referred to as ARMA modelling, which attempts to find the coefficients of the rational transfer function that approximates a given time-domain linear system. This method is based on identifying the unknown system from given input and output sequences that describe the system's behaviour. The method solves for the numerator and denominator coefficients simultaneously in an attempt to minimize the signal error between the output and the given output signal. To include the effect of exogenous inputs to the system in a time series model, the basic ARMA model can be further extended to a number of more elaborate variants, like ARMAX (Autoregressive Moving Average with Exogenous Input). A general input-output linear model for a single output system with input u and output y can be written [7]

$$A(q)y(t) = \sum_{i=1}^{nu} [B_i(q)/F_i(q)]u_i(t - nk_i) + [C(q)/D(q)]e(t) \quad (1)$$

In the above equation, u_i is the input i and A , B_i , C , D and F_i are polynomials in the shift operator (z or q). The general model is defined by setting the time delays nk and the orders of the above polynomials. According to the variance in use, the above equation can be reduced down to a special case, *e.g.*, for ARMAX Eq. (1) can be confined to

$$A(q)y(t) = B(q)u(t - nk) + C(q)e(t) \quad (2)$$

2.4 Gaze-Contingent Decoupling of Cardiac and Respiratory Motion

One of the basic problems in intraoperative registration between 3D tomographic data and cardiac tissues is the large deformations that can occur during the MIS procedures. If deformation is significant, convergence of any iterative registration algorithm is not guaranteed. One of the largest components of cardiac deformation is the respiratory motion, which is typically coupled with the higher frequency cardiac motion. We show in this paper that gaze-contingent motion recovery of the composite

cardiac motion is practical and *in vivo* decoupling of the superimposed components is possible. Given that the dominant variation of the gaze recovered signal is due to cardiac motion and it has a constant direction during a given time window, then Wavelet and Principal Component Analysis can be employed to decouple the motion components [6]. The principal vector is expected to lie on the same plane over time, thus projecting the signal on this plane can isolate any variations due to respiration.

3 Results

3.1 Binocular System Frequency Response

To assess the binocular system frequency response, experiments were carried out involving six subjects. The first set of experiments investigates the oscillatory response of the binocular visual system over a frequency range. The subjects involved, were asked to keep fixating on a feature of their preference on the surface of the heart model. While eyetracking is performed, the robot is set to oscillation of gradually increasing frequency in the z-axis (depth). While the 3D fixation point of a subject is tracked, the position of the robot cameras is also recorded by the Polaris optical tracker. After data collection, ARMAX modelling is used to derive the coefficients of the parametric system that describes the transfer function of the system. Table 1 summarises the response of the visual system in oscillation on the z-axis, which indicates that the response of the system is good up to frequencies of about 1.8Hz. Beyond this limit, there is considerable attenuation and noticeable phase shift.

Table 1. Error analysis comparing the oculomotor response of the 6 subjects over a range of frequencies

Frequency (rad/s)	Amplitude		Phase (rad)	
	Mean	std	Mean	std
0.1	0.950	0.118	-0.031	0.016
1.0	0.918	0.132	-0.307	0.120
3.0	0.822	0.204	-0.871	0.098
6.0	0.650	0.156	-1.494	0.150
10.0	0.577	0.125	-2.394	0.278
12.0	0.520	0.126	-2.955	0.316

3.2 Gaze-Contingent Depth Extraction and Tissue Deformation Tracking

The second part of our experiment was to extract gaze-contingent depth information from the phantom. For this experiment, an Aurora catheter tip was positioned on the epicardial surface of the phantom heart. Though slightly less accurate than the Polaris system, the Aurora tools were used because of their small size. These tools can be easily embedded over or under the epicardial surface and not influence the fixation point position as this would be the case with the much larger Polaris tools. The subjects were asked to keep fixating on a surface feature close to the position sensor. While the phantom was subjected to different levels of deformation, both the fixation

point and the position of the electromagnetic sensor were tracked. In Fig. 3, the gaze-contingent recovered deformation is compared with the actual deformation levels.

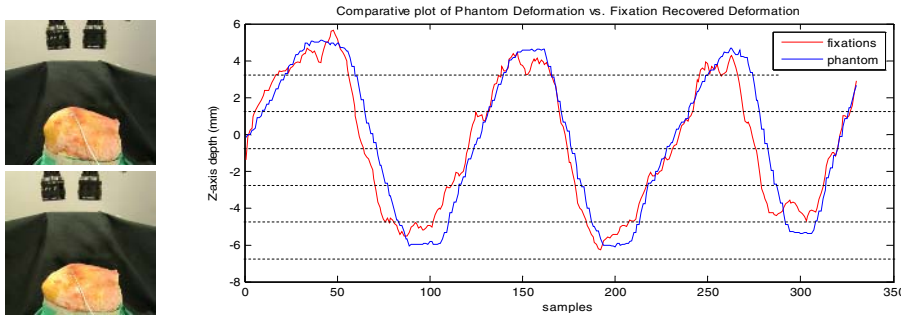


Fig. 3. The recovered deformation from ocular vergence. On the right, actual deformation of the phantom heart surface vs. the gaze-contingent reconstructed deformation. On the left, the setup for the experiment.

3.3 Gaze-Contingent In-Vivo Cardiac and Respiratory Motion Decoupling

To assess the *in vivo* accuracy of the proposed ocular-vergence 3D depth recovery system, data from a real robotic assisted operation where used. The operation was a totally endoscopic coronary artery bypass graft (TECAB) performed with the DaVinci robot. The stereo footage was played back in the stereoscopic viewer while a subject was eyetracked. The purpose of the experiment was to demonstrate how cardiac and respiratory motion can be recovered and decoupled using eyetracking. Fig. 4 shows a snapshot of the two views from a 40sec long footage. A large portion of the view was filled with the robotic endowrist, grasping the cardiac stabilizer just before positioned in place. During the entire footage, the laparoscope was stationary. The deformed tissue under foveation is at the bottom left portion of the image. What appears on the video sequence is a deforming epicardial surface with the respiratory motion principally manifested along the horizontal axis.

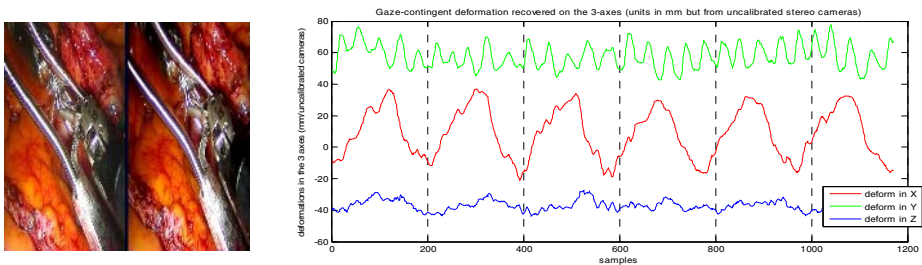


Fig. 4. The recovered epicardial surface deformation for a TECAB procedure. On the left, a snapshot of the binocular views used on the experiment is shown. On the right, the eyetracking acquired tissue deformations on the x, y and z-axes are plotted.

The cardiac motion is superimposed on top of this motion throughout the respiratory periods. The subject was asked to keep fixating a landmark on the tissue. Fig. 4 on the right, shows the collected eyetracking data on the x , y and z -axis (depth). Wavelet and PCA analysis is used to successfully decouple respiratory from cardiac motion as it can be seen on Fig. 5.

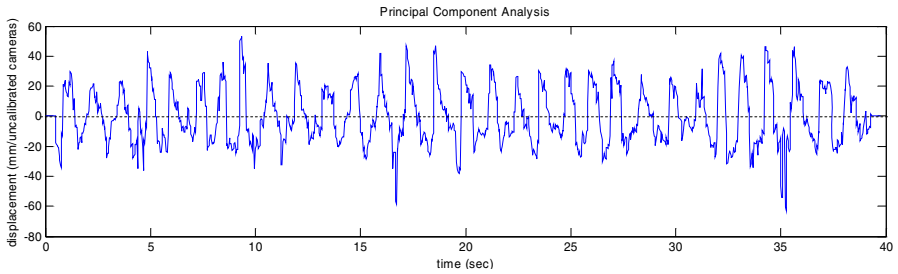


Fig. 5. The extracted cardiac motion by using Wavelet and PCA analysis of the gaze data

4 Discussion and Conclusions

In this paper, we have demonstrated the potential of gaze-contingent soft tissue motion and deformation recovery in MIRS. In this study, it is shown that the cut-off of 3D depth recovery with the proposed method reached at frequencies of about 1.5-1.8Hz when oscillations occur on the z -axis (depth). This corresponds to a heartbeat rate in the range of 100bbs. It is expected that the response of the visual system in oscillations on the x and y axes will be better. In this case, and when depth recovery is not required, it would be possible to decouple the two eyes during eyetracking and use just the saccadic information of one of them preferably by using the dominant eye. In this way, it is possible to eliminate a large amount of noise caused by the intrinsic behaviour of the visual system during binocular coordination. It is also possible to get a better overall response by re-positioning the robotic endoscope in such a way that the principal oscillation component is shifted from z to the other two axes (x or y).

It is worth noting that there are also other important factors that can adversely affect the dynamic response of the visual system. These factors are accumulated through the video relay chain from the endoscope cameras up to the monitors of the stereoscope. It is worth noting that the human visual system is highly dynamic, none of the fitted models could accurately describe the response of the visual system to the oscillatory stimulus. We have demonstrated in this paper that it is practically possible to determine the 3D deformation of the surface at an oscillation frequency of about 100bbs. Potential applications of the proposed approach for MIRS include motion stabilization, active constraints and intra-operative image registration under large tissue deformation.

Acknowledgements

The authors would like to thank Eddy Kavanagh, Bob Pullen and Danny Sharpe for their invaluable help with the phantom experiment.

References

1. Mon-Williams, M., Tresilian, JR, Roberts, A., Vergence provides veridical depth perception from horizontal retinal image disparities. *Exp Brain Res* **133** (2000) 407-413
2. Yang, G.-Z., Dempere-Marco, L., Hu, X.-P., Rowe, A., Visual search: psychophysical models and practical applications. *Image and Vision Computing* **20** (2002) 291-305
3. Andrew T. Duchowski, Eric Medlin, Nathan Cournia, Anand Gramopadhye, Brian Melloy, Santosh Nair, 3D eye movement analysis for VR visual inspection training, Proceedings of the symposium on ETRA 2002: eye tracking research & applications symposium, March 25-27, 2002, New Orleans, Louisiana 103-110
4. Mylonas, P.G., Darzi, A., Yang, G.-Z., Gaze Contingent Depth Recovery and Motion Stabilisation for Minimally Invasive Robotic Surgery. *MIAR* (2004), **LNCS 3150**, 311-319
5. Steiglitz, K., McBride, L.E., "A Technique for the Identification of Linear Systems." *IEEE Trans. Automatic Control*. Vol. **AC-10** (1965) 461-464
6. Deligianni, F., Chung, A., Yang, G.-Z., Decoupling of Respiratory Motion with Wavelet and Principal Component Analysis. Proceedings of the eighth Annual MIUA, (2004) 13-16
7. Ljung, L., System Identification Toolbox for use with MATLAB, User's Guide, www.mathworks.com

Registration and Integration for Fluoroscopy Device Enhancement

James C. Ross¹, David Langan¹, Ravi Manjeshwar¹, John Kaufhold¹,
Joseph Manak¹, and David Wilson²

¹ GE Research Center, Niskayuna NY 12309, USA

² Case Western Reserve University, Cleveland OH 44106, USA

Abstract. We investigated a method, motion compensated integration (MCI), for enhancing stent Contrast-to-Noise Ratio (CNR) such that stent deployment may be more easily assessed. MCI registers fluoroscopic frames on the basis of stent motion and performs pixel-wise integration to reduce noise. Registration is based on marker balls, high contrast interventional devices which guide the clinician in stent placement. It is assumed that stent motion is identical to that of the marker balls. Detecting marker balls and identifying their centroids with a high degree of accuracy is a non-trivial task. To address the required registration accuracy, in this work we examine MCI's visualization benefit as a function of registration error. We employ adaptive forced choice experiments to quantify human discrimination fidelity. Perception results are contrasted with CNR measurements. For each level of registration inaccuracy investigated, MCI conferred a benefit ($p < 0.05$) on stent deployment assessment suggesting the technique is tolerant of modest registration error. We also consider the blurring effect of cardiac motion during the x-ray pulse and select frames for integration as a function of cardiac phase imaged.

1 Introduction

Coronary artery disease (CAD) afflicts approximately thirteen million Americans [1]. It is caused by the buildup of plaques and fatty deposits on the interior walls of coronary arteries. These narrowed (stenotic) vessels restrict blood flow to the heart and can ultimately lead to myocardial infarction. One of the standard techniques used to restore function to an impaired coronary vessel is stent deployment. Using x-ray fluoroscopy, the operating cardiologist guides a stent, compressed on a guidewire, through a catheter in the patient's vasculature to the site of the stenosis. The position of the stent on the guidewire is delineated by two radio-opaque marker balls. The marker balls serve as high contrast visual cues in the x-ray imagery for the clinician to position the stent prior to deployment. Once the stent is positioned to span the site of the stenosis, the balloon is inflated to expand the stenotic vessel and simultaneously deploy the stent. Figure 1 depicts an example fluoroscopic image containing a catheter, guidewire, marker balls, stent, and balloon. The deployed stent is often barely perceptible, even with the visual cues provided by the marker balls, making it difficult to ascertain whether the stent has been fully deployed. Partial deployment occurs when some portions of the stent fail to fully expand. This is problematic

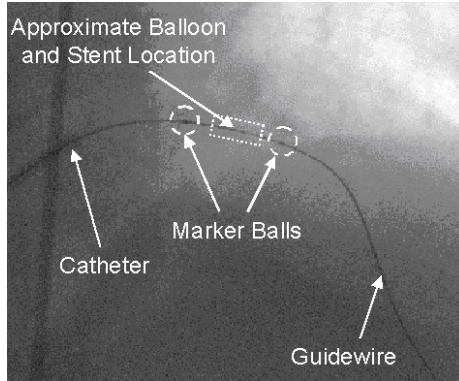


Fig. 1. Sample fluoroscopic image. The marker balls on the guidewire assist in visualizing the stent location.

since stent malapposition is believed to increase the chance of restenosis. In an effort to determine whether a stent is completely deployed, a cardiologist will typically administer a bolus of contrast agent to visualize the flow through the stented vessel via fluoroscopy. A uniform flow profile along the length of the stent indicates proper dialation of the stenosis and can be used to infer full stent deployment. In some cases the ends of an otherwise fully deployed stent are “pinched” (i.e. improperly deployed) while the flow profile of the bolus suggests successful deployment. The flow profile of a contrast bolus is an indirect and incomplete measurement of stent deployment. There exists a clinical need for direct assessment of stent deployment.

We overview a method, motion compensated integration (MCI), for enhancing the stent CNR so that stent deployment may be assessed directly. Here, CNR is defined as the difference between the stent mean intensity and the background mean intensity divided by the background noise standard deviation. The background noise standard deviation is computed over a flat image region. The method parallels the approach taken by Koolen [2]. In this work we aim to measure the benefit of applying MCI to fluoroscopic x-ray sequences as a function of registration accuracy. We measure MCI performance in terms of human perception studies and contrast these results with computed CNR. These results may be used to specify registration performance. In section 2 we survey device properties, common imaging geometries and techniques leading to expected device CNR. In section 3 we briefly overview our method for improving stent CNR. In sections 5 through 7 we present our experimental approach, results, and discussion. Finally, in section 8 conclusions will be drawn.

2 Devices, Acquisition Geometries, and Expected CNRs

In order to ascertain the proper deployment of a stent, the shape of the stent must be resolved. In a projection x-ray image of a stent, multiple struts often

Table 1. CNR Values for Patient Thicknesses and System Parameters

Patient Thickness (cm)	kVp	mAs	Cu Filtration (um)	Marker Ball CNR	Stent CNR
20	75	1.1	200	9.21	2.57
25	85	0.767	200	5.20	1.31
30	120	0.303	200	2.75	0.56
35	120	0.303	200	1.54	0.31
40	120	0.303	200	0.87	0.19

occupy the same ray path from source to detector. Here, as a baseline, we will determine the CNR of an individual stent strut for a range of common patient thicknesses subject to typical x-ray system performance for fluoroscopy. Stent struts are commonly made of thin (~ 0.16 mm) wires of a cobalt-chromium or stainless steel alloy. In the table below are shown the calculated CNRs for a 0.16mm thick cobalt-chromium stent strut and a 1.0mm Tungsten markerball for a variety of patient thicknesses. The simulated X-ray system was operating in fluoroscopic mode (< 10 Rad/min) filtered by a 200 micron layer of copper at a frame rate of 30 frames per second (20cm FOV, 70cm SOD, 90cm SID). The background noise term was calculated based on the quantum noise and an electronic noise term to simulate a digital flat-panel detector. As can be seen from the table, there is an order of magnitude variation in stent and markerball CNR as a function of projected patient thickness.

3 Motion Compensated Integration (MCI)

The objective of MCI is to reduce the quantum noise in the x-ray image while retaining stent contrast. The strategy is to take a series of x-ray images, co-register the frames on the basis of stent location, and average the frames. We assume the stent motion is rigid to that of the markerballs and that the markerballs are present in each of the x-ray images. Due to their relatively high CNR, the markerballs are more easily detected and are used as a surrogate motion estimate for the stent. Each frame in the fluoroscopic sequence is registered to the first frame on the basis of the markerball centroids using a similarity transform [3].

X-ray frames with significant out of plane motion, identified by a foreshortening of the distance between the markerballs may be ignored. Similarly, frames exhibiting significant blurring, identified by ECG or extended markerball footprint, may also be discarded.

The quantum x-ray noise is modelled as a Poisson process. For regions with constant mean, the noise may be equivalently modelled as an independent, identically distributed (i.i.d.) Gaussian process. It is assumed that for a given pixel location, the mean is constant among the registered frames. Therefore, the noise may be temporally averaged resulting in a noise reduction by \sqrt{N} , where N is the number of frames averaged.

Once the sequence frames have been registered to the reference frame, they are temporally averaged according to the following expression

$$\text{MCI} = \frac{1}{n+1} \left[\mathbf{I}_0 + \sum_{i=1}^n \mathbf{T}_i(\mathbf{I}_i) \right] \quad (1)$$

where $n+1$ is the number of frames in the sequence, \mathbf{I}_0 is the reference frame, \mathbf{I}_i is the i^{th} frame in the sequence, and \mathbf{T}_i is the transformation that registers frame \mathbf{I}_i to the reference frame. The result is referred to as the motion compensated integration image, **MCI**.

4 Marker Ball Detection and Tracking

A critical issue for MCI to be routinely used in a catheterization laboratory is automated unsupervised detection and tracking of the marker balls. This topic is beyond the scope of this paper, and here we briefly outline our dynamic programming based approach for marker ball detection and tracking.

The first step is to design a matched filter for a single marker ball on the basis prior knowledge of the device and acquisition geometry. Detecting the marker balls individually avoids the need to address the orientation of marker ball pairs.

The marker ball matched filter is applied to each frame. The processed frame is then thresholded for a specified constant false alarm rate, N . The N detections are paired in accordance with a distance criterion determined by the device geometry and imaging magnification factor. Detects which are not successfully paired are discarded. Paired detects are tracked through multiple frames on the basis of ECG with an allowance for respiratory motion. Paired detects which fail to track are discarded. We further discard detects whose motion deviates from an ECG periodicity assumption. After a small number of cardiac cycles, the marker balls are identified and tracked through the frames.

5 Experimental Methods

The objective of our work was to quantify the benefit of MCI, as a function of registration accuracy, relative to the standard fluoroscopic sequence for stent deployment assessment. We employed a four alternative forced-choice (4-AFC) discrimination paradigm to measure the necessary stent CNR for the observer to correctly assess stent deployment 90 percent of the time [4] for both our baseline fluoroscopic sequence and for MCI subject to a specified registration accuracy.

At a high level, AFC experiments involve visually discriminating between two objects, a reference object and a test object, fully and partially deployed stents respectively. For a given trial, an observer is presented with four images – one contains the partially deployed stent. The contrast with which the stents are displayed is adjusted over the course of many trials such that, at the experiment’s conclusion, the contrast level for which the observer achieves 90 percent detection accuracy of the test object is reliably identified.

The AFC experiments required a database of fluoroscopic sequences and stent templates. We created a database of clinically acquired sequences that had a relatively uniform level of noise and background clutter.

Next, we digitally added a guidewire with marker balls to each database sequence with periodic elliptic interframe motion. A 0.75mm (3 pixel) guidewire with 1.5mm (6 pixel) marker balls was synthetically generated and low pass filtered to mimic detector MTF. Our fully deployed stent template was based on an x-ray image of a Cordis drug-eluting stent with 160μ strut diameter. This stent template was manipulated by downsampling both ends to simulate partial deployment. The intensity of the downsampled ends was adjusted to maintain the level of total x-ray attenuation by the stent.

Registration error was modelled in polar coordinates with uniform distribution in θ and Gaussian, $N(0, \sigma^2)$, in r . We measured MCI performance for $\sigma = [0.25, 0.75, 1.25, 2.0, 4.0]$.

We followed the reference/test technique described in [4] to conduct our experiments. The stent contrast was adapted so as to maintain observer performance at 90% probability of making the correct choice corresponding to a detectability index, d' , of 2.4517 [5]. Results were obtained after 100 reference trials interlaced with 100 test trials. Contrast levels were determined from maximum-likelihood estimates, and standard errors were estimated by a method that accounted for adaptation. We used the Michelson contrast for our experiments:

$$C = \frac{|\mu_b - \mu_s|}{\mu_b + \mu_s} \quad (2)$$

where μ_b and μ_s were the mean gray-scale value of the background and stent, respectively. The normalized stent templates were added to the sequences according to the following expression:

$$\mathbf{A}' = \mathbf{A} - \frac{2\mu_b \mathbf{C}}{1 + \mathbf{C}} (\mathbf{1} - \mathbf{S}) \quad (3)$$

where \mathbf{A} is the anatomical sequence, \mathbf{S} is the stent template sequence, and \mathbf{A}' is the resultant sequence.

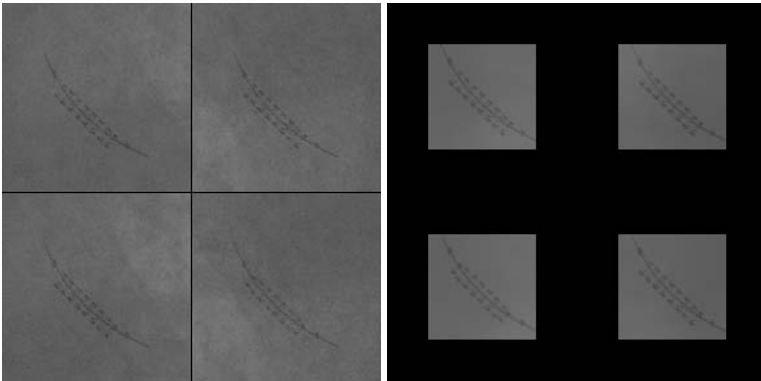


Fig. 2. Left: Example reference experiment trial. Right: Example test experiment trial. The stents were added to the fluoroscopy sequences with a high level of contrast for display purposes.

For the reference experiment, four sequences were randomly selected from the database, and stent templates (three fully deployed, one partially deployed) were added to them according to (3). They were then simultaneously displayed on a medical quality gray-scale monitor with the location of the test sequence randomly varied. The observer selected the sequence most likely to contain the partially deployed stent.

Each reference trial was followed by a test trial. Interlacing the reference (fluoroscopic sequence) and test (MCI) trials minimized observer effects such as fatigue, lapses in attention, or possible physiological changes that could occur during the experiments [4]. Figure 2 shows example reference and test trials.

6 Results

Contrast sensitivity is defined as the inverse of the final Michelson contrast values following adaptation. A higher sensitivity indicates an improved ability to detect the partially deployed stent. The ratio of test sensitivity to reference sensitivity for a given experiment is used in order to compensate for inter-observer variation. A ratio greater than one indicates an improved discrimination ability using MCI.

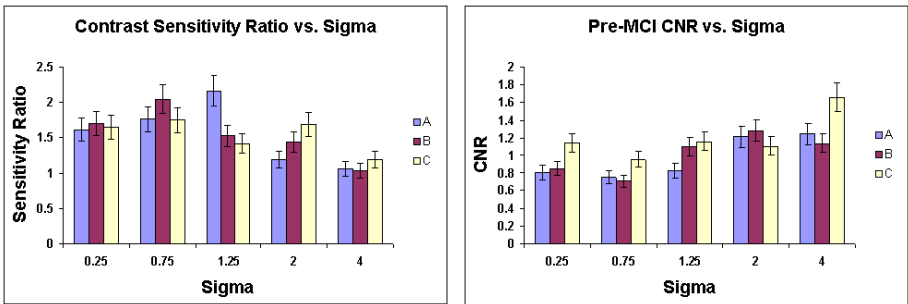


Fig. 3. Left: contrast sensitivity results for observers A, B, and C. Right: pre-MCI CNR results for observers A, B, and C.

Figure 3 gives contrast sensitivity ratios for each observer. A one-sample t-test shows that MCI provides the observer a significant ($p < 0.05$) advantage for performing the discrimination task for each sigma value investigated. There is not a statistically significant trend over sigma values of 0.25, 0.75, and 1.25, but there is a fall-off in contrast ratio levels exceeding 1.25. The decrease in ratios over this range indicates the blurring due to registration error is impacting perception. Figure 3 also shows pre-MCI CNR values for each observer (corresponding to test experiment final contrast values). It is reassuring to see that these values are representative of what we can expect to observe as presented in table 1.

7 Discussion

MCI enhances the visibility of stents through two mechanisms: the integration of multiple image frames to reduce noise; and removing motion enabling the observer to inspect a spatially static stent. It has been shown that the spatio-temporal bandwidth of the human visual system degrades the detection of low contrast targets that are moving [6]. Whiting et al have shown that the detection of vessel defects is enhanced when the vessels are stabilized [7].

MCI performance degrades with registration error, but more slowly than perhaps expected. Although a stent strut width is on the order of one pixel, the projected stent scaffolding width often spans several pixels. This multi-pixel projection footprint coupled with the knowledge that the observer is concerned with the overall shape of the stent, not individual struts, leads to the expectation that a sub-pixel registration accuracy is not required. Indeed, our perception studies shows MCI to be beneficial for a registration error sigma up to a value of 2 (the benefit of MCI at a sigma value of 4 is marginal). Given that a marker ball may be reliably detected, and it's radius is typically 3 pixels, a registration error sigma of 2 should be achievable. This alone motivates investigation of real-time unsupervised markerball detection and tracking approaches.



Fig. 4. MCI enhancement applied to a clinical sequence. Left: frame taken from original sequence. Middle: all fluoroscopy frames used for enhancement. Right: frames with motion blurring omitted from integration.

MCI has two opportunities to improve upon the perception performance reported. First, more frames may be used in the integration. However, noise decreases at a rate of \sqrt{N} , so the law of diminishing marginal returns applies. Second, we made no attempt to manipulate the displayed contrast of the MCI image relative to the standard fluoroscopic sequence. However, the opportunity to increase the displayed contrast of the MCI image exists as a result of the significantly reduced noise. Finally, we need to be cautious in that cardiac motion during the x-ray pulse can significantly degrade the stent and markerball signatures. In the application of MCI, x-ray frames associated with significant motion should be discarded. In Figure 4 we present a clinical example of the improved performance achieved with MCI by discarding such frames.

8 Conclusion

Motion Compensated Integration (MCI) improves discrimination between partially deployed and fully deployed stents aiding deployment assessment. Proper deployment is believed important to maximize patient outcome. The required registration accuracy (~ 2 pixels or less) relative to the markerball radius (~ 3 pixels) appears achievable given that markerballs may be reliably detected.

MCI provides measurable, but modest improvement on observer assessment of stent deployment. For MCI to gain wide clinical use requires that it be appropriately integrated into the interventional cardiologist's workflow. This implies near real-time robust unsupervised markerball detection and tracking. Also, contrast enhancement of the displayed MCI image remains to be explored.

References

1. Heart disease and stroke statistics – 2005 update. Technical report, American Heart Association, Dallas, TX (2004)
2. Koolen, J.J., Ziekenhuis, C.: Drug eluting stents: The importance of deployment. In: EuroPCR. (2003)
3. Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, Cambridge, UK (2000)
4. Gilmore, G.C., Xue, P., Thomas, C.W., Wilson, D.L.: An adaptive reference/test paradigm: Application to pulsed fluoroscopy perception. *Behav. Res. Methods Instrum. Comput.* **30** (1998) 332–348
5. Green, D., Swets, J.A.: Signal Detection Theory and Psychophysics. Krieger, New York (1974)
6. Xue, P., Wilson, D.L.: Detection of moving objects in pulsed-x-ray fluoroscopy. *Journal of the Optical Society of America - A* **15** (1998) 375–388
7. Whiting, J.S., Eckstein, M.P., Morioka, C.A., Eigler, N.L.: Effect of additive noise, signal contrast, and feature motion on visual detection in structured noise. In Kundel, H.L., ed.: *Proc. SPIE Medical Imaging 1996: Image Preception*. (1996) 26–38

Computer Aided Detection for Low-Dose CT Colonography

Gabriel Kiss, Johan Van Cleynenbreugel, Stylianos Drisis,
Didier Bielen, Guy Marchal, and Paul Suetens

Faculties of Medicine & Engineering,
Medical Image Computing (Radiology - ESAT/PSI),
University Hospital Gasthuisberg,
Herestraat 49, B3000 Leuven, Belgium

Abstract. The paper describes a method for automatic detection of colonic polyps, robust enough to be directly applied to low-dose CT colonographic datasets. Polyps are modeled using gray level intensity profiles and extended Gaussian images. Spherical harmonic decompositions ensure an easy comparison between a polyp candidate and a set of polypoid models, found in a previously built database. The detection sensitivity and specificity values are evaluated at different dose levels. Starting from the original raw-data (acquired at 55mAs), 5 patient datasets (prone and supine scans) are reconstructed at different dose levels (down to 5mAs), using different kernel filters and slice increments. Although the image quality decreases when lowering the acquisition mAs, all polyps above 6mm are successfully detected even at 15mAs. Accordingly the effective dose can be reduced from 4.93mSv to 1.61mSv, without affecting detection capabilities, particularly important when thinking of population screening.

1 Introduction

Screening for colorectal polyps (precursors of colonic cancer in 90% of cases [1]) is currently advocated, based on the proven hypothesis that early detection and removal of colonic polyps results in a reduced mortality rate [2]. Amongst other methods Computed Tomographic Colonography (CTC) has been proposed as a screening candidate. It's main advantage compared to standard fiber optic colonoscopy is its minimal invasiveness and better patient tolerance [3]. Recent studies proved that CT colonography is an accurate method for detecting polyps above 5 mm in diameter.

The presence of ionizing radiation, inherent to CT acquisitions, hampers the usability of CT Colonography as a screening tool. The repeated patient scans (as part of the screening protocol) lead to dose accumulation, which can have harmful effects. Fortunately, the contrast between the colonic wall and the insufflated air or CO_2 (Figure 1), employed to distend the colon is high enough and can be exploited to reduce the amount of dose per patient. Several studies [4] [5] have shown that despite significant image quality decreases, the

polyp detection results remain unimpaired as compared to current protocols. One study [4] reported a dose reduction from 8.8 mSv (median effective dose across several centers) to 3.6 mSv.

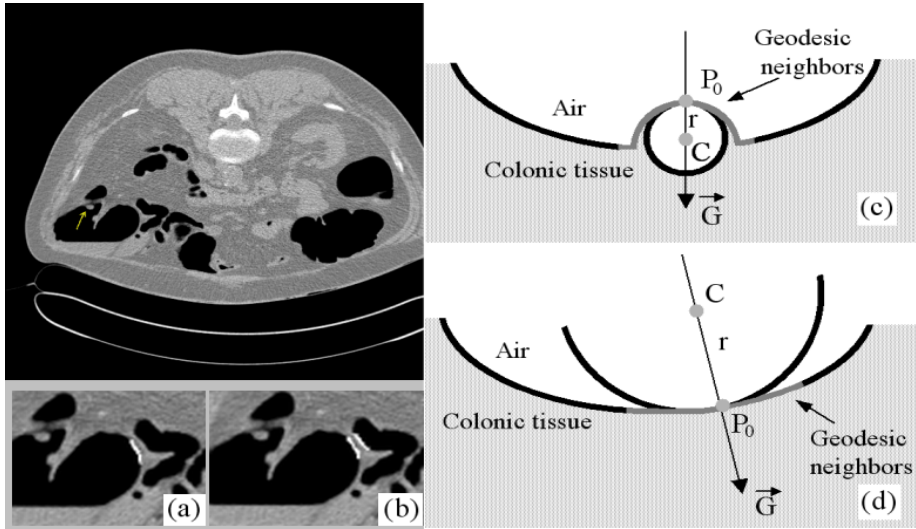


Fig. 1. CTC slice containing an 8 mm polyp. The difference between geodesic (a) and spatial neighbors (b) on a haustral fold is illustrated; the neighbors are depicted in white. The right column shows the fitting process and the resulting value of radius r for a colonic polyp (c) and for normal wall (d).

During the last decade, Computer Aided Detection (CAD) methods have been proposed [6] [7] [8] [9] in order to increase reading sensitivity and to decrease interpretation errors and reading time associated with CTC. The input data for these algorithms is acquired using scan protocols predefined by the software of the scanner. To our knowledge, no study has established a relationship between the acquisition dose and the performance of the CAD technique. In this paper, the applicability of CAD to low-dose colonographic data is assessed. Additionally an alternative CAD method, directly applicable to low-dose datasets, is described.

Some of the previous methods look at the curvature along the colonic wall to detect initial polyp candidates [6] [7] [9], taking advantage of the ellipsoidal shape of polyps. Alternatively statistical pattern recognition and geometrical modeling [8] have been employed to detect polyps in CTC data. As described in [9] curvature based methods tend to be susceptible to noise, a balance between regularization (required to compute the curvature accurately) and the preservation of local image structures (improving overall sensitivity of the CAD method) has to be found. Finding an optimal setup is challenging, even for normal colonoscopic cases. The image noise also affects the computation process of features, thus negatively influencing statistically based approaches.

2 Methods

The noise present in the low-dose data is handled by using robust fitting for generating initial polyp candidates and spherical harmonics for further analysis. Spherical harmonics [10] were chosen for detecting spatial trends and due to their ability to compensate for local perturbations given by noise.

The following subsections present the main steps of the proposed method; including a brief description of the segmentation step to obtain the colonic wall (section 2.1), the candidates generation (section 2.2), modeling of candidates using spherical harmonics and their classification into positive and negative cases (section 2.3).

2.1 Segmentation

The purpose of segmentation is to identify the colonic wall (found as the interface between colonic tissue and the air/ CO_2 inside the colon). The method detailed in [8], consisting of classic region growing [11], with multiple seed points (to overcome collapsed regions), is used. The cumulative Laplacian histogram [12] has the ability to reveal optimal threshold values between air and colonic mucosa, especially important when using low-dose datasets (optimal threshold changes as the dose is lowered).

The neighboring voxels of a given structure have to be determined at each position along the colonic wall. A possible solution is given in [9] where a similarity weight between neighboring voxels is computed based on the vectorial inner product of normals. Our segmentation result is represented as an N-ary tree, based on neighboring relations between voxels. A local breadth-first transversal of the tree is sufficient to retrieve all neighbors with a geodesic distance (distance along the colonic wall) smaller than a threshold T_{dist} .

2.2 Robust Sphere Fitting

The first step of the classification algorithm is the identification of normal colonic wall, approximately 90% of the wall's surface. Assuming that colonic polyps have spherical shapes, we propose a sphere fitting method to discriminate normal colonic wall from colonic polyps.

The sphere fitting can be reduced to a linear process by introducing the following constraint: the center of the fitted sphere at any given location has to be along the gradient of the current point. Figure 1 presents the fitting process for a colonic polyp (c) and normal colonic wall (d). Analytical formulation yields:

$$\begin{aligned} x_c &= x_0 - r * g_x \\ y_c &= y_0 - r * g_y \\ z_c &= z_0 - r * g_z \end{aligned} \tag{1}$$

considering $P_0(x_0, y_0, z_0)$ - the point on the colon wall (currently investigated), $C(x_c, y_c, z_c)$ - center point of the fitted sphere, r - sphere radius (to be computed), $P(x, y, z)$ - neighboring voxels and $\mathbf{G}(g_x, g_y, g_z)$ normalized gradient in $P_0(x_0, y_0, z_0)$.

The sphere equation and the normalization assumption are written as:

$$\begin{aligned} (x - x_c)^2 + (y - y_c)^2 + (z - z_c)^2 &= r^2 \\ g_x^2 + g_y^2 + g_z^2 &= 1 \end{aligned} \quad (2)$$

From equations 1 and 2 r is computed as:

$$r = \frac{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}{2 * ((x - x_0) * g_x + (y - y_0) * g_y + (z - z_0) * g_z)} \quad (3)$$

Equation 3 can be seen as fitting a straight line through a set of data points, $N_i = r * D_i + e$, where N_i and D_i are the numerator and denominator of equation 3, r - radius to be computed and e - fitting error. Instead of solving equation 3 by a least square approach (which consistently overestimated the value of r), robust fitting was preferred. The following merit function is minimized iteratively, as described in [13]:

$$\Upsilon(e, r) = \sum_{i=1}^n |N_i - e - rD_i| \quad (4)$$

Using robust estimation, outliers (generated by noise) are disregarded [13]; additionally the size of the current structure is estimated implicitly. Geometric modeling will take the computed radius into account by adapting its parameters accordingly.

Each location on the colonic wall is labeled as normal if the value of the fitted radius is negative (concave wall patch) or greater than T_{radius} (convex patch, however size is not interesting). The next step is to construct a Hough map, based on the remaining candidates, by back projecting with a radius given by r . The local maxima of the Hough map (identifying spherical shaped neighborhoods) are considered as input for the geometric modeling step.

2.3 Geometric Modeling Using Spherical Harmonics

To decide if a candidate location is a true polyp the most convenient approach is to match it to a set of known colonic polyps. However direct matching is not possible in practice, that is why polyps are modeled using shape descriptors and the corresponding descriptors are compared instead; the challenge becomes to find an accurate and fast measure of similarity with regards to all possible transformations. If translation and scaling can be easily addressed by normalization (to center of mass and to an average radius value for example), rotation is more time consuming. One possibility is to use PCA, however as mentioned in [10] PCA-alignment is negatively influenced by signal noise, since it mainly relies on the second order derivatives (the eigen values of the covariance matrix are computed using only second order information).

Due to their 'almost' spherical nature, polyps can be conveniently described by spherical functions. We have chosen a gray level appearance model (intensity profiles sampled along a sphere) and extended Gaussian images (mapping gradient information onto the unit sphere) to describe the shape of the polyp.

Both descriptors are spherical functions defined on the S^2 sphere, unfortunately rotationally dependant ones. Using the framework proposed in [10], a spherical function $f(\theta, \phi)$, defined on the S^2 can always be decomposed into a sum of its harmonics:

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l a_{lm} Y_l^m(\theta, \phi) \tag{5}$$

The L_2 norm of the spherical harmonics is rotationally invariant, forming a complete orthonormal basis. The vector:

$$SH(f) = \{ \|f_0(\theta, \phi)\|, \|f_1(\theta, \phi)\|, \dots, \|f_l(\theta, \phi)\|, \dots \} \tag{6}$$

with $f_l(\theta, \phi) = \sum_{m=-l}^l a_{lm} Y_l^m(\theta, \phi)$

is invariant to all transformations, allowing to take the L_2 difference between two harmonic representations (up to a given bandwidth bw) of two different functions f and g as their similarity measure:

$$SH(f) - SH(g) = \sum_{l=0}^{bw} (\|f_l(\theta, \phi)\| - \|g_l(\theta, \phi)\|)^2 \tag{7}$$

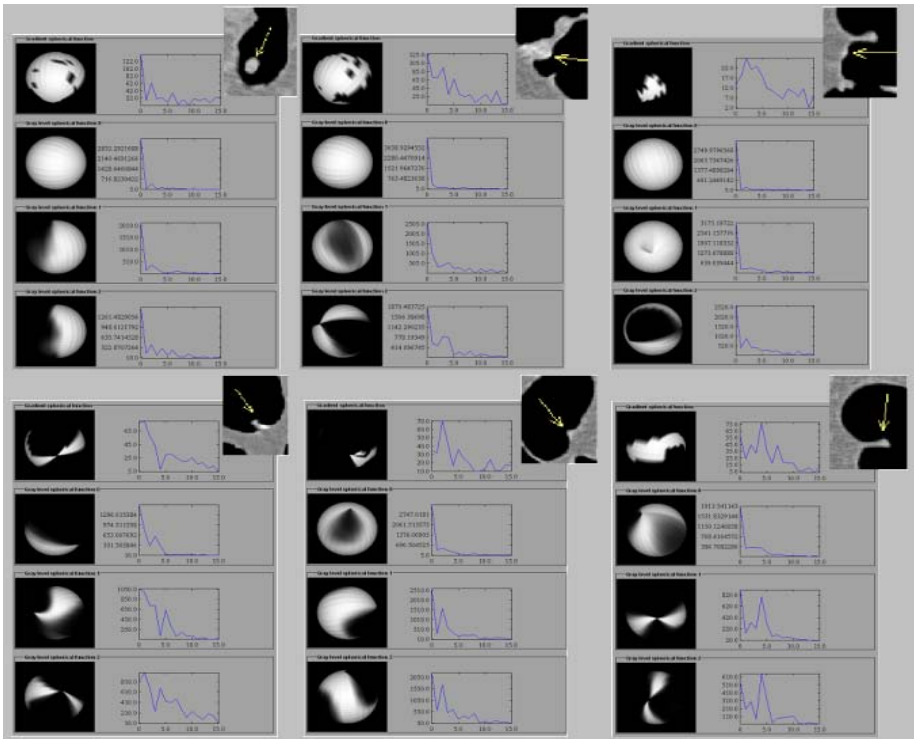


Fig. 2. Spherical harmonic decompositions for polyps (top row) and false positive cases (bottom row); for each case the spherical function and the harmonic decomposition are presented. The four spherical functions correspond, from top to bottom, to the extended Gaussian image and to the three intensity profiles at $r - 2$, r , $r + 2$.

Based on our previously acquired cases (see Patient data) a database of model vectors as given by equation 6 was built. For each polyp three intensity profile decompositions were computed, one sampled on a sphere with radius r (computed as described in section 2.2), and the other two at radii $r - 2$ and $r + 2$ respectively. The bandwidth bw was taken as 16, the first 16 frequency components were included into our harmonic vector. This value was considered sufficient to capture the overall shape of colonic polyps. Additionally the extended Gaussian image around the polyp was computed and its spherical harmonic decomposition determined. Figure 2 presents typical feature vectors (gray level and extended Gaussian images) and their feature vectors given by equation 6.

The database of models was clustered using hierarchical cluster trees (Matlab, Statistics toolbox). Clustering was required since polyps can have stalk, can be attached to the colonic wall or can lie on haustral folds, each having different signatures (as visible in figure 2).

3 Patient Data and Results

Fifty data-sets (cases D1) belonging to 26 patients, 25 normal and 25 with 28 polyps of various sizes (6 below 6 mm, 7 between 6-9 mm, 11 above 9 mm and 4 tumors) were considered as input for our CAD scheme. Additionally a set of 5 patients (cases D2) for which raw data was available and containing 12 polyps (4 between 6-9 mm and 8 above 9mm) formed the low-dose testing cases. All patients underwent CT colonography prior to conventional colonoscopy; standard colonoscopic preparation was given. Informed consent was obtained from all patients.

Cases D1 were acquired using a multi-detector CT (Multi Slice Helical CT; Volume Zoom, Siemens, Erlangen, Germany) using 4x1 mm detector configuration, 7 mm table feed per 0.5 s tube rotation, 0.8 mm reconstruction increment as well as 60 effective mAs and 120 kV. Somatom Sensation 16 (Siemens, Erlangen, Germany) scanned all the cases in D2, using a spiral mode with the

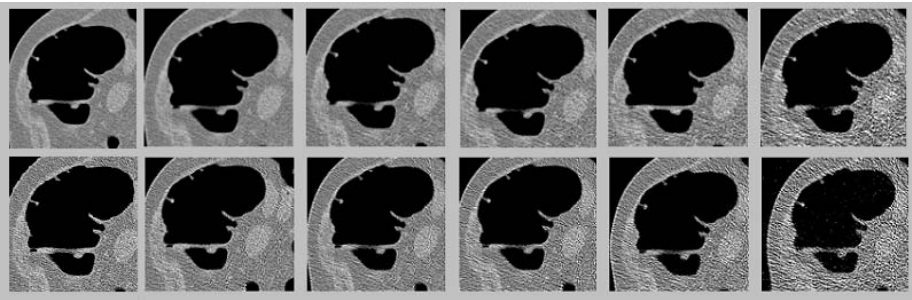


Fig. 3. Part of a CTC slice containing a 7 mm polyp, reconstructed at different mAs values: 55, 45, 35, 25, 15, 5 mAs (from left to right) and using two different kernels B40f-smoother (top row) and B30f-sharper (bottom row)

Table 1. On the left simulated acquisition and reconstruction parameters. $CTDI_w$ (weighted CTDI value) and the effective dose are averaged over the 5 patients and taking into account both prone and supine acquisitions. On the right, results of the presented CAD algorithm on low-dose data, taking into account the best possible acquisition (0.8/1.0 Bf40) and all the reconstructed datasets (Overall).

Voltage (kV)	Dose level (mAs)	$CTDI_w$ (mGy)	Eff. dose (mSv)	Dataset	Acquisition	Sensitivity	Specificity	False positives
120	55	8.60	4.93	0.8/1.0 B40f	55mAs	100%	98.65%	0.7
120	45	7.06	4.05	0.8/1.0 B40f	45mAs	100%	98.39%	0.8
120	35	5.54	3.18	0.8/1.0 B40f	35mAs	100%	96.10%	2.0
120	25	3.92	2.25	0.8/1.0 B40f	25mAs	100%	94.67%	2.8
120	15	2.80	1.61	0.8/1.0 B40f	15mAs	100%	93.42%	3.4
120	05	2.80	1.61	0.8/1.0 B40f	05mAs	100%	76.34%	12.3
Overall				Overall	55mAs	100%	95.94%	2.7
Overlap	Slice increment	Kernel		Overall	45mAs	100%	94.88%	2.8
1.0mm	0.8mm	B40f		Overall	35mAs	100%	94.84%	2.8
2.0mm	1.5mm	B40f		Overall	25mAs	100%	85.80%	7.6
2.0mm	1.5mm	B50f		Overall	15mAs	100%	80.07%	10.2
Overall				Overall	05mAs	100%	73.96%	13.9

following parameters: 16x0.75mm slice collimation, 11mm table feed, 0.5s rotation time, 55 effective mAs and 120kV. All patients were scanned in both supine and prone positions, in average breath-holds of 19 seconds (16 slice) to 30 seconds (4 slice). The CAD process was carried out as follows: after the CT data was transferred to an offline workstation (Intel Pentium 2.4 GHz system), seed points were selected manually to ensure complete colonic segmentation. The segmentation threshold was determined automatically (section 2.1). The CAD pipeline presented in Section 2 was employed on the D1 cases to estimate optimal values for T_{dist} , T_{radius} and also for computing a set of model spherical harmonics. The SpharmonicKit package (<http://www.cs.dartmouth.edu/geelong/sphere/>) is used to compute the spherical harmonic decomposition. Taking conventional colonoscopy as standard of reference, ROC curves were computed. Parameter optimization consisted in maximizing the area under the ROC curve for values of T_{dist} and T_{radius} . A leave one patient out technique was preferred to compute the ROC curves for each possible parameter combination. The maximal value for the area under the ROC curve was: 0.995, for $T_{dist} = 3mm$ and $T_{radius} = 9mm$. These optimal values of T_{dist} and T_{radius} were applied to the cases in D2.

For the database D2 the influence of effective mAs, reconstruction kernel, slice overlap and increment, on the performance of CAD, was studied. Using the Somatom Noise+(V2.0) program (provided by Siemens, Forchheim, Germany), artificial noise (simulating a low-dose acquisition) was added to the original raw-data. The resulting raw datasets were imported into the scanner and 30 reconstructions per patient generated (Table 1). Figure 3 presents the same CT slice reconstructed at different dose levels. Due to limited space only results for the optimal setup (0.8 slice increment, 1.0 overlap, kernel B40f) and overall results (taking into account all 150 reconstructions) and are shown in table 1.

The average computation time for the method is 78.18 seconds and can be further decomposed into: segmentation 7.97s, N-ary tree building 2.13s, candidates generation 27.38s, spherical harmonic computation and polyp extraction 40.70s

4 Discussion and Conclusions

The novelty of the paper is two-fold; first it proposes robust 3D models to describe colonic polyps, using the same framework, namely spherical harmonic decomposition, to capture both gray level and geometrical information. Second, the method was tested on low-dose data and proved the assumption that colonic polyps can be retrieved, without impaired sensitivity and specificity, even on data sets acquired at 15 mAs. Accordingly the dose reduction is significant, the effective dose can be reduced by a factor of 3, very important when thinking of population screening and the associated risk of inducing cancer [4].

Several improvements to the proposed scheme can be introduced. First, our single scale statistical modeling (single bandwidth $bw = 16$) can be transformed into multi-scale pyramidal modeling, to further reduce computation times. A modeling scheme based on spherical wavelets can be considered. Although better for describing smaller polyps, it is expected to be more sensitive to image noise [10].

Regarding low-dose data generation, the effects of reducing the kV of the original acquisition have to be tested. This will further reduce the effective dose. Phantom and real patient data can be used to evaluate this acquisition parameter. Finally, the use of denoising filters as proposed by Rust et al. [14] can be considered prior to CAD. While increasing the computation time, it could allow for even greater dose reductions.

To conclude, our results indicate that low-dose CTC CAD is viable. Sensitivity and specificity values remain high for the detection of clinically interesting cases (polyps larger than 5mm). Low-dose CTC CAD is a possible approach for colonic polyp screening, significantly reducing the chances of inducing tumors due to dose accumulation.

Acknowledgement. This work is part of the GOA/99/05: "Variability in Human Shape and Speech", financed by the Research Fund, K.U. Leuven, Belgium and supported by Siemens AG, Forchheim, Germany. The authors would like to thank Dr. Lutz Guendel for kindly providing the Somatom Noise+(V2.0) program.

References

1. Morson, B.C.: Factors Influencing the Prognosis of Early Cancer in the Rectum. Proc R Soc Med (1966) 59:607-8
2. Gazelle, G.S., et al.: Screening for colorectal cancer Radiology (2000) 215:327-335
3. Svensson, M.H., et al.: Patient acceptance of CT colonography and conventional colonoscopy: prospective comparative study in patients with or suspected of having colorectal disease. Radiology (2002) 222(2):337-345.
4. van Gelder, R.E., et al.: CT colonography at different radiation dose levels: feasibility of dose reduction. Radiology (2002) 224(1):25-33
5. Iannaccone R., et al.: Detection of colorectal lesions: lower-dose multi-detector row helical CT colonography compared with conventional colonoscopy. Radiology (2003) 229(3):775-781

6. Summers, R.M., et al.: An Automated Polyp Detector for CT Colonography - Feasibility Study. *Radiology* (2000) 284-290
7. Yoshida, H., Nappi, J.: 3-D Computer-Aided Diagnosis Scheme for Detection of Colonic Polyps. *IEEE Transactions on Medical Imaging* (2001) 1261-1274
8. Kiss, G., et al.: Computer Aided Detection in CT Colonography. LNCS 2878, MICCAI (2003) 746-753
9. van Wijk, C., et al.: On Normalized Convolution to Measure Curvature Features for Automatic Polyp Detection. LNCS 3216, MICCAI (2004), 200-208
10. D. Healy Jr., et al.: FFTs for the 2-Sphere - Improvements and Variations. *Journal of Fourier Analysis and Applications* 9:4 (2003), pp. 341 - 385.
11. Ballard, D.M., Brown, C.M.: *Computer Vision*. Prentice Hall (1982) 123-166
12. Wiemker R., Pekar, V.: Fast Computation of Isosurface Contour Spectra for Volume Visualization. *Proceedings Computer Assisted Radiology and Surgery CARS* (2001)
13. Press, W.H., et al.: *Numerical Recipes in C - Chapter 15* Cambridge University Press (1992)
14. Rust, G.F., et al.: Noise/dose reduction and image improvements in screening virtual colonoscopy with tube currents of 20 mAs with nonlinear Gaussian filter chains *SPIE - Volume 4683* (2002) 186-197

Photo-Realistic Tissue Reflectance Modelling for Minimally Invasive Surgical Simulation

Mohamed A. ElHelw¹, Stella Atkins², Marios Nicolaou¹,
Adrian Chung¹, and Guang-Zhong Yang¹

¹ Department of Computing, Imperial College London, UK
{mohammed.elhelw, m.nicolaou, a.chung, g.z.yang}@imperial.ac.uk

² School of Computer Science, Simon Fraser University, Canada
stella@cs.sfu.ca

Abstract. Computer-based simulation is an important tool for surgical skills training and assessment. In general, the degree of realism experienced by the trainees is determined by the visual and biomechanical fidelity of the simulator. In minimally invasive surgery, specular reflections provide an important visual cue for tissue deformation, depth and orientation. This paper describes a novel image-based lighting technique that is particularly suitable for modeling mucous-covered tissue surfaces. We describe how noise functions can be used to control the shape of the specular highlights, and how texture noise is generated and encoded in image-based structure at a pre-processing stage. The proposed technique can be implemented at run-time by using the graphics processor to efficiently attain pixel-level control and photo-realism. The practical value of the technique is assessed with detailed visual scoring and cross comparison experiments by two groups of observers.

1 Introduction

With the maturity of Virtual Reality (VR) and Augmented Reality (AR), there is now a wide spread use of the techniques in Minimally Invasive Surgery (MIS). Intra-operatively, AR enables navigation in image-guided surgery where the surgeon's view of the operating field is augmented by the addition of computer-generated data. VR, on the other hand, is increasingly being used in a mixed reality environment for surgical simulation. Pre-operatively, VR simulators are used to train practitioners in basic surgical tasks as well as complete interventions. They have been found to be particularly useful for MIS procedures because of the complexity of instrument controls, restricted vision and mobility, difficult hand-eye co-ordination, and the lack of tactile perception require a high degree of operator dexterity [1]. MIS simulations provide an economical and time saving solution for acquiring, as well as assessing, basic surgical skills [2]. They offer the opportunity for safe, repeated practice and objective measurement of surgical performance.

One of the key issues to be addressed in surgical simulation is visual realism and mechanical fidelity. The creation of realistic lighting can greatly affect visual realism and hence the overall quality of the simulation. Traditional methods have found limited use in advanced surgical training, particularly for those involving patient

specific models. To this end, a number of physically-based approaches for computing diffuse and specular light components have been proposed. Diffuse reflection modelling considers diffuse reflection as the result of light entering the surface, scattering, interacting with its material and then exiting in random directions [3]. Relative contributions of surface reflection and sub-surface scattering and transmission are computed, therefore allowing for simulating layered materials and biological tissues. For specular reflection, a number of micro-facet models are used for simulating specular surfaces [4], [5]. These models consider rough surface topology with perfect micro-reflectors accounting for the spectral composition of specular highlights. They are further extended to handle wide-ranging surface types [6] and anisotropic distributions with multiple scattering. Despite the visual realism achieved, physically-based reflection modelling is not particularly suitable for interactive surgical simulation because of its inherent high computational costs.

Existing research has shown that specular highlights constitute a major visual cue for gauging tissue deformation, depth and orientation during MIS procedures [7]. Reproducing the same effect in a simulated environment is difficult due to the nature of tissue, which is comprised of multiple layers covered with translucent mucous. The purpose of this paper is to present a novel photorealistic lighting method for simulating the appearance of specular highlights reflected from the internal lumen. The method is based on using noise functions to modulate the shape of specular highlights and imitate the effects of composite multi-layer tissue types on reflected lighting. We will describe how texture noise is generated and encoded in image-based structures at a pre-processing stage and show how this information is used during run-time by the graphics processor to efficiently attain pixel-level control and achieve photorealistic tissue appearance. The visual realism achieved by the technique is assessed by two groups of observers with detailed visual scoring and cross comparison.

2 Method

2.1 Surface Reflectance Modeling

The use of computer-generated noise is a well-established method for representing natural scenes such as terrain topography, clouds, vegetation distributions, fur/hair, as well as natural effects such as fire, fluid flow and erosion patterns [8-11]. In this study, the modeling of specular highlights is achieved by using 2D Perlin noise [12], which first creates a noise image followed by converting it to a reflectance map. For each point in the noise image, the associated tangent vectors to the surface are defined, and each pixel in a reflectance map encodes the direction that the corresponding point is facing by using three channels of information that are conveniently mapped to a standard Red-Green-Blue (RGB) image.

Since a reflectance map is essentially an RGB image, per-triangle noise information is extracted during run-time by means of texture mapping for calculating the specular highlights. The extracted noise vectors are defined in image coordinate space, so they have to be transformed into a coordinate that is local to the triangle being processed. The basis vectors of the triangle local coordinate system, known as the surface-local coordinate space, can be defined by using the surface tangent (T), bi-tangent (B), and the normal (N), where the first two vectors can be computed from the

partial derivatives of the object-space coordinates of the triangle in terms of its texture coordinates [13], *i.e.*,

$$T = (\partial x / \partial u, \partial y / \partial u, \partial z / \partial u) = (-T_0, -T_1, -T_2) \tag{1}$$

where

$$\begin{aligned} T_0 &= [(v_1 - v_0)(x_2 - x_0) - (x_1 - x_0)(v_2 - v_0)] / [(u_1 - u_0)(v_2 - v_0) - (v_1 - v_0)(u_2 - u_0)] \\ T_1 &= [(v_1 - v_0)(y_2 - y_0) - (y_1 - y_0)(v_2 - v_0)] / [(u_1 - u_0)(v_2 - v_0) - (v_1 - v_0)(u_2 - u_0)] \\ T_2 &= [(v_1 - v_0)(z_2 - z_0) - (z_1 - z_0)(v_2 - v_0)] / [(u_1 - u_0)(v_2 - v_0) - (v_1 - v_0)(u_2 - u_0)] \end{aligned}$$

and

$$B = (\partial x / \partial v, \partial y / \partial v, \partial z / \partial v) = (-B_0, -B_1, -B_2) \tag{2}$$

where

$$\begin{aligned} B_0 &= [(x_1 - x_0)(u_2 - u_0) - (u_1 - u_0)(x_2 - x_0)] / [(u_1 - u_0)(v_2 - v_0) - (v_1 - v_0)(u_2 - u_0)] \\ B_1 &= [(y_1 - y_0)(u_2 - u_0) - (u_1 - u_0)(y_2 - y_0)] / [(u_1 - u_0)(v_2 - v_0) - (v_1 - v_0)(u_2 - u_0)] \\ B_2 &= [(z_1 - z_0)(u_2 - u_0) - (u_1 - u_0)(z_2 - z_0)] / [(u_1 - u_0)(v_2 - v_0) - (v_1 - v_0)(u_2 - u_0)] \end{aligned}$$

In the above equations, (x_0, y_0, z_0) , (x_1, y_1, z_1) , (x_2, y_2, z_2) and (u_0, v_0) , (u_1, v_1) , (u_2, v_2) represent the object- and texture-space coordinates respectively for the prospective triangle. Subsequently, (N) can be calculated from the cross product of (T) and (B) or alternatively the normal supplied by the original model can be used. By computing the basis vectors of the surface-local coordinate space, the GPU can then be used to efficiently transform the extracted image-space noise vectors to the surface-local space and carry out specular highlight computations.

In order to evaluate the effects of adding specular highlights on visual realism, subject-specific textures of 3D models were derived. Video bronchoscope images (Olympus BF Type; with field of view 120°) were registered with CT scans (Siemens Somatom Volume Zoom four-channel multi-detector) for five patients. This enabled the exact camera pose of the bronchoscope examinations to be identified. The surface details including texture and shading parameters are also extracted. The texture map is derived directly from the video bronchoscope images. The shading parameters are recovered by modeling the bidirectional reflectance distribution function (BRDF) ρ of the visible surfaces by using a cubic curve parameterized on γ , the cosine of the angle between the viewing vector (V) and surface normal (N) [14]:

$$\rho(V, N) = \sum_0^3 c_i B_i^3(\gamma)$$

where

$$\gamma = \frac{V \cdot N}{\|V\| \|N\|}, B_i^n(t) = \frac{n!}{i!(n-i)!} (1-t)^{(n-i)} t^i \tag{3}$$

2.2 Assessment of Visual Realism

A user study was conducted to evaluate the effects of the proposed surface reflectance modeling for added visual realism. Participants from two subject groups were

considered, where the first group consisted of 16 computing science students, and the second group consisted of 7 experienced bronchoscopists who had each performed between 150 and 2000 endoscopy procedures. All subjects had normal or corrected vision and were assumed to have good comprehension and analytical skills.

Static images on a computer monitor were presented to each subject for evaluation, one image at a time. The images show views of the bronchial tree from 8 poses from 5 different patients (estimated from camera parameters and used for generating synthetic views). Five images were created for each pose comprising five different categories ranging from least realistic to real. Category 1 uses the conventional shading algorithm that represents the lowest quality in the evaluation scale, whereas Category 5 is a real captured video bronchoscope image. Category 3 represents images rendered with subject-specific texture extracted with the BRDF method described above after 2D/3D registration. Category 4 is as category 3 but with specular highlights added by using the proposed technique to improve visual realism. For Category 2, low-resolution surface texture was used. The participants were first shown two examples of category 1 (most unrealistic) and category 5 (real) images displayed side-by-side for visual calibration. Then the subjects were presented with series of 15 images and asked to rank each image in terms of visual realism by following the Likert scale (1 to 5). A two alternative forced choice (2AFC) experiment was also conducted. In this experiment, the subjects viewed 23 side-by-side image pairs from categories 3, 4 and 5, always from different poses, and were asked to choose which of the two (left or right) was the most realistic. It is worth noting that during the experiments, no time limits were imposed and the images were displayed in a random order. The subjects were not told which the real images were, nor were they informed how the images were obtained.

3 Results

Figure 1 depicts the results obtained by implementing the described method in Cg [15] on NVIDIA FX graphics hardware compared with the conventional OpenGL multi-texturing approach. It can be seen that the method effectively avoids the problem of plastic-like surface appearance and provides realistic specular highlights. Furthermore, by varying the colour of the mucous layer and using different noise types, tissue appearance can be modified. The bottom row of Figure 1 further demonstrates the effect of different noise functions on the visual appearance of the rendered surface, where the frequency of the noise function is increased (left to right).

In Figure 2, the general appearance of the five category images used in the user study for assessing visual realism is illustrated. This demonstrates the increased visual realism that is achieved by using the proposed algorithm. As mentioned earlier, the last image category represents the real bronchoscope image.

Figure 3 summarizes the mean score for all the images of each category, averaged over all the participants for each group. It is seen that the overall score of all the subjects shows a steady increase in the scored realism using the proposed method for enhancing visual realism. It is also evident that the expert group is not significantly different to the naïve group in judging realism. This shows that results using naïve subjects to test for realism of tissue samples may transfer to training simulators for physicians.

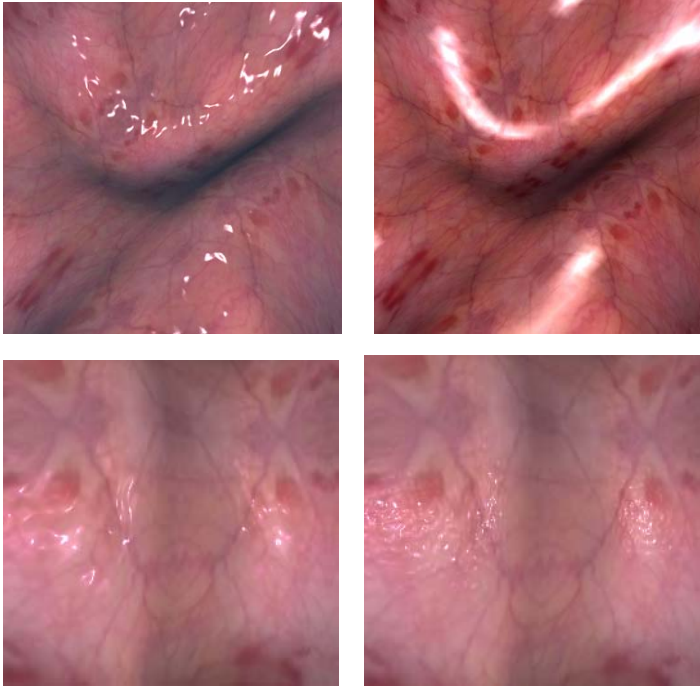


Fig. 1. (Top) Two views of the same surface rendered by using the proposed method (left) versus OpenGL multi-texturing (right). Notice the plastic-like surface and the hexagonal shape of the specular highlights rendered with the multi-texturing method (Bottom) By varying the frequency of the noise function, different specular appearance of the tissue can be simulated.

A more detailed statistical summary is provided in Table 1, which uses the combined realism scores from all the participants for images of Categories 3, 4 and 5. By applying a paired two-tailed *t*-test to the data, synthetic images of Category 3 without specular highlights had a significantly lower score than the images with specular highlights added ($p=0.057$) and also a significantly lower score than the real images ($p=0.026$). It is also shown that there is no significant difference between synthetic images with highlights and real images ($p=0.780$) which indicates that the Category 4 specular highlights can be perceived as close to photorealism.

For the 2AFC test mentioned above, it was found that when comparing a Category 3 image side-by-side with a Category 4 image, the Category 4 image was selected as the most realistic 68% of the time. Therefore missing specular highlights reduced the users' perception of reality. This result was confirmed when comparing Category 3 images with real (Category 5) images, where the real images were selected as being more realistic 73% of the time. However, when comparing Category 4 images with real images, the real images were selected as being more realistic only 48% of the time. This shows how the addition of noise-based specular highlights positively affected users' judgment of realism.

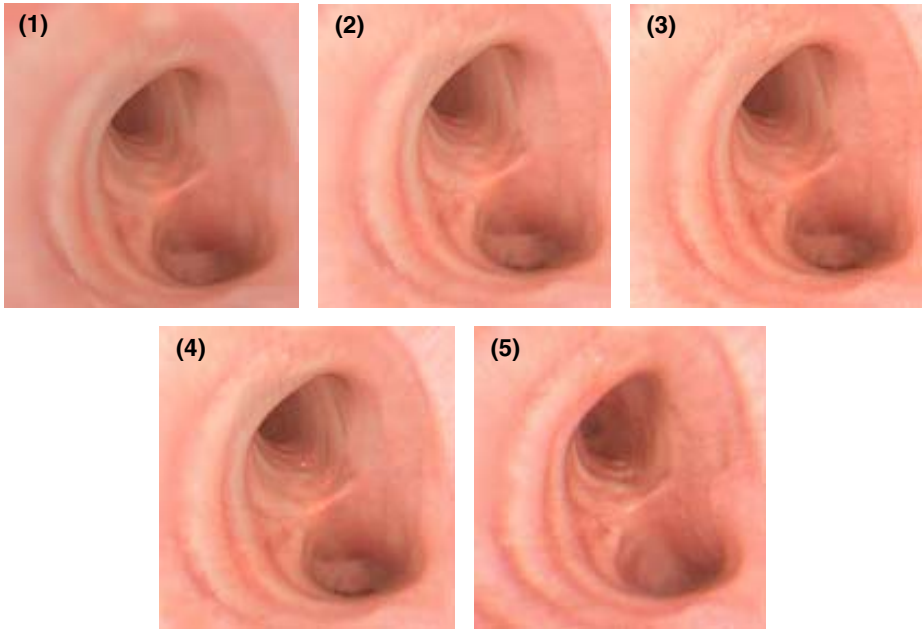


Fig. 2. The five different categories used as the stimuli for one pose are defined as follows (1) Unreal (2) BRDF-low resolution (3) BRDF (4) BRDF-Specular (5) Real bronchoscope image

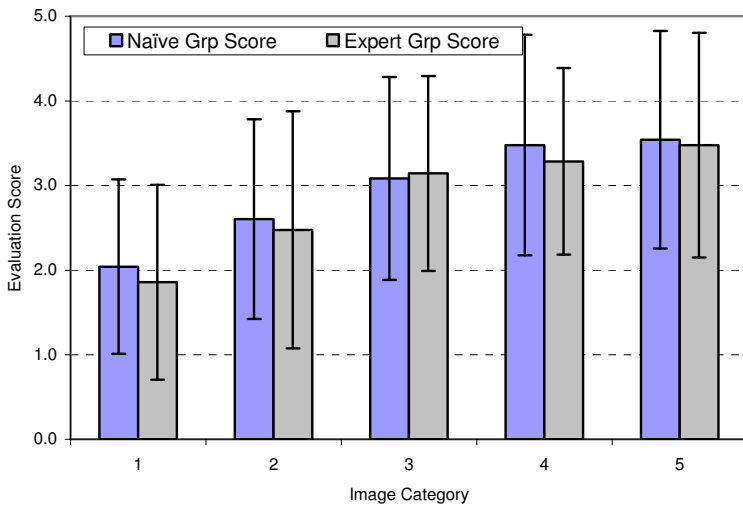


Fig. 3. Average score for all the images of each category, averaged over all the participants for each group. Error bars show one standard deviation

Table 1. Details of the *t*-test with the scores of images from Categories 3, 4 and 5

	Cat 3 Vs. Cat 4		Cat 3 Vs. Cat 5		Cat 4 Vs. Cat 5	
	Cat3	Cat4	Cat3	Cat5	Cat4	Cat5
Mean	3.08	3.48	3.08	3.54	3.48	3.54
Variance	1.44	1.70	1.44	1.66	1.70	1.66
Hypothesized Mean Diff	0.00		0.00		0.00	
<i>t</i> Stat	-1.94		-2.30		-0.28	
<i>p</i> ($T \leq t$) two-tail	0.057		0.026		0.78	
<i>t</i> Critical two-tail	2.01		2.01		2.01	

4 Conclusion

In this paper, a novel photorealistic rendering method suitable for MIS simulation is described. The rendering method is based on using reflectance maps to model the effect of surface details and mucous layer on overall tissue appearance. For efficiency, reflectance map generation is achieved in pre-processing. During run-time, the graphics processor is used to allow for per-pixel control and rendering efficiency. The visual assessment results from the user study have shown that when combined with subject specific texture extraction through a BRDF model, the difference between images derived from the proposed method and real video bronchoscope is minimal. The results have also shown that the visual score derived had little difference between the naïve and expert groups, thus highlighting the potential value of the technique for both basic and advanced surgical skills training and assessment.

References

1. Bro-Nielsen, M.: Simulation Techniques for Minimally Invasive Surgery. *Journal of Minimally Invasive Therapy & Allied Technologies* (1997) 106-110
2. Shah, J., Darzi, A.: Simulation and Skills Assessment. *International Workshop on Medical Imaging and Augmented Reality (MIAR '01)*, Hong Kong (2001) 5-9
3. Hanrahan, P., Wolfgang, K.: Reflection from Layered Surfaces due to Subsurface Scattering. In Kajiyu, J.T. (ed.): *Computer Graphics (Proceedings of ACM SIGGRAPH '93 Conference)*, Vol. 27. Annual Conference Series, ACM Press, New York (1993) 165-174
4. Blinn, J.F.: Light Reflection Functions for Simulation of Clouds and Dusty Surfaces. In Bergeron, R.D. (ed.): *Computer Graphics (Proceedings of ACM SIGGRAPH '82 Conference)*, Vol. 16. Annual Conference Series, ACM Press, New York (1982) 21-29
5. Cook, R.L., Torrance, K.E.: A Reflection Model for Computer Graphics. *ACM Transactions on Graphics*, Vol. 1, No.1 (1982) 7-24
6. He, X.D., Torrance, K.E., Sillion, F.X., Greenberg, D.P.: A Comprehensive Physical Model for Light Reflection. In Thomas, J.J.(ed.): *Computer Graphics (Proceedings of ACM SIGGRAPH '91 Conference)*, Vol. 25. Annual Conference Series, ACM Press, New York (1991) 175-186

7. Neyret, F., Heiss, R. Senegas, F.: Realistic Rendering of an Organ Surface in Real-Time for Laparoscopic Surgery Simulation. *The Visual Computer*. Vol.18, No.3 (2002) 135-149
8. Perlin, K., Hoffert, E.: Hypertexture. In Thomas, J.J.(ed.): *Computer Graphics (Proceedings of ACM SIGGRAPH '89 Conference)*, Vol. 22. Annual Conference Series, ACM Press, New York (1989) 253 - 262
9. David, S.E. (ed.): *Texturing and Modeling: A Procedural Approach*. Morgan Kaufmann (2002)
10. Prusinkiewicz, P. Lindenmayer, A.: *The Algorithmic Beauty of Plants*. Springer-Verlag, New York (1990)
11. Weber, J., Joseph, P.: Creation and rendering of realistic trees. In Mair, S.G., Cook, R. (eds.): *Computer Graphics (Proceedings of ACM SIGGRAPH '95 Conference)*. Annual Conference Series, ACM Press, New York (1995) 119-128
12. Perlin, K.: An Image Synthesizer. In Cole, P., Heilman, R., Barsky, B.A. (eds.): *Computer Graphics (Proceedings of ACM SIGGRAPH '85 Conference)*, Vol.19, No.3. Annual Conference Series, ACM Press, New York (1985) 287-296
13. Fernando, R., Kilgard, M.J.: *The Cg Tutorial*. Addison Wesley (2003)
14. Chung, A.J., Deligianni, F., Shah, P., Wells, A., Yang, G.Z.: Enhancement of Visual Realism with BRDF for Patient Specific Bronchoscopy Simulation. In Barillot, C., Haynor, D.R., Hellier, P. (eds.): *Lecture Notes in Computer Science*, Vol. 3217. Springer-Verlag, Berlin Heidelberg New York (2004) 486 – 493
15. Mark, W.R., Glanville, R.S., Akeley K., Kilgard, M.J.: Cg: A System for Programming Graphics Hardware in a C-Like Language. *ACM Transactions on Graphics*, Vol. 22, No.3 (2003) 896-907

Motion Tracking and Intensity Surface Recovery in Microscopic Nuclear Images

Lopamudra Mukherjee¹, Mingen Lin¹, Jinhui Xu¹, and Ronald Berezney²

¹ Department of Computer Science, State University of New York at Buffalo, USA
{lm37, mlin6, jinhui}@cse.buffalo.edu

² Department of Biological Sciences, State University of New York at Buffalo, USA
berezney@buffalo.edu

Abstract. Current techniques for microscopic imaging do not provide necessary spatial and temporal resolutions for real time visualization of the nucleus. Images can only be acquired in time lapse mode, leading to significant loss of information between image frames. Such data, if available, can be extremely helpful in the study of nuclear organization and function. In this paper, we present a gamut of geometric-technique-based approaches for solving the problem. Our techniques, working together, can effectively recover complicated motion and deformation as well as the change of intensity surfaces from pairs of images in a microscopic image sequence, and has low time complexity, particularly desirable by many biological applications where large amount of DNA need to be processed. These techniques are also readily applicable to other types of images for reconstructing motion and intensity surfaces of deformable objects.

1 Introduction

A clear understanding of the mammalian cell nucleus and its exquisite architecture which governs gene expression and regulation, remains the final frontier in modern nuclear biology. To this end, recent developments in microscopy, imaging and labelling have led to a significantly improved view of the cell nucleus. This technology has enabled the observation of DNA replication and transcription sites and their organization into higher order units or “zones”. Researchers believe that this dynamic interplay and “re-zoning” may form the structural basis for elaborate global coordination of replicational and transcriptional programs within the mammalian nucleus[1]. However, despite significant progress in microscopic imaging technologies, the task of analyzing the nuclear images to derive semantic (and biologically relevant) information is rather difficult. One key impediment is the presence of enormous amount of chromatin in the nucleus and their complicated movement (non-uniform translation, rotation and deformation). This makes it extremely difficult for biologists to do real time visualization of such nuclear activities. This problem is further compounded by the limited spatial and temporal resolution of current microscopic techniques which introduce under-sampling or “gaps” between acquired images. Biological analysis of such data is generally the result of long sessions involving manual calculations and hypothesis (based on experience), making it human error prone

and very tedious. These problems motivate the design of a fully-automated set of techniques to provide biologists real time visualization of nuclear activities within the scope of existing imaging modalities.

The problem of recovering motion and deformation of moving objects, given only a few snapshots of the objects during their motion, has been a challenging problem in various fields such as computer vision and pattern recognition, and robotics[2, 3]. Most of these approaches provide interesting insights into possible interpolation strategies and work well for the specific objects, images or assumptions they were designed for. Unfortunately, these techniques do not extend well to microscopic images for a number of reasons. They were primarily designed for large sized objects where sufficient information about the mechanics and physics behind the motion is available or can be assumed. For genomic structures, however, where no such data is available, it becomes necessary to exploit the limited information hidden in the images in an effort to recreate an accurate motion of the objects under consideration.

In this paper, we present a suite of geometric techniques that address the problem of “gaps” between microscopic images. We suggest techniques to determine the shape and motion of genomic structures for any number of intermediate (or unavailable) time points in a 2D image sequence. We further extend our idea by not only regenerating the contours but also the continuous changes of intensity surfaces from one image to another. This allows us to generate as many “missing” in-between image frames as desired given as few as two time separated image frames. When a number of nuclear images acquired in time-lapse mode are available, our techniques allow an online recreation of a “movie” sequence, enabling a real-time visualization of nuclear processes with negligible or no manual interaction. These provide a far less tedious biological interpretation of the imaging data.

2 Method

In this section, we overview our method for motion tracking and intensity surface recovery of a set of objects with prior knowledge of only their initial and final positions, shapes and intensities. We illustrate our results on microscopic images of the cell nucleus.

2.1 Simplification and Segmentation

We begin with a sequence of 2D microscopic images of the same set of genomic structures (representing chromatin domains or chromatin foci) taken at different time points. One major problem with microscopic images is that they are blurred in portions by clouds of quasi-foci (formed by overlapping of neighboring foci) which complicates the process of segmentation and intensity recovery. Therefore our first goal is to simplify the images to make segmentation easier. In a recent work [4], we reported our observations of an interesting phenomena that showed that image profiles of chromatin foci, even in the presence of noise,

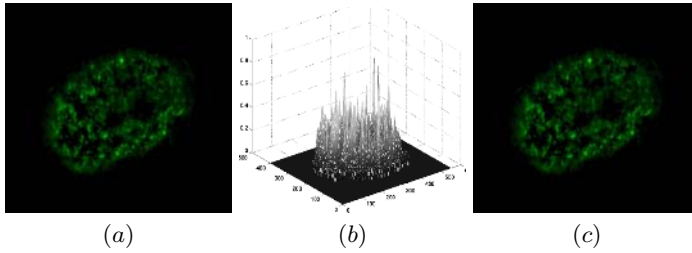


Fig. 1. (a) Original cell nuclear image. (b) The corresponding intensity profile. (c) Approximated cell nuclear image.

resembled certain exponential distributions. We extend that idea to make the task of segmentation easier by preceding it with a simplification process.

The main idea of the simplification is to determine a set of *approximation functions* for approximating the intensity surface of the image within a certain error tolerance ϵ of its original intensity. More specifically, we solve the following problem: Given a set $P = \{p_1, p_2, \dots, p_n\}$ of pixels with each p_i associated with an intensity value $I_i > 0$, find a small set $F = \{f_1(\cdot), f_2(\cdot), \dots, f_k(\cdot)\}$ of functions with each $f_i(\cdot)$ defined by a generalized normal distribution surface such that the intensity I_i of each pixel p_i is closely approximated by the maximum value of the k functions at p_i , i.e., $|I_i - \max_{j=1}^k f_j(p_i)| \leq \epsilon$ (or $|I_i - \max_{j=1}^k f_j(p_i)| \leq \epsilon I_i$) for some small constant $\epsilon > 0$. Each of these functions $f_i(\cdot)$ follows a generalized 2-D normal distribution $f_i(X) = \sigma e^{-(X-\mu)^T Q (X-\mu)}$, where Q is a positive definite matrix (i.e., the intersection of $f_i(\cdot)$ with a plane $z = c$ is an arbitrarily oriented ellipse). The details of the algorithm are omitted due to lack of space. Experimental results have shown that the approximated image generated not only closely resembles the original image (see Fig 1) but makes segmentation and further processing much more tractable since it is just represented by a set of mathematical functions.

The approximated image, free of diffused signals, is then segmented using relative thresholds (w.r.t. the peak value of the normal distribution functions) to shrink the domain of each approximation function and then a merge segmentation method [5] is used to generate the boundary of each foci. This approach shows significantly better results over commonly used segmentation techniques (such as threshold based, edge-detection based, watershed, etc.), especially for images which contain many small-sized objects or whose intensity varies significantly over the scope of the images. However, even after the images are approximated and segmented, the contours may still contain many undesired furrows (undulations or small recesses) caused by other types of unavoidable noise on the periphery. These furrows not only hide the actual boundary of the genomic structures but also increase the processing time dramatically. Taking this into account, we apply our contour simplification algorithm, proposed in [6] to simplify and smooth the boundary within a predecided error tolerance. This process yields the set of contours of both the initial and final images modeled as polygons.

2.2 Motion Tracking

Once the genomic structures have been represented as polygonal contours, the next step is to determine the correspondence between the initial and final polygonal sets. Previous methods to solve this problem have relied on manual estimation or heuristic techniques based on additional feature information [7]. Such approaches do not work directly for our problem, as we need to globally optimize the results for large sets of polygons, as well as consider the shape of individual objects while determining the correspondences. To overcome this difficulty, we propose a technique which takes both the shapes and relative positions of polygons into account and is also tailor-made for the problem at hand.

In determining correspondences between the two polygonal sets, our technique first tries to find a global transformation which moves every polygon in the initial set, P , within a small neighborhood of some polygon in the final set, Q . This step is useful because once such a transformation is determined, local neighborhoods can be examined for proximity. For example, if a transformed polygon p_i of P , overlaps with q_j of Q (or the neighborhood of q_j), it is an indication that q_j may be a possible correspondence for p_i . The notion of neighborhood can be thought of as a square of radius R (or any bounded region) centered roughly in the middle of each $q_j \in Q$. Clearly, when the neighborhood R is large, a global transformation satisfying the criteria above can be easily found, hence a large R is always feasible. To obtain better accuracy, we employ a binary search procedure on R , to determine the smallest *feasible* R , R_{min} .

The above procedure yields a reasonably small value of R_{min} . It also gives for each $p_i \in P$ (or $q_i \in Q$), one or more probable candidates for correspondence in Q (or P) based on R_{min} and the overlap criteria. We follow up this procedure with another refinement process by modelling the problem as the well-known Maximum Bipartite Graph Matching problem. In our case, the vertex sets are corresponding polygons of P and Q and the edge weights are the distances between individual polygons in these two subsets.

Correspondences established using the above techniques are generally accurate and comparable with the results obtained from other well-established techniques like Nearest Neighbour. Correspondences obtained are mostly one-to-one, but in some cases, they may be one-to-many or many-to-one. We attribute this to the chromatin domains disappearing or one domain splitting into a few smaller chromatin domains. These exceptional cases can be reduced to a general case of a one-to-one correspondence of polygons by employing additional routines such as Voronoi diagrams.

Once the global transformations have been determined, we have essentially taken care of capturing the movement of the cell nucleus as a whole. However, each individual chromatin foci have their independent position (rigid) and shape (non-rigid) change which our technique still needs to address.

To account for the rigid transformations, we first scale the two polygons so that they have comparable sizes. We then use a curve matching algorithm to compute the best translation and rotation so that after these transformations, the contours of the two corresponding chromatin domains have the least differ-

ence in shape, measured by the Hausdorff distance metric[8]. Given two polygons P and Q , we find the Minkowski sum (union) of a square of size r with Q , which denotes an inner and outer boundary around Q enclosing a region S . The image $T(P)$ of P , under an appropriate transformation T , should always lie in S . In essence, we set a limit on the value of Hausdorff distance, denoted by r and then determine a transformation T such that the Hausdorff distance between $T(P)$ and Q falls within that limit. After rotating Q by an appropriate angle, finding the adequate translation is realized by determining for each edge e of P a set of translations, under which e will always lie within S and then finding the intersection of all such sets in translation space. When no such transformation exists i.e., if there is no intersection, we modify the initial value of r and repeat the above procedure till a transformation is obtained for the minimum possible value of r .

Having found the best match under rigid body transformations, we then proceed to determine the non-rigid shape changes by a routine called shape-deformation. Shape-deformation procedure first computes one or a few central points in the common intersection of Q and $T(P)$ so that all points in $Q \cup T(P)$ can be guarded (i.e., visible to) by at least one of the central points (such points always exists, otherwise we can transform one of the polygons to overlap the other, obtaining a match with a smaller Hausdorff distance, which is not possible). For each central point o , we draw line segments from o to all visible vertices and intersections of the polygons. Any pair of such line segments, say the i^{th} pair, encloses a portion of the boundary, P_i and $T(Q)_i$, of the polygons Q and $T(P)$, respectively. The technique then proceeds to determine the image of $T(P)_i$ under the transformation $S_t(T(P)_i, s, \theta)$, where t is an intermediate time point between the initial and final time point and s and θ are the scaling factor and angle of rotation for the local scaling and rotation which resulted in $T(P)_i$ in being transformed to Q_i . Ultimately, all such local transformations $S_t(T(P)_i)$ are combined to yield the deformed shape of $T(P)$ at time point t .

Combining the rigid body transformations with the shape deformation, our approach obtains a continuous motion and deformation of the whole set of polygons (or contours) in two consecutive images (see Figure 2).

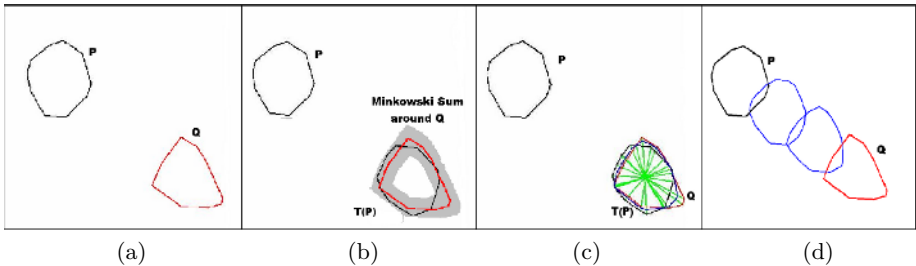


Fig. 2. (a) Initial and Final positions of a polygon in case of one-to-one correspondence (b) Rigid Transformations (c) Shape deformation around a central point (d) Shapes and positions of the deforming object at a sequence of intermediate time points

2.3 Intensity Surface Recovery

Once the contours for all desired intermediate images have been generated, we need to recover the intensity surfaces corresponding for each pair of chromatin domains in P and Q . Recall that the intensities of these chromatin domains can be approximated by a set of normal distribution functions. (see Section 2.1). Our objective is to reconstruct for each intermediate contour, the set of normal distribution functions that should represent its intensity surface. In doing so, we derive inspiration from the fact that most 2D functions retain their spatial properties under rigid transformations like translation, rotation or scaling. Also, the shape deformation routine for a particular region can be reduced to a set of unique rigid transformations localized to smaller sub-regions, this holds true for the mathematical functions defining every sub-region. The context of this statement will become clear when we describe the algorithm.

Before discussing the details of the algorithm, we will formally define the notations. Let F_P and F_Q be the set of normal distribution functions for a corresponding pair of chromatin domain P and Q respectively. Consider $T_{P \rightarrow Q}$ to be the transformation which when applied to P , yields Q . This transformation is symmetrical i.e., $T_{P \rightarrow Q}^{-1}$ when applied to Q will give P . By the same token, the transformations for generating the k th intermediate image P_k from P can be represented as $T_{kP \rightarrow Q}$.

Our algorithm works on the principle that the domains of normal distributions are equally affected by transformations such as rotation, translation, scaling and shape-deformations, as are the contours containing them. Hence we use the transformations generated previously, to obtain the domains of normal distributions at the intermediate time points. Also, their intensities are likely to be influenced by the intensities of both the initial and final images. Therefore, we generate the intermediate domains of the normal distributions from either side.

The distribution functions of F_P (and F_Q) are first projected on to the xy plane to yield a set of regions, represented as D_P (and D_Q), as shown in Fig. 3. Though this procedure is “lossy”, the outer periphery of the union of the members of D_P and D_Q makes up P and Q . $T_{kP \rightarrow Q}$ is then applied to each member of D_P which results in a set of 2D domains. Similarly, $T_{kQ \rightarrow P}$ is applied to D_Q to obtain another set of domains. These two set of domains can be thought

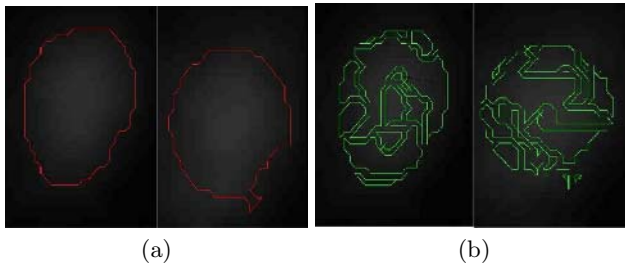


Fig. 3. (a) Two corresponding contours (b) Their normal distribution domains projected on a x-y plane(z axis represent the intensity values)

of as “clones” of D_P and D_Q under transformations in the forward and reverse directions and cover the same region bounded by the contour of $T_{k_{P \rightarrow Q}}(P)$. At this stage, for each pixel inside the contour $T_{k_{P \rightarrow Q}}(P)$, we run a simple check for membership in both $T_{k_{P \rightarrow Q}}(D_P)$ and $T_{k_{Q \rightarrow P}}(D_Q)$. Note, that owing to the construction, a pixel has to belong to at least one domain in both of these domain sets. Once the membership is established, we refer to their corresponding “clones” in D_P and D_Q and the pixel intensity is then expressed as an affine combination of the intensity values originally associated with the untransformed clones i.e. for a pixel p_{k_l} , if $I_{D(P)_i}$ and $I_{D(Q)_j}$ be the intensity values associated with the clones $D(P)_i$ and $D(Q)_j$ of its parent domain, then $I_{p_{k_l}} = \alpha * I_{D(P)_i} + (1 - \alpha) * I_{D(Q)_j}$ where α is the ratio of time intervals from the initial to the intermediate time point and the intermediate to the final time point.

Preliminary results indicate that intensity surfaces can be efficiently recovered using this technique and matches the actual intensity surface very well.

3 Results

We implemented our algorithms in C++ using CGAL-2.4 and LEDA Version 4.4 on a 1.6 GHz machine running RedHat Linux Enterprise Version 2.0. We tested our technique using images of chromatin domains taken at an interval of 3 secs. Analysis of the results reveal that for 90% of the chromatin domains, the average movement is in the range of 2-4 pixels(Fig 4a) and deformation of the domains measured by the Hausdorff distance is very small(in the range of .01 pixel unit) for about 98% of the chromatin domains (Fig 4b). These results obtained are consistent with results of independent experiments performed on same types of images and hence pass a quantitative evaluation of the correctness of our technique. Here we can show only small portion of the results due to lack of space. Figures 5 (a) and (g) show the segmented images of chromatin domains in the nucleus of a living cell at time 0 sec and time 3 sec, which are the initial and final time points respectively. The images generated at intermediate

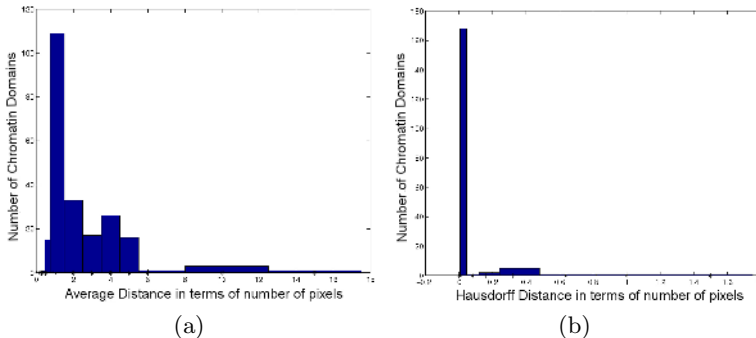


Fig. 4. Histograms showing (a) Average movement and (b) Hausdorff Distance calculated, vs number of Chromatin Domains for images at an interval of 3 secs

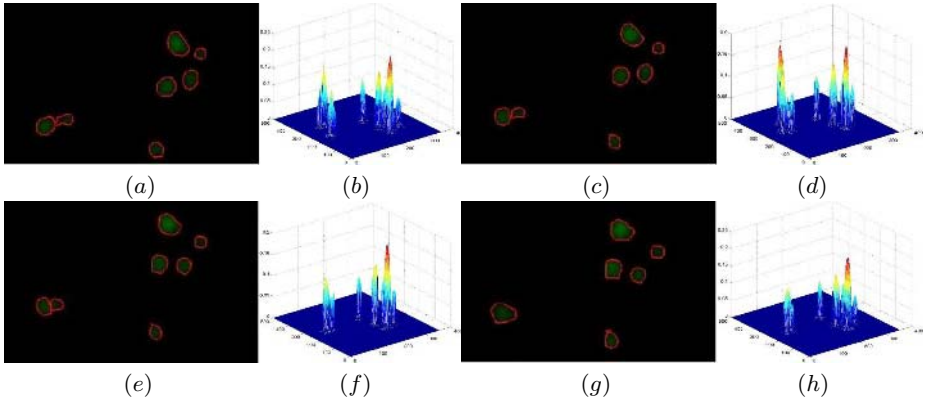


Fig. 5. (a-b) Segmented image at 0 sec with intensity plot (c-d) Image generated at 1 sec with intensity plot (e-f) Image generated at 2 sec with intensity plot (g-h) Segmented image at 3 sec with intensity plot

time points 1 sec and 2 sec are shown in Figures 5(c) and (e). The algorithm can be used to generate continuous motion by generating as many intermediate images as desired. This serves the objective of providing biologists with movies depicting the processes involving genomic structures taking place inside the cell nucleus in an online real time fashion.

4 Discussion

We have presented a novel geometric-technique-based approach for recovering continuous motion, deformation and intensity surface for genomic structures in microscopic nuclear images of living cells. A key feature of our approach is to use normal distribution functions to capture the main structure of each nuclear image, to segment genomic structures from diffused signals, and to recover the intensity surfaces of the segmented genomic structures. Our technique considers both rigid and non-rigid motion and does not require detailed knowledge about the physics and mechanics of the chromatin domains considered. Also, our techniques have low time complexity which is particularly desirable for biological images involving a large amount of data.

References

1. R. Berezney, X.W.: The new paradigm: integrating genomic function and the nuclear architecture. In: *J. Cell Biochem Suppl.* Volume 31. (1998) 238–342
2. D. Metaxas, D.T.: Shape and nonrigid motion estimation through physics-based synthesis. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence.* (1993) 580–591

3. L. Zhou, C.K.: Hierarchical structure and nonrigid motion recovery from 2d monocular views. In: CVPR' 00. Volume 2. (2000) 752–759
4. Mingen Lin, Jinhui Xu, S.B.R.B.: Surface approximation based image simplification and applications. In: IEEE Systems, Man and Cybernetics. (2004)
5. Sonka, M., Hlavac, V., Boyle, R.: Image Processing, Analysis and Machine Vision. Brooks/Cole, (Pacific Grove, CA, USA)
6. Lopamudra Mukherjee, Jinhui Xu, A.P.R.B.: Efficient algorithms for motion and deformation recovery with biological applications. In: IEEE Systems, Man and Cybernetics. (2004)
7. Longin Jan Latecki, R.L.: Shape similarity measure based on correspondence of visual parts. IEEE transactions on Pattern Analysis and MACHine Intelligence **22** (2000) 1185–1190
8. H. Alt, B.B., Blomer., J.: Approximate matching of polygonal shapes. In: Ann. Math. Artif. Intell. Volume 13. (1995) 251–266

Towards Automated Cellular Image Segmentation for RNAi Genome-Wide Screening

Xiaobo Zhou^{1,2}, K.-Y. Liu^{1,2}, P. Bradley³, N. Perrimon³,
and Stephen TC Wong^{1,2,*}

¹ Center for Bioinformatics, Harvard Center for Neurodegeneration and Repair,
Harvard Medical School, 3rd floor, 1249 Boylston, Boston, MA 02215
² Department of Radiology, Brigham and Women's Hospital, Boston, MA 02115
³ Department of Genetics, Howard Hughes Medical Institute,
Harvard Medical School, MA 02115, USA
wong@crystal.harvard.edu

Abstract. The Rho family of small GTPases is essential for morphological changes during normal cell development and migration, as well as during disease states such as cancer. Our goal is to identify novel effectors of Rho proteins using a cell-based assay for Rho activity to perform genome-wide functional screens using double stranded RNA (dsRNAs) interference. We aim to discover genes could cause the cell phenotype changed dramatically. Biologists currently attempt to perform the genome-wide RNAi screening to identify various image phenotypes. RNAi genome-wide screening, however, could easily generate more than a million of images per study, manual analysis is thus prohibitive. Image analysis becomes a bottleneck in realizing high content imaging screens. We propose a two-step segmentation approach to solve this problem. First, we determine the center of a cell using the information in the DNA-channel by segmenting the DNA nuclei and the dissimilarity function is employed to attenuate the over-segmentation problem, then we estimate a rough boundary for each cell using a polygon. Second, we apply fuzzy c-means based multi-threshold segmentation and sharpening technology; for isolation of touching spots, marker-controlled watershed is employed to remove touching cells. Furthermore, Voronoi diagrams are employed to correct the segmentation errors caused by overlapping cells. Image features are extracted for each cell. K-nearest neighbor classifier (KNN) is employed to perform cell phenotype classification. Experimental results indicate that the proposed approach can be used to identify cell phenotypes of RNAi genome-wide screens.

1 Introduction

High content screening by automated fluorescence microscopy is becoming an important and widely used research tools to assist scientists in understanding the complex cellular processes such as mitosis and apoptosis, as well as in disease diagnosis and prognosis, drug target validation, and compound lead selection [1]. Using images

* Corresponding author.

acquired by automated microscopy, biologists visualize phenotypic changes resulting from reverse-functional analysis by the treatment of *Drosophila* cells in culture with gene-specific double-stranded RNAs (dsRNAs) [2]. In a small scale study by manual analysis [3], biologists were able to observe a wide range of phenotypes with affected cytoskeletal organization and cell shape. Nonetheless, without the aid of computerized image analysis tools, it becomes an intractable problem to characterize morphological phenotypes quantitatively and to identify genes as well as their dynamic relationships required for distinct cell morphologies on a genome-wide scale.

In this paper, we will study image-based morphological analysis in automatic high-content genome-wide RNAi screening for novel effectors of Rho family GTPases in *Drosophila* cells. About 21,000 dsRNAs specific to predicted *Drosophila* genes are robotically arrayed in 384-well plates. *Drosophila* cells are plated and take up dsRNA from culture media. After incubation with the dsRNA, expression of Rac1V12, RhoAV14 or Cdc42V12 is induced. Cells are fixed, stained, and imaged by automated microscopy. Each screen will generate ~400,000 images, or millions if including replicates. Clearly, there is a growing need for automated image analysis as high throughput technologies are extended to visual screens. Biologists have developed a cell-based assay for Rho GTPase activity using the *Drosophila* Kc167 embryonic cell line. Three-channel images are obtained by labeling F-actin, GFP-Rac and DNA. Fig. 1 gives an example of RNAi cell images of one well acquired with three channels for phenotypes of (a) DNA, (b) Actin, and (c) Rac. It is tough to segment the cells from (b) or (c). The three phenotypes are shown in Fig. 2. They are: S-spikey, R-ruffling, and A-actin acceleration at edge. The question is how to identify the three phenotypes automatically for each image. To reach this aim, we propose the following three steps: first, we segment each cell, then we calculate the morphological and textural features for each cell and built training data sets, finally we perform feature reduction and classify cellular phenotypes.

The key issue is how to automatically segment cells of cell-based assays in a cost-effective manner, as such fast screening generate rather poor image quality and tens and hundreds of millions of cells in each study. There exist a number of publications on nuclei segmentation and cell segmentation. For example, Wahlby, et. al., [4] proposed a cytoplasm segmentation based on watershed segmentation and rule-based merging and splitting of over-segmented and under-segmented objects. Marker-controlled watershed segmentation is a popular method in cell segmentation [5-7]. In the literature, watershed methods with or without seeds are extensively studied. Although the oversegmentation caused by watershed can be reduced by rule-based merging of fragmented objects, it is difficult to devise reliable rules to merge the example which consists of one cell with three nuclei inside the cytoplasm. Lindblad, et. al., recently studied a similar problem about automatic segmentation of cytoplasm and classification of Rac1 activation [7]. There are several different points between our work and Lindblad's work. First, their data source is Chinese hamster ovary hIR (human insulin receptor) cells transfected with GFP-Rac1 reporter protein, and ours is *Drosophila* Kc167 embryonic cell line transfected with an inducible GFP-RacV12 protein. Second their data is two-channel (nucleus and GFP-Rac1 channels) 3D images, while our data is three-channel (DNA, F-actin and GFP-RacV12 channels) 2D images from larger scale genome-wide screening. Third, the quality of their images is better as they employed automated confocal laser scanning microscopy (see Fig.2 in [7]) while we used more commonly available, standard automated epi-fluorescence

microscopy. So it is much more challenging to segment RNAi genome-wide images. To address this hard problem, we propose a two-step segmentation approach. We then quantitate the tens of millions of cells and classify the cell phenotypes.

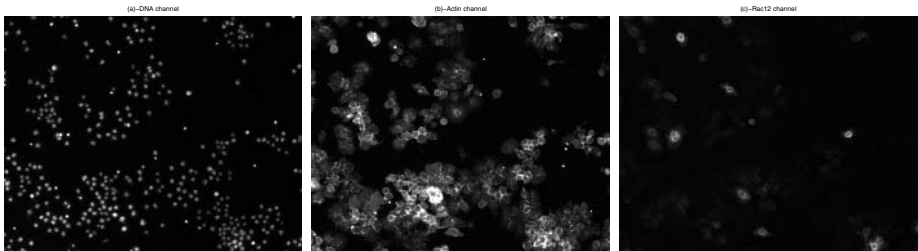


Fig. 1. RNAi cell images with three channels

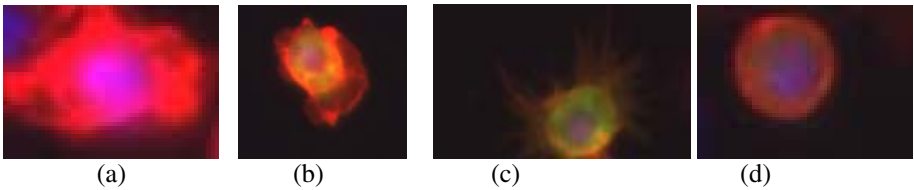


Fig. 2. Four different RNAi cell phenotypes: (a)-A-A, (b) N, (c) R and (d)

2 A Two-Step Segmentation Approach

Extracting rough boundary for each cell is the first step of our approach. It consists of two sub-steps: determine the center of each cell and determine the polygon for each cell. We then propose a fuzzy *c*-means based segmentation and sharpening; marker-controlled watershed is employed to extract each cell, and the Voronoi diagrams are further employed to correct errors due to some overlapping cells.

2.1 Extracting Rough Boundary for Each Cell

Large scale intensity variations as well as shading and shadowing effects in our images are often caused by uneven illumination over the field of view. A data-driven approach is employed to deal with this problem [4]. The algorithm works by iteratively making better distinction of the background of the image. A cubic B-spline surface is employed to model the background shading. After removing the shading, we adopt morphological transformation to enhance the image's contrast. Morphological filtering for enhancing images has been proposed by [8]. The first step of this method is to find peaks and valleys from original images. Peaks represent brighter spots of original image, and valleys represent the darker spots. Peaks are obtained by subtracting the morphologically opened image from the original image and valleys by subtracting the original image from morphologically closed image. The former is the Top-Hat transformation, and the later is the Bottom-Hat transformation. The contrast-enhanced image is obtained by the summation of the original, the peak and the nega-

tive valley image. Once the above two steps are completed, we then apply the ISODATA algorithm [9] to segment the nuclei in the DNA channel. After the above image processing, the touching spots are reduced, compared with the threshold methods traditionally used after preprocessing.

However, some nuclei may remain touching each other, we propose a method to segment touching cells. The aspect ratio of a particle image p is defined as $r(p) = w_{\min}(p)/w_{\max}(p)$, where $w_{\min}(p)$ and $w_{\max}(p)$ are the minimum and maximum diameters of the particle area. Denote the average size of the particles in the whole image to be \aleph , and size of cell p to be $s(p)$. The following condition is employed to isolate the touching cells: $r(p) < 0.5$, $s(p) > \aleph$, and $w_{\max}(p)$ is bigger than or equal to the pre-defined value 90 by experience. If the above condition is satisfied, then split the touching spots into two spots at the location of $w_{\min}(p)$. The merging criterion is different from the traditional approach. As an example, take a case where three nuclei can be seen within a single original cell. However, when watershed over-segments the original object, it is extremely difficult to find a proper rule to merge the fragments back into a single object. From the watershed point of view, the large object should be separated into three smaller objects. On the other hand, biologically, the three small nuclei belong to one single cell such that the large object should not be separated at the first place. Here we adopt the Hue transformation. Hue is a useful attribute in color segmentation since it is influenced by non-uniform illumination such as shade and shadow. The objective function used here is the square error of the piecewise constant approximation of the observed hue image H which is the hue transformation of the original three channels. Denote two cells regions as Ω_i and Ω_j . Define the mean Hue value as $\mu(\Omega_i)$ and $\mu(\Omega_j)$ and the following dissimilarity function:

$$f(\Omega_i, \Omega_j) = \frac{\|\Omega_i\| \cdot \|\Omega_j\|}{\|\Omega_i\| + \|\Omega_j\|} \left[\mu(\Omega_i) - \mu(\Omega_j) \right]^2 I(i, j)$$

where $I(i, j)$ is 1 if Ω_i, Ω_j are neighbors; infinity otherwise. Finally if the $f(\Omega_i, \Omega_j)$ is less than a threshed 1.5, then merge the two cells. Fig. 2 shows an example the segmentation result of the DNA channel based on the original image.

Our aim in this step is to find a rough boundary or polygon that encloses the entire cell whose center is the nuclei. The assumption is that one cell shape area cannot reach other nuclei's area. Assume that we are studying the polygon of one cell whose center of the nuclei is denoted by $P_0(x_0, y_0)$, we pick up those cells whose distances between the centers of their nuclei and the (x_0, y_0) is less than a pre-defined threshold $T_0 = 100$. Denote those centers as $P_1(x_1, y_1), \dots, P_N(x_N, y_N)$. Now the question is how to find certain points to be composed as vertices of a polygon. We define eight regions along the center $P_0(x_0, y_0)$ as follows:

$$R_k = \left\{ \alpha \mid -\frac{\pi}{8} + \frac{\pi}{4}(k-1) \leq \alpha < \frac{\pi}{8} + \frac{\pi}{4}(k-1) \right\}, \quad k = 1, 2, \dots, 8$$

We then calculate the slope and the angle between $P_i(x_i, y_i)$ and $P_0(x_0, y_0)$ as:

$$\theta_i = \begin{cases} \arctan\left(\frac{y_i - y_0}{x_i - x_0}\right) & \text{if } x_i - x_0 \geq 0 \\ \pi + \arctan\left(\frac{y_i - y_0}{x_i - x_0}\right) & \text{if } x_i - x_0 < 0 \end{cases}, \quad i = 1, 2, \dots, N$$

Then it is easy to determine in which region each point of P_1, P_2, \dots, P_N is located. Without loss of generality, assume P_1, P_2, \dots, P_M points are located in region 7. Pick up the center whose distance is the closest to $P_0(x_0, y_0)$, say $P_1(x_1, y_1)$. Denote the Euclidean distance between the two centers as d_{p_0, p_1} , and the fitted radius of nuclei of $P_1(x_1, y_1)$ as r_1 . Denote a new radius d as $d = d_{p_0, p_1} - r_1 / 2$. The point whose distance being d to the center point $P_0(x_0, y_0)$ in direction $3\pi/2$ is the right vertex in region 7. The obtained 8 vertex points are the vertex points of the expected polygon. The method proposed in this section gives us a rough boundary of each cell. Next we focus on how to segment each cell in this boundary.

2.2 RNAi Cell Segmentation

After using the above method to determine the boundary of each cell, then we can focus on this region and try to extract this cell. For our goal, we will first binarize the gray-level image. To effectively binarize the RNAi cells, we proposed the fuzzy c-means segmentation with sharpening method. Obviously, the segmentation can be treated as an unsupervised classification problem. In all clustering algorithms, the fuzzy c-means [10] is an attractive algorithm because its convergence property and low complexity, and thus is efficient to implement to screen large number of images. We first use the fuzzy c-mean clustering to partition the image pixels into $K = 3$ classes. Assume the class k is the right class we are interested in, and then we sharpen the pixels in this class by using fuzzy c-means clustering again. Here we present the sharpening technique. Because of the low contrast, it is necessary to adjust the membership values of $u_{i,k}$ (the i th pixel in the k th class) of the output from fuzzy system. Let $u_p(y)$ be the fuzzy membership value that indicates how a possible pixel y belongs to the set containing the notion of the measure of fuzziness to sharpen the fuzzy region of interest defined as 1 when $u_p(y) \geq u_0$, and 0 otherwise, where u_0 is a fuzzy membership threshold. Pham, et. al., [11] proposed to select $u_0 = u_{c^*}(y^*)$, where y^* is the pixel with maximum intensity value, and c^* is the right class which we are interested in. This choice of u_0 does not work in our images. Principally, the misclassified pixels mainly come from the closed membership values between the biggest and the second biggest values. If they are too close, the classification results are not reliable. Denote the difference between the biggest and the second biggest membership values in the class c^* to be ν , i.e.,

$v_i = \max_k(u_{i,k}) - \max_{k \neq k^0}(u_{i,k})$, $i = 1, \dots, \zeta$, where $k^0 = \arg \max_k(u_{i,k})$ and ζ is the total number of pixels in class c^* . Here we propose to estimate the threshold u_0 using fuzzy c-means again. We first partition the values of $v_i, i = 1, \dots, \zeta$ into two classes. Denote the minimal of the bigger class as u_1 , and the maximal of the smaller class as u_2 . Then the u_0 is defined as the average of these two parameters.

Isolating the touched cells is extremely challenging in automatic RNAi screening. If we adopt the watershed method, it can easily generate many false positives due to oversegmentation. It is difficult to remove the oversegmentation simply by applying certain heuristic rules or criteria [4, 6]. Since we already know the rough region and center of each cell, we propose to modify the marker function so that pseudo minima can be removed while the center of each cell can be kept. The catchment's basins associated to the pseudo minima are filled by morphological reconstruction and then transformed into plateaus. Then those minima will not generate different regions when watersheds are obtained. This method is called marker-controlled watershed algorithm. The markers are the centers of cells which are obtained in the first step. As the boundary of many cells is weak, it is hard to extract their boundary by using intensity gradient of pixels. Thus, after we extract the binarized cells, we then calculate the Euclidian Distance map of the binary image developed with the proposed fuzzy c-means segmentation and sharpening. We impose the markers to this Euclidian distance map, and then we applied watershed algorithm to segment the cells. Note that the above procedure is for segmenting a single cell. After we process all cells, it still could cause overlapping between a small number of cells. We finally apply the Voronoi diagrams [12] to correct the overlapped cells. Fig. 3 is the segmentation result by using the proposed approach.

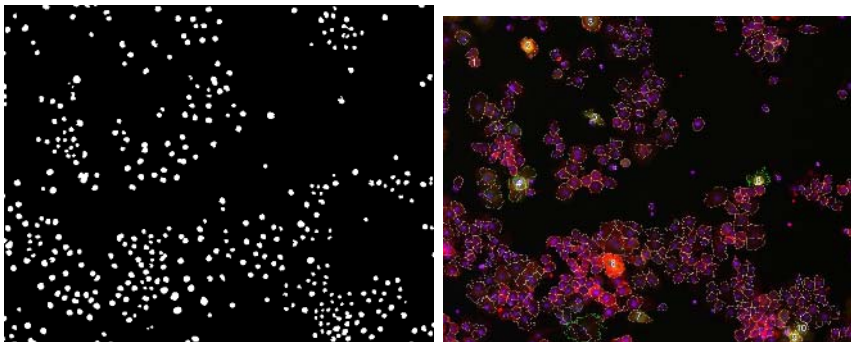


Fig. 3. left-DNA channel segmentation result; right- segmentation result

3 RNAi Cell Phenotype Recognition

In RNAi genome-wide screen, our goal is to identify all the possible cell phenotypes derived in the screening. In the meantime, we also need to consider some location features such as roundness and eccentricity. The first set of features was based on

Zerike moments. We extracted 49 features. The second set are co-occurrence contexture features. We extracted 28 textural features. The third set was based on cell hull location features [13]. Here we choose the three features which are the fraction of the convex hull area occupied by protein fluorescence, the roundness of the convex hull, and the eccentricity of the convex hull. We totally extracted 15 common image features. Since the feature values have completely different ranges, an objective scaling of features was achieved by calculating z-scores. Principle component analysis (PCA), commonly used in microarray research as a cluster analysis tool, is employed to reduce features. KNN classifier is a non-parametric classifier, which is easy to implement. Hence we use KNN to finish up the cell phenotype classification. Four classes are defined which are shown in Fig. 2. The four classes are Spikey, Ruffling, Actin acceleration at edge, and Normal cells. A training set is obtained for each class. We use KNN to classify test data.

4 Results and Discussions

In our study, four images of cells that had been visually classified are used to establish the training data sets. The limited amount of training data was due mainly to the tedious visual classification. We first used the proposed automated segmentation algorithm to segment each image, and then we asked the biologists to mark the cell with different phenotypes. Note that we actually have four classes in the training data sets as the other one is the class of normal cell which does not have significant change in their morphological shape. Cross-validation is employed to test the performance of the proposed automatic screening approach. We randomly split the data sets into training data set and test data sets, where 70% of cells are treated as training data and 30% of cells are treated as testing data. The feature reduction and normalization are first done with the training data sets, and then applied the transformation matrix of feature reduction to the test data set and similar procedure applied to normalization. Then we run 100 times and calculate the mean of the recognition accuracy based on the test data. The recognition accuracy is listed in Table 1. It is seen that the recognition accuracy is between 62% and 75%. Our biological collaborators reckoned that 70% accuracy should be adequate for the purpose of automatic screening. To further improve the accuracy and specificity, we continue to improve the segmentation algorithm, the phenotype definition, and the specific image features for extraction.

In this study, we proposed a two-step segmentation approach to segment high content cell images automatically. Certain regular image features, Heraik contextual feature and Zerike moment features are extracted for each cell. KNN is employed to perform cell phenotype classification. Experiments show that the proposed approach can automatically and effectively identify cell phenotype. Although we have built an initial workable system for automated RNAi genome-wide screening, there are certain problems remains. For example, we conjecture that the snake model in segmentation might be more effective than the marker-controlled watershed algorithm and Voronoi diagrams as the cells in GFP-Rac12 and F-actin have many different kinds of closed curves. We are in the processing of testing this conjecture. For feature extraction, additional image features specific to the different cell phenotypes, such as spiky region, ruffling region and actin acceleration region, would need to be identified. Furthermore, we would study how to automatically extract phenotypes hidden in cell

shapes by using cluster analysis. The ultimate goal is to score each image by establishing robust mathematic models to map the number of different phenotypes in each image to a scoring system which would let biologists easily find the novel candidate genes in their screens.

Table 1. Recognition accuracy for RNAi cell phenotypes

Phenotype	Actin-A	Ruffling-R	Spiky-S	Normal-N
Actin	72.87%	8.6%	1.69%	16.8%
Ruffling	8.68%	75.14%	6.13%	10.06%
Spikey	6.97%	20.36%	62.83%	9.84%
Normal	14.5%	6.54%	4.79%	74.17%

References

1. Z.E. Perlman, et.al., "Multidimensional drug profiling by automated microscopy," *Science*, 306 (2004) 1194-1198.
2. M. Boutros, et.al., "Genome-Wide RNAi analysis of growth and viability in *Drosophila* cells," *Science*, 303 (2004), 832-835.
3. A.A. Kiger, et. al., "A functional genomic analysis of cell morphology using RNA-interference," *Journal of Biology*, 2(4): 27, 2003.
4. C. Wählby, et.al., "Algorithms for cytoplasm segmentation of fluorescence labeled cells," *Analytical Cellular Pathology*, 24(2002):101-111.
5. N. Malpica, et. al., "Applying watershed algorithms to the segmentation of clustered nuclei," *Cytometry*, 28(1997):289-97.
6. G. Lin, et.al., "Hierarchical, model-based merging of multiple fragments for improved three-dimensional segmentation of nuclei," *Cytometry Part A*. 63(2005):20-33, 2005.
7. J. Lindblad, et.al., "Image Analysis for Automatic Segmentation of Cytoplasm and Classification of Rac1 Activation," *Cytometry Part A*, 57A(1):22-33, 2004.
8. S. Chen and M. Haralick, "Recursive erosion, dilation opening and closing transform," *IEEE Transactions on Image Processing*, 4(1995):335-345.
9. N. Otsu, "A threshold selection method from gray level histogram," *IEEE Trans. on System man Cybernetics*, 8(1978): 62-66.
10. 10 X. Zhou, et.al., "Gene clustering based on cluster-wide mutual information," *Journal of Computational Biology*, 11(2004):151-165.
11. 11 T.D. Pham, et.al., "Extraction of fluorescent cell puncta by adaptive fuzzy segmentation," *Bioinformatics*, 20(2004):2189-2196.
12. A. Okabe, et. al., "Nearest neighbourhood operations with generalized Voronoi diagrams: a review," *International Journal on Geographical Information Systems*, 8(1994):43-71.
13. M. V. Boland and R. Murphy, *Bioinformatics*, 17(2001):1213-1223.

Adaptive Spatio-Temporal Restoration for 4D Fluorescence Microscopic Imaging

Jérôme Boulanger^{1,2}, Charles Kervrann^{1,2}, and Patrick Bouthemy¹

¹ IRISA - INRIA, Campus Universitaire de Beaulieu, 35042 Rennes, France

² INRA - MIA Domaine de Vilvert 78352 Jouy en Josas, France

Abstract. We present a spatio-temporal filtering method for significantly increasing the signal-to-noise ratio (SNR) in noisy fluorescence microscopic image sequences where small particles have to be tracked from frame to frame. Image sequence restoration is achieved using a statistical approach involving an appropriate on-line window geometry specification. We have applied this method to noisy synthetic and real microscopic image sequences where a large number of small fluorescently labeled vesicles are moving in regions close to the Golgi apparatus. The SNR is shown to be drastically improved and the enhanced vesicles can be segmented. This novel approach can be further exploited for biological studies where the dynamics of small objects of interest have to be analyzed in molecular and sub-cellular bio-imaging.

1 Introduction

Time-lapse confocal microscopy is now a popular technology in biomedicine and biophysics for analysis of dynamics at sub-cellular or molecular levels. Several applications in biology are concerned with the acquisition of sequences of $2D$ images or $3D$ volumes representing small fluorescently-tagged particles with highly variable velocities. In order to analyze these moving objects, the acquisition time is reduced and consequently the signal-to-noise ratio becomes low. As a consequence, motion estimation and trajectography of such small objects remain difficult tasks.

Several algorithms have been developed for $3D$ image sequences restoration. Traditionally, these $4D$ data sets are processed, volume by volume, using $3D$ spatial filters, in order to save computation time, even if spatio-temporal filtering would appear more appropriate. In contrast, spatio-temporal filtering has been largely investigated in video processing, and usually relies on a motion-compensated approach [1]. Most of these spatio-temporal filters can nevertheless be considered as an extension of well-known principles developed for $2D$ image restoration or smoothing. To our knowledge, non-linear anisotropic diffusion [2, 3], recent PDE-based methods [4], Wiener filters [5], and wavelet shrinkage [6], have been used for processing $2D$ image sequences. Finally, $3D$ image sequence smoothing algorithms have been specifically developed for ultrasound imaging [7, 8].

In this paper we address the problem of $3D$ image sequence restoration by significantly extending the framework described for smoothing $2D$ still images

in [9, 10, 11]. Unlike robust anisotropic diffusion [12] and non-linear Gaussian filtering [13], this local adaptive estimation approach provides a scale selection for each pixel by locally estimating the appropriate spatial filtering window [9]. Additionally, a confidence level (i.e. variance) attached to the each pixel is provided. Furthermore, the *Total Variation* (TV) minimization method [14], commonly-used for 2D image restoration, cannot be easily extended to a space-time domain.

The novel method introduced in this paper is able to jointly estimate the spatial and temporal discontinuities in 3D image sequences. Also, the proposed algorithm is non-iterative and relatively fast since its complexity can be bounded by the user, which is necessary for processing large 4D data sets. Finally, the theoretical properties of such an estimator are well known in non-parametric statistics for adaptive smoothing [15, 11, 10].

While being general, the described method has been designed for analyzing¹ the role of fluorescence-tagged proteins moving around the Golgi apparatus and participating to the intra-cellular traffic. These proteins are embedded into vesicles whose movement is supposed to be dependent on a microtubules network. These vesicles propelled by motor proteins follow these polarized “cables”. This mechanism explains the observed high velocities which could not be accounted for a simple diffusion. Let us point out that the method is data-driven and requires no motion estimation, which is known to be problematic when images are heavily corrupted by noise.

2 Proposed Approach

We consider the following statistical image model: $Y_i = u(\mathbf{x}_i) + \xi_i$, where $\mathbf{x}_i \in \Omega$ represents the pixel location in the image domain $\Omega \subset \mathbb{R}^d$ ($d = 3$ for 2D and $d = 4$ for 3D image sequences). The image function $u_i = u(\mathbf{x}_i)$ has to be recovered from observations Y_i . The errors ξ_i are assumed to be independent zero-mean Gaussian variables with unknown variances τ_i^2 .

We need minimal prior assumptions on the structure of the image for recovering u . In what follows, we assume that $u(\mathbf{x}_i)$ is a locally piecewise constant function in the neighborhood of the pixel \mathbf{x}_i . However, the size and shape of these neighborhoods is not constant over the image sequence and must be estimated too. Once such a neighborhood has been determined, the regression function u can be estimated by optimizing a local maximum likelihood (ML) criterion. The proposed method addresses these two issues as described below.

One important feature of our approach is to define a sequence of increasing nested spatio-temporal neighborhoods $(W_{i,n})_{n \in [0:N]}$ at each point \mathbf{x}_i , i.e. $W_{i,n} \subset W_{i,n+1}$ with N indicating the largest window. At the initialization, we choose the 26 nearest neighbors in the 3D space domain as the *pilot* (starting) window $W_{i,0}$ at point \mathbf{x}_i , $\forall \mathbf{x}_i \in \Omega$. Then, we can compute an initial estimate $\hat{u}_{i,0}$ of $u(\mathbf{x}_i)$ and its associated variance $\hat{\sigma}_{i,0}^2$ as:

¹ The authors would like to thank the ACI-IMPBio for its support and the Curie Institute for the image sequence data-base.

$$\hat{u}_{i,0} = \frac{1}{|W_{i,0}|} \sum_{\mathbf{x}_j \in W_{i,0}} Y_j \quad \text{and} \quad \hat{\sigma}_{i,0}^2 = \frac{1}{|W_{i,0}|^2} \sum_{\mathbf{x}_j \in W_{i,0}} \hat{\tau}_j^2 \quad (1)$$

where $|W_{i,0}|$ denotes the number of pixels in $W_{i,0}$ and $\hat{\tau}_j^2$ is an empirical estimate of the local noise variance τ_j^2 as described in Section 3. This initialization step provides the first estimates of the two sequences $(\hat{u}_{i,n})_{n \in [0:N]}$ and $(\hat{\sigma}_{i,n}^2)_{n \in [0:N]}$. The next step consists in considering a larger window $W_{i,1}$ such that $W_{i,0} \subset W_{i,1}$ and calculating new estimates $\hat{u}_{i,1}$ and $\hat{\sigma}_{i,1}^2$ over $W_{i,1}$. At iteration n we define the estimator as:

$$\hat{u}_{i,n} = \sum_{\mathbf{x}_j \in W_{i,n}} \omega_{ij} Y_j \quad \text{and} \quad \hat{\sigma}_{i,n}^2 = \sum_{\mathbf{x}_j \in W_{i,n}} \omega_{ij}^2 \hat{\tau}_j^2, \quad (2)$$

where the weights ω_{ij} are defined as a function of the contrast between the estimate $\hat{u}_{i,n-1}$ at point \mathbf{x}_i and the estimates $\hat{u}_{j,n-1}$ at points \mathbf{x}_j belonging to the neighborhood $W_{i,n}$. The weights ω_{ij} and the geometry and size of the windows $W_{i,n}$ will be formally defined in the next Sections.

2.1 Adaptive Weights

At iteration $n \geq 1$, the weights are calculated as follows:

$$\omega_{ij} = \frac{f(\delta_{ij})}{\sum_{\mathbf{x}_j \in W_{i,n}} f(\delta_{ij})} \quad \text{and} \quad \delta_{ij} = \frac{\hat{u}_{i,n-1} - \hat{u}_{j,n-1}}{\lambda \hat{\sigma}_{i,n-1}} \quad (3)$$

where f is a function of the normalized contrast δ_{ij} . Considering its robustness and smoothing properties, we have chosen the influence function of the Huber M-estimator defined as:

$$f(x) = \begin{cases} 1 & \text{if } |x| \leq 1 \\ \frac{1}{|x|} & \text{otherwise} \end{cases} \quad (4)$$

but other influence functions are possible [12]. Therefore, the weights decide which points \mathbf{x}_j in the spatio-temporal neighborhood $W_{i,n}$ should contribute to the estimation of $u(\mathbf{x}_i)$. This decision is made under the hypothesis that the contrast is Gaussian distributed. Hence, the parameter λ controls the probability of false alarm (satisfy the test when it should not) to include a point \mathbf{x}_j in $W_{i,n}$. In our experiments, we set $\lambda = 3$ which corresponds to a commonly-used probability of error of type I of 0.036.

2.2 On-Line Window Geometry Specification

One of the main contributions of this work is the on-line adaptation of the neighborhood sequence $(W_{i,n})_{n \in [0:N]}$. It is worth noting that points $\mathbf{x}_j \in W_{i,n}$ that contribute to the estimation of $u(\mathbf{x}_i)$ are thus selected by weights in a quite flexible and efficient way unlike usual methods which exploit geometry-based design of the neighborhood. This allows us to use a simple hyper-cubic spatio-temporal

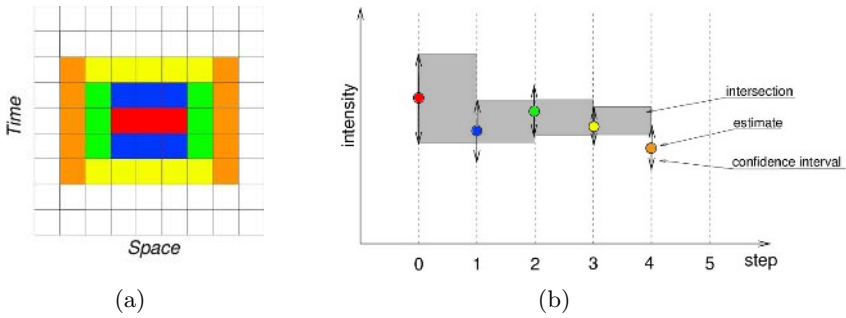


Fig. 1. (a) Spatio-temporal neighborhood: colors correspond to iterations plotted in (b); (b) confidence intervals: circles represent estimates $\hat{u}_{i,n}$ obtained at each iteration n . The grey rectangles represent the intersection between the current confidence interval and the previous one. As long as the estimate belongs to this intersection, the estimation process is updated.

volume for the window shape. Nevertheless, we separate the space dimension from the time dimension and parametrize the spatio-temporal volume by introducing two extents. Figure 1 shows the considered neighborhood sequence and how neighborhoods are enlarged along with the iterations. It is worth stressing that this figure only represents a particular case and that the sequence of neighborhoods is not defined *a priori* but locally estimated simultaneously to the image filtering procedure. Thus, each point is assigned an adapted smoothing support. This support is estimated according to the stopping rules described in the next section.

2.3 Window Estimation

Two pointwise rules are used to drive the window estimation process. The first one is designed to estimate the optimal window at \mathbf{x}_i and is based on the measure of the closeness of the estimator \hat{u} to the unknown function u given by the local L_2 risk. This measure of performance can be decomposed in two terms, that is the squared bias and the variance as $\mathbb{E}[\hat{u}_{i,n} - u_i]^2 = [\text{bias}(\hat{u}_{i,n})]^2 + \hat{\sigma}_{i,n}^2$, where $\mathbb{E}(\cdot)$ denotes the mathematical expectation. In what follows, we reasonably assume the squared bias is an increasing function of the window size and the variance is a decreasing function of the window size [9, 15, 11, 10]. Then, the selected window will be the window which achieves an optimal compromise between these two terms, i.e. the squared bias and variance terms must be of the same order: $\text{bias}(\hat{u}_{i,n}) \asymp \hat{\sigma}_{i,n}$. A closed-form optimal solution is not available for our non-linear estimator and we arbitrarily define, as usually done, the optimal window as the window for which the squared bias and the variance are nearly the same. Now let us consider the sequence of increasing windows $(W_{i,n})_{n \in [0:N]}$ and the sequence of estimates $(\hat{u}_{i,n})_{n \in [0:N]}$ of variance $(\hat{\sigma}_{i,n}^2)_{n \in [0:N]}$. A practical rule based on pairwise comparison of successive estimates can be derived to detect the optimal window. We define the largest window satisfying the following pointwise statistical rule [9, 15, 11]:

$$|\hat{u}_{i,n} - \hat{u}_{i,n'}| < \eta \hat{\sigma}_{i,n'}, \quad \forall n' < n, \quad (5)$$

as the optimal window since $W_{i,n'} \subset W_{i,n}$ at \mathbf{x}_i . Strictly speaking, while the estimates $(\hat{u}_{i,n'})_{n' \leq n}$ are sufficiently close to each other, then we decide to continue the estimation process. The factor η can be easily chosen in the range $[2, 4]$ in order to adapt the decision rule as justified in [9, 11, 10]. As shown in Figure 1, this rule amounts to continue the pointwise estimation process, while new estimates belong to the intersection of estimated confidence intervals $[\hat{u}_{i,n} - \eta \hat{\sigma}_{i,n}, \hat{u}_{i,n} + \eta \hat{\sigma}_{i,n}]$. This representation brings to the fore that we don't have to store all the previous estimates $(\hat{u}_{i,n'})_{n' \leq n}$ but only the intersection of confidence intervals, the previous estimate and its variance for each point. We also introduce a second rule which consists in locally stopping the pointwise estimation if the confidence intervals is small enough: when the output $\hat{u}_{i,N}$ is stored in an integer format, it is useless to continue to improve the estimation when the confidence associated to the estimate is below the unit. This second rule enables to autonomously stop the estimation process and makes the algorithm more data-driven.

3 Algorithm Implementation

As noticed in Section 2, an estimation of the noise variance is required. Here, we assume that the noise variance is constant over the whole 3D image sequence. It can be robustly estimated by calculating pseudo-residuals ε_i as described in [16]. If we choose a 4D 8-neighborhood, pseudo-residuals are compactly represented by $\varepsilon_i = (8Y_i - \Delta Y_i) / \sqrt{72}$ where ΔY_i is the discrete Laplacian at \mathbf{x}_i and the constant $\sqrt{72}$ is introduced to insure that $\mathbb{E}[\varepsilon_i^2] = \tau^2$. Given the residuals ε_i , we can then robustly estimate the noise variance τ^2 by: $\tau = 1.4826 \text{ med}_i (|\varepsilon_i - \text{med}_j |\varepsilon_j| |)$. A local estimation of the noise variance is proposed in [9] and can also be used when the noise model proposed in Section 2 is not appropriate to describe the spatio-temporal inhomogeneity of the image sequence. Moreover, for Poisson distributed noises, the Anscombe transform can be applied [17]. We have also seen in Sections 2.1 and 2.2 that λ and η can be well calibrated using statistical arguments. In our experiments, λ is set to 3 and η to $2\sqrt{2}$. During the estimation, we alternate the increasing of the spatial and temporal extents of the 4D spatio-temporal neighborhoods. Furthermore, the algorithm can be easily parallelized. Actually, estimation steps use only local information and have been distributed over several CPUs. Finally, another possibility to speed up the algorithm is to use a dyadic scheme when increasing the extent of the neighborhood.

4 Experiments

The proposed method has been applied to both synthetic and experimental sequences of 2D and 3D images obtained by fast 3D deconvolution microscopy [18]. Since the dynamics of vesicles are unknown, we first simulate a noisy sequence of 2D images by adding a realistic white Gaussian noise to a real denoised

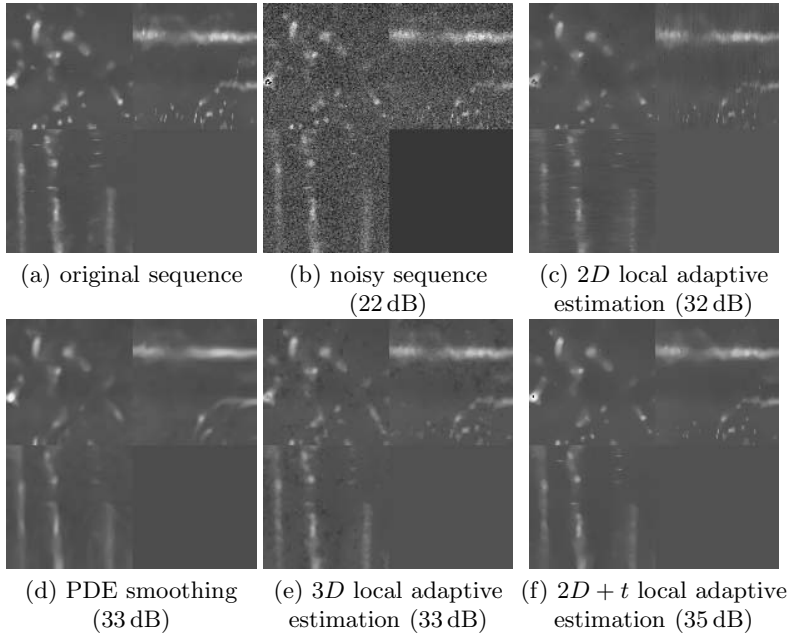


Fig. 2. Comparison with other method on a synthetic sequence. Six $2D$ image sequences are represented as $100 \times 100 \times 100$ volumes. First three quadrants correspond slice XY , TY and XT . The last one is left empty. (a) original sequence used as the ground truth ; (b) artificially noisy sequence ; (c) spatial adaptive estimation [9], i.e. each image of the sequence is processed independently ; (d) $3D$ anisotropic diffusion [4] when the sequence is considered as a homogeneous $3D$ volume ; (e) adaptive estimation algorithm applied the sequence considered as a $3D$ volume ; (f) our $2D + t$ adaptive estimation (see text).

image sequence (Figs 2(a) – 2(b)). We will also compare our method to other methods on this simulated sequence.

In Fig. 2(c), some blinking effects or strides are visible on spatio-temporal slices when each $2D$ frame of the sequence are processed independently with the adaptive estimation described in [9]. Anisotropic diffusion [4] has been applied on the $2D$ sequence considered as a $3D$ volume. In Fig.2(d), the sequence is strongly smoothed and some details are lost. In Fig.2(e), the $3D$ adaptive estimation algorithm described in [9] has been applied to the $3D$ volume. As expected, considering the temporal dimension as a spatial dimension is not appropriate to cope with temporal discontinuities. Finally, the results of our method are shown in Fig.2(f). The peak-signal-to-noise-ratio defined in decibels (dB) as $PSNR = 20 \log_{10}(255/mse)$ where mse denotes the mean squared error between the original sequence and the result of filtering process, is drastically improved and the dynamics of particles are well preserved. Figure 3(a) shows a detail of a real $2D$ image sequence. It is first processed by equally considering the three space-time dimensions. Undesirable blurring over three successive images of a

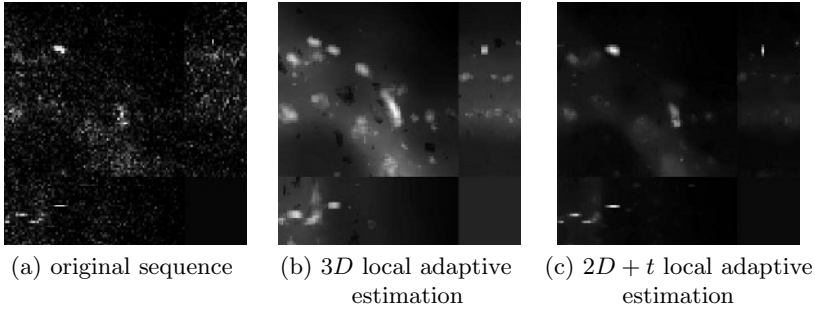


Fig. 3. Detail of a sequence denoised with two versions of the filter: 3D estimator and the proposed 2D + t estimator. First three quadrants correspond slice XY , TY and XT . The last one is left empty.

single spot along its track are visible (Fig. 3(b)). In the case of spatio-temporal filtering (Section 2), these artifacts are removed as shown in Fig. 3(c).

We have also applied our adaptive spatio-temporal filtering method on a database of video-microscopy sequences in order to facilitate the automatic analysis of the dynamics of two proteins Rab6a and Rab6a' involved in particular steps of membrane trafficking. These proteins are fluorescently-tagged and move around the Golgi apparatus. The function of this organelle is to package materials for export from the cell. Figures 4(a) and (b) show respectively a real noisy image extracted from a video-microscopy sequence ($360 \times 445 \times 10$ pixels) and the corresponding restored image for the fifth 2D slice in depth. The algorithm provides visually satisfying results after ten iterations as demonstrated in Figs. 5 (a) and (b). that illustrate a cropped region of a 3D image. Note that, the point-wise window extents shown in Figs. 4 (c) and (d) roughly correspond to the size

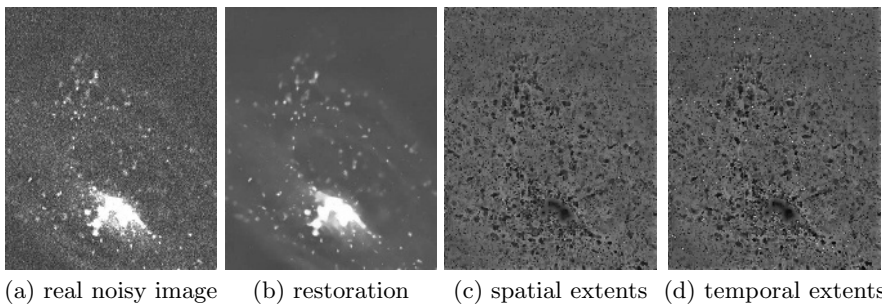


Fig. 4. One frame depicting the spot positions at time $t = 2s$. The dynamics of fluorescent tags has been recorded by fast 3D deconvolution microscopy. The large white region corresponds to the Golgi apparatus while small spots are vesicles moving with a high average speed (~ 10 pixels/frame). (a) real noisy image (2D slice of 3D frame); (b) denoised image using the spatio-temporal adaptive estimation ; (c) spatial extents of windows ; (d) time extents of windows (small extents are coded in black and are located in the vicinity of static or moving spots).



Fig. 5. Volume rendering of a cropped region of a 3D image extracted from the sequence in Fig 4. (a) original real noisy 3D frame; (b) results using our adaptive spatio-temporal method.

of space-time structures in image sequences. These spatio-temporal features will be further exploited for analysis of these biological data. The processing time of a sequence of 50 2D frames (360×445) is typically 10 min

5 Conclusion

We have described an original, adaptive and efficient method for 2D and 3D image sequence restoration. Experiments on synthetic and real video-microscopy image sequences have demonstrated its ability to smooth images while preserving both spatial and temporal discontinuities. Moreover, the presented method does not require a fine tuning of the parameters which are well calibrated using statistical arguments. This approach does not require motion compensation and can be further used for biological studies where dynamics have to be analyzed in noisy sequences. In particular, we could exploit spatio-temporal features (window size) provided by the algorithm, which can be regarded as a preliminary step for analysis of the lifetime kinetics of specific Rab proteins. Furthermore, this method is not restricted to biological applications and could be applied for smoothing three dimensional ultrasound images using an estimator adapted to the Rayleigh noise model.

References

1. Brailean, J., Kleihorst, R., Efstratiadis, S., Katsaggelos, A., Lagendijk, R.: Noise reduction filters for dynamic image sequences: A review. *Proc. of the IEEE* **83** (1995) 1272–1291
2. Lee, S.H., Kang, M.G.: Spatio-temporal video filtering algorithm based on 3-D anisotropic diffusion equation. In: *ICIP'1998*. Volume 3(2)., Chicago, (1998) 447–450
3. Uttenweiler, D., Weber, C., Jähne, B., Fink, R., Scharr, H.: Spatiotemporal anisotropic diffusion filtering to improve signal to noise ratios and object restoration in fluorescence microscopic image sequences. *Jour. of Biomedical Optics* **8** (2003) 40–47
4. Tschumperlé, D., Deriche, R.: Vector-valued image regularization with pde's : A common framework for different applications. *IEEE Trans. on Pattern Analysis and Machine Intelligence* **27** (2005) 506–517

5. Dekeyser, F., Bouthemy, P., Pérez, P.: Spatio-temporal wiener filtering of image sequences using a parametric motion model. In: ICIP'2000, Vancouver, Canada (2000)
6. Rajpoot, N., Yao, Z., Wilson, R.: Adaptive wavelet restoration of noisy video sequences. In: ICIP'2004, Singapore (2004)
7. Angelini, E., Laine, A., Donis, J., Gersony, D., Homma, S.: Quantification of right and left ventricular function with real time three-dimensional ultrasound. In: Proc. of the 23rd Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society. (2001)
8. Mikula, M., Preusser, T., Rumpf, M.: Morphological image sequence processing. *Computing and Visualization in Science* **6** (2004) 197–209
9. Kervrann, C.: An adaptive window approach for image smoothing and structures preserving. In: Proc. of 8th Eur. Conf. on Computer Vision, ECCV'2004, Prague, Czech Republic (2004) 132–144
10. Katkovnik, V., Egiazarian, K., Astola, J.: Adaptive window size image denoising based on intersection of confidence intervals (ici) rule. *Jour. of Mathematical Imaging and Vision* **16** (2002) 223–235
11. Polzehl, J., Spokoiny, V.: Adaptive weights smoothing with applications to image restoration. *Jour. of the Roy. Stat. Soc.: Series B (Stat. Meth.)* **62** (2000) 335–354
12. Black, M., Sapiro, G., Marimont, D., Heeger, D.: Robust anisotropic diffusion. *IEEE Trans. on Image Processing* **7** (1998) 421–432
13. Barash, D.: A fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation. *IEEE Trans. on PAMI* **24** (2002) 844–847
14. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Phys. D* **60** (1992) 259–268
15. Lepski, O.: Asymptotically minimax adaptive estimation 1: upper bounds. *SIAM Jour. Theory of Probability and Application* **36** (1991) 654–659
16. Gasser, T., Sroka, L., Jennen Steinmetz, C.: Residual variance and residual pattern in nonlinear regression. *Biometrika* (1986) 625–633
17. Starck, J., Murtagh, F., Bijaoui, A.: *Image Processing and Data Analysis, the Multiscale Approach*. Cambridge University Press (2000)
18. Sibarita, J.B., Magnin, H., De Mey, J.: Ultra-fast 4D microscopy and high throughput distributed deconvolution. In: Proc. of IEEE Int. Symp. on Biomedical Imaging, ISBI'2002: Macro to Nano, Washington, (2002) 769–772

Kinematic Geometry of Osteotomies

Erin J. Smith¹, J. Tim Bryant^{1,3}, and Randy E. Ellis^{1,2,3}

¹ Mechanical Engineering, Queen's University at Kingston, Canada K7L 3N6

² School of Computing, Queen's University at Kingston, Canada K7L 3N6

³ Department of Surgery, Queen's University at Kingston, Canada K7L 3N6

Abstract. This paper presents a novel method for defining an osteotomy that can be used to represent all types of osteotomy procedures. In essence, we model an osteotomy as a lower-pair mechanical joint to derive the kinematic geometry of the osteotomy. This method was implemented using a commercially available animation software suite in order to simulate a variety of osteotomy procedures. Two osteotomy procedures are presented for a femoral malunion in order to demonstrate the advantages of our kinematic model in developing optimal osteotomy plans. The benefits of this kinematic model include the ability to evaluate the effects of various kinds of osteotomy and the elimination of potentially error-prone radiographic assessment of deformities.

1 Introduction

An osteotomy, the surgical cutting of bone, is used in many orthopedic procedures for the purpose of correcting mechanical and aesthetical deformities of bones and joints. Bone deformities may arise from many causes including congenital defects, poor nutrition, disease, and post-traumatic healing. These deformities can be described by two elemental deformities, in rotation and in translation. It is common to further subdivide these into rotation and translation about/along the long axis of a bone, and rotation and translation about/along some other (transverse) axis. Simple deformities consist of only one elemental deformity, whereas complex deformities are a combination of simple deformities. These deformities are illustrated through the use of an elongated cuboid in Fig. 1.

Numerous osteotomy procedures have been described in the orthopedics literature, differing according to the anatomical site of the osteotomy, orientation and geometry of the cut, and the type of movement of the osteotomized segments [1]. Common osteotomy techniques include rotational, transverse displacement, longitudinal displacement, wedge, and oblique osteotomies. A variety of osteotomy plane geometries, such as Z-cut, step-cut and chevron, have been developed in order to improve the surface contact of the bone segments in some of these procedures.

Typically, radiographic and CT images are used to identify orthopedic deformities and select an appropriate osteotomy procedure. Computer-assisted osteotomy planning has, to date, targeted a specific type of osteotomy procedure [2, 3, 4, 5, 6, 7, 8]. Unfortunately, these methods do not allow surgeons to readily compare the results of using different procedures to correct a deformity and further require that surgeons be familiar with a number of software programs.

This paper presents a novel method for defining an osteotomy procedure that can be used to represent all types of osteotomy. This method can be readily incorporated into

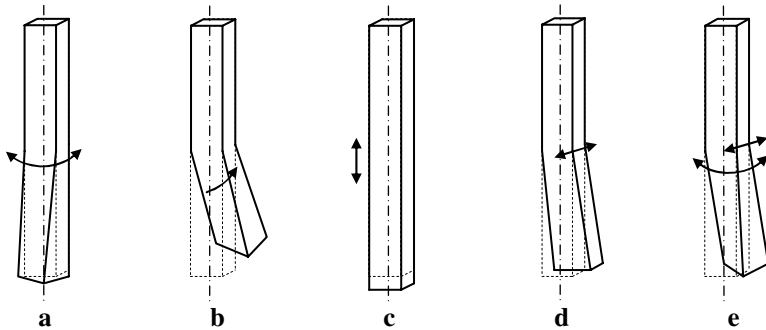


Fig. 1. Common orthopedic deformities: (a) rotational deformation about the long axis; (b) angular deformation about a transverse axis; (c) translational deformation along the long axis; (d) translational deformation along a transverse axis; and (e) complex, both rotational and translational

computer graphics in order to simulate a variety of osteotomy procedures. In essence, we model an osteotomy as a lower-pair mechanical joint to derive the kinematic geometry of the procedure.

We were inspired to consider kinematic geometry because of a common requirement of mechanical joints and the healing of an osteotomy: *surface congruence*. Bearing surfaces must be congruent; otherwise, excessive wear or catastrophic failure may result. For bone to heal, the surfaces of the osteotomized bone should be [roughly] congruent, both for weight-bearing resistance and because mechanical strain is a necessary element in bone remodeling.

2 Kinematic Representation of an Osteotomy

An osteotomy permits relative 3D motion of the osteotomized bone segments. If one segment is selected as the base segment, the other segment has up to six degrees of freedom of relative motion: three translations and three rotations. Simple representations of relative motion, such as homogeneous coordinate transformations, are commonly used in computer graphics and robotics [9]. It is straightforward to represent any given type of osteotomy by such a transformation; this was reported for the oblique osteotomy as early as 1952 [10] and as recently as 2005 [11]. However, these general representations do not elegantly capture the relative motion induced by a particular osteotomy.

The literature on robotics and mechanism theory is rich with descriptions of relative motion of rigid members [12]. Because any particular osteotomy procedure restricts the transformation allowed between the bone segments, it may be useful to consider the osteotomy as a virtual joint.

2.1 A Virtual-Joint Model

In robotics, joints connect successive links and constrain the motion that can occur between the links. The number of degrees of freedom of a joint is known as the joint

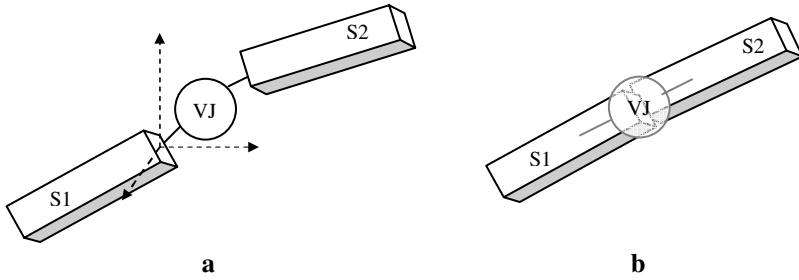


Fig. 2. Virtual-joint model of an osteotomy. (a) A virtual joint (VJ) connects the osteotomized segments, S1 and S2. Motion of S2 can be defined with respect to a coordinate frame attached to S1 at the position of the osteotomy site, having the same orientation as the joint axes; (b) The joint is not a physical presence as it is in robotic systems and, therefore, does not actually compromise the proximity of the osteotomized surfaces as the graphical representation in (a) may suggest.

mobility m . In real mechanisms the mobility can take on values of $0 < m \leq 6$. The simplest joint has a joint mobility of $m = 1$, allowing movement in a single direction only; in robotics the common mobility-one joints are either prismatic (translation only in a fixed direction) or revolute (rotation only about a fixed axis). Complex joints consisting of higher degrees of freedom can be made by combining revolute and prismatic joints, and allow motion in multiple directions.

A distinction made early in the twentieth century divided joint kinematics into *lower-pair* and *higher-pair* motions [13]. A kinematic higher pair has non-congruent joint surfaces; in the real world, such a bearing surface is not durable. An example of a higher pair is a point loading of a plane, for which transverse motion would either wear the point to a locally planar surface or break the point.

In general there are six classes of kinematic lower pairs, each with a specify type of movement: prismatic, revolute, helical ¹, cylindrical, planar, and spherical. If the osteotomized bone surfaces are congruent or nearly so then it must be the case that, for any given type of osteotomy, the relative motion between the bone segments is constrained to be a kinematic lower pair.

We have found a direct mapping from kinematic lower pairs to common osteotomy procedures, which we call virtual joints (VJs). A VJ is a kinematic constraint on the relative segment motion, shown in Fig. 2. A VJ has two important uses in planning an osteotomy. First, a planning system can use a VJ to restrict the allowable motion of bone models so that a surgeon can determine beforehand the effects of various kinds of osteotomy. Second, once a specific osteotomy is selected, the congruent surfaces of the corresponding kinematic lower pair can be calculated and provided to the surgeon for intraoperative guidance of the osteotomy procedure.

We have searched the orthopedic literature extensively and categorized osteotomy procedures as various kinds of VJs. As shown in Table 1, the kinds of osteotomy used to correct simple deformities are either prismatic or revolute VJs, each with mobility

¹ Although helical joints are useful for screws, nuts, bolts, etc., they seem to have little utility in describing osteotomy.

Table 1. Osteotomy procedures modeled as virtual $m = 1$ lower-pair joints

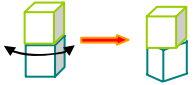



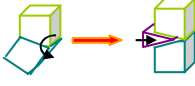





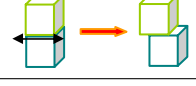

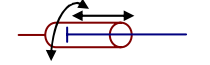


Osteotomy	Common Examples	Diagram	Equivalent Motion
Transverse Rotational	Rotational/derotational osteotomy; Transtrochanteric rotational osteotomy of the femoral head; Rotational acetabular osteotomy		revolute 
Closing Wedge	Coventry high tibial osteotomy; varus/valgus correction		revolute 
Opening Wedge	High tibial opening wedge osteotomy; varus/valgus correction; Distal radius osteotomy		revolute 
Oblique	Tibial/femoral malunions		revolute 
Longitudinal Displacement	Lengthening/shortening of long bones		prismatic 
Transverse Displacement	Pelvic osteotomy; Hallux valgus correction		prismatic 

Table 2. Use of $m \geq 2$ lower-pair joints in describing osteotomy

Osteotomy	VJ Model
Combined wedge and transverse displacement osteotomy; Combined longitudinal displacement and rotational osteotomy	 cylindrical
Combined rotational and transverse displacement osteotomy	 planar
Combined wedge and rotational osteotomy; oblique osteotomy	 spherical

$m = 1$. VJs of higher mobility, $m \geq 2$, can be used to model osteotomy procedures that correct complex deformities, and are shown in Table 2.

3 Implementation and Assessment

The VJ model can be applied to the computer-assisted planning of an osteotomy by creating constraints that define how computer images can be moved relative to one another. Commercial computer-graphics animation software (Maya 6 Complete, Alias/Wavefront, Toronto, Canada) was used to implement VJs. This software package was

selected because of its powerful graphics and development capabilities. Algorithms were developed to constrain motion for all five lower-pair joints of the VJ model, and both forward and inverse planning techniques were implemented.

Virtual osteotomy simulations were conducted using 3D surface meshes of a deformed bone and a normal counterpart. These surface meshes were obtained by laser-scanning foam models of bones (Pacific Research laboratories, Bellinham, USA). Osteotomy simulations were performed on bone models having both simple and complex deformities. For brevity, only two VJ osteotomies will be presented here.

3.1 Forward Planning: An Opening-Wedge Osteotomy

The forward virtual-osteotomy process simulates a surgical procedure. Consider, for example, an opening-wedge osteotomy of a malunited femur with a 30-degree angulation deformity modeled using a revolute virtual joint. This osteotomy was simulated in four steps:

1. The normal and deformed surface meshes were manually aligned in order to visualize the deformity in 3 dimensions, shown in Fig 3(a);
2. A graphical plane was used to define the position and orientation of the osteotomy, shown in Fig 3(b). The coordinate axes of the plane were used to indicate the orientation of the VJ axis;
3. The surface mesh of the deformed bone was separated to create two independent meshes along the osteotomy plane; and
4. A revolute joint constraint was applied to the osteotomized bone which restricted the movement of the mesh images to rotation about the joint axis. A graphical user interface facilitated movement of the bone image according to the constraint model, shown in Fig 3(c). A separate user interface element was created for each degree of freedom depending on the type of joint constraint.

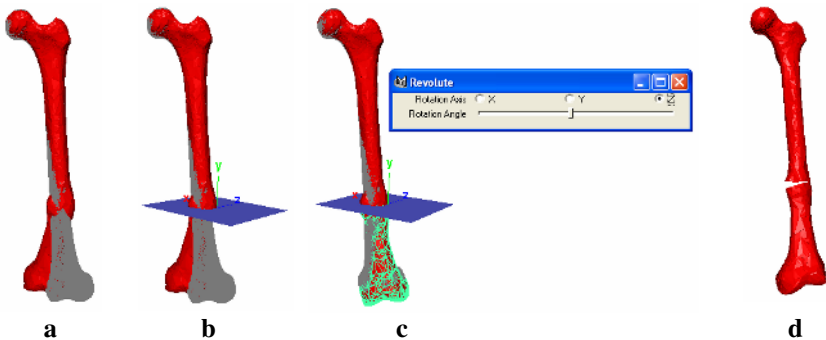


Fig. 3. Virtual opening-wedge osteotomy: (a) Alignment of normal and deformed surface meshes; (b) A graphical plane was used to specify the location of the osteotomy; (c) A slider interface was used to adjust the rotation about the joint axis; (d) Result of the opening-wedge osteotomy

3.2 Inverse Planning: An Oblique Osteotomy

An oblique, single-cut, osteotomy can be used to correct simple angulation deformities as well as complex deformities involving rotational and angular elements. The surgeon makes a single cut such that rotation in the plane of the cut corrects the deformity. The oblique osteotomy, when indicated, is preferred over a wedge osteotomy because it eliminates changes in bone length due to the removal or addition of a wedge, and results in a stable interface for fixation.

Although the theory of the oblique plane is well established [10], determining the orientation of the plane can be quite difficult. Efforts have been made to determine the orientation of the oblique plane from mathematical formulas based on angles of the deformity measured from radiographs [14]. As an extension of these formulas, look-up tables were developed to facilitate the mathematical calculations [15]. Quite recently, a mechanical device was constructed to determine the orientation of the oblique plane [11]; interestingly, these authors noted that the use of a computer would be ideally suited to solving this problem.

An inverse planning algorithm was implemented to determine the orientation of the oblique osteotomy plane. The process was the same as that described for the forward virtual-joint osteotomy; The osteotomy was modeled as a spherical joint constraint, and the user manipulated the spherical degrees of freedom to correct the deformity. The result of the rotational transformation was used to determine the orientation of the oblique osteotomy plane using an equivalent angle-axis representation. Test models were created to validate the algorithm, comparing the orientation of the oblique plane to the results obtained by Sangeorzan *et al.* [14, 15].

The inverse planning algorithm was used to determine the orientation of the oblique plane for a malunited femur with 30-degree angulation, presented earlier as an opening-wedge osteotomy (Fig. 4).



Fig. 4. Virtual osteotomy of a femur using the inverse model to determine the orientation of the osteotomy plane. The display window indicates the orientation of the oblique plane and the rotation angle about the equivalent rotation axis. In this case, the oblique plane that corrects the deformity is not clinically feasible since it would result in an extremely oblique and long osteotomy plane.

4 Discussion

The VJ model provides the mathematical foundation for a novel computer-based preoperative osteotomy planner with the ability to represent all types of osteotomy procedures. By using the VJ method, multiple osteotomy techniques can be evaluated in order to select an optimal plan. For the examples presented, it is possible to compare the use of an oblique plane (that would not completely correct the deformity) to an opening-wedge osteotomy (that may involve leg-lengthening).

A clear advantage of the VJ model is the elimination to initially characterize orthopedic deformities in terms of radiographic measurements. In the VJ model, representation of the osteotomy is directly related to the relative motion that occurs between the osteotomized segments instead of the nature of the deformity itself. This novel approach is especially advantageous in determining the orientation of an oblique osteotomy plane since existing methods involving mathematical formulas [14], look-up tables [15] and a mechanical device [11] rely on initial radiographic assessment, and limit the ability to create optimal preoperative plans. A direct transformation-based solution can increase the accuracy of preoperative osteotomy plans by eliminating errors associated with radiographic measurements that may not be perfectly aligned to the anatomical planes from which the deformities are referenced. A transformation-based approach also allows surgeons to experiment with the outcome of the oblique procedure to produce an optimal plan with both feasible geometry and sufficient deformity correction. It is difficult to do this with the existing approaches, as the relationship between the parameters used to represent the deformity and the orientation of the osteotomy plane is not straightforward.

The accuracy of the current implementation is dependent on the alignment of the deformed bone with its normal counterpart. In this implementation, alignment was performed manually; however, the authors recognize that automated registration techniques could be used for alignment.

For simplicity, the VJ model was used here to implement only those osteotomy procedures that could be represented by planar cuts and motion defined by a single kinematic lower pair. It should be noted, however, that the fundamentals of the VJ model can be easily extended to represent osteotomy procedures with more complex geometries and motions. For instance, the Z-cut, chevron, and crescentic osteotomy could be implemented by creating a library of osteotomy geometries to supplement the lower-pair-joint constraints. The effects of external fixators, such as Taylor and Ilivarov frames, could be included by considering systems of kinematic lower pairs. Procedures with complex motion patterns can be incorporated into the model by extending the library of virtual constraints.

5 Conclusion

A novel method for representing an osteotomy as a constrained transformation was presented. This method was used to simulate a number of osteotomy procedures whose kinematic geometry can be described by a lower pair joint. The proposed virtual-joint model has the capability to represent any type of osteotomy procedure. Potential advantages of this method include the ability to evaluate the effects of various kinds of

osteotomy and the elimination of potentially error-prone radiographic assessment of deformities. The VJ model provides an improved means to define optimal preoperative plans that can be used intraoperatively to guide the osteotomy procedure.

Acknowledgements

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada, the Institute for Robotics and Intelligent Systems, Communications and Information Technology Ontario, and the Ontario Research and Development Challenge Fund.

References

- [1] Canale, S.T., ed.: *Campbell's Operative Orthopaedics Tenth Edition*. Mosby Inc., Philadelphia PA (2003)
- [2] Bilic, R., Zdravkovic, V., Boljevic, Z.: Osteotomy for deformity of the radius: Computer-assisted three-dimensional modelling. *Journal of Bone and Joint Surgery* **76** (1994) 150–154
- [3] Moctezuma, J.L., Gossé, F., Schulz, H.J.: A computer and robot aided surgery system for accomplishing osteotomies. In: *International Symposium on Medical Robotics and Computer Assisted Surgery*. (1994) 31–35
- [4] Chao, E.Y.S., Sim, F.H.: Computer-aided preoperative planning in knee osteotomy. *Iowa Orthopaedic Journal* **15** (1995) 4–18
- [5] Ellis, R.E., Tso, C.Y., Rudan, J.F., Harrison, M.M.: A surgical planning and guidance system for high tibial osteotomy. *Computer Aided Surgery* **4** (1999) 264–274
- [6] Croitoru, H., Ellis, R.E., Small, C.F., Prihar, R., Pichora, D.R.: Fixation-based surgery: A new technique for distal radius osteotomy. *Computer Aided Surgery* **6** (2001) 160–169
- [7] Mayman, D., Rudan, J.F., Yach, J., Ellis, R.E.: The Kingston periacetabular osteotomy: A new computer-enhanced technique. *Computer Aided Surgery* **7** (2002) 179–186
- [8] Athwal, G.S., Ellis, R.E., Small, C.F., Pichora, D.R.: Outcomes of computer-assisted distal radius osteotomy. *Journal of Hand Surgery* **28** (2003) 951–958
- [9] Craig, J.J.: *Introduction to Robotics: Mechanics and Control*. Addison-Wesley, Reading MA (1989)
- [10] d'Aubigne, R.M., Descamps, L.: L'osteotomie plane oblique dans la correction des deformations des membres. *Memoires: Academie de Chirurgie* **78** (1952) 271–276
- [11] Meyer, D., Siebenrock, K., Schiele, B., Gerber, C.: A new methodology for the planning of single-cut corrective osteotomies of mal-aligned bones. *Clinical Biomechanics* **20** (2005) 223–227
- [12] McCarthy, J.M.: *An Introduction to Theoretical Kinematics*. MIT Press, Cambridge MA (1990)
- [13] Hunt, K.H.: *Kinematic Geometry of Mechanisms*. Oxford University Press, New York (1990)
- [14] Sangeorzan, B.P., Judd, R., Sangeorzan, B.J.: Mathematical analysis of single-cut osteotomy for complex long bone deformity. *Journal of Biomechanics* **22** (1989) 1271–1278
- [15] Sangeorzan, B.P., Sangeorzan, B.J., Hansen, S., Judd, R.: Mathematically directed single-cut osteotomy for correction of tibial malunion. *Journal of Orthopaedic Trauma* **3** (1989) 267–275

Predictive Camera Tracking for Bronchoscope Simulation with CONDensation

Fani Deligianni, Adrian Chung, and Guang-Zhong Yang

Department of Computing, Imperial College London
{fani.deligianni, adrian.chung}@ic.ac.uk,
g.z.yang@imperial.ac.uk
<http://vip.doc.ic.ac.uk>
<http://www.doc.ic.ac.uk/~ajchung/VISaVE>

Abstract. This paper exploits the use of temporal information to minimize the ambiguity of camera motion tracking in bronchoscope simulation. The condensation algorithm (Sequential Monte Carlo) has been used to propagate the probability distribution of the state space. For motion prediction, a second-order auto-regressive model has been used to characterize camera motion in a bounded lumen as encountered in bronchoscope examination. The method caters for multi-modal probability distributions, and experimental results from both phantom and patient data demonstrate a significant improvement in tracking accuracy especially in cases where there is airway deformation and image artefacts.

1 Introduction

In surgery, the value of minimally invasive procedures in terms of reduced patient trauma and recovery time has been established for many years. The limitation of the technique due to the complexity of instrument control and a loss of 3D vision and tactile feedback means effective training of visual spatial perception and hand-eye coordination is crucial to its safe practice. Flexible fiber-optic bronchoscopy, for example, is normally performed on patients who are fully awake or with light conscious sedation. The procedure can therefore entail considerable discomfort if it is not handled properly. Training according to the classical apprenticeship scheme is useful but can result in prolonged surgical procedures with increased patient discomfort and a potential risk for further complications. The use of computer simulation, particularly the reliance on patient specific data for building anatomical models both in terms of biomechanical fidelity and photorealism has attracted extensive interests in recent years [1],[2, 3]. Existing work has shown that by fusing real-bronchoscopy video with 3D tomographic data with the same patient, it is possible to generate photorealistic models that allow high fidelity, patient specific bronchoscope simulation [3]. The prerequisite of the technique, however, is accurate 2D/3D registration for recovering the pose of the camera in video bronchoscope sequences [2, 4]. Since bronchoscope images only provide localized views of the inner lumen, image-based technique cannot guarantee the convergence of the registration algorithm. To circumvent this problem, temporal constraints can be used to provide a predictive model by exploiting the temporal coherence of the camera movement.

For video bronchoscope navigation, Naga *et al* used Kalman filtering to increase the speed and accuracy of the registration algorithm [5]. Kalman filter, however, is generally restricted to situations where the probability distribution of the state variables is unimodal [6]. In bronchoscopy, tissue deformation, inter-reflection and view dependent specularly due to mucosa can limit the accuracy of image-based algorithms. The resultant probability density function of the state vector is typically multi-modal. Therefore, the observation probabilistic model cannot be approximated as a Gaussian distribution. The purpose of this paper is to develop a predictive tracking algorithm based on the Condensation algorithm [7]. The method is designed to cater for the general situation when several competing observations forming a non-Gaussian state-density. It uses a stochastic approach that has no restriction on the system/measurement models used and the distribution of error sources. An autoregressive algorithm is used as a predictive model [8], which is based on the fact that during bronchoscope navigation, the motion is restricted within a bounded area and a rapidly moving camera is more likely to slow down or change direction, rather than accelerate further. The proposed method provides a systematic learning procedure with modular training from the ground truth data such that information from different subjects are integrated for creating a dynamical model that accommodates the learnt behavior [9], [10].

2 Method

2.1 Temporal Tracking with the Condensation Algorithm

The main idea of the statistical framework is to maintain a time-evolving probability distribution $p(x_t | Z_t)$ of the tracker state x_t over time t . The *a posteriori* density $p(x_t | Z_t)$ represents knowledge about x_t deducible from the observation history $Z_t = \{z_1, \dots, z_t\}$. The Condensation algorithm, also known as Sequential Monte Carlo Tracking and particle filter, does not explicitly represent the density function. It uses instead Statistical Factor Sampling, which provides a way of approximating $p(x_t | Z_{t-1})$ by using a random number generator for sampling $p(x_t)$. The sampling technique is based on the Bayesian theory which gives [11]:

$$p(x | z) \propto p(z | x)p(x) \quad (1)$$

During statistical factored sampling, a sample set $\{s^{(1)}, \dots, s^{(N)}\}$ is generated from the prior density $p(x_t)$, where N is the number of sample sets. A weight is subsequently assigned to each particle according to the observation density of

$$\pi^{(n)} = \frac{p(Z | x = s^{(n)})}{\sum_{j=1}^N p(Z | x = s^{(j)})} \quad (2)$$

The weighted point set is then used as a representation of the *a posteriori* density, which is becoming more accurate as N increases. By evaluating given moments of the state density, it is possible to estimate the most probable state. During prediction, the method involves the evolution of the state vector with time. Sampling methods based on spatial

Markov processes are generally used for the Condensation Method to represent the dynamics of a moving object/camera. As an observation/measurement model, a pq -space based 2D/3D registration technique developed by Deligianni *et al* has been used [2]. The state of the system has been defined as the 6DoF pose of the camera.

2.2 Prediction Model

In order to construct a motion model for the endoscope camera that moves freely in the 3D tracheo-bronchial tree, an auto-regressive model is used [8]. This is in contrast to the ‘constant acceleration’ model used in [5], which effectively implies that the camera acceleration is expected to be constant during bronchoscope tracking. In this study, the auto-regressive model takes into account that during bronchoscope navigation, motion occurs within a bounded area, and a rapidly moving camera is expected to slow down or change in direction rather than accelerate further [12]. This ensures smooth navigation of the camera, and with modular training [9] multiple training sets can be used to obtain a more general model of the motion behavior. For a K^{th} -order auto-regressive model, the following equation can be used:

$$x_t = \sum_{k=1}^K A_k x_{t-k} + d + Bwt \tag{3}$$

where A_k represents the collection of damped harmonic oscillators associated with vibrational modes, d is a drift per unit time, and w the white noise with covariance coefficient B .

2.3 Training

In practice, it is possible to build a tracking model by approximating its general behavior to intuitive expectations of the observed motion. However, a hand-built model is a difficult in this study due to the high-dimensionality and complex motion involved. Mathematically, learning motion characteristics from a training sequence is to estimate the coefficients A_k , the mean value \bar{X} , and the random component B of an autoregressive process that best model the motion in a training sequence involving camera poses of x_1, \dots, x_M . The estimated pose x_t can be treated as the exact observation of the physical process, and by following the multi-variate algorithm of [8], the auto-correlation coefficients $R_{i,j}$ and $R'_{i,j}$ can be computed for $i, j = 0, 1, 2$ as

$$R_i = \sum_{t=3}^M x_{t-i}, \quad R_{i,j} = \sum_{t=3}^M x_{t-i} x_{t-j}^T, \quad R'_{i,j} = R_{i,j} - \frac{1}{M-2} R_i R_j^T \tag{4}$$

Subsequently, A_1, A_2 and D are given by:

$$\left\{ \begin{array}{l} A_2 = (R'_{02} - R'_{01} R'_{11}{}^{-1} R'_{12}) (R'_{22} - R'_{21} R'_{11}{}^{-1} R'_{12})^{-1} \\ A_1 = (R'_{01} - A_2 R'_{21}) R'_{11}{}^{-1} \\ D = \frac{1}{M-2} (R_0 - A_2 R_2 - A_1 R_1) \end{array} \right. \tag{5}$$

The mean of the AR process, on the other hand, is calculated as

$$\bar{X} = (\mathbf{I} - A_2 - A_1)^{-1} D \quad (6)$$

Finally, the covariance coefficient B is estimated as a matrix square root of C

$$C = \frac{1}{M-2} (R_{00} - A_2 R_{20} - A_1 R_{10} - D R_0^T) \quad (7)$$

For the purpose of bronchoscope simulation, it is more meaningful to collect several training sets from the same as well as different operators in order to construct a more representative dynamic model. To this end, the auto-correlation coefficients of each training set have been calculated individually and then combined in a linear fashion. Since each of these dynamic systems may have a different mean value, the use of pre-estimated mean value of the system can result in a prediction strongly biased. In this study, \bar{X} is estimated on-line as part of the state vector.

2.4 Validation

In order to assess the accuracy of the proposed algorithm, an airway phantom made of silicone rubber and painted with acrylics was constructed. The phantom has a cross sectional diameter of 12cm at the opening and narrows down to 5cm at the far end. The inside face was created such to give the surface a specular finish that looks similar to the surface of the lumen. A real-time, six degrees-of-freedom Electro-Magnetic (EM) motion tracker (FASTRAK, Polhemus) was used to validate the 3D camera position and orientation, as illustrated in Fig. 2(c). The EM-tracker has an accuracy of 0.762mm RMS. The tomographic model of the phantom was scanned with a Siemens Somatom Volume Zoom four-channel multi-detector CT scanner with a slice thickness of 3mm and in-plane resolution of 1mm. A CMOS camera and NTSC standard with frame rate of 30fps was used.

For *in vivo* validation, bronchoscopy examination was performed in five patients according to a standard clinical protocol. During the bronchoscope procedure two similar type videoscopes (Olympus BF Type; with field of view 120°) were used. Video images from the bronchoscope examination were transferred to digital videotapes in PAL format at 25fps. Since the original endoscopic video frames contain both the endoscopic image and redundant black background, only the endoscopic view was digitized and cropped to images of 454×487 pixels. All images were converted to grayscale before the pq -space analysis. Similar to the phantom study, the CT images were acquired from the Siemens Somatom Volume Zoom four-channel multi-detector CT scanner with a slice width of 3mm and collimation of 1mm, and the acquisition volume covered from the aortic arch to the dome of hemidiaphragm. Pre-processing of the video images was necessary in order to alleviate the effects of interlacing, lens distortion and unnecessary texture information. To remove noise and image artifacts, anisotropic filtering was applied to each image.

3 Results

Figure 1 illustrates the effectiveness of the training process involved in this study. The ground truth data of the camera pose from four different patients have been used to

train the auto-regressive model. Subsequently, the performance of the trained model was evaluated on the fifth patient data. The Euclidean distance between the first and subsequent camera positions predicted from the condensation algorithm was used for error analysis. Similar analysis for the error in orientation was also performed. In Figure 1(a), the system mean has been predefined according to the training sequences, whereas for (b), it was included in the state vector of the predictive model such that it was updated in real-time. It is evident that without continuous updating of \bar{X} , the derived tracking value is heavily biased.

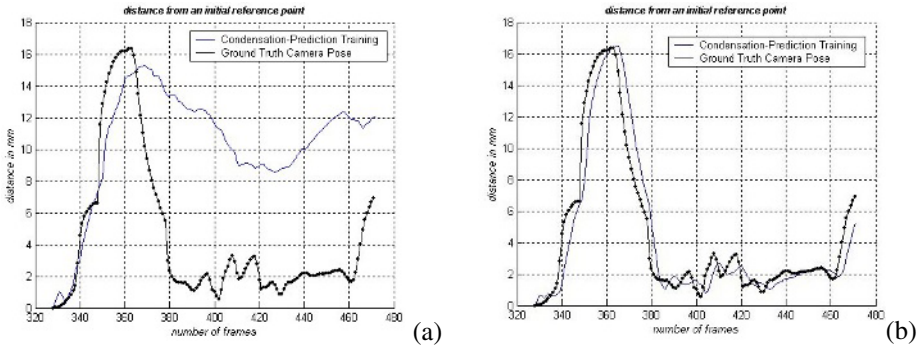


Fig. 1. Assessment of the accuracy of the training model and the effect of excluding (a), and including (b), mean value estimation of \bar{X} as part of the state vector

Table 1. Quantitative assessment of the pq -space based registration with and without the condensation algorithm (pq - CD), respectively

Case	Frames	Absolute Error							
		Position (mm)				Angle (rad)			
		Pq		CD		Pq		CD	
		Mean	\pm Std	Mean	\pm Std	Mean	\pm Std	Mean	\pm Std
Phantom	300	36.5	22.7	5.89	5.5	0.14	0.1	0.12	0.07
Pat1	100	7.0	2.0	3.3	2.2	0.8	0.2	0.03	0.02
Pat2	120	13.3	5.8	2.6	1.7	0.35	0.14	0.11	0.06
Pat3	243	26.98	13.0	1.9	1.52	1.02	0.6	0.11	0.06
Pat4	144	5.87	5.5	3.15	2.1	0.06	0.1	0.05	0.04
Pat5	100	14.4	10.12	2.26	1.65	0.7	0.4	0.19	0.15

The detailed assessment results for the phantom and patient studies are summarized in Table 1. It is evident that in both cases the 2D/3D registration accuracy has been increased significantly by the use of the proposed predictive tracking algorithm. More notably, the method permits more stable tracking results in patients where image artifact (*e.g.* partial occlusion of the images due to mucosa or bleeding) and sudden airway deformation due to coughing can introduce large propagational errors to the original pq -space registration technique. In Figure 2, we demonstrate the extent of this

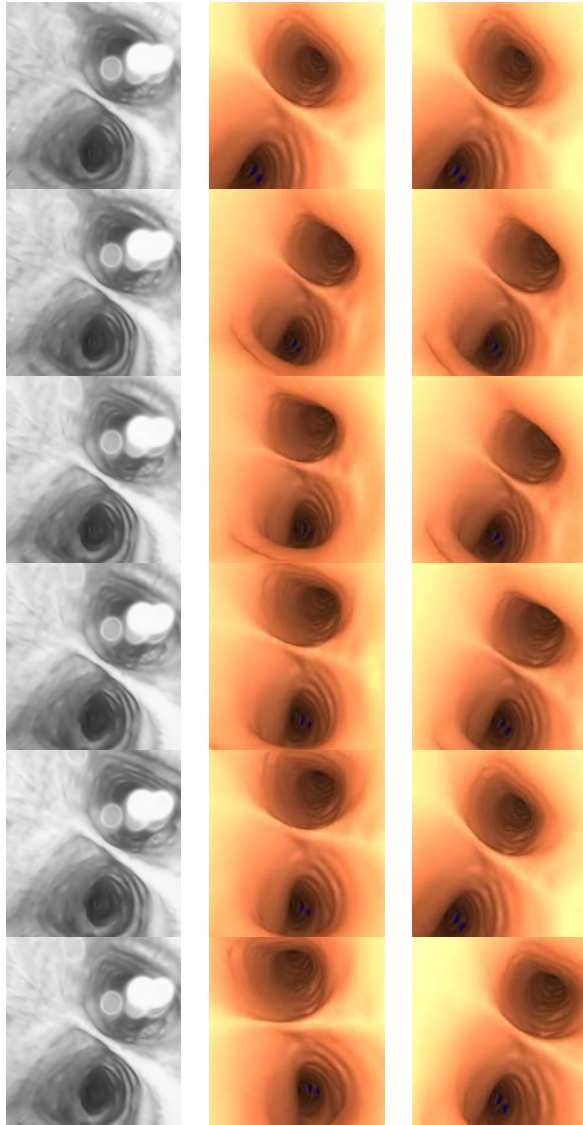


Fig. 2. The effect of airway deformation and partial occlusion of the image due to mucosa and blood on the accuracy of the 2D/3D registration technique without (mid-column) and with (right-column) predictive camera tracking

effect on the actual bronchoscope views. The left column shows the original frames from the bronchoscope video, whereas the middle and right columns are the virtual views of the 3D model by using pq -space registration without and with predictive camera pose tracking. It is worth noting that pre-processing, including radial distortion correction, de-interlacing and anisotropic filtering, has been applied to the real bronchoscope images before the registration step.

4 Discussion and Conclusions

In this paper, we have described the use of predictive camera tracking for increasing the general accuracy and robustness of 2D/3D registration involved in virtual bronchoscope modeling. A stochastic filter is used to resolve the inherent global ambiguity in tracking by exploiting the temporal coherence of the camera tip. The use of the condensation algorithm permits the use of multi-modal probability distributions, and our results from both phantom and patient data demonstrate a significant improvement in tracking accuracy especially in cases where there is airway deformation and image artifacts. The method effectively avoids the registration algorithm being trapped in local minima. The use of auto-regressive model based on the principles of the maximum likelihood learning and extension to modular learning has facilitated the incorporation of multiple sequences from different patients. The proposed method can be further extended to multi-class motion description such that the dynamic behavior of camera navigation in different parts of the tracheo-bronchial tree can be incorporated.

References

1. Indelicato, D., "Virtual Reality in Surgical Training", *The Dartmouth Undergraduate Journal of Science*, vol. I, no. 1, 1999.
2. Deligianni, F., Chung, A., and Yang, G.-Z., "pq-Space 2D/3D Registration for Endoscope Tracking", in Proceedings of Medical Image Computing & Computer Assisted Intervention (MICCAI03), Montréal, Québec, Canada, 2003, pp. 311-318.
3. Chung, A. J., Deligianni, F., Shah, P., Wells, A., and Yang, G.-Z., "Enhancement of Visual Realism with BRDF for Patient Specific Bronchoscopy Simulation", in Proceedings of MICCAI, Rennes, France, 2004, pp. 486-493.
4. Mori, K., Deguchi, D., Sugiyama, J., Suenaga, Y., Toriwaki, J.-i., Jr., C. R. M., Takabatake, H., and Natori, H., "Tracking of a bronchoscope using epipolar geometry analysis and intensity-based image registration of real and virtual endoscopic images", *Medical Image Analysis*, vol. 6, no. 3, 2002, pp. 181-336.
5. Naga, J., Mori, K., Enjouji, T., Deguchi, D., Kitasaka, T., Suenaga, Y., Hasegawa, J.-i., Toriwak, J.-i., Takabatake, H., and Natori, H., "Fast and Accurate Bronchoscope Tracking Using Image Registration and Motion Prediction", in Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI04), Saint-Malo, France, 2004, pp. 551-558.
6. Welch, G. and Bishop, G., "An Introduction to the Kalman Filter", in Proceedings of ACM SIGGRAPH, 2001.
7. Isard, M. and Blake, A., "CONDENSATION-Conditional Density Propagation for Visual Tracking", *International Journal of Computer Vision*, vol. 29, no. 1, 1998, pp. 5-28.
8. Blake, A. and Isard, M., *Active Contours*, London, Springer, 1998.
9. Wildenberg, A. P., *Learning and Initialisation for Visual Tracking*, University of Oxford, Oxford, UK, Modular Training for Visual Tracking Thesis, 1998.
10. North, B., Blake, A., Isard, M., and Rittscher, J., "Learning and classification of complex dynamics", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 9, 2000, pp. 1016-1034.
11. Arulampalam, M. S., Maskell, S., Gordon, N., and Clapp, T., "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking", *IEEE Transactions on Signal Processing*, vol. 50, no. 2, 2002, pp. 174-188.
12. Davison, A. J., "Simultaneous localization and map-building using active vision", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, 2002, pp. 865-880.

Experimental Validation of a 3D Dynamic Finite-Element Model of a Total Knee Replacement

Joel L. Lanovaz¹ and Randy E. Ellis^{1,2,3}

¹ Department of Mechanical and Materials Engineering,
Human Mobility Research Center, Queen's University, Canada
lanovaz@me.queensu.ca

² School of Computing, Queen's University, Canada
ellis@cs.queensu.ca

³ Division of Orthopaedics, Department of Surgery, Queen's University, Canada

Abstract. A 3D forward-dynamics model of a total knee replacement was developed using an explicit finite-element package. The model incorporated both a tibiofemoral and a patellofemoral joint and allowed full 6-DOF kinematics for both joints. Simulated quadriceps contraction was used to drive the model. For validation, a unique experimental apparatus was constructed to simulate an open-chain extension motion under quadriceps control. The ligamentous constraints of the MCL and LCL were simulated using tension springs. The kinematics of the tibia and patella were recorded along with the net forces and moments applied to the femur. Several ligament and inertial configurations were simulated. The RMS differences between the experimental data and model predictions across all simulations were excellent for both the kinematics (angles: 0.3 - 1.6°, displacements: 0.1 - 0.8 mm) and kinetics (forces: 5 - 11 N, moments: 0.2 - 0.6 Nm). The validated model will be extended with physiologically realistic ligaments and utilized in surgical planning simulations.

1 Introduction

Even though total knee replacement (TKR) is a relatively routine orthopedic procedure, it has been demonstrated that the introduction of image guided and computer assisted surgery (CAS) techniques to this task can decrease the variability of standard operations and increase the potential for good outcomes in complicated situations [1, 2].

Preoperative planning with current CAS systems utilize component manufacturer's guidelines, standard orthopedic practices and the surgeon's experience to select optimal component positioning. These criteria can be difficult to apply in complex procedures and may not fully take into account dynamic or loading effects that could accelerate component wear.

Intraoperative placement with CAS systems precisely reproduces the preoperative plan. However, the final *in vivo* performance of the components is still difficult to assess. Ligaments and soft tissue balancing can change the expected kinematics, leading to increased wear and decreased component life.

The focus of this research is to develop a realistic, functionally based dynamic TKR model that can assist in both the planning and execution of the surgery. Ultimately, the goal is to provide the surgeon with the ability to predict the postoperative behavior of TKR components based on preoperative and intraoperative data.

Dynamic finite element (FE) analysis is becoming a popular modeling technique for examining various aspects of TKR mechanics [3, 4, 5, 6, 7]. Most current models focus on simulations for the purpose assessing wear behavior, usually by modeling a wear testing standard. All of the dynamic FE models presented in the literature to date utilize some prescribed kinematics as a control method. Few FE models incorporate both the tibiofemoral and patellofemoral joints [7].

This paper presents an explicit dynamic FE model of a TKR. The model incorporates the femoral, tibial and patellar components and is driven by simulated quadriceps action. A unique experimental validation procedure is described which simulates ligament constraints and assesses both the kinematic and kinetic performance of the model.

2 Materials and Methods

This study modeled posterior-cruciate-retaining TKR components (Sigma PFC, Depuy Orthopaedics, Indiana, U.S.A.). A standard right side, size 3 femoral component was used with a corresponding 10 mm thick tibial UHMWPE insert. A standard 10 mm thick UHMWPE patellar resurfacing button was also used.

2.1 Experimental Apparatus

An experimental apparatus was designed to simulate an open-chain knee extension motion (Fig. 1). The device incorporated a fully mobile patella and shank, with an actuator-controlled quadriceps and a simulated patellar tendon. The medial and lateral ligament constraints were also simulated.

The base of the apparatus consisted of a 6-DOF force transducer (MC3A-250, AMTI, Massachusetts, U.S.A.) mounted on a rigid aluminum frame. The femoral component was fixed to a mounting block which was secured to the free end of the force transducer such that the femoral long axis was parallel to the horizontal and the trochlear groove was facing up.

To remove the confounding effects of back-side motion, the tibial component was fixed directly to a mounting block which was bolted to the proximal end of a simulated shank segment. The shank segment consisted of 4 threaded rods, each 300 mm long, held in place by Plexiglas end plates. A 2.2 kg mass was fixed on the center line of the shank at the distal end.

Steel tension springs were used to simulate ligament constraints of the knee joint. Spherical rod-end bearings were bolted to Plexiglas plates that were rigidly attached to the frame on the medial and lateral sides of the femoral component. Bearings were also bolted to custom-fabricated ABS plastic mounting blocks fixed to the shank. Each tension spring was attached to a femur bearing on one end and a shank bearing on the other. In this way, ligament constraints similar to the MCL and LCL could be represented.

The springs were nearly identical and calibration showed a linear stiffness throughout the operating range (average 1.36 N/mm, SD 0.01 N/mm). Each spring produced tensile forces in the range of 40 N to 60 N during the various experimental trials.

The quadriceps and patellar tendon were simulated using 1.6 mm diameter steel cables. One end of the quadriceps cable was attached, via low friction pulleys, to the

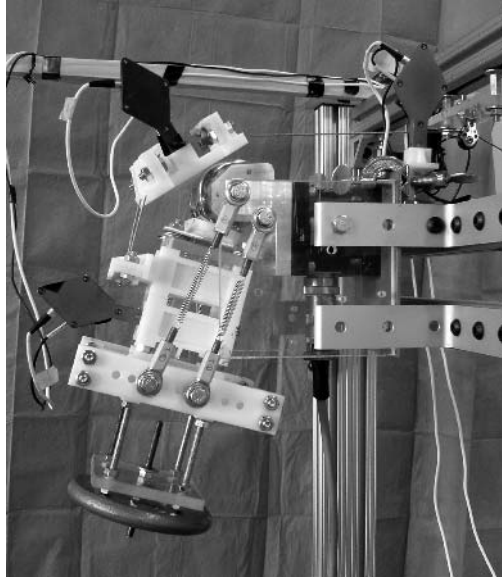


Fig. 1. Experimental apparatus in a 4 spring symmetrical configuration

cross head of a computer-controlled material testing machine (Model 5500R, Instron, Massachusetts, U.S.A.). Two 90 mm lengths of cable were used to simulate the patellar tendon.

A custom mounting block for the patellar button was fabricated from ABS plastic, which included attachment points for the quadriceps cable and the patella tendon cables. The cable attachments were lubricated ABS spherical bearings. Attachment points for the patellar tendon cables were also mounted on the shank segment. Once assembled, the motion of the joint complex was restricted so that the patella button remained in contact with the trochlear groove on the femoral component.

In order to validate the FE model, both kinematic and kinetic variables were experimentally measured. Optical tracking arrays were rigidly fixed to the shank, the patella mounting block and to the apparatus frame. An Optotrak 3020 system (NDI, Waterloo, Canada) was used to record the 3D kinematic data. The components were located relative to their respective tracking arrays through fiducial and shape matching registration based on points obtained with a calibrated optical tracking probe. The location and orientation of the force transducer, along with the line of action of the quadriceps cable, were also obtained with the probe. Insertion locations of the springs and cables were known in the coordinate frames of the respective rigid bodies to which they were attached.

For the dynamic experimental trials, the kinematic data and the conditioned voltage signals from the force transducer were collected simultaneously with the Optotrak system at a sampling rate of 100 Hz.

For comparison with the FE model, all kinematic data were transformed into the transducer coordinate frame with the positive X-axis corresponding to the femoral

medial direction, the positive Y-axis to the femoral anterior direction and the positive Z-axis to the femoral distal direction.

2.2 Finite Element Model

Three dimensional digital point cloud representations of the TKR component geometries were obtained by laser scanning (nominal scan resolution 0.05 mm). Commercial software (Imageware, EDS Corp., Texas, U.S.A.) was used to fit surfaces (fit tolerance of 0.1 mm) to the 6 contact regions (3 for the femur, 2 for the tibia and 1 for the patella) of each of the 3 components. The contact surfaces were saved in IGES format and transferred to meshing software (CUBIT, Sandia National Laboratories, New Mexico, U.S.A.).

Quadrilateral meshes were constructed for each contact surface. Pilot studies examining model convergence showed that a stable solution could be found for this model using a nominal element edge size of 1.0 mm. This resulted in approximately 925 quadrilaterals representing each of the tibial contact regions, 1550 quadrilaterals representing each of the femoral contact regions and 950 quadrilaterals representing the patella. Hexahedral meshes for all components were generated from the quadrilateral meshes using custom routines written in Matlab (Mathworks, Massachusetts, U.S.A.).

The dynamic explicit FE package LS-DYNA (Livermore Software Technology Corp., California, U.S.A.) was utilized as the solver. A non-linear elastic-plastic deformable material model was used for all of the UHMWPE components [5, 6]. Because the elastic modulus of the CoCrMo femoral component was more than two orders of magnitude greater than the UHMWPE, the femoral component was modeled as a rigid material. The full thickness of the UHMWPE components were modeled, whereas the femoral components were made 1.5 mm thick (for visualization purposes only). A rigid layer of elements was fixed to the bottom layer of the UHMWPE elements in both the tibial and patella meshes to simulate the boundary conditions in the experimental apparatus.

The springs and cables were modeled with linear tension-only beam elements. The actuation of the quadriceps cable was modeled using a linear motor joint.

2.3 Model Validation

Four configurations were tested experimentally. The first had one medial and one lateral spring with symmetrical insertions. The second was a two-spring asymmetrical configuration with the shank insertions placed anterior to the femur insertion on the medial side and posterior to the femur insertion on the lateral side. The third configuration had the same spring insertions as the first but with an additional 0.5 kg mass fixed at the distal end of the shank, offset 150 mm to the medial side. The final configuration was a symmetrical four-spring design.

Average results were obtained from four trials for each configuration. For each trial, the knee joint started in the most flexed position with the patella button still in contact with the femur. The material testing machine was programmed to pull the quadriceps cable at a constant rate of 16 mm/s for 2.5 seconds. This produced an extension of the

knee joint from 60° to 25° of flexion over that time period. For all trials, the contact surfaces were lubricated with 20°C distilled water.

The FE model simulated the complete inertial properties of the shank and patella, including the tracking arrays. The solution was run in explicit mode with a time step of 10 μs. The initial poses of the component models were obtained from the experimental data. A 500 ms period was given at the start of each solution run to stabilize contact prior to initiating the cable actuator. A penalty-based contact algorithm was selected in LS-DYNA that calculated reaction forces based on nodal penetration and material characteristics. A dynamic coefficient of friction of 0.10 was applied to each of the three CoCrMo/UHMWPE contact regions in the model; this value was obtained through pilot studies and is consistent with most values from the literature [3, 4, 5].

3 Results

The FE model was first evaluated by comparing the predicted 3D kinematics of the tibia (Fig. 2) and patella (Fig. 3) to the experimental values. The kinematic data were all calculated with respect to the force transducer coordinates. The rotations were expressed using three Cardan angles while the displacements were expressed relative to the transducer origin.

The nodal contact forces applied to the femur model were used to predict the net forces and moments recorded by the force transducer (Fig. 4).

The RMS differences between the experimental data and the FE predictions for all 4 configurations were calculated (Table 1). The maximum relative RMS differences occurred when predicting the rotational kinematics of the patella and these were mainly due to offsets that occurred during the initialization phase.

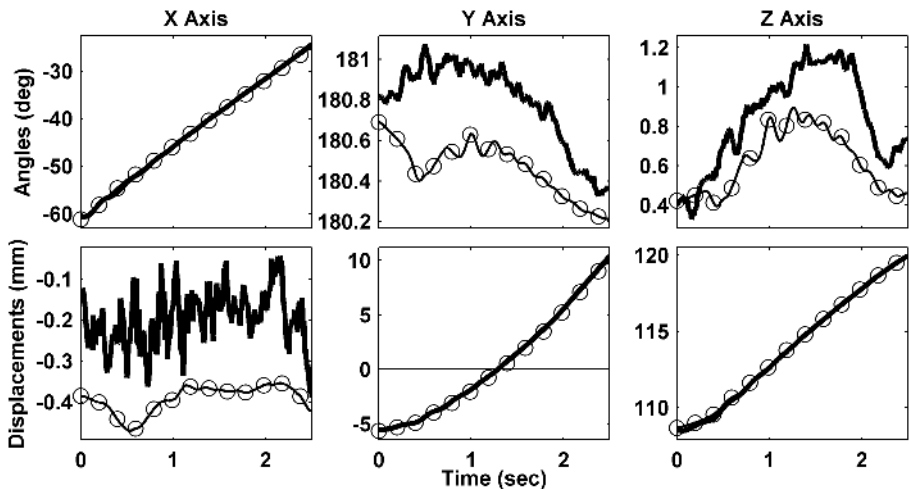


Fig. 2. Tibial kinematics comparing experimental (solid lines) and FE results (circles) from the two spring symmetrical configuration

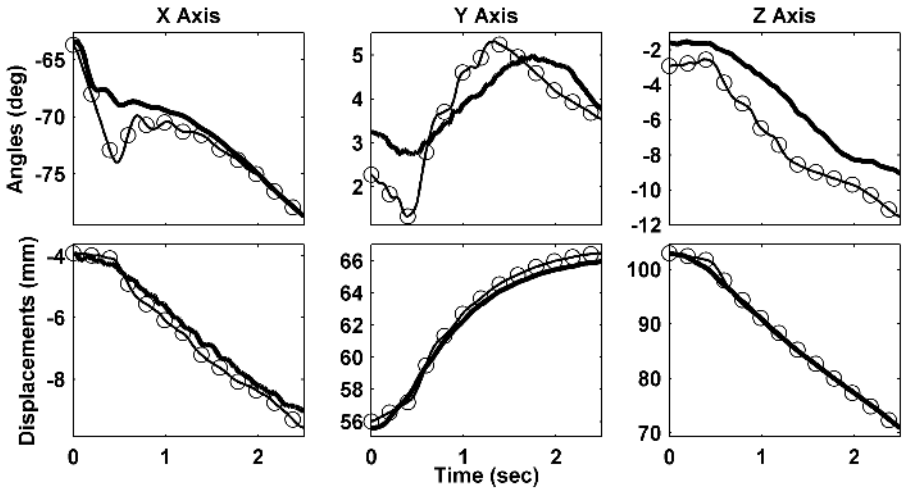


Fig. 3. Patellar kinematics comparing experimental (*solid lines*) and FE results (*circles*) from the two spring symmetrical configuration

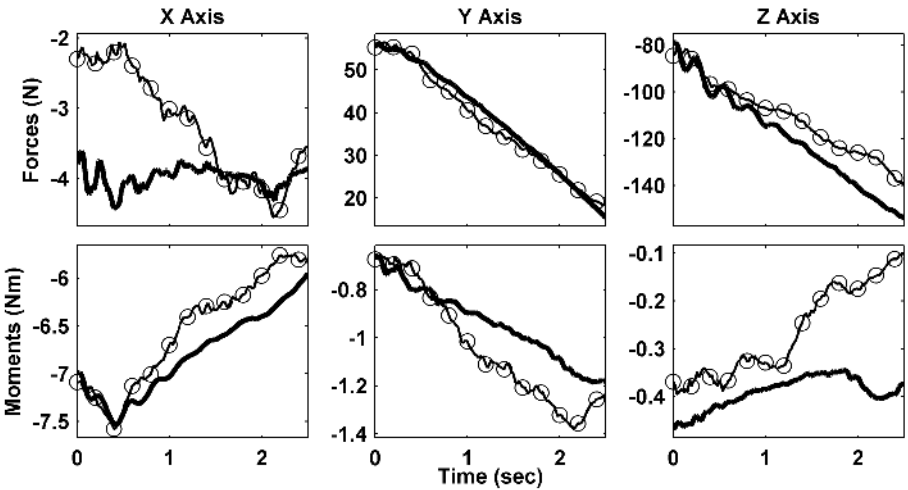


Fig. 4. Comparison between experimentally measured net forces and moments applied to the femur (*solid lines*) and the FE results (*circles*) from the two spring symmetrical configuration

4 Discussion

We wished to validate the FE model by simulating an open-chain knee extension motion under simulated quadriceps control. The difficulty in obtaining accurate geometric data and material properties during *in vitro* studies suggested that a more controlled mechanical simulation would be a better alternative. The experimental apparatus used

Table 1. The RMS differences between the experimental data and FE models. The angle data are the means of the 3 axes. All other data are vector sums. Model 1: 2 symmetrical springs, Model 2: 2 asymmetrical springs, Model 3: 2 symmetrical springs with offset mass, Model 4: 4 symmetrical springs.

Model	Tibia		Patella		Femoral Transducer	
	Angle (deg)	Position (mm)	Angle (deg)	Position (mm)	Force (N)	Moment (Nm)
1	0.31	0.16	1.58	0.50	10.04	0.29
2	0.46	0.20	1.23	0.48	11.14	0.26
3	0.47	0.27	1.05	0.41	5.04	0.34
4	0.43	0.13	1.05	0.74	6.27	0.56

in this study enabled the simulation of ligament-like constraints with the advantage of having precise material characteristics.

The FE model performed well under all 4 test conditions. The tibial kinematics were predicted with a high degree of fidelity. The RMS differences for the tibial displacements and angles is at the limit of the accuracy of the Optotrak system. The patellar angle predictions showed the highest RMS differences but were still under 2° and closely followed the shape of the experimental data. It is not unreasonable to expect less accuracy predicting the patellar kinematics due to its low relative mass. Under the higher loading conditions of the 4 spring configuration, the patellar predictions improved.

Direct validation of dynamic contact models is difficult. Pressure measurement systems are available but can alter the contact characteristics [8]. One of the motivations in collecting the net forces and moments applied to the femur was to provide an indirect assessment of the contact data in the model. Because both the kinematics and kinetics were reproduced faithfully, we can assume that the contact information in the model is valid. This is a key component in this work because, in the future, the model will be used to calculate forward dynamics based on functional control inputs (such as muscle forces) as opposed to prescribed kinematic parameters (such as flexion angle).

LS-DYNA has the capability, similar to other FE packages [7], of switching to an all-rigid analysis without having to change the model. In this mode, kinematics and contact pressures are still calculated, but in less than 10% of the computational time. This provides a useful method of obtaining an efficient simulation if internal stress information is not required.

5 Conclusions and Future Work

A dynamic FE model of a TKR was presented that incorporates the femoral, tibial and patellar components. The model simulated a dynamic extension motion under quadriceps control. A validation experiment was conducted and the model was able to accurately predict the kinematics of both the tibiofemoral and patellofemoral joints along with predicting the net forces and moments applied to the femur. To the best of our knowledge, this is the first experimental validation of a dynamic FE model of a TKR to employ simultaneous kinematic and force data.

Such models are valuable because they can be used to predict physiological motion prior to surgery. This is of particular importance for patients with significant deformities or poor overall health, both of which are predictors for poor surgical outcomes. A dynamic knee model is also of great use in the design of new prostheses, as it makes it possible to reduce or eliminate expensive, time-consuming physical tests.

Future work will include the addition of physiologically realistic ligament structures and bone/ligament interaction. Further experimental validation will be conducted *in vitro*, including simulation of TKR surgeries and predicting outcomes of simulated muscle-controlled loading from MRI-derived ligament data.

References

1. Stulberg, S.D., Loan, P., Sarin, V.: Computer-assisted navigation in total knee replacement: Results of an initial experience in thirty-five patients. *Journal of Bone and Joint Surgery* **84** (2002) 90–98
2. Victor, J., Hoste, D.: Image-based computer-assisted total knee arthroplasty leads to lower variability in coronal alignment. *Clinical Orthopaedics and Related Research* **428** (2004) 131–139
3. Giddings, V.L., Kurtz, S.M., Edidin, A.A.: Total knee replacement polyethylene stresses during loading in a knee simulator. *Journal of Tribology-Transactions of the ASME* **123** (2001) 842–847
4. Otto, J.K., Callaghan, J.J., Brown, T.D.: The coventry award paper - Mobility and contact mechanics of a rotating platform total knee replacement. *Clinical Orthopaedics and Related Research* **392** (2001) 24–37
5. Godest, A.C., Beaugonin, M., Haug, E., Taylor, M., Gregson, P.J.: Simulation of a knee joint replacement during a gait cycle using explicit finite element analysis. *Journal of Biomechanics* **35** (2002) 267–275
6. Taylor, M., Barrett, D.S.: Explicit finite element simulation of eccentric loading in total knee replacement. *Clinical Orthopaedics and Related Research* **414** (2003) 162–171
7. Halloran, J.P., Petrella, A.J., Rullkoetter, P.J.: Explicit finite element modeling of total knee replacement mechanics. *Journal of Biomechanics* **38** (2005) 323–331
8. Wilson, D.R., Apreleva, M.V., Eichler, M.J., Harrold, F.R.: Accuracy and repeatability of a pressure measurement system in the patellofemoral joint. *Journal of Biomechanics* **36** (2003) 1909–1915

An In Vitro Patient-Tailored Model of Human Cerebral Artery for Simulating Endovascular Intervention

Seiichi Ikeda¹, Fumihito Arai¹, Toshio Fukuda¹, Makoto Negoro²,
Keiko Irie², and Ikuo Takahashi³

¹ Dept. of Micro-Nano Systems Engineering, Nagoya Iniverisity,
Furo-cho 1, Chikusa-ku, Nagoya, 464-8603, Japan
ikeda@robo.mein.nagoya-u.ac.jp,
{arai, fukuda}@mein.nagoya-u.ac.jp,
<http://www.mein.nagoya-u.ac.jp/index.html>

² Dept. of Neurosurgery, Fujita Health University,
Kutsukake-cho, Toyoake, 470-1192, Japan

³ Dept. of Neurosurgery, Anjo Kosei Hospital,
Anjo-cho 28, Anjo, 446-8602, Japan

Abstract. An in vitro patient-tailored reproduction model of cerebral artery, a hardware platform for simulating endovascular intervention for making diagnoses and surgical trainings is presented. 3-D configuration of vessel lumen is reproduced as vessel model with 13 μm modeling resolution, using CT and MRI information. Physical characteristics of cerebral artery, such as elastic modulus and friction coefficient, are also reproduced. We also propose a novel method to visualize stress condition on vessel wall using photoelastic effect. Consequently, it should be helpful for clinical applications, academic researches and other various purposes.

1 Introduction

Recently established endovascular intervention is a potent treatment modality for various vascular diseases, which reduces the invasiveness for patients [1][2]. Many devices for endovascular intervention have been developed since late 1980 to treat with various diseases. However, the tortuosity and diversity of cerebral vasculature disturbs maneuvering of those sophisticated devices, so it requires an advanced surgical skill and experience. However current training environment for this surgery is inadequate, and this situation makes surgical skills more difficult to acquire [5].

Traditionally, interventionalists have solely depended on images projected on displays or papers as an only means to provide medical information obtained by angiography or other imaging modalities. Although, it is now possible to reconstruct the 3-D configuration of vascular system from 2-D fluoroscopic images, even so, accurate recognition of 3-D vascular configuration is still not easy and it may lead to misinterpret information.

As a solution for these difficulties, we propose an in vitro vessel model of cerebral artery that reproduces the 3-D configuration of patient's vascular lumens as membra-

nous silicone structure, utilizing CT and MRI information. With this model, it becomes possible for young trainees to practice surgical skills of endovascular intervention. And it becomes also possible with this model to preliminary simulate neurovascular procedures before surgery.

So far, dissection based modeling method for vasculature, (a method that utilize vessel as casting mold) was developed by P. Guilloud et al. And, this method is now largely utilized for various purposes [3][4]. However, the dissection model is not applicable for preliminary simulation and requires rather high time cost and labor cost for respective modeling.

2 Required Condition for Surgical Simulator

2.1 Patient-Tailored High Precision Modeling

Cerebral artery and cerebrovascular disease take very different configuration for respective patients (Fig.1), and the diversity of configuration contributes to make neurovascular procedures more difficult. Therefore conventional surgical simulator which imitates the general vascular configuration is inadequate for comprehensive surgical training. Here, it becomes very helpful to construct a patient-tailored vessel model using CT and MRI information which are usually available for all patient. Furthermore, as a simulator for neurovascular procedures, vessel lumen with less than 1mm in diameter need be reproduced

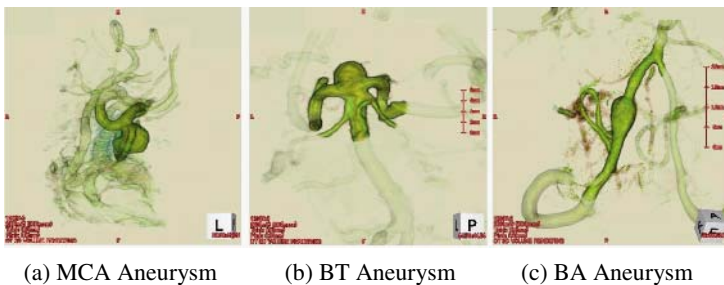


Fig. 1. Some typical configuration of cerebral aneurysm

2.2 Reproduction of Physical Characteristics

Operational feel plays an important role in treating cerebrovascular diseases safely without applying excessive piercing stress on vessel structure. Therefore, reproduction of physical characteristics of vesculature becomes important, to reproduce the difficulties of manipulating medial instruments and the behaviors of these medical instruments (e.g. slip / stick motion and unwinding motion of catheter) Meanwhile, vessel model should be compatible with medial imaging modalities (especially DAS), to allow simulating neurovascular procedures under practical IVR (interventional radiology) condition.

2.3 Reproduction of Visco-Elastic Vascular Deformation

Membranous vessel configuration and surrounding soft brain tissue allow cerebral artery to deform against surgical treatments, and some neurovascular treatments affirmatively utilize this characteristics. Expansion of stenosis with balloons and stents is an instance of this treatment. Meanwhile, this characteristics simultaneously results in technical difficulties. Therefore, to reproduce these important features, reproduction of visco-elastic deformation becomes essential

2.4 Summary of Required Conditions

As the required conditions for an in vitro vasculature model for simulating neurovascular procedures, we summarize the above contents as follows;

- 1) Patient-tailored with sub-millimeter resolution
- 2) Reproduced with thin membranous vacular configuration
- 3) Reproduced with material properties of vascular tissue
- 4) Reproduced with visco-elastic vascular deformation
- 5) Compatible with medical imaging modalities

3 An In Vitro Cerebral Arterial Model Reproduced with Human-Like Physical Characteristics

3.1 Rapid Prototyping for Patient-Tailored Modeling

Lately developed laminating modeling modality (rapid prototyping (RP) modality) has made it possible to construct organ models based on individual information, such as CT and MRI information. This RP-based organic model had proved its effectiveness for anatomical studies and for some clinical applications.

However, since the material used in RP modeling is limited, it is currently impossible to express the physical characteristics of organic tissues (current RP model is too rigid, fragile and anisotropic). Furthermore the RP modeling is usually accompanied tiered up rough surface (disturbs manipulating medical instruments, and deteriorates visibility). Consequently, the direct PR modeling modality is still not adequate for organic reproduction.

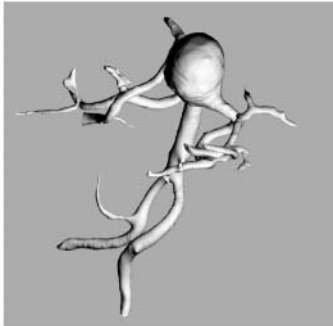
3.2 Production Methodology for Patient-Tailored Modeling

Here in after, we introduce a production method for an in vitro patient-tailored vasculature model that satisfies all the requirements described. Proposed production modality consists of the following 4 technical procedures; (1) reconstruction of 3-D vascular configuration, (2) fabrication of solid lumen model with RP modality, (3) fabrication of thin silicone membrane and (4) hollow construction of vascular lumen.

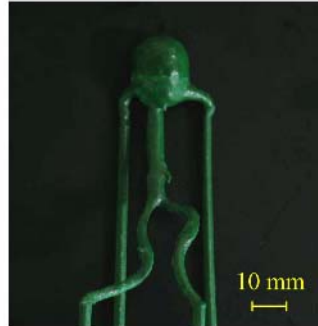
Firstly, we reconstructed the 3-D configuration of cerebral artery, using fluoroscopic image obtained with CT angiography (resolution: 0.3 mm, imaging pitch: 0.5 mm). We reconstructed the 3-D field of imaged area by heaping up these 2-D images, then extracted 3-D surface of cerebral artery by extracting area with identical CT value. Then we eliminated unimportant branches and discontinuous segment from this structure to

simplify to leave only basilar artery and aneurysm. By this means, we finally obtained a 3-D structure of basilar tip artery from the CT information (Fig. 2 (a)).

Then we fabricated solid structure of the reconstructed vascular lumen by RP modality (Fig. 2 (b)). Laminating thickness was 13 μm . We adopted fused deposition RP modality, since it allows the use of sulfonamide resin as construction material. Its material characteristics (melts at low temperature into high fluidity, easily dissolves in acetone) are considerably favorable in following modeling process.



(a) Reconstructed 3D figure of basilar artery reconstructed with CT angiography



(b) Solid lumen structure fabricated by RP modality (layering pitch: 13 μm)

Fig. 2. Materialization of individual vasculature structure based on CT / MRI angiography (small uninterested branches are ablated)

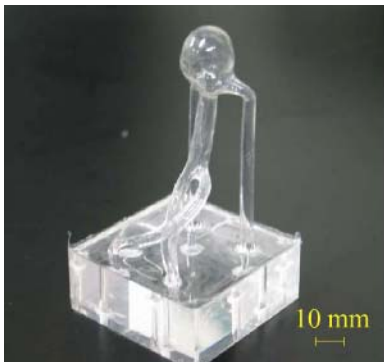


Fig. 3. Elastic membranous model of basilar artery fabricated based on individual information (inner RP model is eliminated)



Fig. 4. Cerebral arterial model with vascular-like membranous configuration and brain-like soft circumferential configuration

We fabricated thin uniform membrane of silicone elastomer (provides artery-like physical characteristics and good transparency) with 600 μm thick around this RP lumen structure, through velocity controlled dipping coating process and following addi-

tion polymerization process (Fig. 3; in this figure, solid lumen model is eliminated). After that, we embedded the structure within silicone gel (provides brain like elastic modulus) inside a cubic casting mold to reproduce circumferential soft brain structure.

Then we eliminated inward RP lumen structure through two independent processes. Firstly, we evacuated most part of the lumen model by selectively melting the RP structure at 120°C (higher than the melting point of RP model and lower than the heatproof temperature of silicone). Then we completely dissolved the remained quantity (especially in tight corners and dead ends) by injecting acetone inside lumens. With this procedure, the evacuation was managed in short period.

Consequently, we fabricated a transparent silicone model of basilar artery reproduced with thin membranous configuration and surrounding soft brain configuration that hollowly reproduces the vascular lumen of living patient with 13 μm modeling resolution (Fig. 4). As the reflection indexes of silicone elastomer and silicone gel are almost identical value (silicone elastomer (n_D^{25}): 1.410, silicone gel (n_D^{25}): 1.404), no optical distortion take place and it realized good visibility.

4 APPLICATION: Photoelastic Stress Visualization and Analysis

Transparent isotropic material, including silicone elastomer and polyurethane elastomer, shows temporal birefringent effect (double refraction) when external load is applied. And making use of this effect, we visualized stress distribution on proposed vessel model as fringe pattern (called photoelastic stress analysis). In this experiment, we applied transparent polyurethane elastomer (photoelastic coefficient: $3.5 \times 10^{-9} \text{ Pa}^{-1}$; Young's modulus: 1.8 MPa) as construction material to increase its sensitivity to birefringent effect. And with a general configuration of circular polariscope, stress condition on vascular membrane was clearly visualized as rainbow-colored interference fringes pattern as shown in Fig. 5.



Fig. 5. Stress distribution against catheter operation visualized by means of photoelastic effect with circular polariscope



Fig. 6. Stress distribution against catheter operation visualized by means of photoelastic effect with reflective circular polariscope

However, it is impossible, with this method, to distinguish the optical effect took place on fore vascular wall and rear wall. As a solution, we developed a new method for quantitative real-time stress analysis. In this method, projected polarized light, passing through fore wall with birefringent effect, is reflected at the surface of inner lumen filled with liquid added with aluminum powder (serves as mirror), then passes back the same point on fore wall with doubled birefringent effect (sensitivity becomes doubled than usual). Fig. 6 shows the visualized stress pattern with this method. As to this method, observed color can be related with phase shift R , and stress condition (max. principal stress σ_1 , min. principal stress σ_2 and direction of principal stress θ) on vasculature membrane is calculated from the following equation:

$$R = \alpha (\sigma_1 \cos^2 \theta + \sigma_2 \sin^2 \theta) D \quad (1)$$

Here, α is photoelastic coefficient, and D is thickness of vascular membrane (Numerical analysis is future work).

Presented method should be helpful for evaluating surgical procedures and the performance of medical systems, and also for various academic studies such as hemodynamic study and pathological studies.

5 Results and Discussion

Physical characteristics of presented model are very similar to arterial tissue. Although arterial tissue shows anisotropic nonlinear characteristics, elastic modulus of human artery normally ranges from 1 to 3 MPa, and applied silicone elastomer takes fairly close 1.9 MPa (Table 1). An this value is adjustable, by mixing silicone oil before polymerization, and it is possible to imitate sclerotic and other conditions. Friction coefficient between the surface of proposed model and LDPE (low-density polyethylene) catheter was 0.041. It is also fairly close to the vascular value 0.039 measured between artery surface and same LDPE catheter by R.A. Caldwell et al. (Table 1).

Proposed modeling method allowed constructing thin vascular lumens, and minimally attainable diameter was 300 μ m (geometrical error: less than 40 μ m). Furthermore, we realized artery-like smooth lumen surface by slightly dissolving the surface of RP lumen model with water (68 C) (Fig. 7). Here, this process less deteriorates its accuracy, as the dissolved material re-stick on the model surface. Furthermore, as shown in Fig. 8 and Fig. 9, thickness of vascular membrane is adjustable from less than 50 μ m to more than 250 μ m thick by adjusting withdrawing velocity in dip-coating process. Here, by layering several membranes with different thickness and different material property, it is possible, for example, to reproduce triple-layered general configuration of artery.

Consequently, proposed patient-tailored vasculature model reproduces artery-like visco-elastic deformation. Interventionalists verified the reproducibility of the visco-elastic motion, the operational feel coming from the deformation, and its usefulness for neurovascular simulations.

Since the applied elastomer and gel are transparent to visible light (transmissivity: > 88% over 10 mm thickness), it allowed clear recognition of inward condition, and this visibility is further improved by injecting glycerol solution (mixture of water (44%) and glycerin (56%)), provides the same reflection index) into vascular lumens.

Meanwhile, proposed vessel model confirmed compatible with major imaging modalities, such as CT, MR and supersonic imaging.

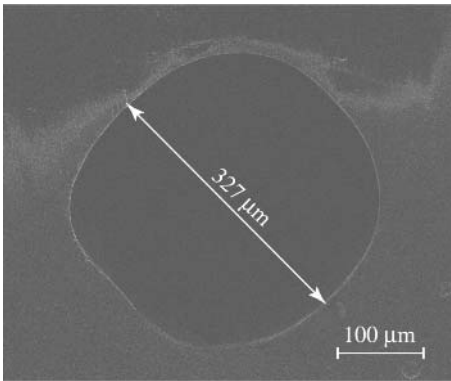


Fig. 7. SEM cross-section of fabricated vascular lumen with 300 μm in diameter (modeling error was less than 40 μm)

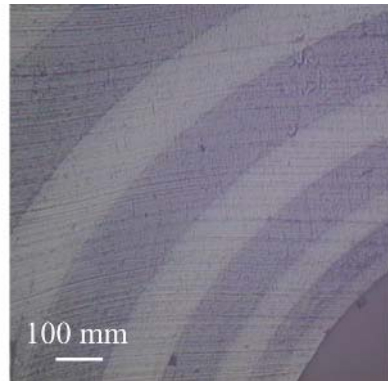


Fig. 8. Cross-section of layered silicone membrane with different wall thickness (from less than 50 μm to more than 250 μm)

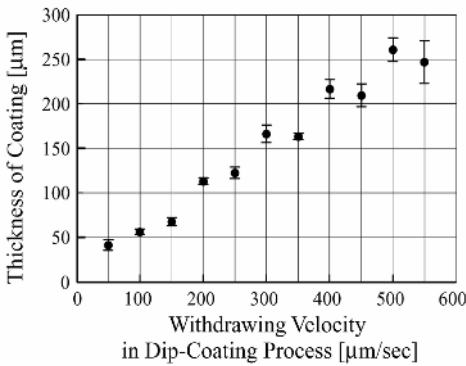


Fig. 9. Correlation between thickness of membrane and withdrawing velocity in dipcoating fabrication process

Table 1. Comparison of material properties between vessel model and human arterial

	Elastic Modulus [MPa]	Friction Coefficient
Cerebral Arterial Model	1.9	0.041
Arterial Tissue	1~3 (Carotid)	0.039

Table 2. Comparison of material properties between vessel model and human brain

	Elastic Modulus [kPa]
Cerebral Arterial Model	5.0
Brain Gray Matter	6.6 (Averaged)

We evaluated its effectiveness for the preliminary simulation. Representative case was female, age of 72, IC aneurysm (15 mm). It initially seemed wide-necked and difficult to treat by endovascular approach. In this case, we constructed a patient-tailored vessel model of IC aneurysm, and tested intervention procedures within this model (Fig. 10). In practical surgery, we could lead microcatheter into IC aneurysm from its orifice locating C2/C3 segment with the same manner as the preliminary

simulation (Fig. 11). We do not confirm any coil compaction in the follow-up after 6 months in this represent case.

Cerebral aneurysms are now widely treated with endovascular approach and scrupulous diagnosis became more and more important for securing safety. However, interventional radiology is currently the only information available. In this sense, proposed patient-tailored vascular model might be very helpful for making diagnosis and preliminary surgical simulations, since it allows swift and safe treatment for respective cases.



Fig. 10. Presurgical simulation (inserting microcatheter and Pt. coils) for surgical planning with patient-tailored cerebral arterial model



Fig. 11. DSA picture of neurovascular procedure, preliminary simulated with proposed patient-tailored vasculature model

References

1. Molyneux A, Kerr R, Stratton I, Sandercock P, et al.: International Subarachnoid Aneurysm Trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomised trial, *Lancet*, Vol. 360 (2002) 1267–74
2. P. Ng, M.S. Khangure, C.C. Phatouros, M. Bynevelt, et al.: Endovascular Treatment of Intracranial Aneurysms With Guglielmi Detachable Coils: Analysis of Midterm Angiographic and Clinical Outcomes, *Stroke*, Vol. 33 (2002) 210–217
3. C.W. Kerber, C.B. Heilman: Flowdynamics in the carotid artery: 1. Preliminary observations using a transparent elastic model, *Am. J. Neuroradiol*, Vol. 13 (1992) 173–180
4. P. Gailloud, J.R. Pray, M. Muster, M. Piotin, et al.: An in vitro anatomic model of the human cerebral arteries with saccular arterial aneurysms, *Surg Radiol Anat*, Vol. 19 (1997) 119–121
5. K. Sugiu, J.B. Martin, B Jean, P. Gailloud, et al.: Artificial Cerebral Aneurysm Model for Medical Testing, Training, and Research”, *Neurol Med Chir*, Vol. 43 (2003) 69–73
6. S. Ikeda, F.Arai, T.Fukuda, and M.Negoro.: An In Vitro Soft Membranous Model of Individual Human Cerebral Artery Reproduced with Visco-Elastic Behavior, *Proc. 2004 IEEE ICRA*, 2511–2516

Author Index

- Abolmaesumi, Purang I-83, II-197
Acar, Burak I-213
Adams, J.E. II-733
Adamson, Chris I-294
Admiraal-Behloul, Faiza II-451
Agid, Yves II-385
Agnus, Vincent II-155
Aharon, Shmuel II-773
Aït-Ali, L.S. I-409
Akdemir, U. II-781
Akiyama, Kenta II-543
Alexander, D.C. I-164
Alia-Klein, Nelly I-679
Aljabar, P. II-247
Andronache, Adrian II-976
An, Jung-ha I-286
Anxionnat, R. II-214
Arai, Fumihito I-925
Arbel, Tal II-163
Archip, Neculai II-862
Arridge, Simon R. II-616, II-649
Arsigny, Vincent I-115, II-927, II-943
Ashton, Edward I-451
Atkins, Stella I-868
Audette, Michel I-9
Auvray, Vincent II-352
Avants, Brian B. I-164, II-328
Avila, Ricardo S. I-688
Awad, Mike I-811, II-526
Awate, Suyash P. II-517
Ayache, Nicholas I-115, I-539, I-753,
II-295, II-385, II-492, II-927, II-943
Aylward, Elizabeth I-238
Aylward, Stephen R. I-59, I-221,
I-523, II-806
Azzabou, Noura I-597
- Bach Cuadra, Meritxell I-417
Baegert, Claire II-49
Bajka, M. II-575
Bammer, Roland I-213
Banovac, Filip II-992
Bardinet, Eric II-385, II-492
Bardyn, T. I-474
- Barillot, Christian I-409, I-597
Barkovich, J. I-548
Barratt, Dean C. I-605, II-1000
Barta, Patrick I-238
Bauernschmitt, Robert II-287
Bazin, Pierre-Louis II-484
Beg, M. Faisal I-238
Bello, Musodiq I-254
Berezney, Ronald I-876
Berger, M-O. II-214
Bergmann, H. II-960
Bernasconi, A. I-375
Bernasconi, N. I-375
Bertrand, Gilles II-394
Bhagwat, Jui G. I-506
Bhatia, Kanwal K. I-99
Bhattacharjee, Manik II-385
Bhotika, Rahul I-688
Bielen, Didier I-859
Birkfellner, W. II-960
Blackband, Stephen J. I-156
Bland, Peyton H. II-419
Blumenkranz, Steve I-811
Boas, David A. II-649
Bobick, Aaron II-459
Boctor, Emad M. I-621
Boesiger, Peter I-786
Boomsma, Kevin I-786
Borschneck, D.P. I-459
Bos, Clemens I-467
Boubertakh, Redha II-894
Boughner, Derek II-171
Bouix, Sylvain I-515
Boukerroui, Djamel I-262
Boulanger, Jérôme I-893
Bourgeat, Pierrick II-813
Bouthemy, Patrick I-893, II-352
Boyes, R.G. II-247
Bradley, P. I-885
Brandel, Jean-Philippe II-492
Broadbent, Kathy A. I-434
Bryant, J. Tim I-902
Buechel, Christian I-498

- Bullitt, Elizabeth I-26, I-59, I-383,
 II-239
 Buonaccorsi, Giovanni A. I-91
 Burschka, Darius I-811
 Bystrov, Daniel I-467

 Cachia, A. II-344
 Calcagni, A. II-509
 Camara, Oscar I-99
 Camara-Rey, Oscar II-616
 Campbell, Jennifer S.W. I-123
 Cardin, M.A. II-9
 Caroli, A. I-392
 Carsin, B. I-409
 Carson, James I-254
 Carter, Timothy J. I-605, II-1000
 Cascio, Carissa II-765
 Castellanos, Joaquín II-320
 Cattin, Philippe II-336, II-976
 Chakravarty, M. Mallar II-394
 Chan, Carolyn S.K. I-605, II-1000
 Chandrashekara, R. I-425
 Chang, S. I-442
 Chan, Tony F. II-657, II-675
 Chapman, Brian E. I-564
 Charnoz, Arnaud II-155
 Chef d'hotel, Christophe I-662
 Chen, Elvis C.S. I-770
 Chen, Nan-Kuei I-148
 Chen, Yunmei I-278, I-286, II-741
 Cheung, Sue I-91
 Chiang, Ming-Chang II-666
 Chinzei, Kiyoyuki I-9
 Chiu, Wah I-254
 Choti, Mike I-811, II-526
 Christiansen, Claus I-327
 Chung, Adrian I-868, I-910
 Chung, Albert C.S. I-229
 Ciccarelli, O. I-164
 Claridge, E. II-509
 Clatz, Olivier II-295
 Cleary, Kevin II-312, II-992
 Clouchoux, C. II-344
 Cloutier, Guy I-319
 Cocquerez, Jean-Pierre I-262
 Cointepas, Y. I-196
 Colchester, A.C.F. II-749
 Collins, D. Louis I-392, II-163, II-394
 Colliot, O. I-375
 Commowick, Olivier II-927

 Cook, P.A. I-164
 Cooperberg, Peter II-862
 Cootes, T.F. II-733
 Cornu, Philippe II-385
 Corouge, Isabelle I-131
 Cotin, Stéphane I-43, II-534
 Cottone, Lisa I-679
 Cotton, S.D. II-509
 Coudière, Y. II-607
 Coulon, O. II-344
 Coupé, Pierrick I-597
 Cozzone, Patrick II-492
 Cremers, Daniel II-757
 Croome, Kris II-57
 Crouch, Earl R. III II-591
 Crouch, Jessica R. II-591, II-624
 Crozier, Stuart II-813
 Crum, William R. I-99, II-616
 Cuisenaire, Olivier I-417
 Cuvillon, Loïc II-551

 Dai, Xiangtian (Donald) I-811
 Dam, Erik I-327
 Danilouchkine, Mikhail G. II-951
 Darcourt, Jacques II-369
 Darzi, Ara I-360, I-843, II-25,
 II-114, II-139
 Davatzikos, Christos I-1, I-400, II-263,
 II-500, II-870
 Davies, Karen I-91
 Davis, B.C. I-442
 Davis, Brad II-411
 Dawant, Benoit M. II-427
 Dawson, Steve I-43, II-534
 Deguchi, Daisuke II-543
 Deligianni, Fani I-843, I-910, II-139
 Delingette, Hervé II-295
 de Mathelin, Michel II-551
 Descoteaux, Maxime I-9
 D'Haese, Pierre-Francois II-427
 Diamond, Solomon Gilbert II-649
 Dijkers, J.J. I-712
 Dikici, Engin II-692
 DiMaio, Simon P. I-827
 DiMarzio, Charles II-935
 Ding, Mingyue II-17
 Dione, Donald P. II-919
 Dixon, Kat I-335
 Dobrucki, Lawrence W. II-684, II-701,
 II-919

- Dold, Christian I-482
 Dormont, Didier II-385, II-492
 Dornheim, Lars I-335
 Downey, Donal II-17
 Drisis, Stylianos I-859, II-361
 Duchesne, S. I-392
 Duncan, James S. I-761, I-778
 Dupont, Pierre II-41
 Duriez, C. II-534
 Dutta, Sandeep I-738
- Edan, G. I-409
 Ede, C. II-960
 Edwards, Phillip J. I-605, II-968,
 II-1000
 Egan, Gary I-294
 Eichele, Gregor I-254
 El-Baz, Ayman I-34, I-720, II-821
 Eldiasty, Tarek I-720, II-821
 El-Ghar, Mohamed Abou I-720,
 II-821
 ElHelw, Mohamed A. I-868
 Ellis, Randy E. I-75, I-459, I-770,
 I-902, I-917, II-223
 Ema, Sinya II-854
 Evelhoch, Jeffrey I-451
 Exner, Philipp I-819
- Fahmi, Rachid I-531
 Fahrig, Rebecca I-631
 Falk, Robert I-654
 Fan, Yong I-1, II-902
 Farag, Aly A. I-34, I-531, I-654,
 I-720, II-821
 Faucheux, Baptiste II-492
 Feldman, Michael D. I-729
 Feng, Jun II-279
 Fenlon, Michael R. II-968
 Fenster, Aaron II-17
 Ferrarini, Luca II-451
 Fetita, Catalin II-830
 Feuerstein, Marco II-287
 Fevens, Thomas I-670
 Fichtinger, Gabor I-621, I-639,
 II-312
 Figl, M. II-960
 Fillard, Pierre I-115, II-927, II-943
 Finnis, Kirk W. I-835
 Firle, Evelyn A. I-482
 Fischer, Gregory S. II-73
- Fisher, John I-310
 Fletcher, P. Thomas I-131
 Florie, J. I-712
 Florin, Charles I-246
 Folkesson, Jenny I-327
 Forest, Clément II-155
 Forgione, Antonello II-551
 Foroughi, Pezhman I-83
 Foskey, M. I-442
 Foster, Norman L. II-517
 Fox, Nick II-616
 Franceschini, Maria Angela II-649
 Frangi, A.F. II-271
 Freiman, M. II-131
 Friman, Ola I-107
 Fripp, Jurgen II-813
 Frisoni, G.B. I-392
 Fujiwara, Yasunori II-97
 Fukuda, Toshio I-925
- Gabrieli, John D. I-205
 Galanaud, Damien II-492
 Galbraith, Natalie II-57
 Gallardo, Greg I-302
 Galloway, Graham II-813
 Gamper, U. II-336
 Gangi, Afshin II-49
 Gangloff, Jacques II-551
 Gautham, Sardar Mal I-738
 Gee, James C. I-164, I-172,
 II-328, II-725
 Geiger, Bernhard I-662
 Geiser, Edward I-286
 Gerig, Guido I-26, I-131, I-383
 Germann, Jurgen II-394
 Geroldi, C. I-392
 Ghoneim, Mohamed A. II-821
 Giddens, Don P. I-490
 Gilmore, John H. I-131
 Gimel'farb, Georgy I-34, I-720
 Gimpel Smith, Rachel II-765
 Ginestet, Chantal I-556
 Glenn, O. I-548
 Glossop, Neil II-992
 Goddard, Dorothy A. I-434
 Goksel, Orcun I-827
 Golby, Alexandra J. II-295
 Goldberger, Jacob II-790
 Goldberg-Zimring, Daniel I-506
 Goldstein, Rita I-679

- Golestani, N. I-196
 González Ballester, Miguel Ángel
 I-474, II-369, II-492
 Gooding, Mark J. I-434
 Gorges, S. II-214
 Goshtasby, Ardeshir II-717
 Goyal, L. I-442
 Grady, Leo I-442, II-773
 Grau, Vicente I-589
 Greenspan, Hayit II-790
 Grenier, Philippe II-830
 Grimson, W. Eric L. I-188, 310
 Gross, M. II-567
 Gu, Xianfeng II-657
 Guehring, Jens I-498
 Guetter, Christoph II-255
 Guiraudon, Gerard II-171
 Guo, Hongyu II-984
 Guo, Ting I-835
- Hadjidemetriou, Stathis I-761
 Hager, Gregory D. I-621, I-802, I-811
 Haïk, Stéphane II-492
 Hajnal, J.V. II-247
 Haker, Steven I-188, I-490, I-506,
 II-147, II-459
 Hakozaki, Yoshihide II-97
 Harders, M. II-567, II-575
 Hasser, Chris I-811
 Hassouna, M. Sabry I-654
 Hattori, Asaki II-559
 Hawkes, David J. I-605, II-968, II-1000
 Hayashibe, Mitsuhiro II-559
 Hayashi, Kiralee M. II-657
 Hayashi, Yuichiro I-696
 Hellier, Pierre I-409, I-597
 Hendriks, Emile A. II-846
 Heng, Pheng-Ann I-572, II-709
 Hennig, Jürgen I-482
 Hero, Alfred O. III II-271, II-419
 Hettleman, Bruce I-67, 704
 He, Weijun I-745
 He, Y. II-468
 Hill, Derek L.G. I-99, I-425, II-607,
 II-616, II-894
 Hill, Nicholas A. II-171
 Ho, Hon Pong I-278, II-741
 Homolka, P. II-960
 Hornegger, Joachim II-255
 Huang, Feng I-286
- Huang, Heng I-67, I-704
 Huang, Ling I-148
 Huang, Xiaolei I-745
 Huang, Xishi II-171
 Hummel, J. II-960
 Huppert, Theodore J. II-649
 Hushek, Stephen G. I-34
- Ianzini, Fiorenza I-302
 Ide, Ichiro II-854
 Ikeda, Seiichi I-925
 Inder, Terrie I-294
 Iordachita, Iulian I-621
 Iordanova, B. I-548
 Ip, Horace H.S. II-279
 Irie, Keiko I-925
- Jackowski, Marcel II-701, II-717
 Jackson, Alan I-91
 Jain, Ameet I-639
 James, Adam II-25
 Janke, Andrew II-813
 Jayson, Gordon C. I-91
 Jenkinson, Mark I-99
 Jiang, T.Z. II-468
 Jin, Chao I-670
 John, Matthias II-798
 Johnston, Leigh I-294
 Jolesz, Ferenc A. II-295
 Jomier, Julien I-59, I-523, II-806
 Jomier, Matthieu II-765
 Joshi, Sarang I-131, I-442,
 II-411, II-984
 Joskowicz, L. II-131
 Ju, Tao I-254
- Kaipio, Jari P. II-649
 Kakadiaris, Ioannis A. I-254, I-343
 Kalender, Willi A. II-33, II-443
 Kameyama, Fumie II-97
 Kanade, Takeo II-1
 Kao, Chris II-427
 Kapoor, Ankur II-89
 Karacali, Bilge II-500
 Kassim, Ashraf A. I-51, I-270
 Katayama, Youichi II-81
 Kaufhold, John I-851
 Kaur, Manmeen II-205
 Kavoussi, L. II-526
 Kerrien, E. II-214

- Kervrann, Charles I-893
 Khamene, Ali II-65
 Khan, Ali I-238
 Kikinis, Ron I-43, I-310,
 II-295, II-583
 Kim, Jung II-599
 Kim, Seyoung I-352
 Kiss, Gabriel I-859
 Kitasaka, Takayuki I-696, II-543,
 II-854
 Klironomos, D. I-375
 Kobayashi, Etsuko II-81
 Kolehmainen, Ville II-649
 Konrad, Peter E. II-427
 Kon, Ryan I-639
 Köstler, W. I-474
 Kouby, V. El I-196
 Koulibaly, Pierre Malick II-369
 Koyama, T. II-105
 Krasinskas, Alyssa M. I-564
 Krishnan, A. II-781
 Krissian, Karl I-43
 Krzyżak, Adam I-670
 Kumar, Rajesh I-811
 Kurtcuoglu, Vartan I-786

 Laine, Andrew F. II-205
 Langan, David I-851
 Lanovaz, Joel L. I-770, I-917
 Le Bihan, D. I-196
 LeDigarcher, Vincent I-59, I-523,
 II-806
 Lee, Su-Lin I-360
 Lee, Vivian S. II-205
 Lelieveldt, Boudewijn P.F. II-951
 Lenoir, J. II-534
 Leskovjan, Andreaa I-679
 Leven, Joshua I-811
 Levitt, James J. I-310
 Lewis, Emma II-616
 Liang, M. II-468
 Liénard, Jean II-352
 Li, Ming II-89
 Li, Shimin II-684
 Li, Shuo I-670
 Li, Song I-670
 Li, X.B. II-468
 Likar, Boštjan II-231
 Lin, Henry C. I-802
 Lin, Mingan I-876

 Lin, Ning I-778
 Linguraru, Marius George II-492
 Litt, Harold II-902
 Littmann, Arne I-498
 Liu, Huafeng I-278, II-741
 Liu, K.-Y. I-885
 Liu, Tianming I-148
 Liu, Yi-Hwa II-684
 Liu, Yuncai II-122
 Lo, Benny P.L. II-25
 Loeckx, Dirk II-361
 Lohmann G. II-749
 Lomax, A. II-336
 Lorenzen, Peter II-411
 Luboz, Vincent I-43
 Lucidarme, Olivier II-830
 Lui, Lok Ming II-675

 Ma, Burton I-75, I-459
 Mackey, Michael A. I-302
 Madabhushi, Anant I-729
 Maddah, Mahnaz I-188
 Maeder, Philippe I-417
 Maeno, Takashi II-633
 Maes, Frederik II-361
 Makedon, Fillia I-67, 704
 Malandain, Grégoire II-155, II-369,
 II-385, II-927
 Manak, Joseph I-851
 Mangin, J.-F. I-196, II-344
 Manjeshwar, Ravi I-851
 Mansi, T. I-375
 Marchal, Guy I-859, II-361
 Mareci, Thomas H. I-156
 Mareels, Iven I-294
 Marohn, Mike I-811
 Marquering, Henk A. II-846
 Marsden, Paul K. II-968
 Martin-Fernandez, Marcos I-515
 Maurer, Calvin R. Jr. II-543, II-838
 Maurice, Roch L. I-319
 Mazilu, D. II-526
 McCarley, Robert W. I-515
 McGurk, Mark II-968
 McShane, Teresa I-451
 Meetz, Kirsten I-467
 Mekada, Yoshito II-854
 Meller, Sebastian II-443
 Mellor, Matthew I-434
 Mendonça, Paulo R.S. I-688, I-738

- Merriam, John C. II-591
 Metaxas, Dimitris I-745
 Meuli, Reto I-417
 Meunier, Jean I-319
 Mewes, Andrea U.J. I-188
 Meyer, Charles R. II-419
 Mian, Asim I-506
 Miller, James V. I-688, I-738
 Miller, Karol II-583
 Miller, Michael I-238
 Moallem, Merhdad II-57
 Moghari, Mehdi Hedjazi II-197
 Mohamed, Ashraf I-400, II-263
 Mohiaddin, Raad II-402
 Momi, Y. II-105
 Morady, Roya II-57
 Moreau-Villéger, V. II-607
 Mori, Kensaku I-696, II-543, II-854
 Mori, Masaki II-543, II-854
 Morimoto, K. II-105
 Morris, James I-827
 Mortamet, Bénédicte I-383
 Mukherjee, Lopamudra I-876
 Murase, Hiroshi II-854
 Murphy, Todd E. I-802
 Mylonas, George P. I-843, II-139
- Naessens, V. I-375
 Nagai, Noritaka II-633
 Nagel, Markus II-33
 Naghavi, Morteza I-343
 Nain, Delphine II-459
 Nakajima, Y. II-105
 Nakata, Norio II-559
 Natori, Hiroshi II-543, II-854
 Navab, Nassir II-287, II-303
 Negoro, Makoto I-925
 Nelson, Bradley J. I-819
 Netsch, Thomas I-467
 Neumann, P. II-534
 Nicolaou, Marios I-868, II-25
 Nicolau, S.A. I-539
 Niederer, P. II-960
 Niermann, Ken II-427
 Niessen, Wiro J. II-911
 Nijmeh, Ayman D. II-968
 Nio, Y.C. I-712
 Nkenke, Emeka II-443
 Noble, J. Alison I-589
- Noble, Nicholas M.I. II-894
 Nolte, L. I-474
- Oda, Masahiro I-696
 Odell, Edward II-968
 O'Donnell, Lauren I-140
 O'Donnell, Thomas II-692
 Oguz, Ipek II-765
 Ohno-Machado, Lucila I-506
 Okada, K. II-781
 Okamura, Allison M. I-802, II-624
 Olofsen, Hans II-451
 Olsen, Ole Fogh I-327
 O'Malley, Sean M. I-343
 Onogi, S. II-105
 Otake, Yoshito II-559
 Oubel, E. II-271
 Ourselin, Sébastien II-813
 Özarslan, Evren I-156
- Padfield, Dirk R. I-738
 Padhani, Anwar R. II-886
 Pallavaram, Srivatsan II-427
 Papademetris, Xenophon II-701, II-919
 Paragios, Nikos I-246
 Parker, Geoff J.M. I-91
 Park, Hyunjin II-419
 Parrent, Andrew G. I-835
 Patel, Rajni II-57
 Pathak, Chetna II-239
 Patriciu, A. II-526
 Pearlman, Justin I-67, I-704
 Pennec, Xavier I-115, I-539, I-753, II-927, II-943
 Penney, Graeme P. I-605, II-911, II-1000
 Perchant, Aymeric I-753
 Pernuš, Franjo II-231
 Perperidis, Dimitrios II-402
 Perrimon, N. I-885
 Pescatore, J. II-214
 Peters, Terry M. I-835, II-171
 Pettersen, Paola I-327
 Petzold, Ralf II-33
 Pham, Dzung L. II-484
 Picard, L. II-214
 Pichon, Eric I-180
 Pidoux, Bernard II-385
 Pike, G. Bruce I-123
 Pinel, Philippe II-476

- Plewes, D.B. II-9
 Pohl, Kilian M. I-310
 Poline, Jean-Baptiste I-196, II-476
 Pollo, Claudio I-417
 Poulikakos, Dimos I-786
 Poupon, C. I-196
 Prastawa, Marcel I-26, I-383
 Preece, S.J. II-509
 Prêteux, Françoise II-830
 Primak, Serguei II-57
 Prima, S. I-409
 Prince, Jerry L. I-631, II-870

 Qian, Xiaoning II-377
 Qi, Feihu II-179

 Rahn, Norbert II-798
 Rajagopalan, Srinivasan I-794
 Ramamurthi, Krishnakumar I-631
 Ramesh, V. II-781
 Rangarajan, Anand II-984
 Rayman, Reiza II-57
 Razavi, Reza S. I-425, II-607,
 II-894
 Rees, Sandra I-294
 Régis, J. II-344
 Reiber, Johan H.C. II-451, II-846
 Ren, Jing II-171
 Reyes, Mauricio II-369
 Rhode, K.S. I-425, II-607
 Richardson, Theodor II-435
 Rit, Simon I-556
 Rivaz, Hassan I-613
 Rivière, D. I-196, II-344
 Robb, Richard A. I-794
 Roberts, Caleb I-91
 Roberts, M.G. II-733
 Rodriguez-Carranza, C. I-548
 Rohlfing, T. II-838
 Rohling, Robert I-613, I-827, II-862
 Rohr, Karl II-320
 Röper, Barbara II-303
 Rosenman, J. I-442
 Rosen, Mark I-729
 Ross, James C. I-738, I-851
 Rousseau, F. I-548
 Rousson, Mikael II-757
 Roy Cardinal, Marie-Hélène I-319
 Rueckert, Daniel I-99, I-425,
 II-247, II-402

 Ruf, Amit II-790
 Rusinek, Henry II-205

 Sadikot, Abbas F. II-394
 Sakas, Georgios I-482
 Sakuma, Ichiro II-81, II-105
 Salcudean, Septimiu E. I-827
 Saleh, Christian II-385
 Samaras, Dimitris I-679
 Sánchez Castro, F. Javier I-417
 Sanchez-Ortiz, G.I. I-425
 Sanderse, Martijn II-846
 Sarrut, David I-556
 Sauer, Frank II-65, 255
 Savadjiev, Peter I-123
 Schiwietz, Thomas II-773
 Schmidt, Gerd II-33
 Schmid, Volker J. II-886
 Schnabel, Julia II-616
 Schneider, Chad M. II-624
 Schreck, Pascal II-49
 Schulz, Heinrich I-467
 Schutysen, Filip I-368
 Schwartzman, David II-1
 Schweiger, Martin II-616
 Scott, Marietta L.J. II-878
 Seemann, R. II-960
 Sermesant, M. I-425, II-607
 Setser, Randolph II-692
 Shafran, Izhak I-802
 Shah, Mubarak I-270
 Shamir, R. II-131
 Shen, Dinggang I-1, II-179, II-263,
 II-500, II-902
 Shen, Li I-67, I-704
 Shenton, Martha E. I-310, I-515
 Shen, Weijia I-270
 Shepherd, Timothy M. I-156
 Shi, Fanhuai II-122
 Shi, Hongjian I-531, II-821
 Shi, Jianbo I-729
 Shi, Lin I-572
 Shimachi, Shigeyuki II-97
 Shi, Pengcheng I-278, II-188,
 II-741
 Shipley, Jacqueline A. I-434
 Shoham, M. II-131
 Shoshan, Y. II-131
 Siddiqi, Kaleem I-9, I-123
 Sidler, R. I-474

- Sierra, R. II-575
 Simpson, A.L. I-459
 Singh, M. II-781
 Sinusas, Albert J. I-778, II-684,
 II-701, II-919
 Sirohey, Saad A. I-688
 Slomczykowski, Mike I-605, II-1000
 Smith, Erin J. I-902
 Smyth, Padhraic I-352
 Soellinger, Michaela I-786
 Soler, Luc I-539, II-49, II-155
 Solomon, S.B. II-526
 Song, Ting II-205
 Sonka, Milan I-302
 Soulez, Gilles I-319
 Speck, Oliver I-482
 Spooner, John II-427
 Srinivasan, Mandayam A. II-599
 Staib, Lawrence H. II-701, II-919
 Stefanescu, Radu II-927, II-943
 Steinemann, D. II-567
 Stern, Hal I-352
 Stetten, George I-647
 Stiehl, Hans-Siegfried I-498
 Stoianovici, D. II-526
 Stoll, Jeffrey II-41
 Stoyanov, Danail I-843, II-114, II-139
 Strobel, Norbert I-631
 Studholme, C. I-548
 Styles, I.B. II-509
 Styner, Martin A. I-474, II-765
 Subotic, Vanja II-57
 Sudarsky, Sandra I-662
 Südkamp, N. I-474
 Suenaga, Yasuhito I-696, II-543,
 II-854
 Suetens, Paul I-368, I-859, II-361
 Sugano, N. II-105
 Sui, M.Q. II-468
 Summers, Paul I-786
 Sundaram, Tessa A. II-328
 Sundar, Hari II-902
 Suprijanto, S. I-580
 Suzuki, Naoki II-559
 Suzuki, Shigeyuki II-559
 Suzuki, Takashi II-81
 Swennen, Gwen I-368
 Szczerba, Dominik II-641
 Székely, Gábor II-336, II-567, II-575,
 II-641, II-976
 Tada, Mitsunori II-633
 Tagare, Hemant D. II-377
 Tahmasebpour, Hamid II-862
 Tajine, Mohamed II-155
 Takabatake, Hirotsugu II-543, II-854
 Takahashi, Ikuo I-925
 Talos, Ion-Florin I-506, II-295
 Tamura, Y. II-105
 Tang, Thomas S.Y. II-223
 Tannenbaum, Allen R. I-180, I-490,
 II-147, II-459
 Tao, Xiaodong II-870
 Tasdizen, Tolga II-517
 Tavakoli, Mahdi II-57
 Taylor, N. Jane II-886
 Taylor, Russell H. I-811, II-73,
 II-89, II-312
 Teoh, Eam Khwang I-221
 Thacker, Neil A. II-878
 Thaller, Christina I-254
 Thérasse, Éric I-319
 Thiran, Jean-Philippe I-417
 Thirion, Bertrand II-476
 Thompson, A.J. I-164
 Thompson, Paul M. II-657, II-666,
 II-675
 Tian, L.X. II-468
 Tobon-Gomez, C. II-271
 Toews, Matthew II-163
 Tolxdorff, Thomas II-320
 Tomasi, Dardo I-679
 Tomaszewski, John E. I-729
 Tomažević, Dejan II-231
 Tönnies, Klaus D. I-335
 Toomre, Derek I-761
 Toriwaki, Jun-ichiro I-696
 Touhami, Wala I-262
 Troussset, Y. II-214
 Truyen, R. I-712
 Tsai, Chia-Ling II-935
 Tsang, Eric C.C. I-572
 Tublin, Mitchell E. I-564
 Turken, And U. I-205
 Turner, Jessica I-352
 Turner, Wesley D. I-688
 Ungar, Lida I-515
 van Assen, Hans C. II-951
 van Buchem, Mark A. II-451

- Van Cleynenbreugel, Johan I-859
 van de Kraats, Everine B. II-911
 van der Lugt, Aad II-846
 Van Horn, Mark II-239
 van Reiber, Johan H.C. II-951
 van Vliet, L.J. I-712
 van Walsum, Theo II-911
 van Wijk, C. I-712
 van Wynsberghe, Aimee II-57
 Vandermeulen, Dirk II-361
 Vavuranakis, Manolis I-343
 Vemuri, Baba C. I-17, I-156
 Venema, H.W. I-712
 Ventikos, Yiannis I-786
 Vercauteren, Tom I-753
 Vigneron, D. I-548
 Villard, Caroline II-49
 Villemure, Jean-Guy I-417
 Vivert, Christelle I-59
 Vogel, M.W. I-580
 Volkow, Nora I-679
 Vollmers, Karl I-819
 von Cramon, D.Y. II-749
 von Siebenthal, M. II-336
 Vos, F.M. I-580, I-712
 Vossepoel, A.M. I-580
 Vrooman, H.A. I-580
- Wagenknecht, Gudrun II-320
 Wainer, Josh II-624
 Wang, David I-647
 Wang, Defeng I-572
 Wang, Fei I-17
 Wang, J.X. II-9
 Wang, Song II-435
 Wang, Yalin II-657, II-666,
 II-675
 Wang, Y.F. II-468
 Warfield, Simon K. I-188, I-506,
 II-295, II-583, II-862
 Warger, William II II-935
 Warren, Joe I-254
 Watson, Yvonne I-91
 Weeks, Susan II-239
 Wein, Wolfgang II-303
 Wei, Zhouping II-17
 Wells, William M. III I-310, I-506
 Westenbergh, Jos J.M. II-951
 Westermann, Rüdiger II-773
- Westin, Carl-Fredrik I-43, I-107,
 I-140, I-180
 Whitaker, Ross T. II-517
 Whitcher, Brandon J. II-886
 White, Richard D. II-692
 Whitfield-Gabrieli, Susan L. I-205
 Wildhirt, Stephen M. II-287
 Williams, Jim I-246
 Wilson, David I-286, I-851
 Wittek, Adam II-583
 Wong, Stephen TC I-148, I-885
 Wong, Tien-Tsin I-572
 Wu, Bing I-647
 Wu, Guorong II-179
 Wu, Huadong I-564
 Wu, Jue I-229
 Wu, Xunlei I-43
- Xia, Yan I-205
 Xie, Jun II-709
 Xu, Chenyang II-255
 Xu, Jinhui I-876
 Xu, Sheng II-312
 Xue, Zhong II-500, II-902
- Yan, Ping I-778
 Yan, Pingkun I-51, 270
 Yang, Fuxing I-302
 Yang, Guang-Zhong I-360, I-843,
 I-868, II-25, II-114, II-139, II-886
 Yang, X. II-960
 Yang, Yan I-490
 Yau, Shing-Tung II-657
 Yelnik, Jérôme II-385
 Yesin, K. Berk I-819
 Yeung, Daniel S. I-572
 Yeung, Sai Kit II-188
 Ying, Xiaoyou I-745
 Yonenobu, S. II-105
 Yörük, Erdem I-213
 Young, Geoffrey I-148
 Yuh, David D. I-802
 Yuksel, Seniha E. II-821
 Yushkevich, Paul A. I-164, I-172,
 II-725
- Zátonyi, J. II-575
 Zaitsev, Maxim I-482
 Zang, Y.F. II-468
 Zehavi, E. II-131

Zeng, Donglin I-383

Zhang, Gary I-811

Zhang, Hui I-164, I-172,
II-725, II-992

Zhang, Jing II-122

Zhang, Lei I-679

Zhang, Rong I-67, I-704

Zhao, Zheen I-221

Zhao, Zijian II-122

Zhong, Guang I-910

Zhong, Hua II-1

Zhou, Xiaobo I-885

Zhou, Yu I-639

Zhu, C.Z. II-468

Zhu, Lei I-490, II-147

Zou, Kelly H. I-506