

Tianzi Jiang
Nassir Navab
Josien P.W. Pluim
Max A. Viergever (Eds.)

LNCS 6361

Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010

13th International Conference
Beijing, China, September 2010
Proceedings, Part I

1
Part I



MICCAI

MyCopy powered by  SpringerLink

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

Alfred Kobsa

University of California, Irvine, CA, 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

TU Dortmund University, Germany

Madhu Sudan

Microsoft Research, Cambridge, MA, USA

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max Planck Institute for Informatics, Saarbruecken, Germany

Tianzi Jiang Nassir Navab
Josien P.W. Pluim Max A. Viergever (Eds.)

Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010

13th International Conference
Beijing, China, September 20-24, 2010
Proceedings, Part I

Volume Editors

Tianzi Jiang

The Chinese Academy of Sciences, Institute of Automation

Beijing 100080, P. R. China

E-mail: jiangtz@nlpr.ia.ac.cn

Nassir Navab

Technische Universität München, Institut für Informatik I16

Boltzmannstr. 3, 85748 Garching, Germany

E-mail: navab@cs.tum.edu

Josien P.W. Pluim

Max A. Viergever

University Medical Center Utrecht, QS.459

Heidelberglaan 100, 3584 CX Utrecht, The Netherlands

E-mail: j.pluim@umcutrecht.nl, max@isi.uu.nl

Library of Congress Control Number: 2010933822

CR Subject Classification (1998): I.4, I.5, I.2.10, I.3.5, J.3, I.6

LNCS Sublibrary: SL 6 – Image Processing, Computer Vision, Pattern Recognition, and Graphics

ISSN 0302-9743

ISBN-10 3-642-15704-1 Springer Berlin Heidelberg New York

ISBN-13 978-3-642-15704-2 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.com

© Springer-Verlag Berlin Heidelberg 2010

Printed in Germany

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

Printed on acid-free paper 06/3180

Preface

The 13th International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI 2010, was held in Beijing, China from 20-24 September, 2010. The venue was the China National Convention Center (CNCC), China's largest and newest conference center with excellent facilities and a prime location in the heart of the Olympic Green, adjacent to characteristic constructions like the Bird's Nest (National Stadium) and the Water Cube (National Aquatics Center).

MICCAI is the foremost international scientific event in the field of medical image computing and computer-assisted interventions. The annual conference has a high scientific standard by virtue of the threshold for acceptance, and accordingly MICCAI has built up a track record of attracting leading scientists, engineers and clinicians from a wide range of technical and biomedical disciplines.

This year, we received 786 submissions, well in line with the previous two conferences in New York and London. Three program chairs and a program committee of 31 scientists, all with a recognized standing in the field of the conference, were responsible for the selection of the papers. The review process was set up such that each paper was considered by the three program chairs, two program committee members, and a minimum of three external reviewers. The review process was double-blind, so the reviewers did not know the identity of the authors of the submission.

After a careful evaluation procedure, in which all controversial and gray area papers were discussed individually, we arrived at a total of 251 accepted papers for MICCAI 2010, of which 45 were selected for podium presentation and 206 for poster presentation. The acceptance percentage (32%) was in keeping with that of previous MICCAI conferences. All 251 papers are included in the three MICCAI 2010 LNCS volumes.

We are greatly indebted to the reviewers and to the members of the program committee for their invaluable efforts in critically assessing and evaluating the submissions in a very short time frame.

The annual MICCAI event has, in addition to its main conference, a rising number of satellite tutorials and workshops, organized on the day before and the day after the main conference. This year's call for submission for tutorials and workshops led to a record number of proposals, of which a significant fraction had to be rejected because of space and time limitations. The final program hosted eight tutorials, which together gave a comprehensive overview of many areas of the field, and provided rich educational material especially aimed at PhD students and postdoctoral researchers.

The 15 workshops gave - mostly younger - researchers the opportunity to present their work, often in an early stage of their investigations, so that they could obtain useful feedback from more experienced scientists in the field. The

workshop subjects highlighted topics that were not all fully covered in the main conference, and thus added to the diversity of the MICCAI program. In particular, several workshops offered so-called challenges in which researchers were in competition to best segment or register a set of clinical images with ground truth provided by medical experts. We are grateful to the tutorial and workshop committees, in particular to the chairs Dinggang Shen and Bram van Ginneken, for making these satellite events a success.

Highlights of the conference were the two keynote lectures. Professor Alan C. Evans of the McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Canada described recent activity in brain network modeling with an emphasis on anatomical correlation analysis in his presentation “Network Analysis of Cortical Anatomy”. Professor Guang-Zhong Yang of the Royal Society/Wolfson Medical Image Computing Laboratory, Imperial College, London, UK outlined key clinical challenges and research opportunities in developing minimally invasive surgery systems in his presentation “Snake and Lobster - A Feast for MICCAI?”.

MICCAI 2010 would not have been feasible without the efforts of many people behind the scenes. We are particularly indebted to the local organizing committee in Beijing, consisting of Nianming Zuo, Yong Liu, Ming Song, Bing Liu, Bizhen Hong, Shaomei Wang, and Gangqin Zhang, all of the Institute of Automation of the Chinese Academy of Sciences, for their excellent work before and during the conference, and to Jacqueline Wermers for her outstanding assistance with the editorial work in compiling the three Springer LNCS books that contain the proceedings of this conference.

We are obliged to the Board of the MICCAI Society for the opportunity to organize this prestigious conference, and to many of the Society Board and Staff members for valuable and continuous advice and support through all phases of the preparation.

A special word of thanks goes to our sponsors, who generously provided financial support of the conference as a whole, or of specific activities. This greatly helped us with the overall organization of the meeting, as well as allowed us to award prizes for best papers in various categories and travel stipends to an appreciable number of student participants.

It was our great pleasure to welcome the attendees to Beijing for this exciting MICCAI 2010 conference and its satellite tutorials and workshops. The 14th International Conference on Medical Image Computing and Computer-Assisted Intervention will be held in Toronto, Canada, from 15-21 September 2011. We look forward to seeing you there.

September 2010

Tianzi Jiang
Nassir Navab
Josien Pluim
Max Viergever

Organization

General Chair and Co-chairs

Tianzi Jiang	Institute of Automation, CAS, China
Alan Colchester	University of Kent, UK
James Duncan	Yale University, USA

Program Chair and Co-chairs

Max Viergever	Utrecht University and UMC Utrecht, The Netherlands
Nassir Navab	Technische Universität München, Germany
Josien Pluim	University Medical Center Utrecht, The Netherlands

Workshop Chair and Co-chairs

Bram van Ginneken	Radboud University Nijmegen, The Netherlands
Yong Fan	Institute of Automation, CAS, China
Polina Golland	Massachusetts Institute of Technology, USA
Tim Salcudean	University of British Columbia, Canada

Tutorial Chair and Co-chairs

Dinggang Shen	University of North Carolina, USA
Alejandro Frangi	Universitat Pompeu Fabra, Spain
Gábor Székely	ETH Zürich, Switzerland

MICCAI Society, Board of Directors

Nicholas Ayache	INRIA Sophia Antipolis, France
Kevin Cleary	Georgetown University, USA
James Duncan (President)	Yale University, USA
Gabor Fichtinger	Queen's University, Canada
Polina Golland	Massachusetts Institute of Technology, USA
Tianzi Jiang	Institute of Automation, CAS, China
Nassir Navab	Technische Universität München, Germany
Alison Noble	University of Oxford, UK
Sébastien Ourselin	University College London, UK
Ichiro Sakuma	University of Tokyo, Japan
Sandy Wells	Harvard Medical School, USA
Guang-Zhong Yang	Imperial College London, UK

Program Committee

Christian Barillot	IRISA Rennes, France
Albert Chung	Hong Kong UST, China
Gabor Fichtinger	Queen's University, Canada
Alejandro Frangi	Universitat Pompeu Fabra, Spain
Jim Gee	University of Pennsylvania, USA
Bram van Ginneken	Radboud University Nijmegen, The Netherlands
Polina Golland	Massachusetts Institute of Technology, USA
David Hawkes	University College London, UK
Xiaoping Hu	Emory University, USA
Hongen Liao	University of Tokyo, Japan
Huafeng Liu	Zhejiang University, China
Cristian Lorenz	Philips Research Lab Hamburg, Germany
Frederik Maes	University of Leuven, Belgium
Anne Martel	University of Toronto, Canada
Kensaku Mori	Nagoya University, Japan
Mads Nielsen	University of Copenhagen, Denmark
Poul Nielsen	University of Auckland, New Zealand
Wiro Niessen	Erasmus MC Rotterdam, The Netherlands
Xiaochuan Pan	University of Chicago, USA
Franjo Pernuš	University of Ljubljana, Slovenia
Terry Peters	Robarts Research Institute, Canada
Daniel Rueckert	Imperial College London, UK
Tim Salcudean	University of British Columbia, Canada
Yoshinobu Sato	Osaka University, Japan
Dinggang Shen	University of North Carolina, USA
Pengcheng Shi	Rochester Institute of Technology, USA
Gábor Székely	ETH Zürich, Switzerland
Jocelyne Troccaz	TIMC-IMAG, Grenoble, France
Simon Warfield	Harvard University, USA
Carl-Fredrik Westin	Harvard University, USA
Guang-Zhong Yang	Imperial College London, UK

Local Organizing Committee

Nianming Zuo	Institute of Automation,
Yong Liu	Chinese Academy of Sciences, China
Ming Song	
Bing Liu	
Bizhen Hong	
Shaomei Wang	
Gangqin Zhang	
Jacqueline Wermers	UMC Utrecht, The Netherlands

Reviewers

Abolmaesumi, Purang
Abugharbieh, Rafeef
Acar, Burak
Aja-Fernández, Santiago
Akselrod-Ballin, Ayelet
Alexander, Andrew
Alexander, Daniel
Aljabar, Paul
Alomari, Raja
Alvino, Christopher
An, Jungha
Angelini, Elsa
Anor, Tomer
Arbel, Tal
Arridge, Simon
Ashburner, John
Astley, Sue
Atkinson, David
Audette, Michel
Avants, Brian
Awate, Suyash
Babalola, Kolawole
Bach Cuadra, Meritxell
Baillet, Sylvain
Banks, Scott
Barratt, Dean
Batchelor, Philip
Baumann, Michael
Bazin, Pierre-Louis
Beckmann, Christian
Beg, Mirza Faisal
Beichel, Reinhard
Berger, Marie-Odile
Bergtholdt, Martin
Berman, Jeffrey
Betke, Margrit
Bhalerao, Abhir
Bhotika, Rahul
Bian, Junguo
Birkfellner, Wolfgang
Birn, Rasmus
Bischof, Horst
Boctor, Emad
Boisvert, Jonathan
Bosch, Johan
Bouix, Sylvain
Boukerroui, Djamel
Bourgeat, Pierrick
Brady, Mike
Bricault, Ivan
Brun, Caroline
Buelow, Thomas
Bullitt, Elizabeth
Burschka, Darius
Butakoff, Constantine
Cahill, Nathan
Cai, Yiyu
Camara, Oscar
Cardenes, Ruben
Cates, Joshua
Cattin, Philippe
Chakravarty, Mallar
Chen, Elvis
Chen, Sheng
Chen, Wei
Chen, Yunmei
Chen, Zhiqiang
Cheriet, Farida
Chinzei, Kiyoyuki
Chou, Yiyu
Christensen, Gary
Chung, Moo
Cinquin, Philippe
Ciuciu, Philippe
Claridge, Ela
Clarysse, Patrick
Cleary, Kevin
Clerc, Maureen
Colchester, Alan
Collins, Louis
Colliot, Olivier
Comaniciu, Dorin
Commowick, Olivier
Cook, Philip
Cootes, Tim
Cotin, Stéphane
Coulon, Olivier
Coupé, Pierrick

Craddock, Cameron
 Crozier, Stuart
 Crum, William
 Darkner, Sune
 Dauguet, Julien
 Dawant, Benoit
 De Bruijne, Marleen
 De Buck, Stijn
 De Craene, Mathieu
 Deguchi, Daisuke
 Dehghan, Ehsan
 Deligianni, Fani
 Demirci, Stefanie
 Deriche, Rachid
 Descoteaux, Maxime
 Desphande, Gopikrishna
 Desvignes, Michel
 Dey, Joyoni
 Dijkstra, Jouke
 DiMaio, Simon
 Doignon, Christophe
 Douiri, Abdel
 Drangova, Maria
 Du, Yiping
 Duan, Qi
 Duchesne, Simon
 Duncan, James
 Dupont, Pierre
 Ebrahimi, Mehran
 Ecabert, Olivier
 Eggers, Georg
 Ehrhardt, Jan
 El-Baz, Ayman
 Ellis, Randy
 Enescu, Monica
 Fabry, Thomas
 Fahrig, Rebecca
 Fan, Yong
 Farag, Aly
 Fenster, Aaron
 Feragen, Aasa
 Ferrari, Ricardo
 Feuerstein, Marco
 Figl, Michael
 Fillard, Pierre

Fischer, Bernd
 Fitzpatrick, Michael
 Fleig, Oliver
 Florack, Luc
 Fouard, Celine
 Freysinger, Wolfgang
 Fuernstahl, Philipp
 Funka-Lea, Gareth
 Gan, Rui
 Ganz, Melanie
 Gao, Fei
 Gee, Andrew
 Gerig, Guido
 Gessat, Michael
 Gholipour, Ali
 Gibaud, Bernard
 Gladilin, Evgeny
 Glocker, Ben
 Goksel, Orcun
 Gonzalez Ballester, Miguel Angel
 Gooding, Mark
 Goodlett, Casey
 Gooya, Ali
 Gorbunova, Vladlena
 Grady, Leo
 Graham, Jim
 Grau, Vicente
 Groher, Martin
 Gu, Lixu
 Guehring, Jens
 Guetter, Christoph
 Haake, Anne
 Hager, Gregory
 Hahn, Horst
 Hamarneh, Ghassan
 Han, Xiao
 Hanson, Dennis
 Harders, Matthias
 Hastreiter, Peter
 Hata, Nobuhiko
 Haynor, David
 He, Yong
 Heimann, Tobias
 Hellier, Pierre
 Heng, Pheng Ann

Hermosillo, Gerardo
Higgins, William
Hipwell, John
Ho, Hon Pong
Hoffmann, Kenneth
Hogeweg, Laurens
Holmes, David
Holz, Dirk
Hoogendoorn, Corné
Hornegger, Joachim
Howe, Robert
Hu, Mingxing
Hu, Zhenghui
Huang, Heng
Huang, Qi-xing
Huang, Xiaolei
Huo, Xiaoming
Hyde, Damon
Ingalhalikar, Madhura
Işgum, Ivana
Jain, Ameet
Janke, Andrew
Jannin, Pierre
Jin, Mingwu
Jomier, Julien
Joshi, Anand
Joshi, Sarang
Kabus, Sven
Kadah, Yasser
Kadir, Timor
Kadoury, Samuel
Kamen, Ali
Kang, Dong-Goo
Karemore, Gopal
Karssemeijer, Nico
Kaus, Michael
Kazanzides, Peter
Keeve, Erwin
Kerrien, Erwan
Kervrann, Charles
Kikinis, Ron
Kim, Boklye
Kindlmann, Gordon
King, Andrew
Kirchberg, Klaus
Kitasaka, Takayuki
Klein, Arno
Klein, Stefan
Klinder, Tobias
Kontos, Despina
Krissian, Karl
Kruggel, Frithjof
Kutter, Oliver
Kybic, Jan
Lai, Shang-Hong
Laine, Andrew
Landman, Bennett
Langs, Georg
Larrabide, Ignacio
Larsen, Rasmus
Lassen, Bianca
Law, Max
Lazar, Mariana
Lee, Junghoon
Leemans, Alexander
Lei, Hao
Lekadir, Karim
Lelieveldt, Boudewijn
Leow, Alex
Lepore, Natasha
Lerch, Jason
Lesage, David
Li, Chunming
Li, Ming
Li, Quanzheng
Li, Shuo
Liang, Jianming
Liao, Rui
Liao, Shu
Likar, Boštjan
Lin, Xiang
Lindseth, Frank
Linguraru, Marius George
Linte, Cristian
Litt, Harold
Liu, Alan
Liu, Tianming
Liu, Yong
Lo, Pechin
Loeckx, Dirk

Loew, Murray
 Lu, Le
 Luan, Kuan
 Luboz, Vincent
 Luo, Yishan
 Ma, Burton
 Madabhushi, Anant
 Maeder, Anthony
 Magee, Derek
 Maier-Hein, Lena
 Mainprize, James
 Malandain, Gregoire
 Manduca, Armando
 Mangin, Jean-François
 Mao, Hongda
 Mao, Hui
 Markelj, Primož
 Martí, Robert
 Martin-Fernandez, Marcos
 Masamune, Ken
 Masutani, Yoshitaka
 Mazza, Edoardo
 McClelland, Jamie
 McCulloch, Andrew
 McGregor, Robert
 Metaxas, Dimitris
 Metz, Coert
 Meyer, Chuck
 Miller, James
 Milles, Julien
 Mohamed, Ashraf
 Moireau, Philippe
 Mollemans, Wouter
 Mungwe, Stanley
 Murgasova, Maria
 Murphy, Keelin
 Mylonas, George
 Naish, Michael
 Nakamoto, Masahiko
 Nash, Martyn
 Nedjati-Gilani, Shahrum
 Nichols, Thomas
 Nicolau, Stephane
 Niemeijer, Meindert
 Niethammer, Marc

Nimura, Yukitaka
 Noble, Alison
 Noël, Peter
 Nolte, Lutz
 Noonan, David
 Oda, Masahiro
 O'Donnell, Lauren
 O'Donnell, Thomas
 Ogier, Arnaud
 Oguz, Ipek
 Olabariaga, Silvia
 Olmos, Salvador
 Olszewski, Mark
 Orkisz, Maciej
 Otake, Yoshito
 Ourselin, Sébastien
 Ozarslan, Evren
 Pang, Wai-Man
 Pantazis, Dimitrios
 Papadopoulo, Théo
 Paragios, Nikos
 Pasternak, Ofer
 Patriciu, Alexandru
 Pavani, Sri Kaushik
 Payan, Yohan
 Peitgen, Heinz-Otto
 Penneç, Xavier
 Penney, Graeme
 Petersen, Kersten
 Petr, Jan
 Peyrat, Jean-Marc
 Pham, Dzung
 Pichon, Eric
 Pike, Bruce
 Pitiot, Alain
 Pizarro, Luis
 Pohl, Kilian Maria
 Poignet, Philippe
 Prager, Richard
 Prastawa, Marcel
 Prause, Guido
 Prima, Sylvain
 Prince, Jerry
 Promayon, Emmanuel
 Qi, Jinyi

Qian, Xiaoning
 Radeva, Petia
 Rajagopal, Vijayaraghavan
 Rajpoot, Nasir
 Rangarajan, Anand
 Rasche, Volker
 Reichl, Tobias
 Reinhardt, Joseph
 Rexilius, Jan
 Reyes, Mauricio
 Rhode, Kawal
 Ribbens, Annemie
 Ridgway, Gerard
 Rittscher, Jens
 Rivaz, Hassan
 Riviere, Cameron
 Robb, Richard
 Robinson, Emma
 Rohlfing, Torsten
 Rohling, Robert
 Rohr, Karl
 Rougon, Nicolas
 Rousseau, François
 Russakoff, Daniel
 Sabuncu, Mert Rory
 Sachse, Frank
 Sakuma, Ichiro
 Salvado, Olivier
 Samani, Abbas
 Sanchez, Clara
 Savadjiev, Peter
 Schaap, Michiel
 Scherrer, Benoit
 Schnabel, Julia
 Schweikard, Achim
 Sebastian, Rafa
 Sermesant, Maxime
 Shams, Ramtin
 Shechter, Guy
 Shi, Yonggang
 Shi, Yundi
 Shimizu, Akinobu
 Siddiqi, Kaleem
 Sidky, Emil
 Siewerdsen, Jeffrey
 Simaan, Nabil
 Skrinjar, Oskar
 Slagmolen, Pieter
 Sled, John
 Smal, Ihor
 Smeets, Dirk
 Smelyanskiy, Mikhail
 So, Wai King
 Sommer, Stefan
 Song, Xubo
 Sonka, Milan
 Sørensen, Lauge
 Spillmann, Jonas
 Sporring, Jon
 Staal, Joes
 Staib, Lawrence
 Staring, Marius
 Stewart, James
 Stoyanov, Danail
 Studholme, Colin
 Styner, Martin
 Suarez, Ralph
 Subramanian, Navneeth
 Sukno, Federico
 Summers, Ronald
 Suzuki, Kenji
 Szczerba, Dominik
 Szilagyi, Laszlo
 Tanner, Christine
 Tao, Xiaodong
 Tasdizen, Tolga
 Taylor, Chris
 Taylor, Russell
 Taylor, Zeike
 Tek, Huseyin
 Ter Haar Romeny, Bart
 Thévenaz, Philippe
 Thiran, Jean-Philippe
 Thiriet, Marc
 Thirion, Bertrand
 Todd Pokropek, Andrew
 Toews, Matthew
 Tomaževič, Dejan
 Tosun, Duygu
 Tristán-Vega, Antonio

Tsechpenakis, Gavriil
Tustison, Nicholas
Tutar, Ismail
Twining, Carole
Unal, Gozde
Vaillant, Regis
Van Leemput, Koen
Van Rikxoort, Eva
Van Stralen, Marijn
Van Walsum, Theo
Vannier, Michael
Vemuri, Baba
Venkataraman, Archana
Vercauteren, Tom
Verma, Ragini
Vidal, Pierre Paul
Vik, Torbjörn
Vilanova, Anna
Villard, Pierre-Frederic
Von Berg, Jens
Voros, Sandrine
Vos, Frans
Vosburgh, Kirby
Vrooman, Henri
Vrtovec, Tomaz
Wachinger, Christian
Wang, Defeng
Wang, Fei
Wang, Junchen
Wang, Linwei
Wang, Yalin
Wang, Yongmei Michelle
Ward, Aaron
Watton, Paul
Weber, Stefan
Weese, Jürgen
Wein, Wolfgang
Weisenfeld, Neil
Wells, William
West, Jay
Whitaker, Ross

Wiemker, Rafael
Wimmer, Andreas
Wolf, Ivo
Wolz, Robin
Wong, Ken
Woolrich, Mark
Wu, Ed
Wu, Guorong
Wu, John Jue
Xia, Dan
Xu, Jianwu
Xu, Qianyi
Xue, Zhong
Yan, Pingkun
Yang, Hua
Yap, Pew-Thian
Yeo, Thomas
Yezzi, Anthony
Yoo, Terry
Yoshida, Hiro
Young, Alistair
Yu, Weichuan
Yushkevich, Paul
Zang, Yufeng
Zhang, Heye
Zhang, Hui
Zhang, Yong
Zhao, Fuqiang
Zheng, Bo
Zheng, Guoyan
Zheng, Yefeng
Zhou, Luping
Zhou, Kevin
Zhou, Xiang
Zhou, Yu
Zhu, Hongtu
Zhu, Yun
Zikic, Darko
Zöllei, Lilla
Zuo, Nianming
Zwiggelaar, Reyer

Awards Presented at the 12th International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI 2009, London

MICCAI Society Enduring Impact Award

The Enduring Impact Award is the highest award of the Medical Image Computing and Computer-Assisted Intervention Society. It is a career award for continued excellence in the MICCAI research field. The 2009 Enduring Impact Award was presented to **Ron Kikinis**, Harvard Medical School, USA.

MICCAI Society Fellowships

MICCAI Fellowships are bestowed annually on a small number of senior members of the Society in recognition of substantial scientific contributions to the MICCAI research field and service to the MICCAI community. The first fellowships were presented in 2009, to

Nicholas Ayache (INRIA Sophia-Antipolis, France)

Alan Colchester (University of Kent, UK)

Takeyoshi Dohi (University of Tokyo, Japan)

Guido Gerig (University of Utah, USA)

David Hawkes (University College London, UK)

Karl Heinz Höhne (University of Hamburg, Germany)

Ron Kikinis (Harvard Medical School, USA)

Terry Peters (Robarts Research Institute, Canada)

Richard Robb (Mayo Clinic, USA)

Chris Taylor (University of Manchester, UK)

Russ Taylor (Johns Hopkins University, USA)

Max Viergever (University Medical Center Utrecht, The Netherlands).

MedIA-MICCAI Prize

The 2009 MedIA-MICCAI Prize for the best paper in the special MICCAI issue of Medical Image Analysis, sponsored by Elsevier, was awarded to

Vicky Wang (University of Auckland, New Zealand)

for the article “Modelling passive diastolic mechanics with quantitative MRI of cardiac structure and function”, authored by Vicky Y. Wang, Hoi I. Lam, Daniel B. Ennis, Brett R. Cowan, Alistair A. Young, and Martyn P. Nash.

Best Paper in Navigation

The prize for the best paper in the MICCAI 2009 conference in the area of navigation, sponsored by Medtronic, was awarded to

Wolfgang Wein (Siemens Corporate Research, Princeton, USA)

for the article: “Towards guidance of electrophysiological procedures with real-time 3D intracardiac echocardiography fusion to C-arm CT”, authored by Wolfgang Wein, Estelle Camus, Matthias John, Mamadou Diallo, Christophe Duong, Amin Al-Ahmad, Rebecca Fahrig, Ali Khamene, and Chenyang Xu.

Best Paper in Computer-Assisted Intervention Systems and Medical Robotics

The prize for the best paper in the MICCAI 2009 conference in the area of computer-assisted intervention systems and medical robotics, sponsored by Intuitive Surgical, was awarded to

Marcin Balicki (Johns Hopkins University, USA)

for the article “Single fiber optical coherence tomography microsurgical instruments for computer and robot-assisted retinal surgery”, authored by Marcin Balicki, Jae-Ho Han, Iulian Iordachita, Peter Gehlbach, James Handa, Russell Taylor, Jin Kang.

MICCAI Young Scientist Awards

The Young Scientist Awards are stimulation prizes awarded to the best first authors of MICCAI contributions in distinct subject areas. The nominees had to be a full-time student at a recognized university at - or within the two years before - the time of submission. The 2009 MICCAI Young Scientist Awards were presented to

Tammy Riklin Raviv (MIT, USA), for the article “Joint segmentation of image ensembles via latent atlases”

Christopher Rohkohl (Friedrich-Alexander University, Germany), for the article “Interventional 4-D motion estimation and reconstruction of cardiac vasculature without motion”

Peter Savadjiev (Harvard Medical School, USA), for the article “Local white matter geometry indices from diffusion tensor gradients”

Lejing Wang (TU Munich, Germany), for the article “Parallax-free long bone X-ray image stitching”

Yiyi Wei (INRIA Lille, France; LIAMA CASIA, China), for the article “Toward real-time simulation of blood-coil interaction during aneurysm embolization”.

Table of Contents – Part I

Computer-Aided Diagnosis

Automated Macular Pathology Diagnosis in Retinal OCT Images Using Multi-Scale Spatial Pyramid with Local Binary Patterns	1
<i>Yu-Ying Liu, Mei Chen, Hiroshi Ishikawa, Gadi Wollstein, Joel S. Schuman, and James M. Rehg</i>	
Non-invasive Image-Based Approach for Early Detection of Acute Renal Rejection	10
<i>Fahmi Khalifa, Ayman El-Baz, Georgy Gimel'farb, and Mohammed Abu El-Ghar</i>	
Hierarchical Segmentation and Identification of Thoracic Vertebra Using Learning-Based Edge Detection and Coarse-to-Fine Deformable Model	19
<i>Jun Ma, Le Lu, Yiqiang Zhan, Xiang Zhou, Marcos Salganicoff, and Arun Krishnan</i>	
Automatic Detection and Segmentation of Axillary Lymph Nodes	28
<i>Adrian Barbu, Michael Suehling, Xun Xu, David Liu, S. Kevin Zhou, and Dorin Comaniciu</i>	
Image Dissimilarity-Based Quantification of Lung Disease from CT	37
<i>Lauge Sørensen, Marco Loog, Pechin Lo, Haseem Ashraf, Asger Dirksen, Robert P.W. Duin, and Marleen de Bruijne</i>	

Planning and Guidance of Interventions

Vessel Segmentation for Ablation Treatment Planning and Simulation	45
<i>Tuomas Alhonnoro, Mika Pollari, Mikko Lilja, Ronan Flanagan, Bernhard Kainz, Judith Muehl, Ursula Mayrhooser, Horst Portugaller, Philipp Stiegler, and Karlheinz Tscheliessnigg</i>	
Integrated Segmentation and Nonrigid Registration for Application in Prostate Image-Guided Radiotherapy	53
<i>Chao Lu, Sudhakar Chelikani, Zhe Chen, Xenophon Papademetris, Lawrence H. Staib, and James S. Duncan</i>	

Anatomically-Driven Soft-Tissue Simulation Strategy for Cranio-Maxillofacial Surgery Using Facial Muscle Template Model	61
<i>Hyungmin Kim, Philipp Jürgens, Lutz-Peter Nolte, and Mauricio Reyes</i>	
Dynamic Shape Instantiation for Intra-Operative Guidance	69
<i>Su-Lin Lee, Adrian Chung, Mirna Lerotic, Maria A. Hawkins, Diana Tait, and Guang-Zhong Yang</i>	
Dynamic Guidance for Robotic Surgery Using Image-Constrained Biomechanical Models	77
<i>Philip Pratt, Danail Stoyanov, Marco Visentini-Scarzanella, and Guang-Zhong Yang</i>	

Image Segmentation

Actin Filament Segmentation Using Spatiotemporal Active-Surface and Active-Contour Models	86
<i>Hongsheng Li, Tian Shen, Dimitrios Vavylonis, and Xiaolei Huang</i>	
Model-Based Esophagus Segmentation from CT Scans Using a Spatial Probability Map	95
<i>Johannes Feulner, S. Kevin Zhou, Martin Huber, Alexander Cavallaro, Joachim Hornegger, and Dorin Comaniciu</i>	
Automatic Segmentation and Components Classification of Optic Pathway Gliomas in MRI	103
<i>Lior Weizman, Liat Ben-Sira, Leo Joskowicz, Ronit Precel, Shlomi Constantini, and Dafna Ben-Bashat</i>	
Spatial Decision Forests for MS Lesion Segmentation in Multi-Channel MR Images	111
<i>Ezequiel Geremia, Bjoern H. Menze, Olivier Clatz, Ender Konukoglu, Antonio Criminisi, and Nicholas Ayache</i>	
Segmentation Subject to Stitching Constraints: Finding Many Small Structures in a Large Image	119
<i>Elena Bernardis and Stella X. Yu</i>	

Image Reconstruction and Restoration

Measurement Selection in Untracked Freehand 3D Ultrasound	127
<i>Catherine Laporte and Tal Arbel</i>	
Efficient MR Image Reconstruction for Compressed MR Imaging	135
<i>Junzhou Huang, Shaoting Zhang, and Dimitris Metaxas</i>	

4D Computed Tomography Reconstruction from Few-Projection Data via Temporal Non-local Regularization	143
<i>Xun Jia, Yifei Lou, Bin Dong, Zhen Tian, and Steve Jiang</i>	
ECG-Gated Interventional Cardiac Reconstruction for Non-periodic Motion	151
<i>Christopher Rohkohl, Günter Lauritsch, Lisa Biller, and Joachim Hornegger</i>	
Physiological Fusion of Functional and Structural Data for Cardiac Deformation Recovery	159
<i>Ken C.L. Wong, Linwei Wang, Heye Zhang, and Pengcheng Shi</i>	

Functional and Diffusion-Weighted MRI

Adaptive Noise Filtering for Accurate and Precise Diffusion Estimation in Fiber Crossings	167
<i>Matthan W.A. Caan, Ganesh Khedoe, Dirk Poot, Arjan den Dekker, Silvia Olabarrriaga, Kees Grimbergen, Lucas van Vliet, and Frans Vos</i>	
A Sobolev Norm Based Distance Measure for HARDI Clustering: A Feasibility Study on Phantom and Real Data	175
<i>Ellen Brunenberg, Remco Duits, Bart ter Haar Romeny, and Bram Platel</i>	
MicroTrack: An Algorithm for Concurrent Projectome and Microstructure Estimation	183
<i>Anthony J. Sherbondy, Matthew C. Rowe, and Daniel C. Alexander</i>	
Joint Generative Model for fMRI/DWI and Its Application to Population Studies	191
<i>Archana Venkataraman, Yogesh Rathi, Marek Kubicki, Carl-Fredrik Westin, and Polina Golland</i>	
Detection of Brain Functional-Connectivity Difference in Post-stroke Patients Using Group-Level Covariance Modeling	200
<i>Gaël Varoquaux, Flore Baronnet, Andreas Kleinschmidt, Pierre Fillard, and Bertrand Thirion</i>	

Modeling and Simulation

Understanding the Optics to Aid Microscopy Image Segmentation	209
<i>Zhaozheng Yin, Kang Li, Takeo Kanade, and Mei Chen</i>	
Complete Valvular Heart Apparatus Model from 4D Cardiac CT	218
<i>Saša Grbić, Razvan Ionasec, Dime Vitanovski, Ingmar Voigt, Yang Wang, Bogdan Georgescu, Nassir Navab, and Dorin Comaniciu</i>	

Improving Deformable Surface Meshes through Omni-Directional Displacements and MRFs	227
<i>D. Kainmueller, H. Lamecker, H. Seim, S. Zachow, and H.-C. Hege</i>	
Multiplicative Jacobian Energy Decomposition Method for Fast Porous Visco-Hyperelastic Soft Tissue Model	235
<i>Stéphanie Marchesseau, Tobias Heimann, Simon Chatelin, Rémy Willinger, and Hervé Delingette</i>	
Fast Ultrasound Image Simulation Using the Westervelt Equation	243
<i>Athanasios Karamalis, Wolfgang Wein, and Nassir Navab</i>	

Instrument and Patient Localization and Tracking

Accounting for Anisotropic Noise in Fine Registration of Time-of-Flight Range Data with High-Resolution Surface Data	251
<i>L. Maier-Hein, M. Schmidt, A.M. Franz, T.R. dos Santos, A. Seitel, B. Jähne, J.M. Fitzpatrick, and H.P. Meinzer</i>	
Self-encoded Marker for Optical Prospective Head Motion Correction in MRI	259
<i>Christoph Forman, Murat Aksoy, Joachim Hornegger, and Roland Bammer</i>	
Robust 3D Visual Tracking for Robotic-Assisted Cardiac Interventions	267
<i>Rogério Richa, Antônio P.L. Bó, and Philippe Pognet</i>	
Real-Time Stereo Reconstruction in Robotically Assisted Minimally Invasive Surgery	275
<i>Danail Stoyanov, Marco Visentini Scarzanella, Philip Pratt, and Guang-Zhong Yang</i>	
Prostate Brachytherapy Seed Reconstruction Using C-Arm Rotation Measurement and Motion Compensation	283
<i>Ehsan Dehghan, Junghoon Lee, Mehdi Moradi, Xu Wen, Gabor Fichtinger, and Septimiu E. Salcudean</i>	

Quantitative Image Analysis

Reconstructing Geometrically Consistent Tree Structures from Noisy Images	291
<i>Engin Türetken, Christian Blum, Germán González, and Pascal Fua</i>	
Dual-Model Automatic Detection of Nerve-Fibres in Corneal Confocal Microscopy Images	300
<i>M.A. Dabbah, J. Graham, I. Petropoulos, M. Tavakoli, and R.A. Malik</i>	

Characterizing the Regional Structural Difference of the Brain between Tau Transgenic (rTg4510) and Wild-Type Mice Using MRI	308
<i>Zhiyong Xie, Dewen Yang, Diane Stephenson, Daniel Morton, Carol Hicks, Tracy Brown, and Thomas Bocan</i>	
Spatially Regularized SVM for the Detection of Brain Areas Associated with Stroke Outcome	316
<i>Rémi Cuingnet, Charlotte Rosso, Stéphane Lehéricy, Didier Dormont, Habib Benali, Yves Samson, and Olivier Colliot</i>	
Sparse Unbiased Analysis of Anatomical Variance in Longitudinal Imaging	324
<i>Brian Avants, Philip A. Cook, Corey McMillan, Murray Grossman, Nicholas J. Tustison, Yuanjie Zheng, and James C. Gee</i>	

Image Registration

Intra-Patient Supine-Prone Colon Registration in CT Colonography Using Shape Spectrum	332
<i>Zhaoqiang Lai, Jiayi Hu, Chang Liu, Vahid Taimouri, Darshan Pai, Jiong Zhu, Jianrong Xu, and Jing Hua</i>	
A Parameterization of Deformation Fields for Diffeomorphic Image Registration and Its Application to Myocardial Delineation	340
<i>Hua-mei Chen, Aashish Goela, Gregory J. Garvin, and Shuo Li</i>	
Multiple Cortical Surface Correspondence Using Pairwise Shape Similarity	349
<i>Pahal Dalal, Feng Shi, Dinggang Shen, and Song Wang</i>	
Cortical Sulcal Atlas Construction Using a Diffeomorphic Mapping Approach	357
<i>Shantanu H. Joshi, Ryan P. Cabeen, Bo Sun, Anand A. Joshi, Boris Gutman, Alen Zamanyan, Shruthi Chakrapani, Ivo Dinov, Roger P. Woods, and Arthur W. Toga</i>	
Non-rigid Registration with Missing Correspondences in Preoperative and Postresection Brain Images	367
<i>Nicha Chitphakdithai and James S. Duncan</i>	

Computational and Interventional Cardiology

System to Guide Transcatheter Aortic Valve Implantations Based on Interventional C-Arm CT Imaging	375
<i>Matthias John, Rui Liao, Yefeng Zheng, Alois Nöttling, Jan Boese, Uwe Kirschstein, Jörg Kempfert, and Thomas Walther</i>	

Cardiac Anchoring in MRI through Context Modeling	383
<i>Xiaoguang Lu, Bogdan Georgescu, Marie-Pierre Jolly, Jens Guehring, Alistair Young, Brett Cowan, Arne Littmann, and Dorin Comaniciu</i>	
Real-Time Respiratory Motion Correction for Cardiac Electrophysiology Procedures Using Image-Based Coronary Sinus Catheter Tracking	391
<i>YingLiang Ma, Andy P. King, Nicolas Gogin, C. Aldo Rinaldi, Jaswinder Gill, Reza Razavi, and Kawal S. Rhode</i>	
Accurate Segmentation of the Left Ventricle in Computed Tomography Images for Local Wall Thickness Assessment	400
<i>J. Peters, J. Lessick, R. Kneser, I. Wächter, M. Vembar, O. Ecabert, and J. Weese</i>	
Regional Heart Motion Abnormality Detection via Information Measures and Unscented Kalman Filtering	409
<i>Kumaradevan Punithakumar, Ismail Ben Ayed, Ali Islam, Ian G. Ross, and Shuo Li</i>	
<i>In vivo</i> Human 3D Cardiac Fibre Architecture: Reconstruction Using Curvilinear Interpolation of Diffusion Tensor Images	418
<i>Nicolas Toussaint, Maxime Sermesant, Christian T. Stoeck, Sebastian Kozerke, and Philip G. Batchelor</i>	
Physics-Based Modeling of Aortic Wall Motion from ECG-Gated 4D Computed Tomography	426
<i>Guanglei Xiong and Charles A. Taylor</i>	
Whole Heart Segmentation of Cardiac MRI Using Multiple Path Propagation Strategy	435
<i>X. Zhuang, K.K. Leung, K.S. Rhode, R. Razavi, D. Hawkes, and S. Ourselin</i>	
Combined Model-Based Segmentation and Elastic Registration for Accurate Quantification of the Aortic Arch	444
<i>Andreas Biesdorf, Karl Rohr, Hendrik von Tengg-Kobligk, and Stefan Wörz</i>	
Conditional Shape Models for Cardiac Motion Estimation	452
<i>Coert Metz, Nora Baka, Hortense Kirisli, Michiel Schaap, Theo van Walsum, Stefan Klein, Lisan Neeffjes, Nico Mollet, Boudewijn Lelieveldt, Marleen de Bruijne, and Wiro Niessen</i>	

Cross-Modality Assessment and Planning for Pulmonary Trunk Treatment Using CT and MRI Imaging	460
<i>Dime Vitanovski, Alexey Tsymbal, Razvan Ioan Ionasec, Bogdan Georgescu, Martin Huber, Andrew Taylor, Silvia Schievano, Shaohua Kevin Zhou, Joachim Hornegger, and Dorin Comaniciu</i>	
Automatic Cardiac MRI Segmentation Using a Biventricular Deformable Medial Model	468
<i>Hui Sun, Alejandro F. Frangi, Hongzhi Wang, Federico M. Sukno, Catalina Tobon-Gomez, and Paul A. Yushkevich</i>	
Automatic Aorta Segmentation and Valve Landmark Detection in C-Arm CT: Application to Aortic Valve Implantation	476
<i>Yefeng Zheng, Matthias John, Rui Liao, Jan Boese, Uwe Kirschstein, Bogdan Georgescu, S. Kevin Zhou, Jörg Kempfert, Thomas Walther, Gernot Brockmann, and Dorin Comaniciu</i>	
Lesion-Specific Coronary Artery Calcium Quantification for Predicting Cardiac Event with Multiple Instance Support Vector Machines	484
<i>Qingshan Liu, Zhen Qian, Idean Marvasty, Sarah Rinehart, Szilard Voros, and Dimitris N. Metaxas</i>	
Joint Registration and Segmentation of Dynamic Cardiac Perfusion Images Using MRFs	493
<i>Dwarikanath Mahapatra and Ying Sun</i>	
3D Radio Frequency Ultrasound Cardiac Segmentation Using a Linear Predictor	502
<i>Paul C. Pearlman, Hemant D. Tagare, Albert J. Sinusas, and James S. Duncan</i>	
Automated Interventricular Septum Thickness Measurement from B-Mode Echocardiograms	510
<i>Navneeth Subramanian, Dirk Padfield, Sheshadri Thiruvekadam, Anand Narasimhamurthy, and Sigmund Frigstad</i>	
Probabilistic-Driven Oriented Speckle Reducing Anisotropic Diffusion with Application to Cardiac Ultrasonic Images	518
<i>G. Vegas-Sanchez-Ferrero, S. Aja-Fernandez, M. Martin-Fernandez, A.F. Frangi, and C. Palencia</i>	
Patient Specific Models for Planning and Guidance of Minimally Invasive Aortic Valve Implantation	526
<i>I. Waechter, R. Kneser, G. Korosoglou, J. Peters, N.H. Bakker, R.v.d. Boomen, and J. Weese</i>	

Diffusion Tensor MR Imaging and Analysis

A Framework For Using Diffusion Weighted Imaging to Improve Cortical Parcellation	534
<i>Matthew J. Clarkson, Ian B. Malone, Marc Modat, Kelvin K. Leung, Natalie Ryan, Daniel C. Alexander, Nick C. Fox, and Sébastien Ourselin</i>	
Tract-Based Probability Densities of Diffusivity Measures in DT-MRI	542
<i>Çağatay Demiralp and David H. Laidlaw</i>	
Inference of a HARDI Fiber Bundle Atlas Using a Two-Level Clustering Strategy	550
<i>Pamela Guevara, Cyril Poupon, Denis Rivière, Yann Cointepas, Linda Marrakchi, Maxime Descoteaux, Pierre Fillard, Bertrand Thirion, and Jean-François Mangin</i>	
DTI Based Diagnostic Prediction of a Disease via Pattern Classification	558
<i>Madhura Ingahalikar, Stathis Kanterakis, Ruben Gur, Timothy P.L. Roberts, and Ragini Verma</i>	
Probabilistic Anatomical Connectivity Using Completion Fields	566
<i>Parya MomayyezSiahkal and Kaleem Siddiqi</i>	
Reconstruction of Scattered Data in Fetal Diffusion MRI	574
<i>Estanislao Oubel, Meriam Koob, Colin Studholme, Jean-Louis Dietemann, and François Rousseau</i>	
Symmetric Positive-Definite Cartesian Tensor Orientation Distribution Functions (CT-ODF)	582
<i>Yonas T. Weldeselassie, Angelos Barmpoutis, and M. Stella Atkins</i>	
Model-Free and Analytical EAP Reconstruction via Spherical Polar Fourier Diffusion MRI	590
<i>Jian Cheng, Aurobrata Ghosh, Tianzi Jiang, and Rachid Deriche</i>	
Group-Wise Diffeomorphic Diffusion Tensor Image Registration	598
<i>Xiujuan Geng, Hong Gu, Wanyong Shin, Thomas J. Ross, and Yihong Yang</i>	
Fast and Accurate Reconstruction of HARDI Data Using Compressed Sensing	607
<i>Oleg Michailovich and Yogesh Rathi</i>	
General and Efficient Super-Resolution Method for Multi-slice MRI	615
<i>D.H.J. Poot, V. Van Meir, and J. Sijbers</i>	

In-Vivo Estimates of Axonal Characteristics Using Optimized Diffusion MRI Protocols for Single Fibre Orientation	623
<i>Torben Schneider, Claudia A.M. Wheeler-Kingshott, and Daniel C. Alexander</i>	
Diffusion-Based Population Statistics Using Tract Probability Maps	631
<i>Demian Wassermann, Efstathios Kanterakis, Ruben C. Gur, Rachid Deriche, and Ragini Verma</i>	
Axon Diameter Mapping in the Presence of Orientation Dispersion with Diffusion MRI	640
<i>Hui Zhang and Daniel C. Alexander</i>	
Model-Free, Regularized, Fast, and Robust Analytical Orientation Distribution Function Estimation	648
<i>Jian Cheng, Aurobrata Ghosh, Rachid Deriche, and Tianzi Jiang</i>	
Biomarkers for Identifying First-Episode Schizophrenia Patients Using Diffusion Weighted Imaging	657
<i>Yogesh Rathi, James Malcolm, Oleg Michailovich, Jill Goldstein, Larry Seidman, Robert W. McCarley, Carl-Fredrik Westin, and Martha E. Shenton</i>	
A Novel White Matter Fibre Tracking Algorithm Using Probabilistic Tractography and Average Curves	666
<i>Nagulan Ratnarajah, Andrew Simmons, Oleg Davydov, and Ali Hojjatoleslami</i>	
Multi-Diffusion-Tensor Fitting via Spherical Deconvolution: A Unifying Framework	674
<i>Thomas Schultz, Carl-Fredrik Westin, and Gordon Kindlmann</i>	
Statistical Analysis of Tensor Fields	682
<i>Yuchen Xie, Baba C. Vemuri, and Jeffrey Ho</i>	
Multivariate Varying Coefficient Models for DTI Tract Statistics	690
<i>Hongtu Zhu, Martin Styner, Yimei Li, Linglong Kong, Yundi Shi, Weili Lin, Christopher Coe, and John H. Gilmore</i>	
Author Index	699

Table of Contents – Part II

Ultrasound Imaging

Temporal Diffeomorphic Free-Form Deformation for Strain Quantification in 3D-US Images	1
<i>Mathieu De Craene, Gemma Piella, Nicolas Duchateau, Etel Silva, Adelina Doltra, Hang Gao, Jan D’hooge, Oscar Camara, Josep Brugada, Marta Sitges, and Alejandro F. Frangi</i>	
Tracked Ultrasound Elastography (TrUE)	9
<i>Pezhman Foroughi, Hassan Rivaz, Ioana N. Fleming, Gregory D. Hager, and Emad M. Boctor</i>	
Evaluation of Inter-session 3D-TRUS to 3D-TRUS Image Registration for Repeat Prostate Biopsies	17
<i>Vaishali V. Karnik, Aaron Fenster, Jeff Bax, Lori Gardi, Igor Gyacskov, Jacques Montreuil, Cesare Romagnoli, and Aaron D. Ward</i>	
Manifold Learning for Image-Based Breathing Gating with Application to 4D Ultrasound	26
<i>Christian Wachinger, Mehmet Yigitsoy, and Nassir Navab</i>	
Measurement of the Skin-Liver Capsule Distance on Ultrasound RF Data for 1D Transient Elastography	34
<i>Stéphane Audière, Maurice Charbit, Elsa D. Angelini, Jennifer Oudry, and Laurent Sandrin</i>	
Incremental Shape Statistics Learning for Prostate Tracking in TRUS	42
<i>Pingkun Yan and Jochen Kruecker</i>	
Fast and Accurate Ultrasonography for Visceral Fat Measurement	50
<i>You Zhou, Norihiro Koizumi, Naoto Kubota, Takaharu Asano, Kazuhito Yuhashi, Takashi Mochizuki, Takashi Kadowaki, Ichiro Sakuma, and Hongen Liao</i>	
Real-Time Gating of IVUS Sequences Based on Motion Blur Analysis: Method and Quantitative Validation	59
<i>Carlo Gatta, Simone Balocco, Francesco Ciompi, Rayyan Hemetsberger, Oriol Rodriguez Leor, and Petia Radeva</i>	

Registration of a Statistical Shape Model of the Lumbar Spine to 3D Ultrasound Images	68
<i>Siavash Khallaghi, Parvin Mousavi, Ren Hui Gong, Sean Gill, Jonathan Boisvert, Gabor Fichtinger, David Pichora, Dan Borschneck, and Purang Abolmaesumi</i>	
Automatic Prostate Segmentation Using Fused Ultrasound B-Mode and Elastography Images	76
<i>S. Sara Mahdavi, Mehdi Moradi, William J. Morris, and Septimiu E. Salcudean</i>	
Neuroimage Analysis	
ODF Maxima Extraction in Spherical Harmonic Representation via Analytical Search Space Reduction	84
<i>Iman Aganj, Christophe Lenglet, and Guillermo Sapiro</i>	
An Anthropomorphic Polyvinyl Alcohol Triple-Modality Brain Phantom Based on Colin27.....	92
<i>Sean Jy-Shyang Chen, Pierre Hellier, Jean-Yves Gauvrit, Maud Marchal, Xavier Morandi, and D. Louis Collins</i>	
Statistical Analysis of Structural Brain Connectivity	101
<i>Renske de Boer, Michiel Schaap, Fedde van der Lijn, Henri A. Vrooman, Marius de Groot, Meike W. Vernooij, M. Arfan Ikram, Evert F.S. van Velsen, Aad van der Lugt, Monique M.B. Breteler, and Wiro J. Niessen</i>	
Maximum A Posteriori Estimation of Isotropic High-Resolution Volumetric MRI from Orthogonal Thick-Slice Scans	109
<i>Ali Gholipour, Judy A. Estroff, Mustafa Sahin, Sanjay P. Prabhu, and Simon K. Warfield</i>	
Change Detection in Diffusion MRI Using Multivariate Statistical Testing on Tensors	117
<i>Antoine Grigis, Vincent Noblet, Félix Renard, Fabrice Heitz, Jean-Paul Armspach, and Lucien Rumbach</i>	
Increasing Power to Predict Mild Cognitive Impairment Conversion to Alzheimer’s Disease Using Hippocampal Atrophy Rate and Statistical Shape Models	125
<i>Kelvin K. Leung, Kai-Kai Shen, Josephine Barnes, Gerard R. Ridgway, Matthew J. Clarkson, Jurgen Fripp, Olivier Salvado, Fabrice Meriaudeau, Nick C. Fox, Pierrick Bourgeat, and Sébastien Ourselin</i>	

Consistent 4D Cortical Thickness Measurement for Longitudinal Neuroimaging Study	133
<i>Yang Li, Yaping Wang, Zhong Xue, Feng Shi, Weili Lin, Dinggang Shen, and The Alzheimer’s Disease Neuroimaging Initiative</i>	
Fiber-Centered Analysis of Brain Connectivities Using DTI and Resting State fMRI Data	143
<i>Jinglei Lv, Lei Guo, Xintao Hu, Tuo Zhang, Kaiming Li, Degang Zhang, Jianfei Yang, and Tianming Liu</i>	
A Generative Model for Brain Tumor Segmentation in Multi-Modal Images	151
<i>Bjoern H. Menze, Koen Van Leemput, Danial Lashkari, Marc-André Weber, Nicholas Ayache, and Polina Golland</i>	
Spatio-temporal Analysis of Brain MRI Images Using Hidden Markov Models	160
<i>Ying Wang, Susan M. Resnick, and Christos Davatzikos</i>	
Estimating Local Surface Complexity Maps Using Spherical Harmonic Reconstructions	169
<i>Rachel Aine Yotter, Paul M. Thompson, Igor Nenadic, and Christian Gaser</i>	
Brain Morphometry by Probabilistic Latent Semantic Analysis	177
<i>U. Castellani, A. Perina, V. Murino, M. Bellani, G. Rambaldelli, M. Tansella, and P. Brambilla</i>	
Joint Factor and Kinetic Analysis of Dynamic FDOPA PET Scans of Brain Cancer Patients	185
<i>N. Dowson, P. Bourgeat, S. Rose, M. Daghish, J. Smith, M. Fay, A. Coulthard, C. Winter, D. MacFarlane, P. Thomas, S. Crozier, and O. Salvado</i>	
Early Detection of Emphysema Progression	193
<i>Vladlena Gorbunova, Sander S.A.M. Jacobs, Pechin Lo, Asger Dirksen, Mads Nielsen, Alireza Bab-Hadiashar, and Marleen de Bruijne</i>	
Unsupervised Learning of Brain States from fMRI Data	201
<i>F. Janoos, R. Machiraju, S. Sammet, M.V. Knopp, and I.Á. Móróc</i>	
Geometrical Consistent 3D Tracing of Neuronal Processes in ssTEM Data	209
<i>Verena Kaynig, Thomas J. Fuchs, and Joachim M. Buhmann</i>	

Analysis of the Striato-Thalamo-Cortical Connectivity on the Cortical Surface to Infer Biomarkers of Huntington’s Disease	217
<i>Linda Marrakchi-Kacem, Christine Delmaire, Alan Tucholka, Pauline Roca, Pamela Guevara, Fabrice Poupon, Jérôme Yelnik, Alexandra Durr, Jean-François Mangin, Stéphane Lehericy, and Cyril Poupon</i>	
The Fiber Laterality Histogram: A New Way to Measure White Matter Asymmetry	225
<i>Lauren J. O’Donnell, Carl-Fredrik Westin, Isaiah Norton, Stephen Whalen, Laura Rigolo, Ruth Propper, and Alexandra J. Golby</i>	
A Geometry-Based Particle Filtering Approach to White Matter Tractography	233
<i>Peter Savadjiev, Yogesh Rathi, James G. Malcolm, Martha E. Shenton, and Carl-Fredrik Westin</i>	
Accurate Definition of Brain Regions Position through the Functional Landmark Approach	241
<i>Bertrand Thirion, Gaël Varoquaux, and Jean-Baptiste Poline</i>	
A Comparison of the Cingulum Tract in ALS-B Patients and Controls Using Kernel Matching	249
<i>Sander van Noorden, Matthan Caan, Maaïke van der Graaff, Lucas van Vliet, and Frans Vos</i>	
Optimally-Discriminative Voxel-Based Analysis	257
<i>Tianhao Zhang and Christos Davatzikos</i>	
Hippocampal Shape Classification Using Redundancy Constrained Feature Selection	266
<i>Luping Zhou, Lei Wang, Chunhua Shen, and Nick Barnes</i>	
Simulation of Brain Mass Effect with an Arbitrary Lagrangian and Eulerian FEM	274
<i>Yasheng Chen, Songbai Ji, Xunlei Wu, Hongyu An, Hongtu Zhu, Dinggang Shen, and Weili Lin</i>	
Relating Structural and Functional Connectivity to Performance in a Communication Task	282
<i>Jeffrey T. Duda, Corey McMillan, Murray Grossman, and James C. Gee</i>	
Bayesian Classification of Multiple Sclerosis Lesions in Longitudinal MRI Using Subtraction Images	290
<i>Colm Elliott, Simon J. Francis, Douglas L. Arnold, D. Louis Collins, and Tal Arbel</i>	

Multivariate Network-Level Approach to Detect Interactions between Large-Scale Functional Systems	298
<i>Wei Gao, Hongtu Zhu, Kelly Giovanello, and Weili Lin</i>	
A Generalized Learning Based Framework for Fast Brain Image Registration	306
<i>Minjeong Kim, Guorong Wu, Pew-Thian Yap, and Dinggang Shen</i>	
Tracking Clathrin Coated Pits with a Multiple Hypothesis Based Method	315
<i>Liang Liang, Hongying Shen, Pietro De Camilli, and James S. Duncan</i>	
Shape-Based Diffeomorphic Registration on Hippocampal Surfaces Using Beltrami Holomorphic Flow	323
<i>Lok Ming Lui, Tsz Wai Wong, Paul Thompson, Tony Chan, Xianfeng Gu, and Shing-Tung Yau</i>	
Detecting Brain Activation in fMRI Using Group Random Walker	331
<i>Bernard Ng, Ghassan Hamarneh, and Rafeef Abugharbieh</i>	
Measures for Characterizing Directionality Specific Volume Changes in TBM of Brain Growth	339
<i>Vidya Rajagopalan, Julia Scott, Piotr A. Habas, Kio Kim, Francois Rousseau, Orit A. Glenn, A. James Barkovich, and Colin Studholme</i>	
Inter-subject Connectivity-Based Parcellation of a Patch of Cerebral Cortex	347
<i>Pauline Roca, Alan Tucholka, Denis Rivière, Pamela Guevara, Cyril Poupon, and Jean-François Mangin</i>	
On Super-Resolution for Fetal Brain MRI	355
<i>F. Rousseau, K. Kim, C. Studholme, M. Koob, and J.-L. Dietemann</i>	
Spatial Regularization of Functional Connectivity Using High-Dimensional Markov Random Fields	363
<i>Wei Liu, Peihong Zhu, Jeffrey S. Anderson, Deborah Yurgelun-Todd, and P. Thomas Fletcher</i>	
Simulation of Anatomical Structures	
Shell Model for Reconstruction and Real-Time Simulation of Thin Anatomical Structures	371
<i>Olivier Comas, Christian Duriez, and Stéphane Cotin</i>	

Personalization of Cubic Hermite Meshes for Efficient Biomechanical Simulations	380
<i>Pablo Lamata, Steven Niederer, David Barber, David Norsletten, Jack Lee, Rod Hose, and Nic Smith</i>	
Real-Time Surgical Simulation Using Reduced Order Finite Element Analysis	388
<i>Zeike A. Taylor, Stuart Crozier, and Sébastien Ourselin</i>	
Simulation of Nodules and Diffuse Infiltrates in Chest Radiographs Using CT Templates	396
<i>G.J.S. Litjens, L. Hogeweg, A.M.R. Schilham, P.A. de Jong, M.A. Viergever, and B. van Ginneken</i>	
High-Fidelity Meshes from Tissue Samples for Diffusion MRI Simulations	404
<i>Eleftheria Panagiotaki, Matt G. Hall, Hui Zhang, Bernard Siow, Mark F. Lythgoe, and Daniel C. Alexander</i>	
A Dynamic Skull Model for Simulation of Cerebral Cortex Folding	412
<i>Hanbo Chen, Lei Guo, Jingxin Nie, Tuo Zhang, Xintao Hu, and Tianming Liu</i>	
Coupled Personalisation of Electrophysiology Models for Simulation of Induced Ischemic Ventricular Tachycardia	420
<i>J. Relan, P. Chinchapatnam, M. Sermesant, K. Rhode, H. Delingette, R. Razavi, and N. Ayache</i>	
Real Time Ultrasound Needle Image Simulation Using Multi-dimensional Interpolation	429
<i>Mengchen Zhu and Septimiu E. Salcudean</i>	
Endoscopic and Microscopic Imaging	
Endoscopic Video Manifolds	437
<i>Selen Atasoy, Diana Mateus, Joe Lallemand, Alexander Meinig, Guang-Zhong Yang, and Nassir Navab</i>	
Automated Training Data Generation for Microscopy Focus Classification	446
<i>Dashan Gao, Dirk Padfield, Jens Rittscher, and Richard McKay</i>	
Augmenting Capsule Endoscopy Diagnosis: A Similarity Learning Approach	454
<i>S. Seshamani, R. Kumar, T. Dassopoulos, G. Mullin, and G. Hager</i>	

A Fully Automated Approach to Segmentation of Irregularly Shaped Cellular Structures in EM Images	463
<i>Aurélien Lucchi, Kevin Smith, Radhakrishna Achanta, Vincent Lepetit, and Pascal Fua</i>	
Automatic Neuron Tracing in Volumetric Microscopy Images with Anisotropic Path Searching	472
<i>Jun Xie, Ting Zhao, Tzumin Lee, Eugene Myers, and Hanchuan Peng</i>	
An Image Retrieval Approach to Setup Difficulty Levels in Training Systems for Endomicroscopy Diagnosis	480
<i>Barbara André, Tom Vercauteren, Anna M. Buchner, Muhammad Waseem Shahid, Michael B. Wallace, and Nicholas Ayache</i>	
3D Localization of Pronuclei of Human Zygotes Using Textures from Multiple Focal Planes	488
<i>A. Giusti, G. Corani, L. Gambardella, C. Magli, and L. Gianaroli</i>	
Motion Compensated SLAM for Image Guided Surgery	496
<i>Peter Mountney and Guang-Zhong Yang</i>	
Region Flow: A Multi-stage Method for Colonoscopy Tracking	505
<i>Jianfei Liu, Kalpathi R. Subramanian, and Terry S. Yoo</i>	
A System for Biopsy Site Re-targeting with Uncertainty in Gastroenterology and Oropharyngeal Examinations	514
<i>Baptiste Allain, Mingxing Hu, Laurence B. Lovat, Richard J. Cook, Tom Vercauteren, Sebastien Ourselin, and David J. Hawkes</i>	
Epitomized Summarization of Wireless Capsule Endoscopic Videos for Efficient Visualization	522
<i>Xinqi Chu, Chee Khun Poh, Liyuan Li, Kap Luk Chan, Shuicheng Yan, Weijia Shen, That Mon Htwe, Jiang Liu, Joo Hwee Lim, Eng Hui Ong, and Khek Yu Ho</i>	
Computing Maximum Association Graph in Microscopic Nucleus Images	530
<i>Branislav Stojkovic, Yongding Zhu, Jinhui Xu, Andrew Fritz, Michael J. Zeitz, Jaromira Vecerova, and Ronald Berezney</i>	
Estimation of 3D Geometry of Microtubules Using Multi-angle Total Internal Reflection Fluorescence Microscopy	538
<i>Qian Yang, Alexander Karpikov, Derek Toomre, and James Duncan</i>	

Image Registration

Recursive Green’s Function Registration	546
<i>Björn Beuthien, Ali Kamen, and Bernd Fischer</i>	
Summarizing and Visualizing Uncertainty in Non-rigid Registration	554
<i>Petter Risholm, Steve Pieper, Eigil Samset, and William M. Wells III</i>	
Coupled Registration-Segmentation: Application to Femur Analysis with Intra-subject Multiple Levels of Detail MRI Data	562
<i>Jérôme Schmid, Jinman Kim, and Nadia Magnenat-Thalmann</i>	
Groupwise Registration with Sharp Mean	570
<i>Guorong Wu, Hongjun Jia, Qian Wang, and Dinggang Shen</i>	
Lung Lobar Slippage Assessed with the Aid of Image Registration	578
<i>Youbing Yin, Eric A. Hoffman, and Ching-Long Lin</i>	
Generalization of Deformable Registration in Riemannian Sobolev Spaces	586
<i>Darko Zikic, Maximilian Baust, Ali Kamen, and Nassir Navab</i>	
An Efficient EM-ICP Algorithm for Symmetric Consistent Non-linear Registration of Point Sets	594
<i>Benoît Combès and Sylvain Prima</i>	
Image Registration Driven by Combined Probabilistic and Geometric Descriptors	602
<i>Linh Ha, Marcel Prastawa, Guido Gerig, John H. Gilmore, Cláudio T. Silva, and Sarang Joshi</i>	
Simultaneous Fine and Coarse Diffeomorphic Registration: Application to Atrophy Measurement in Alzheimer’s Disease	610
<i>Laurent Risser, François-Xavier Vialard, Robin Wolz, Darryl D. Holm, and Daniel Rueckert</i>	
Registration of Longitudinal Image Sequences with Implicit Template and Spatial-Temporal Heuristics	618
<i>Guorong Wu, Qian Wang, Hongjun Jia, and Dinggang Shen</i>	
3D Ultrasound to Stereoscopic Camera Registration through an Air-Tissue Boundary	626
<i>Michael C. Yip, Troy K. Adebar, Robert N. Rohling, Septimiu E. Salcudean, and Christopher Y. Nguan</i>	
Automatic Learning Sparse Correspondences for Initialising Groupwise Registration	635
<i>Pei Zhang, Steve A. Adeshina, and Timothy F. Cootes</i>	

Hierarchical Multimodal Image Registration Based on Adaptive Local Mutual Information	643
<i>Dante De Nigris, Laurence Mercier, Rolando Del Maestro, D. Louis Collins, and Tal Arbel</i>	
LogDemons Revisited: Consistent Regularisation and Incompressibility Constraint for Soft Tissue Tracking in Medical Images	652
<i>T. Mansi, X. Pennec, M. Sermesant, H. Delingette, and N. Ayache</i>	
Correspondences Search for Surface-Based Intra-Operative Registration	660
<i>Thiago R. dos Santos, Alexander Seitel, Hans-Peter Meinzer, and Lena Maier-Hein</i>	
Model-Based Multi-view Fusion of Cinematic Flow and Optical Imaging	668
<i>Mickael Savinaud, Martin de La Gorce, Serge Maitrejean, and Nikos Paragios</i>	
Simultaneous Geometric - Iconic Registration	676
<i>Aristeidis Sotiras, Yangming Ou, Ben Glocker, Christos Davatzikos, and Nikos Paragios</i>	
Groupwise Registration by Hierarchical Anatomical Correspondence Detection	684
<i>Guorong Wu, Qian Wang, Hongjun Jia, and Dinggang Shen</i>	
Author Index	693

Table of Contents – Part III

Segmentation and Modeling

Combining Morphological Information in a Manifold Learning Framework: Application to Neonatal MRI	1
<i>P. Aljabar, R. Wolz, L. Srinivasan, S. Counsell, J.P. Boardman, M. Murgasova, V. Doria, M.A. Rutherford, A.D. Edwards, J.V. Hajnal, and D. Rueckert</i>	
Fast Random Walker with Priors Using Precomputation for Interactive Medical Image Segmentation	9
<i>Shawn Andrews, Ghassan Hamarneh, and Ahmed Saad</i>	
Extraction of the Plane of Minimal Cross-Sectional Area of the Corpus Callosum Using Template-Driven Segmentation	17
<i>Neda Changizi, Ghassan Hamarneh, Omer Ishaq, Aaron Ward, and Roger Tam</i>	
Incorporating Priors on Expert Performance Parameters for Segmentation Validation and Label Fusion: A Maximum a Posteriori STAPLE	25
<i>Olivier Commowick and Simon K. Warfield</i>	
Automated Segmentation of 3-D Spectral OCT Retinal Blood Vessels by Neural Canal Opening False Positive Suppression	33
<i>Zhihong Hu, Meindert Niemeijer, Michael D. Abramoff, Kyungmoo Lee, and Mona K. Garvin</i>	
Detection of Gad-Enhancing Lesions in Multiple Sclerosis Using Conditional Random Fields	41
<i>Zahra Karimaghloo, Mohak Shah, Simon J. Francis, Douglas L. Arnold, D. Louis Collins, and Tal Arbel</i>	
Automated Sulci Identification via Intrinsic Modeling of Cortical Anatomy	49
<i>Yonggang Shi, Bo Sun, Rongjie Lai, Ivo Dinov, and Arthur W. Toga</i>	
In Vivo MRI Assessment of Knee Cartilage in the Medial Meniscal Tear Model of Osteoarthritis in Rats	57
<i>Zhiyong Xie, Serguei Liachenko, Ping-Chun Chiao, Santos Carvajal-Gonzalez, Susan Bove, and Thomas Bocan</i>	

Construction of Neuroanatomical Shape Complex Atlas from 3D Brain MRI	65
<i>Ting Chen, Anand Rangarajan, Stephan J. Eisenschenk, and Baba C. Vemuri</i>	
Non-parametric Iterative Model Constraint Graph Min-Cut for Automatic Kidney Segmentation	73
<i>M. Freiman, A. Kronman, S.J. Esses, L. Joskowicz, and J. Sosna</i>	
Synthetic MRI Signal Standardization: Application to Multi-atlas Analysis	81
<i>Juan Eugenio Iglesias, Ivo Dinov, Jaskaran Singh, Gregory Tong, and Zhuowen Tu</i>	
Multi-organ Segmentation from Multi-phase Abdominal CT via 4D Graphs Using Enhancement, Shape and Location Optimization	89
<i>Marius George Linguraru, John A. Pura, Ananda S. Chowdhury, and Ronald M. Summers</i>	
A Semi-automatic Method for Segmentation of the Carotid Bifurcation and Bifurcation Angle Quantification on Black Blood MRA	97
<i>Hui Tang, Robbert S. van Onkelen, Theo van Walsum, Reinhard Hameeteman, Michiel Schaap, Fufa. L. Tori, Quirijn J.A. van den Bouwhuijsen, Jacqueline C.M. Witteman, Aad van der Lugt, Lucas J. van Vliet, and Wiro J. Niessen</i>	
<i>Standing on the Shoulders of Giants: Improving Medical Image Segmentation via Bias Correction</i>	105
<i>Hongzhi Wang, Sandhitsu Das, John Pluta, Caryne Craige, Murat Altinay, Brian Avants, Michael Weiner, Susanne Mueller, and Paul Yushkevich</i>	
Layout Consistent Segmentation of 3-D Meshes via Conditional Random Fields and Spatial Ordering Constraints	113
<i>Alexander Zouhar, Sajjad Baloch, Yanghai Tsin, Tong Fang, and Siegfried Fuchs</i>	
Cross-Visit Tumor Sub-segmentation and Registration with Outlier Rejection for Dynamic Contrast-Enhanced MRI Time Series Data	121
<i>G.A. Buonaccorsi, C.J. Rose, J.P.B. O'Connor, C. Roberts, Y. Watson, A. Jackson, G.C. Jayson, and G.J.M. Parker</i>	
Nonlocal Patch-Based Label Fusion for Hippocampus Segmentation	129
<i>Pierrick Coupé, José V. Manjón, Vladimir Fonov, Jens Pruessner, Montserrat Robles, and D. Louis Collins</i>	
Cellular Automata Segmentation of Brain Tumors on Post Contrast MR Images	137
<i>Andac Hamamci, Gozde Unal, Nadir Kucuk, and Kayihan Engin</i>	

Agreement-Based Semi-supervised Learning for Skull Stripping	147
<i>Juan Eugenio Iglesias, Cheng-Yi Liu, Paul Thompson, and Zhuowen Tu</i>	
Construction of Patient Specific Atlases from Locally Most Similar Anatomical Pieces	155
<i>Liliane Ramus, Olivier Commowick, and Grégoire Malandain</i>	
Automatic Lung Lobe Segmentation Using Particles, Thin Plate Splines, and Maximum a Posteriori Estimation	163
<i>James C. Ross, Raúl San José Estépar, Gordon Kindlmann, Alejandro Díaz, Carl-Fredrik Westin, Edwin K. Silverman, and George R. Washko</i>	
Graph Search with Appearance and Shape Information for 3-D Prostate and Bladder Segmentation	172
<i>Qi Song, Yinxiao Liu, Yunlong Liu, Punam K. Saha, Milan Sonka, and Xiaodong Wu</i>	
Segmentation of Cortical MS Lesions on MRI Using Automated Lamina Profile Shape Analysis	181
<i>Christine L. Tardif, D. Louis Collins, Simon F. Eskildsen, John B. Richardson, and G. Bruce Pike</i>	
3D Knowledge-Based Segmentation Using Pose-Invariant Higher-Order Graphs	189
<i>Chaohui Wang, Olivier Teboul, Fabrice Michel, Salma Essafi, and Nikos Paragios</i>	
Markov Random Field Driven Region-Based Active Contour Model (MaRACel): Application to Medical Image Segmentation	197
<i>Jun Xu, James P. Monaco, and Anant Madabhushi</i>	
Robotics, Motion Modeling and Computer-Assisted Interventions	
Predicting Target Vessel Location for Improved Planning of Robot-Assisted CABG Procedures	205
<i>Daniel S. Cho, Cristian A. Linte, Elvis Chen, Chris Wedlake, John Moore, John Barron, Rajni Patel, and Terry M. Peters</i>	
Quantification of Prostate Deformation due to Needle Insertion during TRUS-guided Biopsy	213
<i>Tharindu De Silva, Aaron Fenster, Jagath Samarabandu, and Aaron D. Ward</i>	
Optimized Anisotropic Rotational Invariant Diffusion Scheme on Cone-Beam CT	221
<i>Dirk-Jan Kroon, Cornelis H. Slump, and Thomas J.J. Maal</i>	

Control of Articulated Snake Robot under Dynamic Active Constraints	229
<i>Ka-Wai Kwok, Valentina Vitiello, and Guang-Zhong Yang</i>	
Estimating Radiation Exposure in Interventional Environments	237
<i>Alexander Ladikos, Cedric Cagniard, Reza Ghotbi, Maximilian Reiser, and Nassir Navab</i>	
Force Adaptive Multi-spectral Imaging with an Articulated Robotic Endoscope	245
<i>David P. Noonan, Christopher J. Payne, Jianzhong Shang, Vincent Sauvage, Richard Newton, Daniel Elson, Ara Darzi, and Guang-Zhong Yang</i>	
Motion Tracking in Narrow Spaces: A Structured Light Approach	253
<i>Oline Vinter Olesen, Rasmus R. Paulsen, Liselotte Højgaard, Bjarne Roed, and Rasmus Larsen</i>	
Tracking of Irregular Graphical Structures for Tissue Deformation Recovery in Minimally Invasive Surgery	261
<i>Marco Visentini-Scarzanella, Robert Merrifield, Danail Stoyanov, and Guang-Zhong Yang</i>	
Graph Based Interactive Detection of Curve Structures in 2D Fluoroscopy	269
<i>Peng Wang, Wei-shing Liao, Terrence Chen, Shaohua K. Zhou, and Dorin Comaniciu</i>	
Automated Digital Dental Articulation	278
<i>James J. Xia, Yu-Bing Chang, Jaime Gateno, Zixiang Xiong, and Xiaobo Zhou</i>	
Image-Based Respiratory Motion Compensation for Fluoroscopic Coronary Roadmapping	287
<i>Ying Zhu, Yanghai Tsing, Hari Sundar, and Frank Sauer</i>	
Surgical Task and Skill Classification from Eye Tracking and Tool Motion in Minimally Invasive Surgery	295
<i>Narges Ahmadi, Gregory D. Hager, Lisa Ishii, Gabor Fichtinger, Gary L. Gallia, and Masaru Ishii</i>	
Micro-force Sensing in Robot Assisted Membrane Peeling for Vitreoretinal Surgery	303
<i>Marcin Balicki, Ali Uneri, Iulian Iordachita, James Handa, Peter Gehlbach, and Russell Taylor</i>	
C-arm Pose Estimation in Prostate Brachytherapy by Registration to Ultrasound	311
<i>Pascal Fallavollita, Clif Burdette, Danny Song, Purang Abolmaesumi, and Gabor Fichtinger</i>	

Cognitive Burden Estimation for Visuomotor Learning with fNIRS	319
<i>David R.C. James, Felipe Orihuela-Espina, Daniel R. Leff, George P. Mylonas, Ka-Wai Kwok, Ara W. Darzi, and Guang-Zhong Yang</i>	
Prediction Framework for Statistical Respiratory Motion Modeling	327
<i>Tobias Klinder, Cristian Lorenz, and Jörn Ostermann</i>	
Image Estimation from Marker Locations for Dose Calculation in Prostate Radiation Therapy	335
<i>Huai-Ping Lee, Mark Foskey, Josh Levy, Rohit Saboo, and Ed Chaney</i>	
A Machine Learning Approach for Deformable Guide-Wire Tracking in Fluoroscopic Sequences	343
<i>Olivier Pauly, Hauke Heibel, and Nassir Navab</i>	
Collaborative Tracking for MRI-Guided Robotic Intervention on the Beating Heart	351
<i>Y. Zhou, E. Yeniaras, P. Tsiamyrtzis, N. Tsekos, and I. Pavlidis</i>	
Calibration and Use of Intraoperative Cone-Beam Computed Tomography: An In-Vitro Study for Wrist Fracture	359
<i>Erin Janine Smith, Anton Oentoro, Hisham Al-Sanawi, Braden Gammon, Paul St. John, David R. Pichora, and Randy E. Ellis</i>	
A Strain Energy Filter for 3D Vessel Enhancement	367
<i>Changyan Xiao, Marius Staring, Denis Shamonin, Johan H.C. Reiber, Jan Stolk, and Berend C. Stoel</i>	
Virtual Stent Grafting in Personalized Surgical Planning for Treatment of Aortic Aneurysms Using Image-Based Computational Fluid Dynamics	375
<i>Guanglei Xiong and Charles A. Taylor</i>	
MRI-Guided Robotic Prostate Biopsy: A Clinical Accuracy Validation	383
<i>Helen Xu, Andras Lasso, Siddharth Vikal, Peter Guion, Axel Krieger, Aradhana Kaushal, Louis L. Whitcomb, and Gabor Fichtinger</i>	
Online 4-D CT Estimation for Patient-Specific Respiratory Motion Based on Real-Time Breathing Signals	392
<i>Tiancheng He, Zhong Xue, Weixin Xie, and Stephen T.C. Wong</i>	
Modeling and Segmentation of Surgical Workflow from Laparoscopic Video	400
<i>Tobias Blum, Hubertus Feußner, and Nassir Navab</i>	

Fused Video and Ultrasound Images for Minimally Invasive Partial Nephrectomy: A Phantom Study	408
<i>Carling L. Cheung, Chris Wedlake, John Moore, Stephen E. Pautler, and Terry M. Peters</i>	
Probabilistic 4D Blood Flow Mapping	416
<i>Ola Friman, Anja Hennemuth, Andreas Harloff, Jelena Bock, Michael Markl, and Heinz-Otto Peitgen</i>	
Rotational Encoding of C-arm Fluoroscope with Tilt Sensing Accelerometer	424
<i>Victor Grzeda and Gabor Fichtinger</i>	
Robotic Hand-Held Surgical Device: Evaluation of End-Effector’s Kinematics and Development of Proof-of-Concept Prototypes	432
<i>Ali Hassan Zahraee, Jérôme Szewczyk, Jamie Kyujin Paik, and Guillaume Morel</i>	
Guide-Wire Extraction through Perceptual Organization of Local Segments in Fluoroscopic Images	440
<i>Nicolas Honnorat, Régis Vaillant, and Nikos Paragios</i>	
Single-Projection Based Volumetric Image Reconstruction and 3D Tumor Localization in Real Time for Lung Cancer Radiotherapy	449
<i>Ruijiang Li, Xun Jia, John H. Lewis, Xuejun Gu, Michael Folkerts, Chunhua Men, and Steve B. Jiang</i>	
A Method for Planning Safe Trajectories in Image-Guided Keyhole Neurosurgery	457
<i>Reuben R. Shamir, Idit Tamir, Elad Dabool, Leo Joskowicz, and Yigal Shoshan</i>	
Adaptive Multispectral Illumination for Retinal Microsurgery	465
<i>Raphael Sznitman, Diego Rother, Jim Handa, Peter Gehlbach, Gregory D. Hager, and Russell Taylor</i>	
Motion Artifact Correction of Multi-Photon Imaging of Awake Mice Models Using Speed Embedded HMM	473
<i>Taoyi Chen, Zhong Xue, Changhong Wang, Zhenshen Qu, Kelvin K. Wong, and Stephen T.C. Wong</i>	

Image Reconstruction, Enhancement and Representation

Diagnostic Radiograph Based 3D Bone Reconstruction Framework: Application to Osteotomy Surgical Planning	481
<i>Pavan Gamage, Sheng Quan Xie, Patrice Delmas, and Wei Liang Xu</i>	

Comparative Analysis of Quasi-Conformal Deformations in Shape Space	489
<i>Vahid Taimouri, Huiquan He, and Jing Hua</i>	
Establishing Spatial Correspondence between the Inner Colon Surfaces from Prone and Supine CT Colonography	497
<i>Holger Roth, Jamie McClelland, Marc Modat, Darren Boone, Mingxing Hu, Sebastien Ourselin, Greg Slabaugh, Steve Halligan, and David Hawkes</i>	
Heat Kernel Smoothing Using Laplace-Beltrami Eigenfunctions	505
<i>Seongho Seo, Moo K. Chung, and Hourri K. Vorperian</i>	
Under-Determined Non-cartesian MR Reconstruction with Non-convex Sparsity Promoting Analysis Prior	513
<i>Angshul Majumdar and Rabab K. Ward</i>	
A Statistical Approach for Achievable Dose Querying in IMRT Planning	521
<i>Patricio Simari, Binbin Wu, Robert Jacques, Alex King, Todd McNutt, Russell Taylor, and Michael Kazhdan</i>	
Multivariate Statistical Analysis of Deformation Momenta Relating Anatomical Shape to Neuropsychological Measures	529
<i>Nikhil Singh, P. Thomas Fletcher, J. Samuel Preston, Linh Ha, Richard King, J. Stephen Marron, Michael Wiener, and Sarang Joshi</i>	
Shape Analysis of Vestibular Systems in Adolescent Idiopathic Scoliosis Using Geodesic Spectra	538
<i>Wei Zeng, Lok Ming Lui, Lin Shi, Defeng Wang, Winnie C.W. Chu, Jack C.Y. Cheng, Jing Hua, Shing-Tung Yau, and Xianfeng Gu</i>	
Value-Based Noise Reduction for Low-Dose Dual-Energy Computed Tomography	547
<i>Michael Balda, Björn Heismann, and Joachim Hornegger</i>	
Automatic Detection of Anatomical Features on 3D Ear Impressions for Canonical Representation	555
<i>Sajjad Baloch, Rupen Melkisetoglu, Simon Flöry, Sergei Azernikov, Greg Slabaugh, Alexander Zouhar, and Tong Fang</i>	
Probabilistic Multi-Shape Representation Using an Isometric Log-Ratio Mapping	563
<i>Neda Changizi and Ghassan Hamarneh</i>	
Efficient Robust Reconstruction of Dynamic PET Activity Maps with Radioisotope Decay Constraints	571
<i>Fei Gao, Huafeng Liu, and Pengcheng Shi</i>	

Nonlinear Embedding towards Articulated Spine Shape Inference Using Higher-Order MRFs 579
Samuel Kadoury and Nikos Paragios

Improved Method for Point-Based Tracking 587
Andrei Danilchenko, Andrew D. Wiles, Ramya Balachandran, and J. Michael Fitzpatrick

Computer Aided Diagnosis

A Texton-Based Approach for the Classification of Lung Parenchyma in CT Images 595
Mehrdad J. Gangeh, Lauge Sørensen, Saher B. Shaker, Mohamed S. Kamel, Marleen de Bruijne, and Marco Loog

Active Learning for an Efficient Training Strategy of Computer-Aided Diagnosis Systems: Application to Diabetic Retinopathy Screening 603
C.I. Sánchez, M. Niemeijer, M.D. Abràmoff, and B. van Ginneken

Sparse Bayesian Learning for Identifying Imaging Biomarkers in AD Prediction 611
Li Shen, Yuan Qi, Sungeun Kim, Kwangsik Nho, Jing Wan, Shannon L. Risacher, Andrew J. Saykin, and ADNI

Computer-Aided Detection of Pulmonary Pathology in Pediatric Chest Radiographs 619
André Mouton, Richard D. Pitcher, and Tania S. Douglas

Toward Precise Pulmonary Nodule Descriptors for Nodule Type Classification 626
Amal Farag, Shireen Elhabian, James Graham, Aly Farag, and Robert Falk

Morphology-Guided Graph Search for Untangling Objects: *C. elegans* Analysis 634
T. Riklin Raviv, V. Ljosa, A.L. Conery, F.M. Ausubel, A.E. Carpenter, P. Golland, and C. Wählby

Automatic Cephalometric Evaluation of Patients Suffering from Sleep-Disordered Breathing 642
Lior Wolf, Tamir Yedidya, Roy Ganor, Michael Chertok, Ariela Nachmani, and Yehuda Finkelstein

Fusion of Local and Global Detection Systems to Detect Tuberculosis in Chest Radiographs 650
Laurens Hogeweg, Christian Mol, Pim A. de Jong, Rodney Dawson, Helen Ayles, and Bram van Ginneken

Novel Morphometric Based Classification via Diffeomorphic Based Shape Representation Using Manifold Learning	658
<i>Rachel Sparks and Anant Madabhushi</i>	
Semi Supervised Multi Kernel (SeSMiK) Graph Embedding: Identifying Aggressive Prostate Cancer via Magnetic Resonance Imaging and Spectroscopy	666
<i>Pallavi Tiwari, John Kurhanewicz, Mark Rosen, and Anant Madabhushi</i>	
Author Index	675

Automated Macular Pathology Diagnosis in Retinal OCT Images Using Multi-Scale Spatial Pyramid with Local Binary Patterns

Yu-Ying Liu¹, Mei Chen², Hiroshi Ishikawa^{3,4}, Gadi Wollstein³,
Joel S. Schuman^{3,4}, and James M. Rehg¹

¹ School of Interactive Computing, Georgia Institute of Technology, Atlanta, GA

² Intel Labs Pittsburgh, Pittsburgh, PA

³ UPMC Eye Center, University of Pittsburgh Medical Center, Pittsburgh, PA

⁴ Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA

Abstract. We address a novel problem domain in the analysis of optical coherence tomography (OCT) images: the diagnosis of multiple macular pathologies in retinal OCT images. The goal is to identify the presence of normal macula and each of three types of macular pathologies, namely, macular hole, macular edema, and age-related macular degeneration, in the OCT slice centered at the fovea. We use a machine learning approach based on global image descriptors formed from a multi-scale spatial pyramid. Our local descriptors are dimension-reduced Local Binary Pattern histograms, which are capable of encoding texture information from OCT images of the retina. Our representation operates at multiple spatial scales and granularities, leading to robust performance. We use 2-class Support Vector Machine classifiers to identify the presence of normal macula and each of the three pathologies. We conducted extensive experiments on a large dataset consisting of 326 OCT scans from 136 patients. The results show that the proposed method is very effective.

1 Introduction

Optical Coherence Tomography (OCT) is a non-contact, non-invasive 3-D imaging technique which performs optical sectioning at micron resolution. It has been widely adopted as the standard of care in ophthalmology for identifying the presence of disease and its progression [1]. This technology measures the optical back scattering of the tissues, making it possible to visualize intraocular structures, such as the macula and the optic nerve. An example 3D ocular OCT scan is given in Fig. 1a. The ability to visualize the internal structures of the retina (the z-axis direction in Fig. 1a) makes it possible to diagnose diseases, such as glaucoma and macular hole, objectively and quantitatively.

Although OCT imaging technology continues to evolve, technology for automated OCT image analysis and interpretation has not kept pace. With OCT data being generated at increasingly higher sampling rates, there is a strong need for automated analysis to support disease screening and diagnosis. This need is amplified by the fact that an ophthalmologist making a diagnosis under

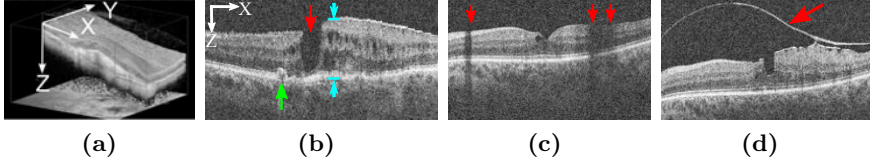


Fig. 1. (a) Example 3D macular OCT scan. (b)(c)(d) Example x-z slice with (b) MH (red), ME (blue), and AMD (green), (c) shadowing effects, (d) a detached tissue.

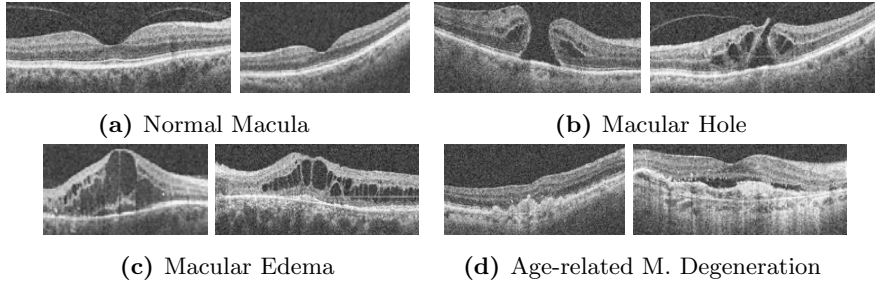


Fig. 2. Characteristics of (a) normal macula: a smooth depression shows at the center (fovea), (b) macular hole (MH): a partial or full hole formation at the fovea, (c) macular edema (ME) : retinal thickening and liquid accumulation appears as black blobs around the fovea, (d) age-related macular degeneration (AMD): irregular contours usually extruded in dome shapes appear at the bottom layer of the retina.

standard clinical conditions does not have the assistance of a radiologist in interpreting OCT data. There have been several prior works addressing topics in ocular OCT image processing, such as intra-retinal layer segmentation [2] and local quality assessment [3]. However, to our knowledge, *there has been no prior work on automated macular pathology identification in OCT images.*

The macula is located at the center of the retina and is responsible for highly-sensitive, accurate vision. Acute maculopathy can lead to the loss of central vision and even blindness. For example, a particular type of maculopathy, called age-related macular degeneration (AMD), is the leading cause of visual loss among elderly persons. The Beaver Dam Eye Study reported that 30% of individuals aged 75 and older have some form of AMD, which has increasingly important social and economic impact in the United States [4]. The screening and diagnosis of maculopathy is a substantial public health problem.

We present a method for automatically identifying macular pathologies given a manually-selected x-z 2D slice centered at the fovea in a retinal OCT scan. Specifically, we identify the presence of the normal macula and each of the following macular pathologies, macular hole (MH)¹, macular edema (ME), and AMD.

¹ For simplicity, we include both macular hole and macular pseudohole in MH category.

Fig. 2 gives example images for each pathology. Note that multiple pathologies can coexist in one eye, as depicted in Fig. 1b. In this case, the automated identification tool should report the existence of each pathology.

2 Approach

Automated pathology identification in ocular OCT images is complicated by four factors. First, the co-existence of pathologies can complicate the overall appearance (see Fig. 1b). Second, the measurement of reflectivity of the tissue is affected by the optical properties of the overlaying tissues [1], e.g. opaque media or the high rate of light absorption by blood vessels will produce shadowing effects (see Fig. 1c). Third, a variety of proliferated or detached tissues can appear, e.g. membrane or floaters (see Fig. 1d). Fourth, a portion of the image may have low quality due to imperfect imaging [3]. As a result of this high variability, attempts to hand-craft a set of features or rules to identify each pathology are unlikely to succeed. Instead, we propose to use a pattern-based global image representation combined with machine learning techniques to automatically discover the discriminative patterns for each pathology from training examples.

Our method consists of the following three steps, which are illustrated in Fig. 3. First, image alignment is performed to reduce the appearance variation across scans. Second, we construct a global descriptor for the aligned image by using multi-scale spatial pyramid (MSSP) and the dimension-reduced Local Binary Pattern (LBP) histogram [5] based on Principle Component Analysis (PCA), in order to represent the retina at different spatial scales and granularities. Finally, for each pathology, we train a 2-class non-linear Support Vector Machine (SVM) [6] with radial basis function (RBF) kernel using a labeled training set. We now describe each step in detail.

2.1 Retina Alignment

Since the imaged retinas have large variations in their inclination angles, positions, and natural curvatures across scans, as shown in Fig. 2, it is desirable

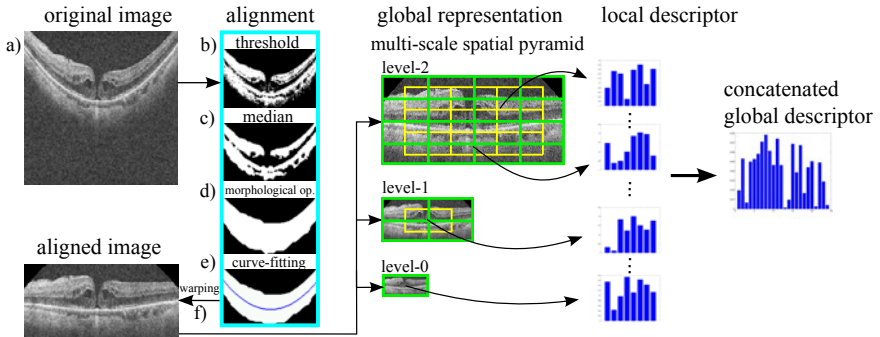


Fig. 3. Stages in the constructing of the global image descriptor

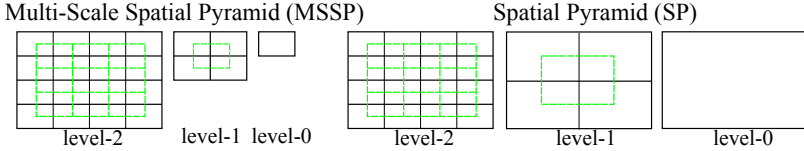


Fig. 4. Structures of a 3-level MSSP and SP. The green lines indicate the overlapped blocks which can also be added in the global representation.

to roughly align the retinas to reduce these variations before constructing our feature representation. To this end, we use a simple procedure as follows: (1) threshold the original image (Fig 3, a) to detect most of the retina structures (Fig 3, b); then, apply a median filter to remove noise and thin detached tissues (Fig 3, c); (2) find the entire retina by using morphological closing and then opening; by closing, we fill-up black blobs inside the retina, and by opening, to remove thin or small objects outside the retina (Fig 3, d); (3) fit the found retina with a quadratic curve by using least-square curve fitting (Fig 3, e); (4) warp the retina to be approximately horizontal by translating each column according to the fitted curve, and crop it in the z-direction (Fig 3, f).

2.2 Multi-Scale Spatial Pyramid (MSSP)

There are three motivations for our choice of a global spatially-distributed feature representation for OCT imagery based on MSSP. First, pathologies are often localized to specific retinal areas, making it important to encode spatial location. Second, the context provided by the overall appearance of the retina is important for correct interpretation; e.g., in Fig. 1c, we can distinguish between a shadow and macular hole only in the context of the entire slice. Third, pathologies can exhibit discriminating characteristics at both small and large scales. Thus, both micro-patterns and macro-patterns should be represented. For these reasons, we use a global image representation which preserves spatial organization in conjunction with multi-scale modeling.

We propose to use the multi-scale version of spatial pyramid (SP) [7], denoted as MSSP, to capture the geometry of the aligned retina at multiple scales and spatial resolutions. This global framework, MSSP, was recently proposed in [8], where it was successfully applied to challenging scene classification tasks. For clarity, in Fig 4, we illustrate the differences between a 3-level MSSP² and SP frameworks. The local features computed from all spatial blocks are concatenated in a predefined order to form an overall global descriptor, as illustrated in Fig. 3. Note that we also add the features from the overlapped blocks (the green blocks in Fig. 4) to reduce boundary effects.

² To form a k -level MSSP, for each level l ($0 \leq l \leq (k-1)$), we rescale the original image by 2^{l-k+1} using bilinear interpolation, and divide the rescaled image into 2^l blocks in each dimension.

2.3 Histogram of LBP and Dimensionality Reduction

Local binary pattern (LBP) [5] is a non-parametric kernel which summarizes the local structure around a pixel. LBP is known to be a highly discriminative operator and has been successfully applied to several tasks in computer vision [9,8] and medical imaging [10,11].

While there are several types of LBP, we follow [9,10,11] in adopting $LBP_{8,1}$ to capture the micro-patterns that reside in each local block. The $LBP_{8,1}$ operator derives an 8 bit binary code by comparing the center pixel to each of its 8 nearest neighbors in a 3×3 neighborhood (see Fig. 5a). The resulting 8 bits are concatenated circularly to form an LBP code in the range [0 255]. For each block of pixels in the MSSP, we compute the histogram of LBP codes to encode the statistical distribution of different micro-patterns, such as spots, edges, corners, and flat areas [5]. Histogram descriptors have proven to be an effective means to aggregate local intensity patterns into global discriminative features. In particular, they avoid the need to precisely localize discriminating image structures, which can be difficult to do in complex and highly variable OCT images. Since we compute LBP histogram in both original and rescaled image blocks, the distribution of both micro-patterns and macro-patterns can be encoded. Note that many previous works applied LBP only in the original image [9,10,11], which may not capture the large-scale patterns effectively.

Although a single $LBP_{8,1}$ histogram has only 256 bins, the concatenation of histograms from each block to form the global feature vector results in an impractically high dimension. We adopt PCA to reduce the dimension of LBP histogram, as proposed in [8]. We denote it as $LBP_{8,1}^{pca}$.

It is important to note that previous works [9,10] employing LBP histogram features have adopted an alternative approach to dimensionality reduction, called uniform LBP [5], which we have found to be less effective than $LBP_{8,1}^{pca}$ for AMD category. An LBP pattern is called *uniform* if it contains at most two bitwise transitions, as demonstrated in Fig. 5b. A histogram of $LBP_{8,1}^{u2}$ is formed by retaining occurrences of each of 58 uniform patterns and putting all occurrences of 198 non-uniform patterns, denoted as $LBP_{8,1}^{-u2}$, to a single bin, resulting in 59 bins in total. It was observed in [5] that $LBP_{8,1}^{u2}$ occupied 90% of all $LBP_{8,1}$ patterns in pixel count, when computed from image textures; however, as recently noted in [12], when LBP codes are computed in the rescaled images, $LBP_{8,1}^{u2}$ may no longer be in the majority. More importantly, the distribution of individual $LBP_{8,1}^{-u2}$ patterns can contain important distinctive information for category discrimination, despite of its low counts (see Fig. 6). To advocate the use of PCA for LBP dimension reduction, we will compare the performance of using $LBP_{8,1}^{pca}$, $LBP_{8,1}^{u2}$, and $LBP_{8,1}$ in Section 3.

2.4 Classification Using Support Vector Machine (SVM)

After computing global features, we train a 2-class non-linear SVM with RBF kernel and probabilistic output [6] for each pathology using a 1 vs. the rest approach. The probability scores for each classifier are compared to a set of decision thresholds to determine the corresponding sensitivities and specificities.

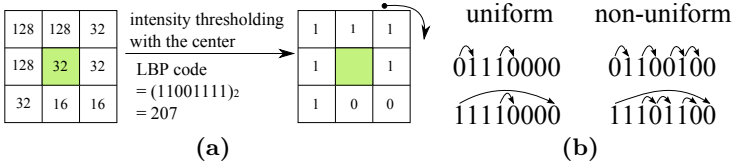


Fig. 5. Examples of (a) $LBP_{8,1}$ code computation, (b) uniform and non-uniform LBP

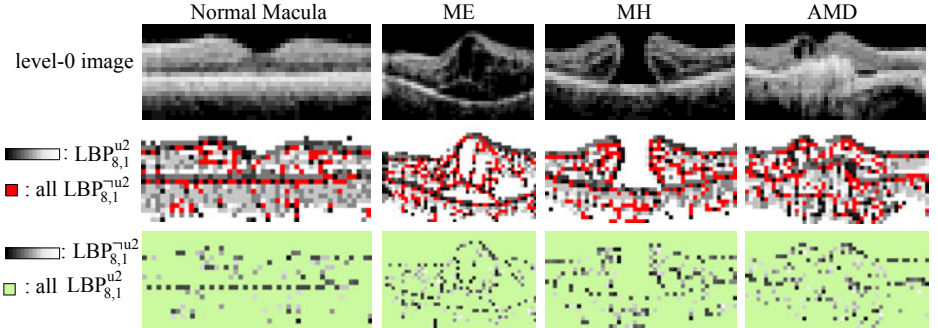


Fig. 6. Visualization of $LBP_{8,1}^{u2}$ and $LBP_{8,1}^{-u2}$ on example level-0 images. All $LBP_{8,1}^{-u2}$ are shown in red in the 2nd row, and the individual $LBP_{8,1}^{-u2}$ codes are shown in gray-level in the 3rd row. We find that although $LBP_{8,1}^{u2}$ patterns have low counts, most of them reside in the important contours and can therefore be useful for discrimination.

3 Experimental Results

We collected a dataset consisting of 326 macular spectral-domain OCT scans from 136 patients. The original resolution of the scans was either $200 \times 200 \times 1024$ or $512 \times 128 \times 1024$ scanned in $6 \times 6 \times 2$ mm volume in width(x), height(y) and depth(z). We rescaled all x-z slices to 200×200 . For each scan, the x-z slice centered at the fovea and the pathologies it contained was identified independently by two expert ophthalmologists. Among all 326 scans, 96.2% of the selected fovea-centered slices from the two experts were within 4 slice distance; for these scans, the labeling agreement for normal macula, ME, MH, AMD is 98.6%, 91.7%, 96.5%, 92.0%, at the patient-level respectively. This imperfect agreement provides further evidence of the need for an objective computer-aided diagnosis tool. We used the labels from one ophthalmologist as the ground truth.

The statistics of the labeled dataset are as follows. We have 67 positive and 259 negative scans for normal macula, 205 and 121 for ME, 81 and 245 for MH, 103 and 223 for AMD. From the standpoint of patient statistics, we have 57 positive and 79 negative patients for normal macula, 87 and 49 for ME, 34 and 102 for MH, 36 and 100 for AMD. We used 10-fold cross validation at the patient level; i.e., 10% of positive and negative patients were put in the testing

set for each run. Here, for a specific pathology, the patients that have at least one positive scan were counted as positive patients; also, all images from the same patient were put together in either the training or testing set. We used one fold of training data to do SVM model selection based on the area under the receiver operator characteristic curve (AUC) performance, and fixed the found best parameters throughout all the experiments. We normalized the sensitivity and specificity values across patients and used AUC as the comparison metric.

To validate the performance of LBP histograms as features, we compared them to several other popular descriptors, including mean and standard deviation of intensity (M,S), intensity histogram (I), and orientation histogram (O). Each feature is employed with a 3-level MSSP with overlapped blocks³. The orientation histogram [13] is formed from the gradient direction and magnitude computed from 2×2 neighborhoods. For I and O features, we used the quantization of 32 bins in neighborhood and angle, respectively, since this produced the best results. For LBP histogram, we quantize the intensity image to 32 levels before LBP encoding in order to suppress pixel noise. This quantization improves the accuracy of LBP by about 0.8% on average. For $LBP_{8,1}^{pca}$ computation, the principle axes were derived from the training images of each fold separately.

Overall, from Table 1, $LBP_{8,1}^{pca}(32)$ achieved the best average performance. In details, we find for normal macula and ME, the AUC are above 0.95 for almost all descriptors; this can be attributed to the reduced variation in appearance for the normal category and to the large amount of positive training data for ME. The AUC results of MH and AMD are lower, which can be due to their greater appearance variation and lack of data. In these two categories, LBP with dimension reduction ($LBP_{8,1}^{u2}$ and $LBP_{8,1}^{pca}$) outperforms the other descriptors by a clear margin. Note that the use of all 256 bins of LBP histogram gives the worst results, presumably due to overfitting in the high dimensional feature space.

We then compare the results of $LBP_{8,1}^{u2}$ and $LBP_{8,1}^{pca}$. From Table 1, $LBP_{8,1}^{u2}$ and $LBP_{8,1}^{pca}$ have similar performance for the first three categories, where the pathologies are described mostly by smooth or global shape changes. In AMD category, however, $LBP_{8,1}^{pca}(32)$ is clearly superior to $LBP_{8,1}^{u2}$. Our conjecture is that since AMD is characterized by the irregular contours of the bottom retinal layer, the removal of individual non-uniform patterns, as used in $LBP_{8,1}^{u2}$, can result in the loss of important discriminative information. Finally, the use of first 32 principal components seems sufficient for category discrimination.

In Table 2, we compare the performance of 3-level multi-scale spatial pyramid (MSSP) with a 3-level spatial pyramid (SP) and single level-2 spatial division (SL)⁴. MSSP achieves the best overall performance; in AMD category, MSSP outperforms the other frameworks with a large margin (0.888 vs. 0.826). When features from the overlapped blocks are also used (denoted as “+O”), the performance of all frameworks is improved.

³ If the feature dimension of the block descriptor is d , a 3-level MSSP with the overlapped blocks will result in a global descriptor of length $d \times 31$, where 31 is the total number of blocks.

⁴ Only $4 \times 4 = 16$ spatial blocks derived from the original image were used.

Table 1. The AUC results obtained from using different local block descriptors. Number in parenthesis represents the feature dimension of each local descriptor.

Pa./Feature	M,S(2)	I(32)	O(32)	$LBP_{8,1}^{u2}(59)$	$LBP_{8,1}^{pca}(32)$	$LBP_{8,1}^{pca}(59)$	$LBP_{8,1}(256)$
Normal M.	0.965	0.970	0.983	0.991	0.987	0.991	0.931
ME	0.951	0.963	0.958	0.965	0.962	0.962	0.845
MH	0.714	0.826	0.845	0.901	0.894	0.879	0.774
AMD	0.784	0.824	0.857	0.867	0.888	0.885	0.693
Ave. of all	0.854	0.895	0.911	0.931	0.933	0.929	0.811
Ave. of MH,AMD	0.749	0.825	0.851	0.884	0.891	0.882	0.734

Table 2. The AUC results obtained from employing different global frameworks with $LBP_{8,1}^{pca}(32)$ local descriptors. “+O”: add the overlapped blocks.

Pa./Framework	MSSP	SP	SL	MSSP+O	SP+O	SL+O
Normal M.	0.985	0.979	0.984	0.987	0.984	0.987
ME	0.957	0.949	0.951	0.962	0.960	0.961
MH	0.891	0.880	0.872	0.894	0.895	0.893
AMD	0.888	0.826	0.826	0.888	0.849	0.843
Ave. of all	0.930	0.909	0.908	0.933	0.922	0.921

4 Conclusion and Future work

In this paper, we propose an effective approach to diagnose multiple macular pathologies in retinal OCT images. First, we align the slices to reduce their appearance variations. Then, we construct our global image descriptor by using multi-scale spatial pyramid (MSSP), combined with dimension-reduced LBP histogram based on PCA as the local descriptors. This approach encodes both the geometry and textures of the retina in a principled way. A binary non-linear SVM classifier is trained for each pathology to identify its presence. We evaluate our approach by comparing its performance with that of other popular global representation and local descriptors. Our results demonstrate the effectiveness and validity of the proposed approach.

There are several future directions we plan to explore. First, we will investigate the effects of including other complimentary features, e.g., encoding shape features from the edge image, so that more discriminative representation can be constructed. Second, we would like to apply a similar method to automatically localize the slice of the anatomic landmark (e.g. fovea) in the 3D scan. Third, we plan to extend our approach to additional retinal pathologies.

References

1. Schuman, J.S., Puliafito, C.A., Fujimoto, J.G.: *Optical Coherence Tomography of Ocular Diseases*, 2nd edn. (2004)
2. Garvin, M.K., Abràmoff, M.D., Kardon, R., Russell, S.R., Wu, X., Sonka, M.: Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-D graph search. *TMI* 27(10), 1495 (2008)
3. Barnum, P., Chen, M., Ishikawa, H., Wollstein, G., Schuman, J.: Local quality assessment for optical coherence tomography. In: *ISBI* (2008)
4. Klein, R., Klein, B.E.K., Linton, K.L.P., DeMets, D.L.: The Beaver Dam Eye Study: The relation of age-related maculopathy to smoking. *American Journal of Epidemiology* 137(2), 190–200 (1993)
5. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. on Pattern Analysis and Machine Intelligence(TPAMI)* 24(7), 971 (2002)
6. Chang, C.C., Lin, C.J.: *LIBSVM: a library for support vector machines* (2001)
7. Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: *CVPR* (2006)
8. Wu, J., Rehg, J.M.: Where am I: Place instance and category recognition using spatial PACT. In: *CVPR* (2008)
9. Ahonen, T., Hadid, A., Pietikäinen, M.: Face description with local binary patterns: Application to face recognition. *TPAMI* 28(12), 2037 (2006)
10. Oliver, A., Lladó, X., Freixenet, J., Martí, J.: False positive reduction in mammographic mass detection using local binary patterns. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part I. LNCS*, vol. 4791, pp. 286–293. Springer, Heidelberg (2007)
11. Sørensen, L., Shaker, S.B., de Bruijne, M.: newblock Texture classification in lung CT using local binary patterns. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part I. LNCS*, vol. 5241, pp. 934–941. Springer, Heidelberg (2008)
12. Liao, S., Zhu, X., Lei, Z., Zhang, L., Li, S.Z.: Learning multi-scale block local binary patterns for face recognition. In: *Int'l. Conf. on Biometrics* (2008)
13. Freeman, W.T., Roth, M.: Orientation histogram for hand gesture recognition. In: *Intl. Workshop on Automatic Face and Gesture Recognition* (1995)

Non-invasive Image-Based Approach for Early Detection of Acute Renal Rejection

Fahmi Khalifa¹, Ayman El-Baz^{1,*},
Georgy Gimel'farb², and Mohammed Abu El-Ghar³

¹Bioimaging Laboratory, Bioengineering Department,
University of Louisville, Louisville, KY, USA

Tel.: (502)-852-5092

aselba01@louisville.edu

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

³ Urology and Nephrology Department, University of Mansoura, Mansoura, Egypt

Abstract. A promising approach for the automatic classification of normal and acute rejection transplants from Dynamic Contrast Enhanced Magnetic Resonance Imaging (DCE-MRI) is proposed. The proposed approach consists of three main steps. The first step segments the kidney from the surrounding abdominal tissues by a level-set based deformable model with a speed function that accounts for a learned spatially variant statistical shape prior, 1st-order visual appearance descriptors of the contour interior and exterior (associated with the object and background, respectively), and a spatially invariant 2nd-order homogeneity descriptor. In the second step, to handle local object deformations due to kidney motion caused by patient breathing, we proposed a new nonrigid approach to align the object by solving Laplace's equation between closed equispaced contours (iso-contours) of the reference and target objects. Finally, the perfusion curves that show the transportation of the contrast agent into the tissue are obtained from the segmented kidneys and used in the classification of normal and acute rejection transplants. Applications of the proposed approach yield promising results that would, in the near future, replace the use of current technologies such as nuclear imaging and ultrasonography, which are not specific enough to determine the type of kidney dysfunction.

1 Introduction

In the United States, approximately 12,000 renal transplants are performed annually [1], and considering the limited supply of donor organs, every effort is made to salvage the transplanted kidney [2]. However, acute rejection - the immunological response of the human immune system to the foreign kidney - is the most important cause of graft failure after renal transplantation, and the differential diagnosis of acute transplant dysfunction remains a difficult clinical problem. Recently, a fairly new imaging technique, Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI), has gained a considerable attention

* Corresponding author.

for detecting acute renal rejection due to its ability to yield anatomical information. However, developing a noninvasive CAD system from DCE-MRI still a challenging problem due to the following reasons: (i) the spatial resolution of the dynamic MR images is low due to fast scanning, (ii) the images suffer from the motion induced by the breathing patient which necessitates advanced registration techniques, and (iii) the intensity of the kidney changes non-uniformly as the contrast agent perfuse into the cortex which complicates the segmentation procedures.

To the best of our knowledge, a limited work on the dynamic MRI to overcome the problems of registration and segmentation has been done. For the registration problem, Gerig et al. [3] proposed using Hough transform to register the edges in an image to the edges of a mask and Giele [4] introduced a phase difference movement detection method to correct kidney displacements. Both of these studies required building a mask manually by drawing the kidney contour on a 2D DCE-MRI image, followed by the registration of the time frames to this mask.

For the segmentation problem, Boykov et al. [5] presented the use of graph cuts using Markov models, where the energy is minimized depending on the manually exerted seed points. Giele [4] used image subtraction to obtain a mask and closed the possible gaps by using hull function.

Following these studies, a multi-step registration approach was introduced by Sun et al. [6]. Initially, the edges are aligned using an image gradient based similarity measure considering only translational motion. Once roughly aligned, a high-contrast image is subtracted from a pre-contrast image to obtain a kidney contour, which is then propagated over the other frames searching for the rigid registration parameters. For the segmentation of the cortex and medulla, a level sets approach was used.

2 Methods

In this paper we introduce a new and automated approach (i) to segment the kidney from its background and (ii) to correct for the motion artifacts caused by breathing and patient motion. The steps of the proposed CAD system is illustrated in Fig. 1, with details outlined below.

2.1 Segmentation

Geometric level-set based deformable models are popular and powerful tools in segmenting medical images because of the flexibility and independence of parameterizations of an evolving contour on the xy -plane [7]. The object-background boundary at each moment t is represented by a zero level $\phi_t(x, y) = 0$ of an implicit level-set function – a distance map $\phi_t(x, y)$ of the signed minimum Euclidean distances from every point (x, y) to the boundary (negative for interior and positive for exterior points). The distance map is evolved iteratively as [8]:

$$\phi_{n+1}(x, y) = \phi_n(x, y) - \tau V_n(x, y) |\nabla \phi_n(x, y)| \quad (1)$$

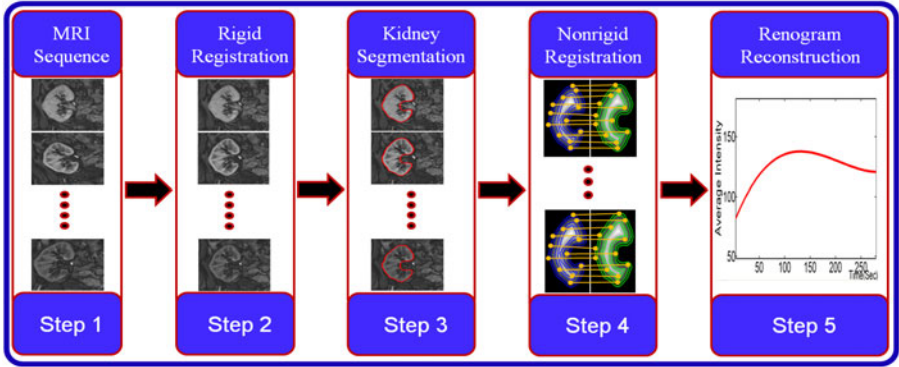


Fig. 1. The proposed CAD system

where n denotes the time instants $t = n\tau$ with step $\tau > 0$; $V_n(x, y)$ is a speed function guiding the evolution, and $\nabla\phi_n = [\frac{\partial\phi_n}{\partial x}, \frac{\partial\phi_n}{\partial y}]$ is the gradient of $\phi_n(x, y)$.

Unfortunately, conventional speed functions accounting for image intensities (e.g., [9]), object edges (e.g., [10]), an intensity gradient vector flow (e.g., [11]), etc., are unsuccessful on very noisy images with low object-background intensity gradients. Additional prior shape knowledge (e.g., [12]) results in more effective segmentation. But, speed functions based solely on the shape priors lead to large segmentation errors under discontinuous object boundaries, large image noise, and image inhomogeneities.

To overcome these limitations, a new speed function (introduced in our recent work [13]) combining the mean curvature of an evolving contour with a spatially variant joint Markov- Gibbs Random Field (MGRF) model of a gray scale image, \mathbf{g} , and its binary object-background map, \mathbf{m} , is used to control the evolution magnitude and direction at every step. The speed function in Eq.(1) is : $V(x, y) = \vartheta(x, y)\kappa$; where κ is the mean contour curvature and $\vartheta(x, y)$ specifies the evolution magnitude and direction [13]:

$$\vartheta(x, y) = \begin{cases} -P_{1:x,y} & \text{if } P_{1:x,y} > P_{0:x,y} \\ P_{0:x,y} & \text{otherwise} \end{cases} \quad (2)$$

where

$$\begin{aligned} P_{1:x,y} &= \frac{W_{1:x,y}}{W_{1:x,y} + W_{0:x,y}}; P_{0:x,y} = 1 - P_{1:x,y} \\ W_{1:x,y} &= p(q|1)p_{h:x,y}(1)p_{s:x,y}(1) \\ W_{0:x,y} &= p(q|0)(1 - p_{h:x,y}(1))(1 - p_{s:x,y}(1)) \end{aligned}$$

and $p(q|1)$ and $p(q|0)$ are the object and background marginal densities obtained by the LCDG model [14], respectively and q is the gray level value at location (x, y) ; $p_{h:x,y}(1)$ is the probability of transition to the object label in the Markov Random Filed (MRF) Potts model at (x, y) at the current evolution step; and

$p_{S:x,y}(1)$ and $p_{S:x,y}(0) = 1 - p_{S:x,y}(1)$ are the empirical object-background probabilities for the manually segmented training images co-aligned by a rigid 2D transformation maximizing their mutual information [15], respectively. (please see [13] for more comprehensive details).

A step by step illustration of estimated densities using the LCDG model [14] is illustrated in Fig. 2 and some segmentation results for different MRI kidney images and their associated errors are shown in Fig. 3. Table 2 presents segmentation results for the MRI kidney images for ten of the subjects employed in this study (sample of the testing data).

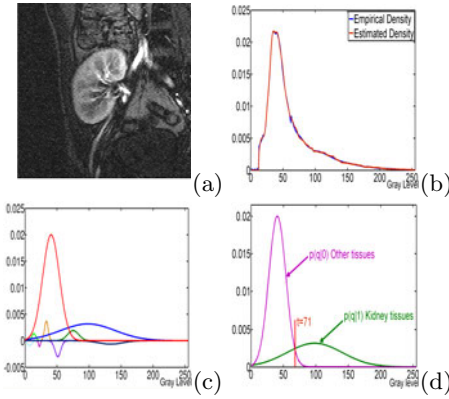


Fig. 2. (a) Typical MR kidney image, its gray level density estimation (the final estimated density (in red) for the empirical density (in blue)), (c) LCDG components, and (d) the final marginal estimated densities using the LCDG model [14](d).

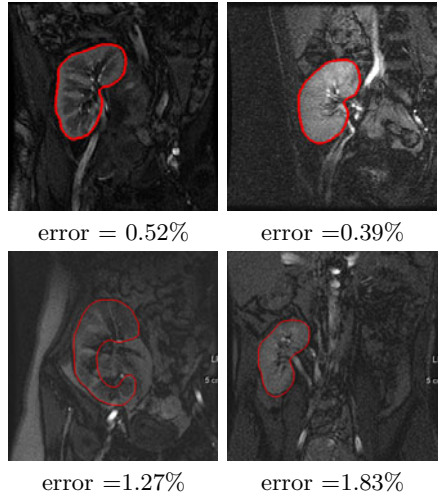


Fig. 3. Segmentation results using the proposed approach. Errors are w.r.t the radiologist segmentation.

2.2 Nonrigid Registration

The registration is the main issue in DCE-MRI sequences. This is due to patient & breathing movements. Although a tremendous number of nonrigid image registration techniques (e.g., [16][17][18]) have been developed, more robust, efficient, and sophisticated registration techniques are required. The centerpiece of this paper is the application of Laplace's equation in a nonrigid registration approach to find the corresponding pixel pairs between target and reference objects. Mathematically, Laplace's equation is a second-order partial differential equation for a scalar field Ψ that is enclosed between two surfaces (e.g., A and B in Fig. 4) and takes the form:

$$\nabla^2 \Psi = \frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} = 0 \quad (3)$$

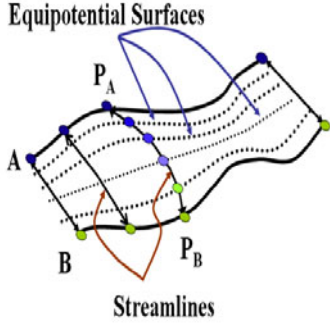


Fig. 4. Two dimensional example of the Laplace method

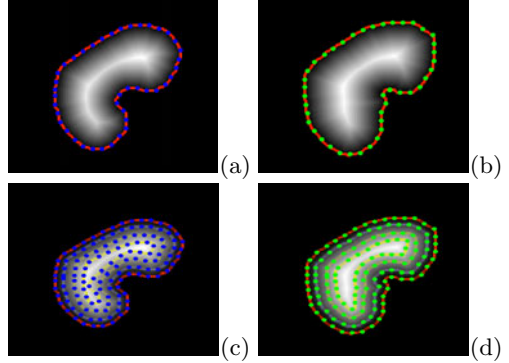


Fig. 5. The distance maps (a, b) and the isocontours (c, d) of the two kidneys

Solving Laplace's equation between the enclosed surfaces results in intermediate equipotential surfaces (dashed lines in Fig. 4) and field lines (streamlines) that connect both surfaces A and B . These field lines are defined as being everywhere orthogonal to all equipotential surfaces (e.g., the line connecting the points P_A and P_B in Fig. 4). Nonrigid registration based on Laplace's equation effectively overcomes patient movement & breathing problems and consists of the following steps:

1. First, generate the distance maps inside the kidney regions as shown in Fig. 5 (a, b).
2. Second, use these distance maps to generate equispaced separated isocontours as shown in Fig. 5 (c, d).
3. Third, solve Eq. (3) between respective iso-contours to find the corresponding pixel pairs in both target and reference objects.

In general, the number of iso-contours depend on the accuracy and the speed which the user needs to achieve. An illustration of the power of Laplace's equation to find the corresponding pixel pairs is shown in Fig. 6. The first case employs no cross-overlap between the target and reference images' edges (Fig. 6 (a, b)), and the second case employs a cross-overlap between the images' edges (Fig. 6 (c, d)). Figure. 7 shows a two dimensional illustration of how the solution of Laplace's equation is used to determine the corresponding pixel pairs. A potential Ψ is found in the interior area enclosed by both surfaces A and B in Fig. 7 by solving Eq. (3) such that Ψ is maximum at A and minimum at B . Then the electric field vectors, $E_x = \frac{\partial \Psi}{\partial x}$, and $E_y = \frac{\partial \Psi}{\partial y}$, in the interior area between A and B are used to find the corresponding pixel pairs as shown in Fig. 7. An enlarged portion of Fig. 7 (a) around a streamline is shown in Fig. 7 (b), and a checkerboard visualization for one kidney object before and after applying the nonrigid registration is shown in Fig. 8 (a) and (b), respectively.

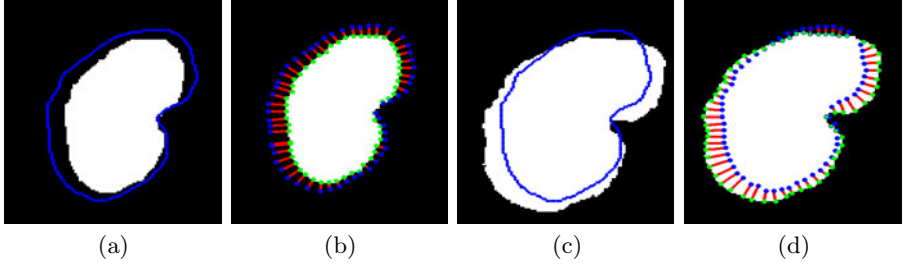


Fig. 6. Rotated kidney objects (a, c) with respect to the reference image edges (in blue), and the solution of Laplace's equation (b, d) to find the corresponding pixels.

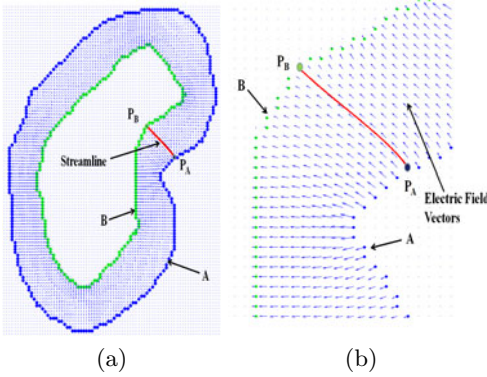


Fig. 7. Electric field vectors in the area between two enclosed regions A (in blue) and B (in green) (a), an enlarged section around the indicated streamline (b)

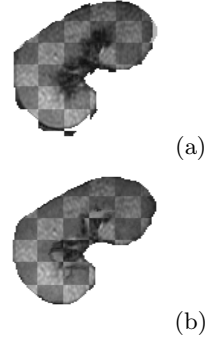


Fig. 8. Checkerboard visualization before (a), and after (b) registration

Table 1. Accuracy of the proposed registration algorithm (all units are in mm)

	Simulated Displacement		
	Small	Moderate	Large
Maximum Simulated Displacement	1.70	0.70	19.99
Average Simulated Displacement	0.60	2.30	9.10
St. dev. of Simulated Displacement	0.40	10.80	1.10
Alignment Error			
Maximum Alignment Error	0.50	1.21	1.71
Average Alignment Error	0.30	0.90	1.15
St. dev. of Alignment Error	0.21	0.31	1.13

Validation of the Proposed Local-Registration Approach: To validate the proposed local registration, we simulated local deformations on the real DCE-MRI data set using the free form deformation (FFD) approach [16] that simulates local displacements with 2D cubic B-splines. To measure the registration accuracy, three different types of deformation fields, namely, with small, moderate, and large deformations indicated in Table II, were generated with the FFD. Our registration model has been applied to each type of deformation, and the accuracy has been quantitatively assessed by comparing the simulated and recovered pixel displacements.

3 Results and Conclusions

The ultimate goal of the proposed algorithms is to successfully construct the renograms (mean intensity signal curves) from the DCE-MRI sequences, showing the behavior of the kidney as the contrast agent perfuses into the transplant. In acute rejection patients, the DCE-MRI images show a delayed perfusion pattern and a reduced cortical enhancement. We tested the above algorithms on twenty-six patients, four of whom are shown in Fig. 9. The normal patient shows the expected abrupt increase to the higher signal intensities and the valley with a small slope. The acute rejection patients show a delay in reaching their peak signal intensities. From these observations, we have been able to conclude that the time-to-peak signal intensity, and the slope between the first peak and the signal measured from the last image in the sequence are the major two features in the renograms of the segmented kidney for classification.

To distinguish between normal and acute renal rejection, we used a k_n -nearest neighbor classifier to learn statistical characteristics from the renogram of the training sets of both normal and acute renal rejection. In our approach we used 50% of the data for the training and the other 50% for testing. For testing data, k_n -nearest neighbor based classification succeeded in classifying 12 out of 13

Table 2. Accuracy of our segmentation algorithm

Subject	Error%			
	Min	Max	Mean	St. dev.
First	0.62	1.20	1.03	0.98
Second	0.88	1.81	1.14	1.05
Third	0.76	2.12	1.35	1.20
Forth	0.90	2.50	1.82	1.54
Fifth	1.01	2.76	2.10	1.78
Sixth	1.20	3.30	2.23	2.01
Seventh	0.56	1.05	0.82	0.90
Eighth	0.30	0.96	0.78	0.85
Ninth	0.23	0.53	0.38	0.43
Tenth	1.10	1.65	1.22	0.65

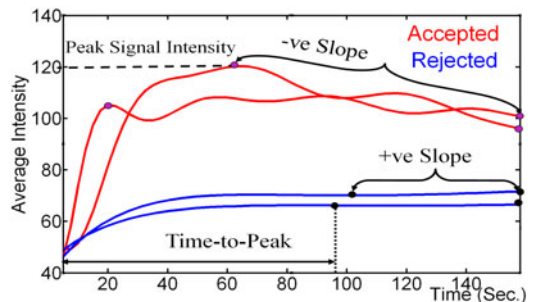


Fig. 9. Normalized average intensity from four subjects w.r.t. scan number. Subjects 1, and 2 (in red) are normal, and Subjects 3, and 4 (in blue) are acute rejection.

correctly (92.31%). For the training data the k_n -nearest neighbor based classification classifies all of them correctly, so the over all accuracy of the proposed approach is 96.15%.

In this paper, we presented a framework for the detection of acute renal rejection from Dynamic Contrast Enhanced Magnetic Resonance Images (DCE-MRI) that includes segmentation of kidneys from abdomen images, nonrigid registration, and k_n -Nearest Neighbor based classification. For segmentation of kidneys, we introduced a new geometric deformable model that evolves with the gray-level information of a given abdomen image, spatial information, and the shape information. Following segmentation, we introduced a nonrigid registration approach that deforms the kidney object on iso-contours instead of a square lattice, which provides more degrees of freedom to obtain accurate deformation. After nonrigid registration, the renogram of the segmented kidney for the whole sequence of a patient is calculated. The features extracted from these renograms undergo k_n -nearest neighbor based classification to understand if the transplanted kidney is undergoing acute rejection or if it is functioning normally. Applications of the proposed approach yield promising results that would, in the near future, replace the use of current technologies to determine the type of kidney dysfunction.

References

1. U.S. Department of Health and Human Services. Annual report of the U.S. scientific registry of transplant recipients and the organ procurement and transplantation network: transplant data: 1990-1999. Bureau of Health Resources Department, Richmond, VA (2000)
2. Neimatallah, M., et al.: Magnetic resonance imaging in renal transplantation. *J. Magn. Reson. Imaging* 10(3), 357–368 (1999)
3. Gerig, G., et al.: Semiautomated ROI analysis in dynamic MRI studies: Part I: image analysis tools for automatic correction of organ displacements. *IEEE Trans. Image Process.* 11(2), 221–232 (1992)
4. Giele, E.: Computer methods for semi-automatic MR renogram determination. Ph.D. dissertation, Dept. Elec. Eng., Eindhoven Univ. of Techno., Eindhoven (2002)
5. Boykov, Y., et al.: Segmentation of dynamic N-D data sets via graph cuts using Markov models. In: Niessen, W.J., Viergever, M.A. (eds.) MICCAI 2001. LNCS, vol. 2208, pp. 1058–1066. Springer, Heidelberg (2001)
6. Sun, Y., et al.: Integrated registration of dynamic renal perfusion MR images. In: *IEEE Int. Conf. Image Process. (ICIP 2004)*, Singapore, pp. 1923–1926 (October 2004)
7. Han, X., Xu, C., Prince, J.L.: A Topology preserving level set method for geometric deformable models. *IEEE Trans. Pattern Anal. Mach. Intell.* 25(6), 61–79 (2009)
8. Osher, S., Fedkiw, R.: *Level Set Methods and Dynamic Implicit Surfaces*. Applied Mathematical Sciences, vol. 153. Springer, Heidelberg (2006)
9. Carson, C., Belongie, S., Greenspan, H., Blobworld, J.M.: Image segmentation using expectation-maximization and its application to image querying. *IEEE Trans. Pattern Anal. and Mach. Intell.* 24(8), 1026–1038 (2002)
10. Ma, W.Y., Manjunath, B.S.: Edgeflow technique for boundary detection and image segmentation. *IEEE Trans. Image Process.* 9(8), 1375–1388 (2000)

11. Xu, C., Prince, J.L.: Snakes, shapes, and gradient vector flow. *IEEE Trans. Image Process.* 7(3), 359–369 (1998)
12. Chen, K., et al.: Using prior shapes in geometric active contours in a variational framework. *Int. J. Comput. Vision* 50(3), 315–328 (2002)
13. Khalifa, F., El-Baz, A., Gimel'farb, G., Ousephand, R., Abu El-Ghar, M.: Shape-Appearance guided level-set deformable model for image segmentation. In: *IEEE Int. Conf. Pattern Recog. (ICPR 2010)*, Istanbul, Turkey (August 2010) (Currently in Press)
14. El-Baz, A., Gimel'farb, G.: EM Based approximation of empirical distributions with linear combinations of discrete Gaussians. In: *Proc. IEEE Int. Conf. Image Process. (ICIP 2007)*, San Antonio, Texas, USA, September 2007, vol. 4, pp. 373–376 (2007)
15. Viola, P., Wells III, W.M.: Alignment by maximization of mutual information. *Int. J. Comput. Vision* 24(2), 137–154 (1997)
16. Rueckert, D., et al.: Nonrigid registration using free-form deformations: Application to breast MR images. *IEEE Trans. Med. Imag.* 18(8), 712–721 (1999)
17. Studholme, C., Constable, R.T., Duncan, J.: Accurate alignment of functional EPI data to anatomical MRI using a physics-based distortion model. *IEEE Trans. Med. Imag.* 19(11), 1115–1127 (2000)
18. Avants, B., Gee, J.C.: Geodesic estimation for large deformation anatomical shape averaging and interpolation. *NeuroImage* 23(1), S139–S150 (2004)

Hierarchical Segmentation and Identification of Thoracic Vertebra Using Learning-Based Edge Detection and Coarse-to-Fine Deformable Model

Jun Ma, Le Lu, Yiqiang Zhan, Xiang Zhou,
Marcos Salganicoff, and Arun Krishnan

CAD & Knowledge Solutions, Siemens Healthcare, Malvern, PA 19355
Center for Imaging Science, Johns Hopkins University, Baltimore, MD 21218

Abstract. Precise segmentation and identification of thoracic vertebrae is important for many medical imaging applications whereas it remains challenging due to vertebra's complex shape and varied neighboring structures. In this paper, a new method based on learned bone-structure edge detectors and a coarse-to-fine deformable surface model is proposed to segment and identify vertebrae in 3D CT thoracic images. In the training stage, a discriminative classifier for object-specific edge detection is trained using steerable features and statistical shape models for 12 thoracic vertebrae are also learned. In the run-time, we design a new coarse-to-fine, two-stage segmentation strategy: subregions of a vertebra first deforms together as a group; then vertebra mesh vertices in a smaller neighborhood move group-wise, to progressively drive the deformable model towards edge response maps by optimizing a probability cost function. In this manner, the smoothness and topology of vertebra's shapes are guaranteed. This algorithm performs successfully with reliable mean point-to-surface errors 0.95 ± 0.91 mm on 40 volumes. Consequently a vertebra identification scheme is also proposed via mean surface meshes matching. We achieve a success rate of 73.1% using a single vertebra, and over 95% for 8 or more vertebra which is comparable or slightly better than state-of-the-art [1].

1 Introduction

A precise vertebra segmentation and identification method is in high demand due to its important impact in many orthopaedic, neurological and oncological applications. In this paper, we focus on thoracic vertebra where accurate segmentation and identification of them can directly eliminate false findings on lung nodules in computer aided diagnosis system [2]. However, this task remains challenging due to vertebra's complexity, i.e., within-class shape variation and different neighboring structures.

Several methods have been reported addressing segmentation and/or identification of vertebra under different modalities, e.g., magnetic resonance imaging (MRI) or computed tomography (CT). Yao et al. [3] present a method to automatically extract and partition the spinal cord in CT images, and a surface-based

registration approach for automatic lumbar vertebra identification is described in [4], where no identification was carried out in both work. Recently, Klinder et al. [1] propose a model-based solution for vertebra detection, segmentation and identification in CT images. They achieved very competitive identification rates of $> 70\%$ for a single vertebra and 100% for 16 or more vertebrae. However, their identification algorithm is based on vertebra appearance model (i.e., an averaged volume block) spatial registration and matching which is very computationally consuming (20 ~ 30 minutes).

In this paper, we present a new automatic vertebra segmentation and identification method. Although this work mainly focuses on thoracic vertebra (for potential lung applications), our approach can be easily extended to cervical and lumbar vertebrae. The main contributions of this paper are summarized as follows. First, we introduce a learning based bone structure edge detection algorithm, including efficient and effective gradient steerable features and robust training data sampling. Second, a hierarchical, coarse-to-fine deformable surface based segmentation method is proposed based on the response maps from the learned edge detector, followed with an efficient vertebra identification method using mean shapes. Finally, the promising results of segmentation and identification are presented, compared with the state-of-the-art [1].

2 Method

Due to complex neighboring structures around vertebra and imaging noise, common edge detectors, e.g., Canny operator, often produce leaking and spurious edge. To achieve robust edge detection, we develop a learning-based object specific edge detection algorithm, similar to semantic object-level boundary lineation in natural images [5].

2.1 Supervised Bone Edge Detection

We manually segmented 12 thoracic vertebrae from 20 CT volumes for training, and generated corresponding triangulated surfaces using Marching Cube algorithm, with about 10,000 triangular faces per vertebra model. It is observed that along the normal direction of the bone boundary, the intensity values roughly form a ridge pattern. Our new set of steerable features is designed to describe the characteristics of boundary appearance, which make it feasible for statistical training.

Gradient steerable features: For each triangle face of the surface mesh, we take 5 sampling points (called a **sampling parcel**) along the face normal direction with one voxel interval. Specially, given x a point on the normal line and n the unit normal vector, the sampling parcel associated with x is

$$\mathcal{P}(x) = \{x - 2n, x - n, x, x + n, x + 2n\}$$

For each of the 5 sampling points we compute three features: intensity I , projections of gradient onto the normal direction $\nabla_1 I \cdot n, \nabla_2 I \cdot n$, where $\nabla_1 I$ and $\nabla_2 I$

are gradient vectors computed using derivative of Gaussian with two different kernel scales. Totally, the **feature vector** of a point x , denoted by $\mathcal{F}(x)$, has 15 elements:

$$\mathcal{F}(x) = \{I(y), \nabla_1 I(y) \cdot n, \nabla_2 I(y) \cdot n | y \in \mathcal{P}(x)\}$$

Fig. 1 illustrates the sampling parcel and its associated features. Our steerable features are indeed oriented-gradient pattern descriptor with easy computation.

Vertebra edge detector training: The training samples of positive and negative boundary voxels are obtained from manually segmented vertebra mesh as below. For a triangle face center c , we define the boundary parcel as

$$\mathcal{P}(c) = \{c - 2n, c - n, c, c + n, c + 2n\}$$

interior parcel as

$$\mathcal{P}(c - 3n) = \{c - 5n, c - 4n, c - 3n, c - 2n, c - n\}$$

and exterior parcel as

$$\mathcal{P}(c + 3n) = \{c + n, c + 2n, c + 3n, c + 4n, c + 5n\}$$

That is, the interior parcel is 3 voxels away backward from boundary parcel while exterior parcel is the 3 voxels forward, where 3 is adjustable. The corresponding feature vectors $\mathcal{F}(c), \mathcal{F}(c - 3n), \mathcal{F}(c + 3n)$ can be also computed. Then we label $\mathcal{F}(c)$ as positive class (i.e., boundary), and assign both $\mathcal{F}(c - 3n)$ and $\mathcal{F}(c + 3n)$ as negative class (i.e., non-boundary), as Fig. 2 (left). Thus, each triangle face provides one positive data and two negative data. Given one vertebra surface mesh with about 10,000 faces, sufficient and adequate training feature vectors are obtained. Note that a single and unified bony edge detector will be learned for all 12 thoracic vertebrae. Compared with implicit, object “inside-outside” learning [6], our boundary/non-boundary delineation strategy directly focuses on modeling the runtime boundary localization process (i.e., explicitly moving towards classified boundary positives), and is expected to have higher precision.

The feature vectors depend on the normal direction of triangle faces so that the edge detector is sensitive to the initialization of the surface template. In our experimental setup, the surface model is first roughly registered with images by automatic detection [7,8] or manual alignment, thus the normal direction of the surface model can not perfectly coincide with the true bony normal. To make the detector more robust to mis-alignment errors

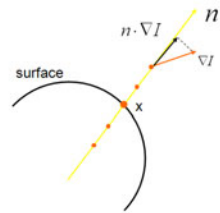


Fig. 1. Steerable features of x . Five red dots indicate sampling parcel associated with x . Yellow arrow indicates the normal direction. Red and black arrows indicate gradient ∇I and projection $\nabla I \cdot n$.

¹ The boundary has to be further inferred from the transition of (object) internal positives and external negatives [6] which may not be trivial.

and the later deformable model convergent, it is important that we synthesize some “noisy” training samples by stress testing. Particularly, we add some random disturbances to the orientations and scales of the template model so that the template surface model does not accurately overlap with the manual segmentation. Considering a similarity transform, a random number between 0.9 and 1.1 for each of the three scales, and a random angle between $-\frac{\pi}{10}$ and $\frac{\pi}{10}$ for each of the three orientation angles are used. The true boundary parcels, as well as interior and exterior parcels are defined using ground truth positions but with disturbed template surface normals. Refer to Fig. 2 (*middle*) for an illustrative example. Their corresponding feature vectors are consequently calculated (with the disturbed face normals) and added into our training sets. The random disturbance process is repeated 10 times for each training mesh to guarantee we get enough noisy samples. We then train an Linear or Quartic Discriminant (LDA, QDA) classifier based on the combined non-disturbed and disturbed feature vectors. Both LDA and QDA are evaluated and we find that LDA yields more robust results. The experiment results are computed with LDA. Finally, given a voxel x and its feature vector $\mathcal{F}(x)$, our classifier will assign a value $\mathcal{L}(x) \in [0, 1.0]$ which indicates the likelihood of x being boundary point.

2.2 Segmentation: Coarse-to-Fine Deformation

The main idea of segmentation is to deform the surface template mesh towards boundary points detected by the learned edge detector. After the surface template is initially positioned into a new volume, (The template can be initialized using similar strategies as marginal space learning [78]) edge detector calculates the edge likelihoods $\mathcal{L}(x)$ for voxels along the normal directions of all mesh faces, where a response map can be generated. As shown in Fig. 2 (**Right**), this response map is informative but unavoidably noisy. To guarantee the surface shape topology and smoothness during deformation/segmentation, we propose a hierarchical deformation scheme of first performing deformation of subregions;

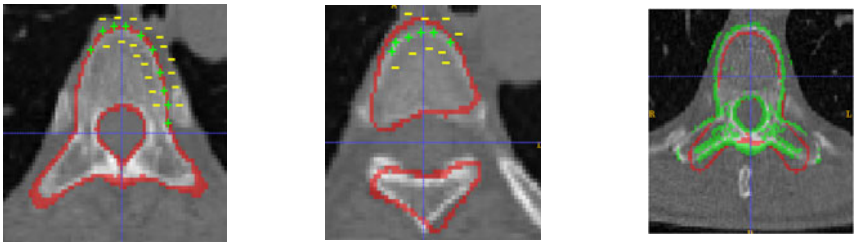


Fig. 2. Left: Surface template perfectly aligned with the true boundary. **Middle:** Disturbed Surface template overlapped within the volume. Green plus and Yellow minus signs are positive or negative sample samples, respectively. **Right:** Response map of vertebra edge detection in the section view of 3D CT volume. The red curve indicates the template surface while the green dots are the voxels classified as boundary points with likelihood values > 0.8 .

then performing patch-wise deformation, i.e., points in the same neighborhood move together.

Deformation of subregions: We manually divide the surface mesh into 12 subregions, as indicated by Fig. 3. In order to maintain the shape of these subregions, a similarity transformation to each subregion is applied such that the total response of edge detection is maximum in the transformed configuration. For a subregion S and some face center f on it, we intend to find a similarity transformation \hat{T} satisfying

$$\hat{T} = \arg \max_{T \in \mathbf{T}} \sum_{f \in S} \mathcal{L}(T(f)) \quad (1)$$

where \mathbf{T} is the set of similarity transformations T . Searching the optimal T involves the 9-dimensional parameters space of $(T_x, T_y, T_z, S_x, S_y, S_z, \theta_x, \theta_y, \theta_z)$. If we perform an exhaustive search with 5 grid steps for each parameter, then possible transformation is 5^9 which is computationally infeasible. To reduce the search space, we perform a three-stage search. First, search for (T_x, T_y, T_z) with displacement $\{-4, -2, 0, 2, 4\}$ voxels for each translation; second, with fixed $(\hat{T}_x, \hat{T}_y, \hat{T}_z)$, search for (S_x, S_y, S_z) with discretization grids of $\{0.8, 0.9, 1.0, 1.1, 1.2\}$ for each scaling; third, with fixed optimal translation and scaling, search for $(\theta_x, \theta_y, \theta_z)$ with intervals of $\{-\pi/10, -\pi/20, 0, \pi/20, \pi/10\}$ for each orientation. In this way, we need to only consider $5^3 \times 3 = 375$ possible poses and select the one with the strongest response as \hat{T} . This heuristic searching strategy turns out to be effective in capturing the true pose of subregions though it might be suboptimal. Fig. 4(a) illustrates the searching process.

After the optimal similarity transformation is found for each subregion, a smooth deformation of the whole surface can be obtained using simple Gaussian smoothing. Let S_1, S_2, \dots, S_{12} denote the twelve subregions, and T_1, T_2, \dots, T_{12} be the corresponding optimal transform. Denote v an arbitrary vertex in the template surface and u a vertex in a certain subregion. Then the new position of v is

$$v' = v + \lambda \sum_{i=1}^{12} \sum_{w \in S_i} (T_i(w) - w) K(w - v)$$

where $K(x) = e^{-\frac{x^2}{2\sigma^2}}$ is the Gaussian kernel and λ is a regulation parameter. Fig. 4(b) shows the result of “deformation of subregion” stage. One can see the surface mesh is more closely aligned with the true boundary through “articulated” similarity moves, although in several areas, the surface mesh still has a certain distance from the true boundary. This will be solved by the finer-scale deformation strategy described below.

Deformation of patches: After deforming the subregions, the surface mesh is approximately overlapped with the vertebra’s boundary in CT volume. Next, we

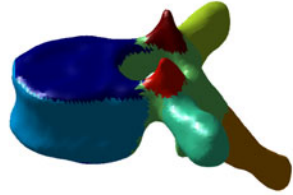


Fig. 3. Subregions of the surface. Subregions are illustrated in different colors.

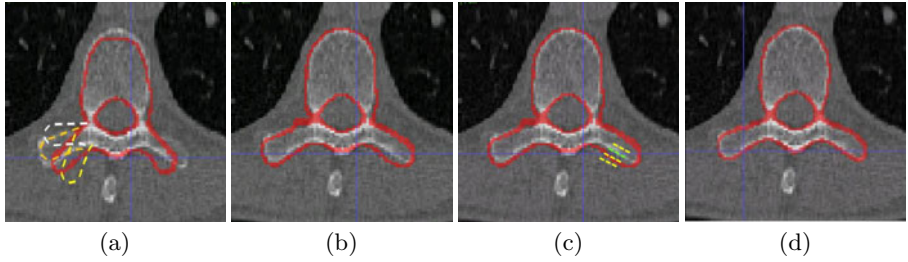


Fig. 4. (a,b) Deformation of left transverse process. (a) Dot curves indicate searching of transformations of this subregion. In this case, the orange curve indicates the optimal position. (b) Subregion deformation result. (c,d) Deformation of patches. (c) Dot curve indicate displacing a patch in the normal direction for search of strongest response. The green dots indicates the optimal displacement. (d) Patch deformation result.

perform deformation on local neighborhoods of 200 patches divided from each vertebra mesh surface (each patch may contain 50 faces approximately). For each patch (denoted as PT), we compute its mean normal by this formula:

$$\bar{n} = \frac{1}{N} \sum_{f \in PT} n(f) \quad (2)$$

where f is a face in the patch and $n(f)$ is the unit normal of the face. Then the patch is moved along its mean normal direction in search of the strongest response, that is, we optimize this term:

$$\hat{i} = \arg \max_i \sum_{f \in S} \mathcal{L}(f + i\bar{n}) \quad (3)$$

where the search range is limited in $i = -6, -5, \dots, 5, 6$. Fig. 4(c) shows the a patch is displaced along its mean normal direction in search for the boundary. After all patches find their optimal displacement, a smooth deformation of surface is again obtained by Gaussian smoothing. Fig. 4(d) shows the segmentation result of “deformation of patches” stage. Clearly, the surface mesh now can accurately capture the true boundary of the vertebra. The two-stage, coarse-to-fine deformation of surface model guarantees the accuracy of segmentation as well as the smoothness of the shapes, using articulated similarity transforms and nonrigid transform respectively.

2.3 Identification Using Mean Shapes

We applied the segmentation algorithm to 40 volumes at 1mm by 1mm by 1mm resolution, and 15 ~ 20 surface meshes are obtained per thoracic vertebra, due to missing vertebra in some volume. Vertex correspondence across meshes for each vertebra is also directly available since surface meshes are deformed by the same template. Therefore we can compute the mean vertebra shapes by

simply taking the arithmetical mean of corresponding vertices' positions. There are 12 thoracic vertebrae, namely T1, T2, ..., T12. Vertebra identification is to label a segmented vertebra to be one of the twelve. In this context, given a single vertebra subvolume, we carry out the identification process by testing which mean shape has the maximum response. Specially, we feed the 12 mean shapes to the vertebra volume one after another, and calculate the supervised edge response scores without deformation. The mean shape with the strongest response is determined as the label of this vertebra.

Let M_1, M_2, \dots, M_{12} denote the twelve mean shapes and f is an arbitrary face center in the mean shapes. One way to calculate the responses is computing the overall likelihood of boundary,

$$\hat{i} = \arg \max_i \sum_{f \in M_i} \mathcal{L}(f) \quad (4)$$

Another way is to count the number of faces with high probability to be boundary point,

$$\hat{i} = \arg \max_i \sum_{f \in M_i} \mathbf{1}_{\mathcal{L}(f) > \alpha} \quad (5)$$

where α is a threshold. We find the second method is more robust against outliers and noise, by tolerating up to $(1 - \alpha)$ portion of data being polluted or not at the correct spatial configuration, and take $\alpha = 0.8$ which is used for following experiments. We also extend the identification method to multiple vertebrae, i.e., a vertebra string. By using more context, multiple vertebrae identification is expected to have higher success rate.

3 Result

We apply our automatic segmentation algorithm to 40 volumes of thoracic scans and the evaluation is performed using four-fold cross validation. In implementation, we run the subregion deformation step multiple (m) times followed by patch-based deformation n times, where m and n are empirically optimized to be 3 and 4, respectively. The supervised edge detection is performed at each iteration to reflect the runtime vertebra mesh surface configuration. In Fig. 5, we show some segmentation examples in axial, sagittal or coronal view, for visual inspection. To quantitatively evaluate our segmentation algorithm, we use the distance of a vertex on the fitted mesh to the closest mesh point (not necessarily a vertex) of the ground truth mesh which is generated from manual segmentation. The mean point-to-surface error and the standard deviation for individual

Table 1. Mean point-to-surface error and standard deviation for individual vertebra

vertebra	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12
mean error (mm)	1.05	1.11	1.03	0.93	0.99	0.92	0.83	0.75	0.89	0.79	0.94	1.21
std deviation(mm)	0.96	0.97	1.04	1.03	1.31	0.92	0.56	0.59	0.68	0.50	0.63	1.16

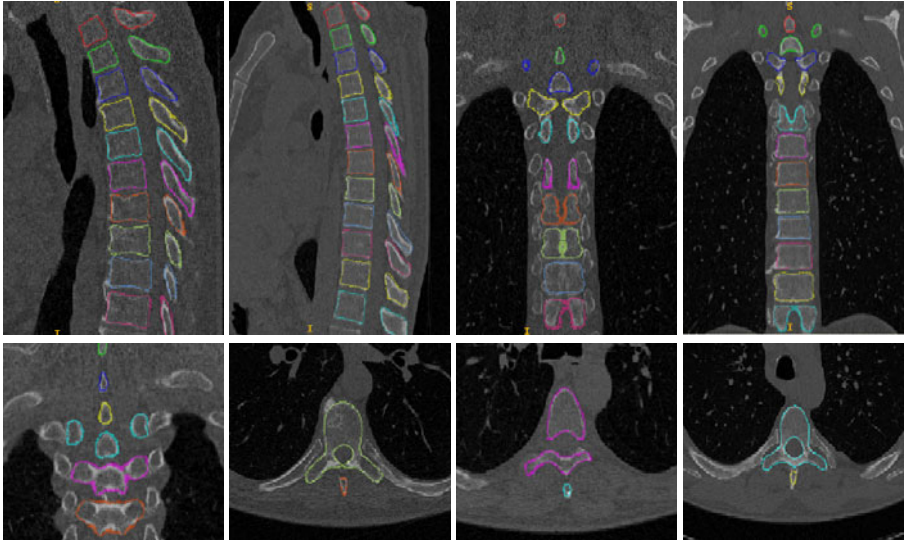


Fig. 5. Segmentation results of chosen volume in axial or sagittal or coronal view. Different vertebrae are featured in different colors.

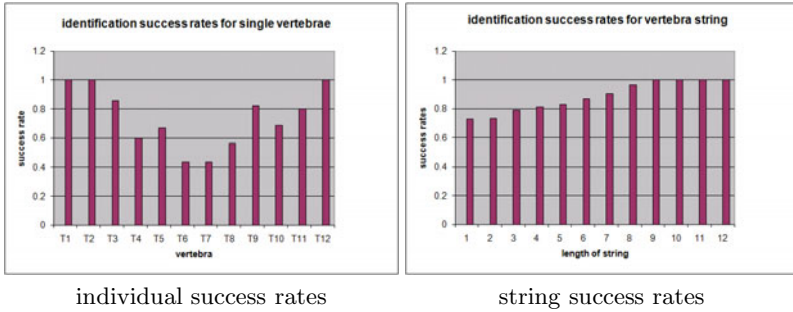


Fig. 6. Identification success rate of individual vertebra and stringed vertebrae

vertebra is shown in Table 1. Highly reliable and accurate segmentation results have been achieved, with the overall final mean error of 0.95 ± 0.91 mm. [1] reports a comparable accuracy level at 1.12 ± 1.04 mm.

For identification, we have an average success rate of 73.1% using single vertebra. This success rate also varies regarding to a specific vertebra where the rates for $T5, T6, T7, T8$ as $\leq 60\%$ are especially lower than others because these four vertebrae look alike. Furthermore, when exploiting vertebra string for identification, the success rate is improved and increases with longer string. With a string of 7 or 8 and more vertebrae, we achieve over 91% or $> 95\%$ success rates, whereas rates are $\approx 71\%$ for one vertebra, $\approx 87\%, 89\%$ for 7 or 8 vertebra strings in [1]. The success rates of individual and stringed vertebra identification

(via mean mesh shapes) are comparable or better than [1] using intensity based matching, as shown in Fig. 6.

A volumetric mean appearance model is used for vertebra identification in [1], which seems more comprehensive than our shape information alone. However we observe that in real cases, the variability of neighboring structures is quite large due to patients' pose variation. The adjacent vertebrae can be so close to each other where the boundary even can not be clearly distinguished; or, successive vertebrae are apart from each other with a large distance. Thus, the neighboring structures are not necessarily positive factors in the identification procedure. A clean shape model without surrounding structures may be of advantage and our identification results are indeed slightly better.

4 Conclusion

In this paper, a hierarchical thoracic vertebra segmentation and identification method is presented. We propose learning-based edge detectors using steerable gradient features. The segmentation applies a surface deformable model by adopting a new two-stage "coarse-to-fine" deformation scheme: first subregion based articulated similarity deformation and then nonrigid local patch deformation. The segmentation result is satisfying with point-to-surface error 0.95 ± 0.91 mm. We also use the generated mean shape model of each thoracic vertebra for identification process. Both our segmentation and identification performance is compared favorably against the state-of-the-art [1].

References

1. Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in ct images. *Medical Image Analysis* 13, 471–481 (2009)
2. Murphy, K., et al.: A large-scale evaluation of automatic pulmonary nodule detection in chest ct using local image features and k-nearest-neighbour classification. *Medical Image Analysis* 13, 757–770 (2009)
3. Yao, J., O'Connor, S., Summers, R.: Automated spinal column extraction and partitioning. In: *Proc. of IEEE ISBI*, pp. 390–393 (2006)
4. Herring, J., Dawant, B.: Automatic lumbar vertebral identification using surface-based registration. *Computers and Biomedical Research* 34(2), 629–642 (2001)
5. Dollar, P., Tu, Z., Belongie, S.: Supervised learning of edges and object boundaries. In: *CVPR* (2006)
6. Zhan, Y., Shen, D.: Deformable segmentation of 3-d ultrasound prostate images using statistical texture matching method. *IEEE Trans. on Medical Imaging* (2006)
7. Lu, L., Barbu, A., Wolf, M., Salganicoff, M., Comaniciu, D.: Simultaneous detection and registration for ileo-cecal valve detection in 3d ct colonography. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) *ECCV 2008, Part IV. LNCS*, vol. 5305, pp. 465–478. Springer, Heidelberg (2008)
8. Zheng, Y., Barbu, A., et al.: Four-chamber heart modeling and automatic segmentation for 3-d cardiac ct volumes using marginal space learning and steerable features. *IEEE Trans. Medical Imaging* 27(11), 1668–1681 (2008)

Automatic Detection and Segmentation of Axillary Lymph Nodes

Adrian Barbu¹, Michael Suehling², Xun Xu², David Liu²,
S. Kevin Zhou², and Dorin Comaniciu²

¹ Statistics Department, Florida State Univ., Tallahassee, FL 32306, USA
abarbu@stat.fsu.edu

<http://stat.fsu.edu/~abarbu>

² Siemens Corporate Research, Princeton, NJ 08540, USA

Abstract. Lymph node detection and measurement is a difficult and important part of cancer treatment. In this paper we present a robust and effective learning-based method for the automatic detection of solid lymph nodes from Computed Tomography data. The contributions of the paper are the following. First, it presents a learning based approach to lymph node detection based on Marginal Space Learning. Second, it presents an efficient MRF-based segmentation method for solid lymph nodes. Third, it presents two new sets of features, one set self-aligning to the local gradients and another set based on the segmentation result. An extensive evaluation on 101 volumes containing 362 lymph nodes shows that this method obtains a 82.3% detection rate at 1 false positive per volume, with an average running time of 5-20 seconds per volume.

1 Introduction

Lymph node (LN) analysis is a difficult task and accounts for a significant part of daily clinical work in Radiology. In particular, automatic lymph node detection and segmentation is important for cancer staging and treatment monitoring. Lymph nodes nearby primary cancer regions are routinely assessed by clinicians to monitor disease progress and effectiveness of the cancer treatment. The assessment is usually based on 3D Computed Tomography (CT) data. When the cancer treatment is successful, the lymph nodes decrease in size. Since finding the lymph nodes is time consuming and highly dependent on the observer's experience, a system for automatic lymph node detection and measurement is desired. For follow-up studies, the system could further report the size change for each major lymph node.

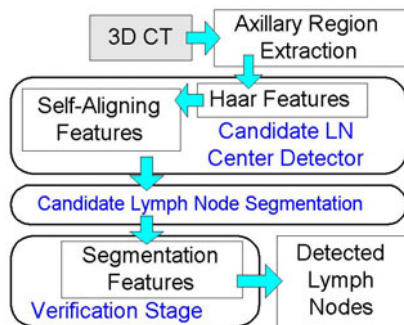


Fig. 1. Diagram of the axillary lymph node detection system

There is a limited amount of work directed to automatic lymph node detection [4,5,7]. These works target mediastinal [5], abdominal [7] and neck [4] lymph nodes while our work targets axillary lymph nodes. The axillary lymph nodes are far from airways or major vessels, so a segmentation of the vessels or airways is not necessary.

Previous work [5] relied on Hessian-based blobness measures and other filters for removing lymph node false positives. Our work comes from a learning perspective, where the appropriate features that discriminate the lymph nodes from the negatives are learned from training examples. Blobness measures could also be included in the feature set of our system.

The idea of coupling segmentation with detection has been recently proposed in the computer vision literature [8]. Our work also combines segmentation with object detection, but in a different way. First, our segmentation method produces a defined object boundary whereas [8] has a fuzzy boundary. Second, our work is oriented towards detecting 3D lymph nodes, which have a high degree of shape variability. In contrast, [8] detects 2D objects of specific shapes such as cars, cows and humans. Third, the segmentation is constructed differently in our work, using a Gaussian MRF and gradient descent as opposed to [8] where it is constructed in a greedy way from a number of patches.

Fourth, our work constructs segmentation-based features that are used to train a classifier, whereas [8] obtains a probability from the segmentation hypotheses by voting.

The diagram of the proposed solid lymph node detection and segmentation method is shown in Fig. 1. For speed and accuracy, the axillary regions are extracted automatically as described in Section 2.1. About 1500 lymph node center candidates per axillary region are generated using a two-stage detector described in Section 2.3. Each candidate is segmented as described in Section 2.4. Finally, a verification stage described in Section 2.5 gives the final result.

2 Proposed Method for Solid Lymph Node Detection and Segmentation

The proposed lymph node detection and segmentation system first detects candidate lymph node centers using a learning based approach. Each candidate is used by a segmentation module to extract a candidate lymph node boundary. A learning-based verification stage uses features obtained from the data and the extracted boundary to score the candidates and keep only the best ones.

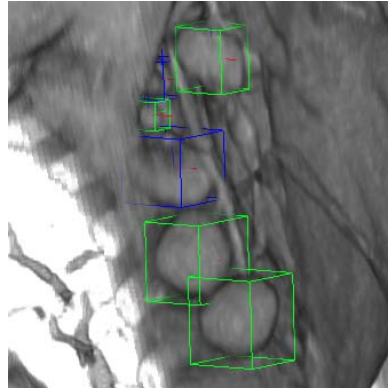


Fig. 2. The lymph nodes are marked with bounding boxes and labeled as solid (green) and non-solid (blue)

2.1 Axillary Region Extraction

To constrain the search, the two axillary regions are detected and cropped automatically; the axillary lymph node detection is performed on these two cropped subvolumes. The axillary subvolumes are obtained by first detecting the lung tips, with an approach similar to [9]. This can be done very reliably and is not the object of this paper. Relative to the two lung tip locations (x, y, z) , subvolumes of size $220 \times 220 \times 220$ voxels (at 1.5mm resolution) are cropped, with the upper-left corner at $(x + 20, y - 135, z - 131)$ for the left lung and upper-right corner at $(x - 20, y - 135, z - 131)$ for the right lung.

2.2 Axillary Lymph Node Annotation

All axillary lymph nodes of size at least 10mm have been annotated by placing bounding boxes around them, as shown in Figure 2. The lymph nodes are labeled as solid or non-solid depending whether they have a homogeneous interior or not. Enlarged lymph nodes with a solid interior are of particular clinical interest since they are believed to have a higher probability of being malignant than lymph nodes that for example have a fatty core.

2.3 Candidate Lymph Node Detection

Lymph node center candidates are detected in the axillary subvolumes ignoring the lymph node size, in the spirit of Marginal Space Learning [12]. Marginal Space Learning is an object detection technique where object candidates are first detected in a subspace where many object parameters (e.g. size, orientation, etc) are ignored. The candidates are refined by searching for the missing parameters using appropriate detectors.

The LN candidates are detected in two stages. The initial set of LN candidates are all voxels in the axillary subvolumes with intensity in the interval $[-100, 200]$ HU. Then, these candidates are evaluated using a fast detector based on Haar features followed by a second detector based on self-aligning gradient features. These two types of features are described below.

1. Haar-Based Features. The first stage of lymph node detection is a trained cascade of Adaboost classifiers trained using 92,000 3D Haar features [10].

2. Self-Aligning Features. The second stage uses a set of features self-aligned to high gradients. The features are computed based on rays casted in 14 directions in 3D space from each candidate location. These 14 directions are $(\pm 1, 0, 0)$, $(0, \pm 1, 0)$, $(0, 0, \pm 1)$, and also $(\pm 1, \pm 1, \pm 1)$. Of the 14 directions, 10 are shown in Fig. 3.

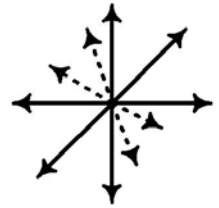


Fig. 3. Self aligning features are computed along 14 directions relative to candidate position

In each direction $d_i, 1 \leq i \leq 14$, local maxima of the gradient above each of 10 thresholds $\tau_j = 10j, 1 \leq j \leq 10$ (see Figure 4), are found at three scales $s_k = 1/2^k, 1 \leq k \leq 3$. Based on them, the following features are evaluated:

- Each of the 24 types of features types (gradient magnitude, angle, intensity value, etc) described in [12] at each of the first three local maxima for each d_i, τ_j, s_k .
- Each of the 24 types of features types computed half way between the candidate location and each of the first three local maxima, for each d_i, τ_j, s_k .
- The distance to each of the first three local maxima for each d_i, τ_j, s_k .
- The differences between distances to the corresponding first three local maxima in any combination of two different directions d_i, d_j for each τ_k, s_l .

This way about 64,000 features are obtained.

The best 1500 candidates above a threshold are kept for each axillary subvolume. An example of detected candidates is shown in Figure 8 (left). For each candidate location, a segmentation is obtained as described below.

2.4 Candidate Lymph Node Segmentation

The segmentation algorithm is specially designed for detecting clinically highly relevant solid lymph nodes. The solid lymph nodes have a blob-like shape that can be described by using a radial function $r : S^2 \rightarrow \mathbb{R}$ defined on the sphere in 3D, representing the distance from the lymph node center to the boundary in all directions. In this work, the sphere has been discretized using a triangulation with 162 vertices, 480 edges and 320 triangles, as shown in Figure 5. Example of lymph node segmentations with this sphere triangulation are shown in Fig. 8. This representation is similar to the shape representation in [3].

Each of the 162 sphere vertices represents a direction d_i . Given a candidate lymph node location C obtained by the candidate detector described in Section 2.3, a segmentation using this location as the center is determined by the radii $r_i, i = 1, \dots, N$ for all directions d_i , where $N = 162$ in our case. These radii form a vector $\mathbf{r} = (r_1, \dots, r_N)$.

The lymph nodes exhibit high shape variability, making it difficult to describe them using generative models such as PCA. To find the segmentation vector \mathbf{r} we adopt an approach similar to the Active Shape Models [2], but using a robust data cost, gradient optimization and a Gaussian MRF shape prior instead of a PCA model.



Fig. 4. In each direction, local gradient maxima above different thresholds τ_j are found

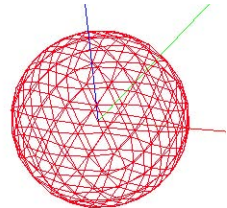


Fig. 5. Sphere triangulation with 162 vertices and 320 triangles

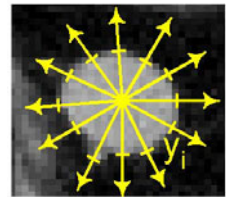


Fig. 6. Measurements y_i are found for each direction d_i as the most probable boundary location

Given the candidate location C , the most likely lymph node boundary locations $y_i, i = 1, \dots, 162$ are found in each direction d_i as

$$y_i = \arg \min_{r \in (0, R_{max})} |I(C) - I(C + (r+1)d_i)| > 50 \quad (1)$$

From the measurement vector $\mathbf{y} = (y_1, \dots, y_{162})$, the segmentation \mathbf{r} is obtained by minimizing the following energy function

$$E(\mathbf{r}) = \alpha \sum_i \rho(r_i - y_i) + \sum_i \frac{1}{2|\partial i|} \sum_{j \in \partial i} (r_i - r_j)^2 \quad (2)$$

where $\rho(x) = \ln(1 + x^2/2)$ and ∂i are the neighbors of i on the sphere mesh.

The first term is the data term, while the second term is the Gaussian MRF prior. If a measurement y_i does not exist, its corresponding term is removed from the data term of eq. (2).

Using the robust cost function ensures that any sporadic outliers in the measurements y_i are ignored. This is illustrated in Figure 7.

Minimization is done by gradient descent, starting with $\mathbf{r} = \mathbf{y}$ as initialization. The energy gradient can be computed analytically, obtaining the update iteration:

$$r_i \leftarrow r_i - \eta \left(\alpha \frac{r_i - y_i}{1 + (r_i - y_i)^2/2} + r_i - \frac{\sum_{j \in \partial i} r_j}{|\partial i|} \right) \quad (3)$$

In practice, we use $\eta = 0.1$ and 1000 gradient update iterations, while $\alpha = 1.6$.

Other segmentation methods such as [3,6,11] could possibly be used, but they lack a robust data term, making them more prone to oversegmentations in low gradient locations.

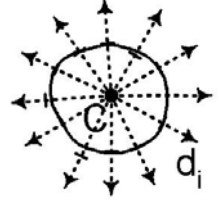


Fig. 7. The segmentation uses a robust cost and a Gaussian MRF to obtain a smooth result that fits most of the y_i

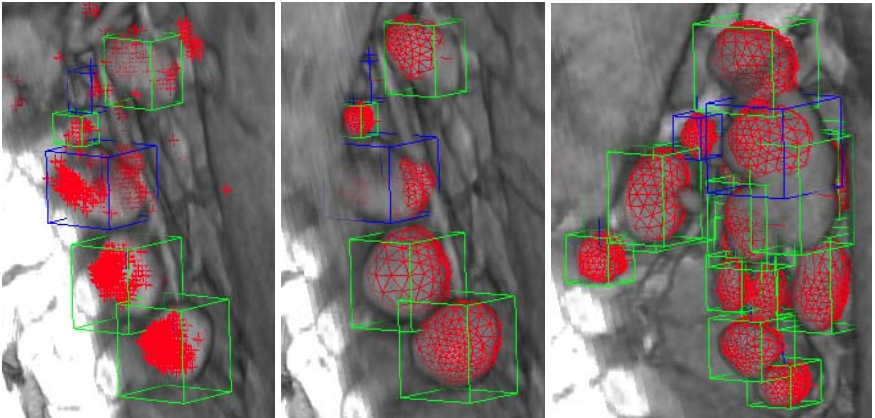


Fig. 8. Left: Detected candidates. middle, right: Detected lymph nodes. right: The method can also handle lymph node conglomerates.

2.5 Final Lymph Node Verification

For each of the candidate lymph node centers obtained using the Candidate detector from Section 2.3, a segmentation with 162 vertices is obtained as described in Section 2.4. The segmentation is used to obtain more informative features for the final evaluation of the lymph node candidates. From the segmentation, a bounding box is extracted for each lymph node. Candidates whose second largest bounding box size is less than 9mm are automatically rejected.

The following features are computed from the segmentation result

- Each of the 24 feature types (gradient magnitude, angle, intensity value, etc) from [12] are computed at the 162 segmentation vertices. For each feature type, the 162 values are sorted in decreasing order.
- For each of the 24 feature types, the 81 sums of feature values at the pairs of opposite vertices are computed and sorted in decreasing order.
- The 81 diameters (distances between opposite vertices relative to the segmentation center) are sorted in decreasing order. For each diameter the following features are computed:
 1. The size of each diameter.
 2. Asymmetry of each diameter, i.e. the ratio of the larger radius over the smaller radius.
 3. The ratio of the i -th sorted diameter and the j -th diameter for all $1 \leq i < j \leq 81$.
 4. For each of the 24 feature types, the max or min of the feature values at the two diameter ends.
 5. For each of the 24 feature types, the max or min of the feature values half way to the diameter ends.

In total there are about 17,000 features.

The classifier trained with these features assigns a score p_i to each candidate i , a higher score meaning a higher likelihood to be a lymph node. All candidates with the score below a threshold τ are automatically removed. On the remaining candidates, a non-maximum suppression scheme is implemented as follows: The algorithm repeats adding the remaining candidate of highest score and removing all candidates close to it. In practice, we chose $N_{max} = 25$ so on each axillary region, a maximum of 25 lymph nodes are detected.

Algorithm 1. Non-maximal Suppression

Input: Candidates $c_i = (x_i, y_i, z_i)$ with scores $p_i > \tau$ and bounding boxes b_i .

Output: Set D of detected lymph nodes.

- 1: Find the candidate c_i with highest score p_i .
 - 2: **if** c_i exists **then** initialize $D = \{i\}$ **else** $D = \emptyset$, **stop**.
 - 3: **for** $n = 2$ **to** N_{max} **do**
 - 4: Remove candidates c_j inside any box $b_i, i \in D$.
 - 5: Find remaining candidate c_j of highest score p .
 - 6: **if** c_j exists **then** add j to detected set: $D \leftarrow D \cup \{j\}$ **else stop**.
 - 7: **end for**
-

Two examples of detected and segmented lymph nodes are shown in red in Figure 8, middle and right. The method can detect parts of lymph node conglomerates as shown in Figure 8, right.

Training details. The training parameters of the three classifiers are given in Table 1. For comparison, a Random Forest [1] classifier with 50 trees was also trained on the segmentation-based features for the verification step.

Table 1. Training details for the three classifiers

Classifier	Features	Type	# Weak	TPR	FPR
1	Haar	AdaBoost cascade	20,50	98%	1%
2	self-aligning	AdaBoost cascade	30,90,270	94%	0.7 – 1%
3	segmentation-based	AdaBoost	27		

3 Experimental Validation

The experiments are performed on a dataset containing 101 CT volumes. All volumes have been converted to 1.5mm isotropic voxel size. In the 101 volumes, a total of 362 solid lymph nodes and 323 non-solid lymph nodes have been found and annotated.

Out of the 101 cases, the region extraction failed only on the left side of one patient that actually had the left lung removed.

The experimental results below are based on a six-fold cross-validation. The CT volumes were divided into six disjoint sets. For each fold, the union of five of the sets was used to train the three classifiers and the remaining set was used for evaluation. Training all three classifiers for each fold took about five hours.

Evaluation Methodology. Since the solid lymph nodes are often very similar and difficult to distinguish from the non-solid ones, we adopted the following evaluation measure for the detection results. A solid lymph node is considered detected if there exists a detection with the center inside the lymph node bounding box. A detection is considered false positive if its center is not inside any annotated solid or non-solid lymph node. Thus any detection on a non-solid lymph node is neither considered a positive nor a false alarm.

Results. Using this evaluation measure, we obtained the ROC curves

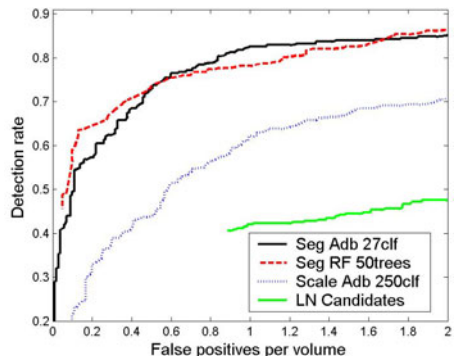


Fig. 9. Detection results with six-fold cross-validation on 101 volumes containing 362 lymph nodes

shown in Figure 9. The solid black curve represents the ROC curve of the system with the verification step having 27 Adaboost classifiers while the interrupted red curve being the ROC of the system with the verification step trained as a Random Forest with 50 trees. For comparison, we show in green color the ROC curve of the LN candidate detector. Thus, the verification step based on segmentation has a great impact on the overall performance of the system.

We also evaluated a system in which the segmentation and verification steps are removed and replaced with a lymph node detector that searches the scale of the lymph node bounding box. This detector is trained using steerable features, as described in [12]. For each of the lymph node candidates, this detector searches 890 combinations of the three lymph node sizes and reports the bounding box of highest score above a threshold. Non-maximal suppression as described in Algorithm 1 is used to further reduce the number of detections. The results using this “scale detector” are shown as a dotted blue curve in Figure 9. This shows that the segmentation-based detector has a much better performance than the scale detector. An added benefit is the fact that the segmentation based detector is about 5 times faster than the scale detector.

Table 2. Detection results and comparison with other methods

Method	Target Area	# cases	Size of nodes	TPR	FP/vol	PPV	Time/vol
Ours	Axillary	101	> 10.0mm	82.3%	1.0	74.9%	5-20sec
Feuerstein [5]	Mediastinum	5	> 1.5mm	82.1%	113.4	13.3%	1-6min
Kitasaka [7]	Abdomen	5	> 5.0mm	57.0%	58	30.3%	2-3h
Dornheim [4]	Neck	1	> 8.0mm	100%	-	76.3%	17min

A comparison with other lymph node detection methods present in the literature is shown in Table 2. Our method achieves a detection rate of 82.3% at 1 false positive per volume, i.e. a 74.9% Positive Predictive Value. This compares favorably with the previous work [5,7]. Dornheim [4] obtains a better detection rate but is evaluated on a single volume, which cannot be considered anywhere close to a thorough evaluation. Moreover, our method is also the fastest, because of the use of the lymph node center detector that ignores the lymph node size and shape and potentially eliminating millions of expensive verifications.

4 Conclusion and Future Work

In this paper, we presented a novel method for automated lymph node analysis based on integrating segmentation with a learning-based detector.

While we address the more restricted problem of solid axillary lymph node detection, we obtain better results with a more thorough evaluation (101 cases compared to 5 cases in [5,7]). At the same time, the proposed method is faster than any of the other existing lymph node detection methods.

In the axillary region, there are no airways or intestines to segment. Nevertheless, a vessel segmentation could further improve the accuracy of our approach.

In the future, we plan to study the improvement brought by using more than one segmentation at each candidate location. We also plan to use the proposed method for segmenting abdominal and mediastinal lymph nodes and also different types of lesions.

References

1. Breiman, L.: Random forests. *Machine Learning* 45(1), 5–32 (2001)
2. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J., et al.: Active shape models—their training and application. *CVIU* 61(1), 38–59 (1995)
3. Dornheim, J., Seim, H., Preim, B., Hertel, I., Strauss, G.: Segmentation of Neck Lymph Nodes in CT Datasets with Stable 3D Mass-Spring Models Segmentation of Neck Lymph Nodes. *Academic Radiology* 14(11), 1389–1399 (2007)
4. Dornheim, L., Dornheim, J.: Automatische Detektion von Lymphknoten in CT-Datensätzen des Halses. In: *BVM* (2008)
5. Feuerstein, M., Deguchi, D., Kitasaka, T., Iwano, S., Imaizumi, K., Hasegawa, Y., Suenaga, Y., Mori, K.: Automatic mediastinal lymph node detection in chest CT. In: *SPIE*, vol. 7260, p. 30 (2009)
6. Kiraly, A.P., Naidich, D.P., Guendel, L., Zhang, L., Novak, C.L.: Novel method and applications for labeling and identifying lymph nodes. In: *SPIE* (2007)
7. Kitasaka, T., Tsujimura, Y., Nakamura, Y., Mori, K., Suenaga, Y., Ito, M., Nawano, S.: Automated extraction of lymph nodes from 3-d abdominal ct images using 3-d minimum directional difference filter. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part II. LNCS*, vol. 4792, p. 336. Springer, Heidelberg (2007)
8. Leibe, B., Leonardis, A., Schiele, B.: Robust object detection with interleaved categorization and segmentation. *IJCV* 77, 259–289 (2008)
9. Seifert, S., Barbu, A., Zhou, S.K., Liu, D., Feulner, J., Huber, M., Suehling, M., Cavallaro, A., Comaniciu, D.: Hierarchical parsing and semantic navigation of full body CT data. In: *SPIE Medical Imaging* (2009)
10. Tu, Z., Zhou, X.S., Barbu, A., Bogoni, L., Comaniciu, D.: Probabilistic 3D polyp detection in CT images: The role of sample alignment. In: *CVPR* (2006)
11. Yan, J., Zhuang, T., Zhao, B., Schwartz, L.H.: Lymph node segmentation from CT images using fast marching method. *Computerized Medical Imaging and Graphics* 28(1-2), 33–38 (2004)
12. Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., Comaniciu, D.: Four-chamber heart modeling and automatic segmentation for 3D cardiac CT volumes using marginal space learning and steerable features. *IEEE TMI* 27(11) (2008)

Image Dissimilarity-Based Quantification of Lung Disease from CT

Lauge Sørensen¹, Marco Loog^{1,2}, Pechin Lo¹, Haseem Ashraf³, Asger Dirksen³, Robert P.W. Duin², and Marleen de Bruijne^{1,4}

¹ The Image Group, Department of Computer Science,
University of Copenhagen, Denmark

lauges@di.ku.dk

² Pattern Recognition Laboratory, Delft University of Technology, The Netherlands

³ Department of Respiratory Medicine, Gentofte University Hospital, Denmark

⁴ Biomedical Imaging Group Rotterdam, Departments of Radiology & Medical Informatics, Erasmus MC, The Netherlands

Abstract. In this paper, we propose to classify medical images using dissimilarities computed between collections of regions of interest. The images are mapped into a dissimilarity space using an image dissimilarity measure, and a standard vector space-based classifier is applied in this space. The classification output of this approach can be used in computer aided-diagnosis problems where the goal is to detect the presence of abnormal regions or to quantify the extent or severity of abnormalities in these regions. The proposed approach is applied to quantify chronic obstructive pulmonary disease in computed tomography (CT) images, achieving an area under the receiver operating characteristic curve of 0.817. This is significantly better compared to combining individual region classifications into an overall image classification, and compared to common computerized quantitative measures in pulmonary CT.

1 Introduction

Quantification of abnormality in medical images often involves classification of regions of interest (ROIs), and combination of individual ROI classification outputs into one global measure of disease for the entire image [1,2,3,4,5,6,7]. These measures may, e.g., express a probability of the presence of certain abnormalities or reflect the extent or severity of disease.

A global image measure based on the fusion of several independent ROI classifications disregards the fact that the ROIs belong to a certain image in the classification step. Moreover, in some cases only global image labels are available, while the images are still represented by ROIs in order to capture localized abnormalities. In some studies, this is handled by propagating the image label to the ROIs within that image, which again allows fusion of individual ROI classifications, to obtain a global image measure [4,5,6]. However, an image showing abnormality will generally comprise both healthy and abnormal regions, and the above approach, incorrectly, labels ROIs without abnormality in such an image as abnormal.

In this paper, we propose to classify medical images using dissimilarities computed directly between the images, where the images are represented by a collection of regions. In this approach, all ROIs in an image contribute when that image is compared to other images, thereby taking into account that the ROIs collectively constitute that image. Further, problems where only a global image label is available are handled automatically since the classification is done at the image level. The images are mapped into a dissimilarity space [8] in which a standard vector space-based classifier can be directly applied, and the soft output of this classifier is used as quantitative measure of disease. The measure used to compute the dissimilarity between two images is the crucial component in this approach, and we evaluate four different image dissimilarity measures in the experiments.

The proposed approach is applied to quantify chronic obstructive pulmonary disease (COPD) in volumetric pulmonary computed tomography (CT) images using texture. Several general purpose classifiers built in the obtained image dissimilarity spaces are evaluated and compared to image classification by fusion of individual ROI classifications as was used in [6].

2 Image Dissimilarity Space

We propose to represent a set of images $\{I_1, \dots, I_n\}$ by their pair-wise dissimilarities $d(I_i, I_j)$ and build classifiers on the obtained dissimilarity representation [8]. From the matrix of pair-wise image dissimilarities $D = [d(I_i, I_j)]_{n \times n}$ computed from the set of images, there exist different ways of arriving at a feature vector space where traditional vector space methods can be applied. In this work, we consider the dissimilarity space approach [8]. An image dissimilarity space is constructed of dimension equal to the size of the training set $|T| = |\{J_1, \dots, J_m\}| = m$, where each dimension corresponds to the dissimilarity to a certain training set image J . All images I are represented as single points in this space, and are positioned according to their dissimilarities to the training set images $D(I, T) = [d(I, J_1), \dots, d(I, J_m)]$. The image dissimilarity measure is a function from two images, represented as sets of ROIs, to a non-negative scalar $d(\cdot, \cdot) : \mathcal{P}(S) \times \mathcal{P}(S) \rightarrow \mathbb{R}_+$ where S is the set of all possible ROIs and $\mathcal{P}(S)$ is the power set of S . It is in this part of the proposed approach that the ROIs are taken collectively into account.

3 Image Dissimilarity Measures

The main issue in obtaining the image dissimilarity space, is the definition of $d(\cdot, \cdot)$. Since the application in this paper is quantification of COPD in pulmonary CT images based on textural appearance in the ROIs, we will focus on image dissimilarity measures suitable for this purpose. In texture-based classification of lung tissue, the texture is sometimes assumed stationary [3,4,6,7]. We will make the same assumption and, therefore, disregard the spatial location of the

ROIs within the lungs. The following are then desirable properties of an image dissimilarity measure for quantification of COPD:

1. Spatial location within the image does not matter. ROIs should be compared solely based on the textural appearance within those regions.
2. The amount of diseased tissue does matter. An image with many abnormal regions is more diseased than an image with few abnormal regions.
3. The appearance of abnormal tissue does matter. Two images with abnormal regions of the same size but with different types of abnormalities should be considered different.

A simple and straightforward image dissimilarity measure between two images, I_1 and I_2 , having the above properties is the sum of all pair-wise ROI dissimilarities:

$$d_{sum}(I_1, I_2) = \sum_{i,j} \Delta(\mathbf{x}_{1i}, \mathbf{x}_{2j}) \quad (1)$$

where \mathbf{x}_{1i} is the i 'th ROI in I_1 and $\Delta(\cdot, \cdot)$ is a texture appearance dissimilarity measure between two ROIs. However, when all ROIs in one image are compared to all ROIs in the other image, the discriminative information of abnormality present in only a few ROIs may be lost. One way to avoid this is to match every ROI in one image with the most similar ROI in the other image. This is the minimum sum distance [9]:

$$d_{ms}(I_1, I_2) = \sum_i \min_j \Delta(\mathbf{x}_{1i}, \mathbf{x}_{2j}) + \sum_j \min_i \Delta(\mathbf{x}_{2j}, \mathbf{x}_{1i}). \quad (2)$$

However, this image dissimilarity measure allows several ROIs in one image to be matched with the same ROI in the other image. This may not be desirable for quantifying COPD since an image with a small abnormal area is considered similar to an image with a large abnormal area. The image dissimilarity measure proposed in the following is a trade-off between d_{sum} and d_{ms} ; it is the sum of several pair-wise ROI dissimilarities, where only one-to-one matchings are allowed, thereby considering images with a small abnormal area as dissimilar to images with a large abnormal area.

3.1 Bipartite Graph Matching-Based Image Dissimilarity Measure

The dissimilarity between two images, or sets of ROIs, $I_1 = \{\mathbf{x}_{1i}\}_n$ and $I_2 = \{\mathbf{x}_{2i}\}_n$, can be expressed as the minimum linear sum assignment between the two sets according to $\Delta(\cdot, \cdot)$. This can be seen as assigning the ROIs in one set to the ROIs in the other set in a way such that the two sets are as similar as possible while only allowing one-to-one matchings. Let $G = (I_1 \cup I_2, E)$ be a weighted undirected bipartite graph with node sets I_1 and I_2 where $|I_1| = |I_2| = n$, edge set $E = \{\{\mathbf{x}_{1i}, \mathbf{x}_{2j}\} : i, j = 1, \dots, n\}$, and with weight $\Delta(\mathbf{x}_{1i}, \mathbf{x}_{2j})$ associated with each edge $\{\mathbf{x}_{1i}, \mathbf{x}_{2j}\} \in E$. The resulting graph is illustrated in Figure 4. A subset M of E is called a perfect matching, or assignment, if every node of

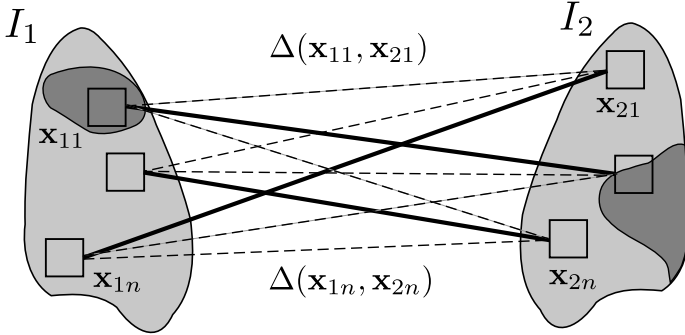


Fig. 1. Illustration of the graph considered when computing the dissimilarity between two images, I_1 and I_2 , in (4). All edges have an associated weight $\Delta(\mathbf{x}_{1i}, \mathbf{x}_{2j})$ that expresses the textural dissimilarity between the two corresponding ROIs \mathbf{x}_{1i} and \mathbf{x}_{2j} . The edges in the perfect matching with minimum weight M^* are shown as solid lines, and the remaining edges, not in M^* , are shown as dashed lines.

G is incident with exactly one edge in M . The perfect matching with minimum weight M^* is given by

$$M^* = \operatorname{argmin}_M \sum_{\{\mathbf{x}_{1i}, \mathbf{x}_{2j}\} \in M} \Delta(\mathbf{x}_{1i}, \mathbf{x}_{2j}) : M \text{ is a perfect matching.} \quad (3)$$

This problem can be solved efficiently using the Hungarian algorithm [10]. The resulting image dissimilarity measure is thus

$$d_{Ia}(I_1, I_2) = \sum_{\{\mathbf{x}_{1i}, \mathbf{x}_{2j}\} \in M^*} \Delta(\mathbf{x}_{1i}, \mathbf{x}_{2j}) \quad (4)$$

where M^* is obtained via (3). No normalization is needed since the images contain an equal amount of ROIs, i.e., n ROIs. Although not used in this work, the formulation can also be relaxed to handle images containing a varying number of ROIs. This will result in an image dissimilarity measure that does not obey the triangle inequality due to partial matches of images. However, this is no problem in the dissimilarity space approach.

4 Experiments

4.1 Data

The data consists of 296 low-dose volumetric CT images from the Danish Lung Cancer Screening Trial with the following scan parameters: tube voltage 120 kV, exposure 40 mAs, slice thickness 1 mm, and in-plane resolution ranging from 0.72 to 0.78 mm. 144 images are from subjects diagnosed as healthy and 152 images are from subjects diagnosed with moderate to very severe COPD. Both groups are diagnosed according to spirometry [11].

4.2 Evaluation

The image dissimilarity-based approach is applied by building classifiers in the CT image dissimilarity spaces obtained using $d(\cdot, \cdot)$. This is compared to using $d(\cdot, \cdot)$ directly as distance in a k nearest neighbor classifier (k NN), which for $k = 1$ corresponds to template matching, and to fusing individual ROI classifications, classified using k NN, for image classification [6]. A posterior probability of each image being positive is obtained using leave-one-out estimation, and receiver operating characteristic (ROC) analysis is used to evaluate the different methods by means of the area under the ROC curve (AUC). The CT image dissimilarity spaces considered in each leave-out trial are of dimension equal to the size of the training set, i.e., 295-dimensional.

Apart from the three image dissimilarity measures described in Section 3, (1), (2), and (4), we also experiment with the Hausdorff distance [9], d_h . This is a classical point set distance measure that do not obey the second property described in Section 3, since it ultimately rely on the dissimilarity between two single ROIs, or points, one from each image. Thus, a total of four different CT image dissimilarity representations are considered in the experiments, one based on each of the four image dissimilarity measures d_{sum} , d_{ms} , d_{la} , and d_h .

4.3 Classifiers

All CT images are represented by a set of 50 ROIs of size $41 \times 41 \times 41$ voxels that each are described by three filter response histograms capturing the local image texture. The filters are: Laplacian of Gaussian (LG) at scale 0.6 mm, gradient magnitude (GM) at scale 4.8 mm, and Gaussian curvature (GC) at scale 4.8 mm. The ROI size as well as the filters are selected based on the results in [6]. The ROI dissimilarity measure used in all experiments is based on the L1-norm between the filter response histograms: $\Delta(\mathbf{x}_1, \mathbf{x}_2) = L_1(h_{LG}(\mathbf{x}_1), h_{LG}(\mathbf{x}_2)) + L_1(h_{GM}(\mathbf{x}_1), h_{GM}(\mathbf{x}_2)) + L_1(h_{GC}(\mathbf{x}_1), h_{GC}(\mathbf{x}_2))$ where $h_i(\mathbf{x})$ is the response histogram of filter i computed in ROI \mathbf{x} .

A SVM with a linear kernel and trade-off parameter $C = 1$ is applied in the obtained CT image dissimilarity spaces. k NN is applied in the following three ways: in the image dissimilarity spaces, using the image dissimilarities directly as distance, and using ROI dissimilarity directly for ROI classification followed by fusion. $k = 1$ is used as well as $k = \sqrt{n}$ where n is the number of prototypes [12]. When classifying CT images, this is $k = \lfloor \sqrt{295} \rfloor = 17$, and when classifying ROIs, this is $k = \lfloor \sqrt{(295 \times 50)} \rfloor = 121$. The following combination rules are considered for fusing individual ROI classifications into image classifications: quantile-based fusion schemes with quantiles ranging from 0.01, i.e., the minimum rule, to 1.00, i.e., the maximum rule, and the mean rule [13]. We also compare to two common CT-based quantitative measures, namely, relative area of emphysema (RA) and percentile density (PD) using the common thresholds of -950 Hounsfield units (HU) and 15% respectively [14]. These measures are computed from the entire lung fields and are denote RA_{950} and PD_{15} , respectively.

4.4 Results

Table 1 shows the estimated AUCs for all the classifiers. The best CT image-dissimilarity based classifier, SVM built in CT image dissimilarity space using d_{la} , achieves an AUC of 0.817. This is better than the best performing mean rule ROI fusion-based classifier, 121NN, which achieves an AUC of 0.751. The common CT-based measures, RA₉₅₀ and PD₁₅, perform worse than all the texture-based measures. The quantile-rule only performed better than the mean rule in the ROI classification fusion in one case, 121NN using maximum rule achieved an AUC of 0.757, and they are therefore not reported in Table 1. SVM in image dissimilarity space using d_{la} or d_{sum} is significantly better, with $p = 0.0028$ and $p = 0.0270$, respectively, than 121NN using the mean rule, while SVM using d_{ms} is not, with $p = 0.085$, according to DeLong, DeLong, and Clarke-Pearson’s test [15].

Table 1. AUCs for COPD diagnosis. **Left:** The results of classification in image dissimilarity space, as well as using the image dissimilarities directly in k NN. **Right:** The results of fusion of individual ROI classification outputs for image classification using the mean rule. The best performing classifier in each approach is marked in bold-face.

	in image dissimilarity space			using $d(\cdot, \cdot)$ directly		fusion of ROI classifications	
	SVM	1NN	17NN	1NN	17NN	1NN	121NN
d_h	0.609	0.522	0.624	0.566	0.668	0.721	
d_{sum} (1)	0.793	0.619	0.643	0.504	0.663	0.751	
d_{ms} (2)	0.795	0.632	0.725	0.600	0.768	0.585	
d_{la} (4)	0.817	0.612	0.671	0.593	0.741	0.589	

5 Discussion

Image dissimilarity measures that match each ROI of one image to an ROI of the other image, under some restrictions, are expected to work well for quantification of COPD within the proposed framework, mainly because more information is taken into account, but also due to increased robustness to noisy ROIs. This is in contrast to measures relying on the match between two ROIs only, such as the Hausdorff distance that is included in the experiments for the sake of completeness. Further, the main argument for building a more global decision rule, such as SVM in a dissimilarity space, instead of applying k NN using the dissimilarity directly as distance is better utilization of the training data and therefore reduced sensitivity to noisy prototypes [8]. This may explain why SVM with a linear kernel built in the dissimilarity space obtained using d_{la} is the best performing of the CT image dissimilarity-based approaches. However, validation on an unseen data set would be needed to draw a final conclusion on this. The experiments showed that SVM with a linear kernel built in the CT image dissimilarity space obtained using d_{la} performed significantly better than using k NN for ROI classification together with the mean rule for CT image classification

($p < 0.05$). This implies that performing the classification at image level, taking into account that an image is in fact a collection of ROIs that collectively constitute that image, is beneficial compared to classifying ROIs individually, while disregarding the fact that they do belong to a certain image.

The computational complexity of the proposed approach using either of the image dissimilarities (1), (2), or (4), in terms of the number of times $\Delta(\cdot, \cdot)$ is evaluated in order to classify a CT image, is the same compared to using the image dissimilarities directly as distance in k NN and to fusion of ROI classifications that are classified using k NN. All approaches require a total of $50 \times 50 \times 295$ evaluations of $\Delta(\cdot, \cdot)$ for classification of a CT image.

When an image is represented by a collection of ROIs and only a label for the entire image is available, the problem of classifying the image can be formulated as a multiple instance learning (MIL) problem [16]. Fusion of independent ROI classifications in order to arrive at an overall image classification can be seen as a “simple” algorithm for solving such a problem. In this paper, we propose to use the dissimilarity-based approach of Pekalska *et al.* [8] on image dissimilarities for solving MIL problems in medical imaging. The approach is similar in spirit to various kernel-based MIL algorithms, such as [17]. The dissimilarity-based approach, however, puts less restrictions on the proximity measure used for comparing objects. Kernel-based approaches require the kernel to be positive definite, which excludes well-known proximity measures such as the Hausdorff distance [9] as well as the bipartite graph matching image dissimilarity measure proposed in this work. Within our framework such measures can be used without any problem.

In conclusion, dissimilarities computed directly between medial images, where the images are represented by a collection of ROIs, was proposed for image classification. This is an alternative to fusion of individual ROI classifications within the images. A SVM built in a dissimilarity space using an image dissimilarity measure based on a minimum sum perfect matching in a weighted bipartite graph, with ROIs as nodes and the textural dissimilarity between two ROIs as edge weight, achieved an AUC of 0.817 on a COPD quantification problem in volumetric pulmonary CT.

Acknowledgements. This work is partly funded by the Danish Council for Strategic Research (NABIIT); the Netherlands Organisation for Scientific Research (NWO); AstraZeneca, Lund, Sweden; and the FET programme within the EU FP7, under the SIMBAD project (contract 213250).

References

1. Müller, N.L., Staples, C.A., Miller, R.R., Abboud, R.T.: “Density mask”. An objective method to quantitate emphysema using computed tomography. *Chest* 94(4), 782–787 (1988)
2. van Ginneken, B., Katsuragawa, S., ter Haar Romeny, B., Doi, K., Viergever, M.: Automatic detection of abnormalities in chest radiographs using local texture analysis. *IEEE Trans. Med. Imag.* 21(2), 139–149 (2002)

3. Park, Y.S., Seo, J.B., Kim, N., Chae, E.J., Oh, Y.M., Lee, S.D., Lee, Y., Kang, S.H.: Texture-based quantification of pulmonary emphysema on high-resolution computed tomography: comparison with density-based quantification and correlation with pulmonary function test. *Invest Radiol.* 43(6), 395–402 (2008)
4. Raundahl, J., Loog, M., Pettersen, P., Tanko, L.B., Nielsen, M.: Automated effect-specific mammographic pattern measures. *IEEE Trans. Med. Imag.* 27(8), 1054–1060 (2008)
5. Arzhaeva, Y., Hogeweg, L., de Jong, P.A., Viergever, M.A., van Ginneken, B.: Global and local multi-valued dissimilarity-based classification: Application to computer-aided detection of tuberculosis. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009, Part II. LNCS*, vol. 5762, pp. 724–731. Springer, Heidelberg (2009)
6. Sørensen, L., Lo, P., Ashraf, H., Sporning, J., Nielsen, M., de Bruijne, M.: Learning COPD sensitive filters in pulmonary CT. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009, Part I. LNCS*, vol. 5761, pp. 699–706. Springer, Heidelberg (2009)
7. Sørensen, L., Shaker, S.B., de Bruijne, M.: Quantitative analysis of pulmonary emphysema using local binary patterns. *IEEE Trans. Med. Imag.* 29(2), 559–569 (2010)
8. Pekalska, E., Duin, R.P.W.: Dissimilarity representations allow for building good classifiers. *Pattern Recog. Lett.* 23(8), 943–956 (2002)
9. Eiter, T., Mannila, H.: Distance measures for point sets and their computation. *Acta Inf.* 34(2), 109–133 (1997)
10. Kuhn, H.W.: The hungarian method for the assignment problem. *Naval Research Logistic Quarterly* 2, 83–97 (1955)
11. Rabe, K.F., Hurd, S., Anzueto, A., Barnes, P.J., Buist, S.A., Calverley, P., Fukuchi, Y., Jenkins, C., Rodriguez-Roisin, R., van Weel, C., Zielinski, J.: Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am. J. Respir. Crit. Care Med.* 176(6), 532–555 (2007)
12. Kittler, J., Alkoot, F.M.: Moderating k-NN classifiers. *Pattern Anal. Appl.* 5(3), 326–332 (2002)
13. Loog, M., Van Ginneken, B.: Static posterior probability fusion for signal detection: applications in the detection of interstitial diseases in chest radiographs. In: *ICPR (1)*, pp. 644–647. IEEE Computer Society, Los Alamitos (2004)
14. Webb, W.R., Müller, N., Naidich, D.: *High-Resolution CT of the Lung*, 3rd edn. Lippincott Williams & Wilkins (2001)
15. DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L.: Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44(3), 837–845 (1988)
16. Dietterich, T.G., Lathrop, R.H., Lozano-Pérez, T.: Solving the multiple instance problem with axis-parallel rectangles. *Artif. Intell.* 89(1-2), 31–71 (1997)
17. Gärtner, T., Flach, P.A., Kowalczyk, A., Smola, A.J.: Multi-instance kernels. In: *ICML*, pp. 179–186. Morgan Kaufmann, San Francisco (2002)

Vessel Segmentation for Ablation Treatment Planning and Simulation^{*,**}

Tuomas Alhonnoro¹, Mika Pollari¹, Mikko Lilja¹, Ronan Flanagan², Bernhard Kainz³, Judith Muehl³, Ursula Mayrhofer⁴, Horst Portugaller⁴, Philipp Stiegler⁴, and Karlheinz Tscheliessnigg⁴

¹ Aalto University School of Science and Technology, Finland

² NUMA Engineering Services Ltd, Ireland

³ Technical University of Graz, Austria

⁴ Medical University of Graz, Austria

Abstract. In this paper, a novel segmentation method for liver vasculature is presented, intended for numerical simulation of radio frequency ablation (RFA). The developed method is a semiautomatic hybrid based on multi-scale vessel enhancement combined with ridge-oriented region growing and skeleton-based postprocessing. In addition, an interactive tool for segmentation refinement was developed. Four instances of three-phase contrast enhanced computed tomography (CT) images of porcine liver were used in the evaluation. The results showed improved accuracy over common approaches and illustrated the method's suitability for simulation purposes.

1 Introduction

Primary liver cancer is the fifth most common cancer worldwide and cause of more than 500 000 deaths a year. While only 5–15% of the patients can be treated surgically, there are others who may benefit from ablation treatments [1]. Radiofrequency ablation (RFA) has recently become the standard treatment for small non-resectable liver tumours, and it can achieve survival rates that are comparable to surgical resection with an improvement over competitive ablation treatments [2].

In RFA, a tumour is destroyed thermally by an electric current passing through a needle-like electrode. Due to the lack of viable planning and limitations in monitoring equipment, assessment of the treatment during or right after the procedure is almost impossible. The problem is pronounced in the vicinity of blood vessels acting as heat sinks and locally interfering the heat propagation [3]. A possible solution is based on numerical simulation of the heat transfer [4].

A subject-specific geometric model is a prerequisite for numerical simulation. Among the most challenging tasks is the segmentation of the hepatic vasculature

* The research leading to these results has received funding from the European Community's Seventh Framework Programme under grant agreement n° 223877, project IMPACT.

** The authors gratefully acknowledge Claire Bost for providing simulations.

including the hepatic artery, portal vein and hepatic vein. A number of vessel segmentation methods have been proposed for different application domains, see review [5]. The most utilized methods for liver applications are based on the work of Selle et al. [6], which successfully combines *region growing-* and *skeleton-based approaches*. An RFA simulation approach based on the method was described in [7]. An alternative approach is based on bottom-up *tracking* [8].

This paper describes a novel, semiautomatic hybrid method for vessel extraction. The method resembles the popular skeleton-based approach first introduced by Selle et al., but it was further augmented with *multi-scale* and *matched filter* approaches [9], tracking schemes and mathematical morphology. Many aspects of the original approach have been revised; most of all, its sensitivity for small vessels has been improved. Furthermore, an interactive tool is introduced for efficient cutting and editing of vessel branches designed especially from the simulation point of view. Finally, a quantitative and qualitative evaluation of the method's performance is presented.

2 Methods

Vessels are segmented from three-phase contrast enhanced CT images, arterial, portal venous and hepatic venous phases distinctly, which are registered, corrected to isotropic according to their smallest dimension and normalized between gray-value range of interest [10].

2.1 Vessel Enhancement

Vessel enhancement filters, a family of multi-scale filters, are a powerful combination of vessel extraction and image denoising, which employ matched differential operators to enhance 3-D tubular structures of different sizes. This work is based on a Hessian tube model [9], in which local scale-dependent second order variations around a tubular structure are characterized by the eigenvalues of the Hessian matrix of a Gaussian filtered image $f_\sigma = G_\sigma * I$ with the standard deviation σ chosen according to the vessel size. Given the eigenvalues $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$, in case of a tube, the condition

$$\begin{aligned} |\lambda_1| &\approx 0 \\ |\lambda_2| &\gg |\lambda_1| \\ |\lambda_2| &\approx |\lambda_3| \end{aligned} \quad (1)$$

should correspond to the maximal probability, according to which the corresponding filter response [11] is given by

$$u(x, \sigma) = \left(1 - \frac{||\lambda_2| - |\lambda_3||}{|\lambda_2| + |\lambda_3|} \right) \left(\frac{2}{3}\lambda_1 - \lambda_2 - \lambda_3 \right). \quad (2)$$

In this work, the computationally intensive multi-scale convolution is replaced by an image-pyramid approach commonly used in solving ordinary differential

equations. The approach scales the image grid instead of the convolution kernel (σ), and the filter response is computed in a top down fashion by propagating the filter responses from each level, and keeping a voxel wise maximum at each, or, more formally, by

$$\begin{aligned} u_d(x) &= u_d(x) \\ u_k(x) &= \max\{u_k(x), L(u_{k+1}(x))\}, k = d - 1, \dots, 0 \\ U(x) &= u_0(x), \end{aligned} \quad (3)$$

where d is the coarsest level, $u_k(x)$ denotes the filter response at level k , and L is a trilinear upsampling operator. Four-point average is used as a restriction operator. The blurring effect resulting from upsampling is compensated by

$$\hat{U}(x) = (1 - \xi \nabla I)U(x), \quad (4)$$

where $\xi \in [0, 1]$ is the sharpening factor.

2.2 Iterative Ridge-Oriented Region Growing

The coarse phase segmentation of the filtered image is obtained by wave-front propagation, the implementation of which is based on the fast marching [12]. Such a wave-front advancing monotonically with a speed $F(x)$ obeys a nonlinear Eikonal equation

$$|\nabla T(\mathbf{x})|F(\mathbf{x}) = 1.0, \quad (5)$$

where $T(\mathbf{x})$ is the time at which the front crosses the point \mathbf{x} . Vice versa, the crossing time t defines the region

$$R(t) = \{\mathbf{x} | T(\mathbf{x}) \leq t\}. \quad (6)$$

The advantages of a wave-front lay on support for complex, e.g., edge-based stopping function, and spatial and directional control over the propagation, e.g., preservation of topology. In this work, a wave-front constrained to i) vessel geometry by a global threshold and ii) topology homeomorphic to a sphere by a topological consistency check [13], is initiated at the seed point close to the vessel opening, and propagated along the vessel tree. In this sense, the speed decreases to zero below the threshold, and thus the stopping time can be defined as a large number.

After coarse phase, discontinuities and small vessels are handled in the detail phase, which is a ridge-oriented alternative to earlier locally adaptive region schemes, e.g., adaptive directional growing [14] or progressive region growing [15]. The method is closely related to minimum-spanning tree and shortest path algorithms, but differs from the usual merging and reconnection schemes [10].

The detail phase is initiated at the boundary of the coarse segmentation. It then follows the steps of the watershed transform [16], but contrary to it, the path is saved as an acyclic graph [17]. It can be thought of as propagating both *at* the watershed line (intensity ridge) and *towards* the watershed minima.

The watershed line can be automatically extracted by using heuristic measures, like branch-length, number of ramifications and ridge curvature. It is intended for small vessels, for which a Gaussian intensity profile holds true. Indeed, the centrelines are likely to be at the locus of intensity maxima after the filtering.

The tracking proceeds until a termination condition or a lower bound for the intensity is met, after which the tracks are labelled. Subsequently, another, marker-controlled watershed transform is applied to the gradient magnitude image $\nabla U(\mathbf{x})$ is applied to extract the boundaries of the tracked vessels. The results are then inspected, and the connected components of the vessel branches are removed if their mass or shape exceeds limits that are considered to correspond to a vessel. The ridge-oriented region growing is a sequence of operations, which, when iterated, is likely to improve the results.

2.3 Post-processing in Skeleton-Domain

The skeleton of the binary image is extracted using a top-down minimum-cost traversal along two distance fields, which preserves the cross-sectional radii (r_k) of the vessel path (given by nodes c_k) [18]. The output is given as a piecewise linear directed tree structure. The skeleton is then processed by parameter-controlled pruning or smoothing according to, e.g., ratio of branch length to radius, Strahler-level [19] (Fig. 2), vessel radius or its derivative along the vessel direction. In addition, an interactive branch cutter and editor was developed, which makes use of intuitive travelling across the image by using a cross-sectional plane perpendicular to a vessel, and two maximum intensity projections. When these are combined with an interactive tool for controlling rotation, it is sufficient to reveal the nearby vessels in their true 3-D nature (Fig. 1). Finally, the processed skeletons are converted into a smooth 0-level set by an inverse distance transform.

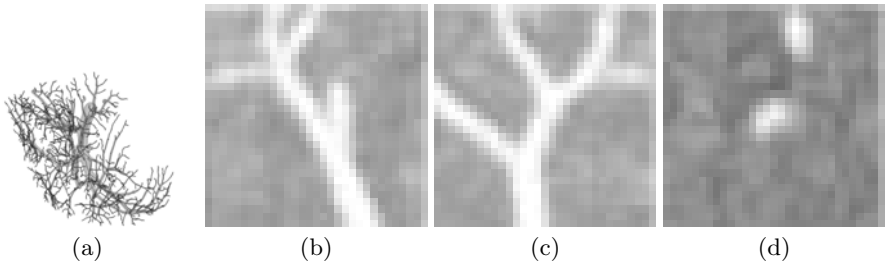


Fig. 1. Tools for interactive refinement and landmark selection. (a) porcine portal vein superimposed on its skeleton. Two maximum intensity projections (b) (in y-direction) (c) (x) inside a local rotated image block, and a cross-section of the block (d)(x-y plane).

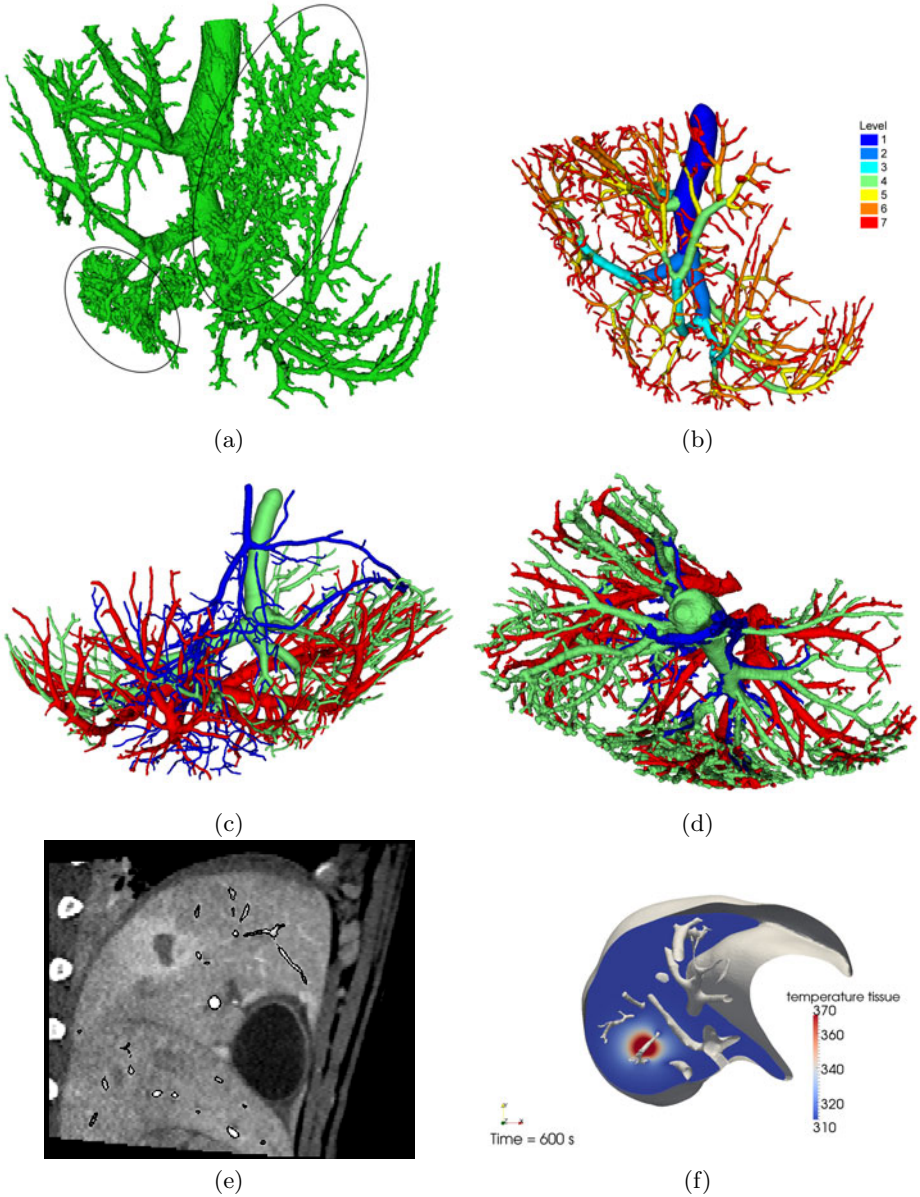


Fig. 2. (a) Unfiltered data (halo errors emphasized) and (b) the result from this method, where the vessel branches are colored according to Strahler scheme. (c) and (d) full vascular models of porcine liver: arteries (blue), portal vein (green) and hepatic veins (red). (e) Portal vein mesh (black) superimposed on CT image. (f) An instance of finite element modelling of heat transfer (temperature in Kelvin). The polygon model has been greatly simplified prior to simulation.

3 Evaluation

The quantitative evaluation is based on a landmark study each landmark consisting of a vessel centreline and a radius $l_i = (c_i, r_i)$ estimated from the original CT image by using one of the tools developed for exploration of vessel images (Fig. 1). Starting at the vessel opening and proceeding towards increasing order of ramification until the vanishing point, a total 1024 landmarks were manually collected. These represent a prescriptive set of ramification pattern. Next, the point correspondence between each landmark and the corresponding semiautomatic vessel tree ($s_k = (c_k, r_k)$) was solved by minimizing the Euclidean distance between each l_i and c_k

$$\{c_j(i), r_j(i)\} = \min_k \|c_i, c_k\|_2. \quad (7)$$

To prevent a false correspondences (e.g. with another branch) from biasing the analysis, we introduced a new measure (hit rate) given by

$$\text{HIT}_i = \begin{cases} 1, & \text{if } \|c_k(i) - c_i\|_2 < r_i \\ 0, & \text{otherwise,} \end{cases} \quad (8)$$

which effectively describes if the c_k lies within the segmented vessel tree. In evaluation, landmarks are compared to results. Missed landmarks are treated as outliers, but considered a superset of other errors. Then the root mean square errors for the radii and centreline position are given by

$$\text{RMSE}(x) = \sqrt{\sum_k \text{HIT}_i (x_k(i) - x_i)^2 / \sum_k \text{HIT}_i}, \quad x = c, r. \quad (9)$$

4 Results

The method was evaluated with four porcine livers using contrast enhanced CT images. For each animal, three images were taken in the same respiration phase to enhance the respective vessels one at a time and to neglect deformations. Helical CT scans were performed at Medical University of Graz using computer-assisted bolus-tracking (ROI surestart) and 320-line Toshiba Aquilion ONE (resolution 512x512x320 @ 0.407–0.463x0.5 mm).

In qualitative evaluation, several visualization techniques were used to present the segmented vasculature. Polygon models of the segmentation depicted in Fig. 2(a)–(b) represent the level of detail extractable by our method compared to conventional thresholding and region growing approaches (from [6]). Figure 2(c) and (d) show arteries (blue), portal vein (green) and hepatic vein (red) in the full vascular model. Figure 2(e) shows polygon model superimposed on the original CT image, and finally Fig 2(f) shows an instance of simulation using our data.

The results of the quantitative evaluation are depicted in Tab. 1. Run-time of the method was less than 4 minutes on a modern quad-core workstation.

Table 1. Hit rate of the landmark study with respect to the vessel radius. Units are all voxels. Error in landmark is of the order of 1 voxel.

radius (R)	hit rate (%)	RMSE(\hat{c}) (of hits)	RMSE(\hat{r}) (of hits)
≤ 9.0	100	1.69	1.9
$6.0 \leq R < 9.0$	100	1.08	0.95
$5.0 \leq R < 6.0$	96	1.17	0.97
$4.0 \leq R < 5.0$	92	1.04	0.87
$3.0 \leq R < 4.0$	97	1.00	0.61
$2.0 \leq R < 3.0$	97	0.91	0.72
$1.0 \leq R < 2.0$	75	0.91	0.49
$0.5 \leq R < 1.0$	25	0.61	0.09
all	87	1.00	0.77

5 Discussion

In literature, a vessel diameter of 3.0 mm has been described a critical threshold for ablation heat propagation [3]. We have shown that our method is capable of extracting 97% of vessels equal or above the threshold (3.0 mm diameter corresponds to 3.0 voxel radius). Furthermore, the accuracy is well below the reasonable resolution for finite element modelling, and does not drop until a subvoxel resolution. The comparison to other vessel segmentation methods is challenging as quantitative evaluation has usually been omitted, but visual evaluation showed improvements over previous methods. Also, comparing the hit rate (sensitivity) results, there were improvements over the only quantitative study published [20].

Our intuitive combination of region growing and ridge tracking is inherently capable of growing over small gaps, and extracting structures down to unit voxel thickness. It resembles live wire approaches and readily supports manual seeding of individual branches, or top-down tracking of false-negative vessel branches near image artifacts, e.g., tumours. Our pyramid approach is a reasonable and robust approximation to vessel enhancement that can be computed on a modern workstation in one minute. The accuracy and suppression of halos of microvasculature are well demonstrated in the Fig. 2(a) and (b). Furthermore, initial trials on human data have shown consistent results.

The skeleton domain post-processing and interactive refinement tools provide an efficient platform not only for error correction but for modifying the vasculature more suitable for simulation. From the RFA point of view, the most important are the nearby vessels, which can be easily isolated. Circular cross-section encoded in both the skeleton and landmarks, and widely used in the field, introduces error in largest oval-shaped vessels (Tab. I). For other vessels, however, this has proven a reasonable assumption, and desired property due to its smoother vessel surface.

6 Conclusion

We have demonstrated a new, efficient and robust hybrid vessel segmentation for RFA ablation simulation. The suitability was carefully evaluated, and its results are already applied to RFA simulation. Further, we believe the method is useful for many other applications domains or imaging modalities.

References

- [1] Gish, R.: Hepatocellular Carcinoma: Overcoming Challenges in Disease Management. *Clinical Gastroenterology and Hepatology* 4(3), 252–261 (2006)
- [2] Garrean, S., et al.: Radiofrequency ablation of primary and metastatic liver tumours: a critical review of the literature. *The American J. of Surgery* (2008)
- [3] Hansen, P., et al.: Radiofrequency ablation lesions in pig liver model. *J. of Magnetic Resonance Imaging* 87(1), 114 (1999)
- [4] Tungjitkusolmun, S., et al.: Three-dimensional finite-element analysis for radio-frequency hepatic tumour ablation. *IEEE, Biomedical Engineering* 49(1), 3–9 (2002)
- [5] Kirbas, C., Quek, F.: A review of vessel extraction techniques and algorithms. *ACM Computing Surveys* 36(2), 81–121 (2004)
- [6] Selle, D., et al.: Analysis of vasculature for liver surgical planning. *IEEE, Medical Imaging* 21(11), 1344–1357 (2002)
- [7] Weihusen, A., et al.: Towards a workflow-oriented software assistance for the radiofrequency ablation. *LN Informatics* 93, 507–513 (2006)
- [8] Flasque, N., et al.: Acquisition, segmentation and tracking of the cerebral vascular tree on 3D magnetic resonance angiography images. *Medical Image Analysis* 5(3), 173–183 (2001)
- [9] Lorenz, C., et al.: Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2D and 3D medical images. *LNCS*, pp. 233–242. Springer, Heidelberg (1997)
- [10] Beichel, R., et al.: Liver segment approximation in CT data for surgical resection planning. In: *Proc. of SPIE*, vol. 5370, p. 1435 (2004)
- [11] Erdt, M., et al.: Hepatic Vessel Segmentation Using Graphics Hardware. In: Dohi, T., Sakuma, I., Liao, H. (eds.) *MIAR 2008. LNCS*, vol. 5128, pp. 403–412. Springer, Heidelberg (2008)
- [12] Sethian, J.: Fast Marching Methods. *SIAM Rev.* 41(2), 199–235 (1999)
- [13] Bertrand, G., Malandain, G.: A new characterization of three-dimensional simple points. *Pattern Recognition L's* 15(2), 169–175 (1994)
- [14] Shang, Q., et al.: Adaptive directional region growing segmentation of the hepatic vasculature. In: *Proc. of SPIE*, vol. 6914, p. 69141F (2008)
- [15] Bock, S., et al.: Robust vessel segmentation. In: *Proc. of SPIE*, vol. 6915, p. 691539 (2008)
- [16] Vincent, L., Soille, P.: Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE, pattern analysis and machine intelligence* 13(6), 583–598 (1991)
- [17] Yim, P., et al.: Gray-scale skeletonization of small vessels in magnetic resonance angiography. *IEEE, Medical Imaging* 19(6), 568–576 (2000)
- [18] Zhou, Y., Toga, A.: Efficient skeletonization of volumetric objects. *IEEE, Visualization and Computer Graphics* 5(3), 196–209 (1999)
- [19] Strahler, A.: Quantitative analysis of watershed geomorphology (1957)
- [20] Lehmann, K., et al.: Portal vein segmentation of a 3-D planning system for liver surgery - in vivo evaluation in a porcine model. *A. Surgical Oncology* 15(7), 1899–1907 (2007)

Integrated Segmentation and Nonrigid Registration for Application in Prostate Image-Guided Radiotherapy

Chao Lu^{1,*}, Sudhakar Chelikani², Zhe Chen², Xenophon Papademetris²,
Lawrence H. Staib^{1,2}, and James S. Duncan^{1,2}

¹ Department of Electrical Engineering

² Department of Diagnostic Radiology
Yale University, New Haven, CT, USA

chao.lu@yale.edu

Abstract. Many current image-guided radiotherapy (IGRT) systems incorporate an in-room cone-beam CT (CBCT) with a radiotherapy linear accelerator for treatment day imaging. Segmentation of key anatomical structures (prostate and surrounding organs) in 3DCBCT images as well as registration between planning and treatment images are essential for determining many important treatment parameters. Due to the image quality of CBCT, previous work typically uses manual segmentation of the soft tissues and then registers the images based on the manual segmentation. In this paper, an integrated automatic segmentation/constrained nonrigid registration is presented, which can achieve these two aims simultaneously. This method is tested using 24 sets of real patient data. Quantitative results show that the automatic segmentation produces results that have an accuracy comparable to manual segmentation, while the registration part significantly outperforms both rigid and non-rigid registration. Clinical application also shows promising results.

1 Introduction

Prostate cancer is the most commonly diagnosed cancer among men in the United States. For the majority, external beam radiotherapy (EBRT) is one of the primary treatment modalities for prostate cancer [1].

Recent advances in EBRT have led to three-dimensional conformal radiotherapy (3DCRT) and intensity modulated radiotherapy (IMRT). Prostate 3DCRT requires a precise delineation of the target volume and the adjacent critical organs in order to deliver an optimal dose to the prostate with minimal side effect on nearby normal tissues. Many current image-guided radiotherapy (IGRT) systems integrate an in-room cone-beam CT (CBCT) with a radiotherapy linear accelerator for treatment day imaging. With both imaging and radiotherapy available on the same platform, daily CBCTs can now be acquired and used for patient positioning.

* This work is supported by NIH/NIBIB Grant R01EB002164.

However, when higher doses are to be delivered, precise and accurate targeting is essential because of unpredictable inter- and intra-fractional organ motions over the process of the daily treatments that often last more than one month. Therefore, a non-rigid registration problem must be solved in order to map the planning setup information in the initial 3DCT data into each treatment day 3DCBCT image. Meanwhile, we must accurately segment the prostate, bladder and rectum from the 3DCBCT images. Due to the quality of CBCT images, these issues can be very tough. Greene *et al.* [2] carried out this by manual segmentation and then involved nonrigid registration. Some initial work has been performed in simultaneously integrating registration and segmentation. Chelikani *et al.* [3] integrated rigid 2D portal to 3D CT registration and pixel classification in an entropy-based formulation. Yezzi *et al.* [4] integrated segmentation using level sets with rigid and affine registration. Chen *et al.* [5] implemented a nonrigid transformation and a hidden Markov random field to improve the segmentation performance.

In this paper, we present an integrated automatic segmentation and softly constrained nonrigid registration algorithm. Our model is based on a *maximum a posteriori*(MAP) framework while the automatic segmentation is performed using level set deformable model with shape prior information, as proposed in [6], and the constrained nonrigid registration part is based on a multi-resolution cubic B-spline Free Form Deformation (FFD) transformation. These two issues are intimately related: by combining segmentation and registration, we can recover the treatment fraction image regions that correspond to the organs of interest (prostate, bladder, rectum) by incorporating transformed planning day organs to guide and constrain the segmentation process; and conversely, accurate knowledge of important soft tissue structures will enable us achieve more precise nonrigid registration which allows the clinician to set tighter planning margins around the target volume in the treatment plan. Escalated dosages can then be administered while maintaining or lowering normal tissue irradiation.

2 Method

The integrated segmentation/registration algorithm was developed using Bayesian analysis to calculate the most likely segmentation in treatment day fractions S_d and the mapping field between the planning day data and treatment day data T_{0d} . This algorithm requires: (1) a planning day 3DCT image I_0 , (2) a treatment day 3DCBCT image I_d , and (3) the segmented planning day organs S_0 .

2.1 MAP Framework

A probabilistic model can be used to realize image segmentation combining prior information and image information, simultaneously, it can also realize

nonrigid registration by incorporating both intensity matching and segmented organ matching constraint. Thus we estimate

$$\widehat{S}_d, \widehat{T}_{0d} = \arg \max_{S_d, T_{0d}} [p(S_d, T_{0d} | I_0, I_d, S_0)] \quad (1)$$

This multiparameter MAP estimation problem in general is difficult to solve, however, we reformulate this problem such that it can be solved in two basic iterative computational stages using an iterative conditional mode (ICM) strategy. With k indexing each iterative step, we have:

$$\widehat{S}_d^k = \arg \max_{S_d^k} [p(S_d^k | T_{0d}^k(S_0), I_d)] \quad (2)$$

$$\widehat{T}_{0d}^{k+1} = \arg \max_{T_{0d}^{k+1}} [p(T_{0d}^{k+1} | S_d^k, S_0, I_d, I_0)] \quad (3)$$

These two equations represent the key problems we are addressing: i.) in equation (2) the estimation of the segmentation of the important day d structures (S_d^k) and ii.) in equation (3), the estimation at the next iterative step of the mapping T_{0d}^{k+1} between the day 0 and day d spaces.

2.2 Segmentation Module

We first apply Bayes rule to equation (2) to get:

$$\widehat{S}_d^k = \arg \max_{S_d^k} [p(S_d^k | T_{0d}^k(S_0), I_d)] = \arg \max_{S_d^k} [p(I_d | S_d^k) p(T_{0d}^k(S_0) | S_d^k) p(S_d)] \quad (4)$$

Here we assume that the priors are stationary over the iterations, so we can drop the k index for that term only, i.e. $p(S_d^k) = p(S_d)$. To build a model for the shape prior, we choose level sets as the representation of the objects. Consider a training set of n aligned images. Each object in the training set is embedded as the zero level set of a higher dimensional level set Ψ . The mean and variance of the boundary of each object can be computed using Principal Component Analysis (PCA). The mean level set, $\bar{\Psi}$, is subtracted from each Ψ to create the deviation. Each such deviation is placed as a column vector in a $N^3 \times n$ -dimensional matrix Q where N^3 is the number of samples of each level set function. Using Singular Value Decomposition (SVD), $Q = U \Sigma V^T$. U is a matrix whose column vectors represent the set of orthogonal modes of shape variation and Σ is a diagonal matrix of corresponding singular values. An estimate of the object shape Ψ_i can be represented by k principal components and a k -dimensional vector of coefficients α_i : $\tilde{\Psi}_i = U_k \alpha_i + \bar{\Psi}$. Under the assumption of a Gaussian distribution of object represented by α_i , we can compute the probability of a certain shape:

$$p(S_d) = p(\alpha_i) = \frac{1}{\sqrt{(2\pi)^k |\Sigma_k|}} \exp \left[-\frac{1}{2} \alpha_i^T \Sigma_k^{-1} \alpha_i \right] \quad (5)$$

Then we impose a key assumption: the segmentation likelihood is separable into two independent data-related likelihoods, requiring that the estimate of the

structure at day d be close to: i.) the same structure segmented at day 0, but mapped to a new estimated position by the current iterative mapping estimate T_{0d}^k and ii.) the intensity-based feature information derived from the day d image.

In equation (4), $p(T_{0d}^k(S_0)|S_d^k)$ constrains the segmentation in day d to be adherent to the transformed day 0 organs by current mapping T_{0d}^k . Thus, the probability density of day 0 segmentation likelihood term can be modeled as:

$$p(T_{0d}^k(S_0)|S_d^k) = \frac{1}{Z} \prod_{(x,y,z)} \exp \left[-(\Psi_{T_{0d}^k(S_0)} - \Psi_{S_d^k})^2 \right] \quad (6)$$

where Z is a normalizing constant. Assuming gray level homogeneity within each object, we use the imaging model defined by Chan [7], where c_1 and σ_1 are the average and standard deviation of I_d inside S_d^k , c_2 and σ_2 outside.

$$p(I_d|S_d^k) = \prod_{inside(S_d^k)} \exp[-(I_d(x,y,z)-c_1)^2/2\sigma_1^2] \cdot \prod_{outside(S_d^k)} \exp[-(I_d(x,y,z)-c_2)^2/2\sigma_2^2] \quad (7)$$

Notice that the MAP estimation of the objects in equation (4), \widehat{S}_d^k , is also the minimizer of the energy function E_{seg} shown below. This minimization problem can be formulated and solved using the level set surface evolving method.

$$\begin{aligned} E_{seg} &= -\ln p(S_d^k|T_{0d}^k(S_0), I_d) = -\ln[p(I_d|S_d^k) \cdot p(T_{0d}^k(S_0)|S_d^k) \cdot p(S_d^k)] \\ &\propto \lambda_1 \int_{inside(S_d^k)} |I_d(x,y,z) - c_1|^2 dx dy dz + \lambda_2 \int_{outside(S_d^k)} |I_d(x,y,z) - c_2|^2 dx dy dz \\ &\quad + \gamma \int_{(x,y,z)} \left| \Psi_{T_{0d}^k(S_0)} - \Psi_{S_d^k} \right|^2 dx dy dz + \omega_i \alpha_i^T \Sigma_k^{-1} \alpha_i \end{aligned} \quad (8)$$

2.3 Registration Module

The goal here is to register the planning day data to the treatment day data and carry the planning information forward, as well as to carry forward segmentation constraints. To do this, the second stage of the ICM strategy described above in equation (3) can be further developed using Bayes rule:

$$\begin{aligned} \widehat{T_{0d}^{k+1}} &= \arg \max_{T_{0d}^{k+1}} [p(T_{0d}^{k+1}|S_d^k, S_0, I_d, I_0)] \\ &= \arg \max_{T_{0d}^{k+1}} [\ln p(S_d^k, S_0|T_{0d}^{k+1}) + \ln p(I_d, I_0|T_{0d}^{k+1}) + \ln p(T_{0d})] \end{aligned} \quad (9)$$

The first two terms on the right hand side represent conditional likelihoods related to first registering the three segmented soft tissue structures at days 0 and d , and second registering the intensities of the images. The third term represents prior assumptions on the overall nonrigid mapping, which is captured with smoothness models and is assumed to be stationary over the iterations.

As discussed in the segmentation section, each object is represented by the zero level set of a higher dimensional level set Ψ . Assuming the objects vary

during the treatment process according to a Gaussian distribution, and given that the different organs can be registered respectively, we further simplifies the organ matching term as

$$\begin{aligned}
 \ln p(S_d^k, S_0 | T_{0d}^{k+1}) &= \sum_{obj=1}^3 \ln p(S_{d-obj}^k, S_{0-obj} | T_{0d}^{k+1}) \\
 &= \sum_{obj=1}^3 \int_{(x,y,z)} \ln \frac{1}{\sqrt{2\pi}\sigma_{obj}} \exp \left[\frac{-(\Psi_{T_{0d}^{k+1}}(S_{0-obj}) - \Psi_{S_{d-obj}^k})^2}{2\sigma_{obj}^2} \right] dx dy dz \\
 &= \sum_{obj=1}^3 -\omega_{obj} \int_{(x,y,z)} \left[\Psi_{T_{0d}^{k+1}}(S_{0-obj}) - \Psi_{S_{d-obj}^k} \right]^2 dx dy dz
 \end{aligned} \tag{10}$$

When minimized, the organ matching term ensures the transformed day 0 organs and the segmented day d organs align over the regions.

The intensities of CT and CBCT differ by a linear transformation and Gaussian noise due to X-ray scatter during CBCT acquisition. In this paper, normalized cross correlation (NCC), which has successfully been tested in prostate IGRT [2], is used to address such intensity matching objective.

$$C_{NCC} = -\ln p(I_d, I_0 | T_{0d}^{k+1}) = -\frac{1}{N} \int_{x,y,z} \frac{[T_{0d}^{k+1}(I_0(x,y,z)) - \bar{I}_0] [I_d(x,y,z) - \bar{I}_d]}{\sigma_{T_{0d}^{k+1}}(I_0) \sigma_{I_d}} dx dy dz \tag{11}$$

where \bar{I}_0 and \bar{I}_d are the mean intensities, σ represents the standard deviation.

To ensure a transformation is smooth, a penalty term is also introduced.

$$C_{smooth} = -\ln p(T_{0d}) = \frac{1}{V} \int_{x,y,z} [(\frac{\partial^2 T_{0d}}{\partial x^2})^2 + (\frac{\partial^2 T_{0d}}{\partial y^2})^2 + (\frac{\partial^2 T_{0d}}{\partial z^2})^2] dx dy dz \tag{12}$$

where V denotes the volume of the image domain. The regularization term penalizes only nonaffine transformations. In addition, we constrain each control point of the tensor B-Spline FFD to move within a local sphere of radius $r < R$ where $R \approx 0.4033$ of the control point spacing. This condition guarantees T_{0d} to be locally injective [2].

Therefore, we introduce the registration energy function that can be minimized using a conjugate gradient optimizer,

$$\begin{aligned}
 E_{reg} &= -\ln p(T_{0d}^{(k+1)} | S_d^k, S_0, I_d, I_0) \\
 &= -\ln p(S_d^k, S_0 | T_{0d}^{(k+1)}) - \ln p(I_d, I_0 | T_{0d}^{(k+1)}) - \ln p(T_{0d}) \\
 &= \sum_{obj=1}^3 \omega_{obj} \int_{x,y,z} \left[\Psi_{T_{0d}^{k+1}}(S_0) - \Psi_{S_d^k} \right]^2 dx dy dz + C_{NCC} + \beta C_{smooth}
 \end{aligned} \tag{13}$$

Equations (8) and (13) run alternatively until convergence. Thereafter, the soft tissue segmentation as well as the nonrigid registration benefit from each other and can be estimated simultaneously.

3 Results

We tested our iterative conditional model (ICM) on 24 sets of real patient data acquired from three different patients. Each of the patient had eight treatment 3D CBCT images, and had an associated planning day 3DCT image and 3DCRT treatment plan. The planning day prostate, rectum, and bladder were hand segmented by a qualified clinician. The experiment results are presented below.

3.1 Segmentation and Registration Results

In Fig. 1(a), we show the sagittal view of the segmentation using only image gray level information from treatment day CBCT data, by which the surfaces cannot lock onto the shapes of the objects, since these structures have very poorly defined boundaries. Fig. 1(b) shows the results using gray level information with the shape prior. The results are better, but the boundaries of the organs overlap a lot where the structures are connected. In Fig. 1(c), we show results of using our ICM model. The three structures can be clearly segmented. The surfaces are able to converge on the desired boundaries even though some parts of the boundaries are too blurred to be detected by gray level information, and there is no overlap of the boundaries due to the use of constraints from transformed planning day organs. Fig. 1(d) shows the clinician’s manual segmentation for comparison.

To validate the segmentation results, we tested our model on 24 treatment day images and then compared with the manual results (used as ground truth) using three metrics: mean absolute distance(MAD), Hausdorff distance (HD), and the percentage of correctly segmented voxels (PTP). While MAD represents the global disagreement between two contours, HD compares their local similarities. Quantitative validation in Table 1 showed remarkable and consistent agreement between the proposed method and manual segmentation by expert.

The registration performance of the proposed ICM algorithm was also evaluated. For comparison, a conventional non-rigid registration (NRR) using only intensity matching and rigid registration (RR) were performed on the same 24 sets of real patient data. Organ overlaps between the ground truth in day d and the transformed organs from day 0, as well as the mutual information between CBCT and transformed CT images were used as metrics to assess the quality of the registration (Table 2). The overlap increase for each object and the increase of mutual information after 10 iterations are presented in Fig. 2.

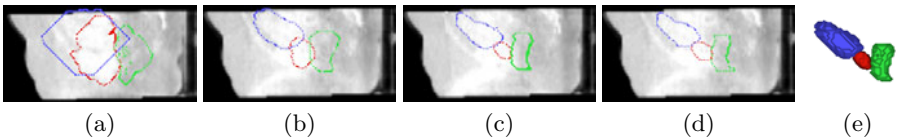


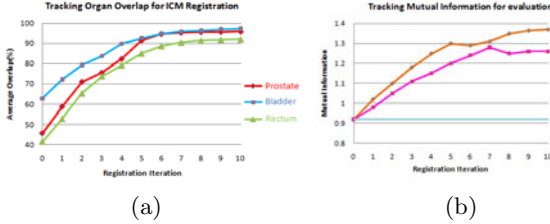
Fig. 1. Segmentation performance. Red:prostate, Blue:bladder, Green:rectum. (a)Using gray level information; (b)With Shape prior; (c)Proposed ICM algorithm; (d)Manual segmentation; (e)Segmented 3D Surfaces using the proposed method.

Table 1. Evaluation of the Segmentation Module

	Method	MAD(mm)	HD(mm)	PTV(%)
Prostate	shape prior	7.84 ± 2.69	9.13 ± 3.18	62.04 ± 4.73
	ICM model	1.92 ± 0.45	2.56 ± 0.75	96.15 ± 2.70
Bladder	shape prior	10.21 ± 5.39	11.28 ± 5.84	63.79 ± 3.72
	ICM model	2.84 ± 0.67	4.23 ± 1.19	94.69 ± 3.59
Rectum	shape prior	9.55 ± 4.76	10.97 ± 5.53	56.66 ± 10.86
	ICM model	2.59 ± 0.58	3.46 ± 1.05	92.72 ± 2.83

Table 2. Evaluation of the Registration Module

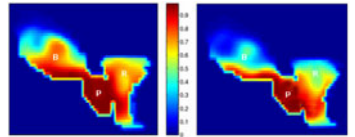
	RR	NRR	ICM
Prostate Overlap(%)	45.32 ± 2.45	77.58 ± 4.14	92.65 ± 1.82
Bladder Overlap(%)	61.85 ± 3.73	73.23 ± 8.28	94.79 ± 2.44
Rectum Overlap(%)	41.96 ± 4.16	64.27 ± 6.32	90.13 ± 2.83
Mutual Information	0.92 ± 0.09	1.26 ± 0.18	1.38 ± 0.25

**Fig. 2.** Registration performance. (a) Average object overlap results at each iteration using the proposed algorithm; (b) Comparison of ICM results to NRR and RR using mutual information at each iteration;

The RR performed the poorest out of all the registrations algorithms, generating an identity transform for all sets of patient data, while the proposed ICM method significantly outperformed the NRR at aligning segmented organs. All results presented here were averaged over the 24 sets of real patient data.

3.2 Treatment Plan Results: Cumulative Dose Distribution

Due to set-up errors and organ motions, treatment plan results varied from day to day. For this reason, it is essential to characterize the dose delivered to each region of interest. The non-rigid transformation achieved using our model was used to warp a fixed dose plan at each treatment day and compute the dose delivered to each voxel of tissue by properly accumulating the warped dose over all treatment fractions. Fig. 3 (left) shows a 2D slice through 3D cumulative dose distribution when dose plans (12mm margin) were mapped and summed across 8 weekly treatment fractions from a single patient. Though it is apparent that the prostate is receiving nearly 90% of the dose, the bladder and the rectum are also receiving substantial amounts. Since our method could achieve a more accurate segmentation and registration, it enables the clinician to set a tighter margin (4mm) prescribed around the Clinical Target Volume (CTV), to ensure accurate delivery of the planned dose to the prostate and to minimize the dose received by the rectum and bladder, as shown in Fig. 3 (right).

**Fig. 3.** Cumulative dose distributions for plans with 12mm(left) and 4mm(right) margin

4 Conclusion

In this paper, we have presented a novel integrated iterative conditional model that could simultaneously achieve automatic segmentation and constrained nonrigid registration. Experiments on 24 sets of real patient data showed that automatic segmentation using the ICM produced results that had an accuracy comparable to that obtained by manual segmentation. For each patient data tested, the proposed method proved to be highly robust and significantly improved the overlap for each soft tissue organ and outperformed the results achieved from the RR and NRR. Updating treatment plans showed promising superiority of the novel method in maintaining radiation dosage to the prostate and lowering radiation dosage to the rectum and bladder. Therefore, the proposed algorithm appears suitable for clinical application in image-guided radiotherapy analysis. Future work would focus on pelvic anatomy incorporating shape prediction. We can potentially adjust the plan to better conform to the dose already delivered and match the predicted variability in the pelvic anatomy for future fractions.

References

1. Potosky, A.L., Legler, J., Albertsen, P.C., Stanford, J.L., Gilliland, F.D., Hamilton, A.S., Stephenson, R.A., Harlan, L.C.: Health outcomes after prostatectomy or radiotherapy for prostate cancer: Results from the prostate cancer outcomes study. *J. Nat. Cancer Inst.* 92, 1582–1592 (2000)
2. Greene, W.H., Chelikani, S., Purushothaman, K., Chen, Z., Knisely, J.P.S., Staib, L.H., Papademetris, X., Duncan, J.S.: A Constrained Non-rigid Registration Algorithm for Use in Prostate Image-Guided Radiotherapy. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part I*. LNCS, vol. 5241, pp. 780–788. Springer, Heidelberg (2008)
3. Chelikani, S., Purushothaman, K., Knisely, J., Chen, Z., Nath, R., Bansal, R., Duncan, J.S.: A gradient feature weighted minima algorithm for registration of multiple portal images to 3DCT volumes in prostate radiotherapy. *Int. J. Radiation Oncology Biol. Phys.* 65, 535–547 (2006)
4. Yezzi, A., Zollei, L., Kapur, T.: A variational framework for integrating segmentation and registration through active contours. *Medical Image Analysis* 7, 171–185 (2000)
5. Chen, X., Brady, M., Rueckert, D.: Simultaneous segmentation and registration for medical image. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) *MICCAI 2004*. LNCS, vol. 3216, pp. 663–670. Springer, Heidelberg (2004)
6. Leventon, M.E., Faugeras, O., Grimson, W.E.L., Wells III, W.M.: Level Set Based Segmentation with Intensity and Curvature Priors. In: *Proc. MMBIA 2000*, pp. 4–11 (2000)
7. Chan, T., Vese, L.: Active Contours Without Edges. *IEEE Transactions on Image Processing* 10(2), 266–277 (2001)

Anatomically-Driven Soft-Tissue Simulation Strategy for Cranio-Maxillofacial Surgery Using Facial Muscle Template Model

Hyungmin Kim¹, Philipp Jürgens², Lutz-Peter Nolte¹, and Mauricio Reyes¹

¹ Institute for Surgical Technology and Biomechanics, University of Bern,
Stauffacherstrasse 78, 3014 Bern, Switzerland
hyungmin.kim@istb.unibe.ch

² Department of Cranio-Maxillofacial Surgery, University Hospital Basel,
Spitalstrasse 21, 4031 Basel, Switzerland

Abstract. We propose a computationally efficient and bio-mechanically relevant soft-tissue simulation method for cranio-maxillofacial (CMF) surgery. A template-based facial muscle reconstruction was introduced to minimize the efforts on preparing a patient-specific model. A transversely isotropic mass-tensor model (MTM) was adopted to realize the effect of directional property of facial muscles in reasonable computation time. Additionally, sliding contact around teeth and mucosa was considered for more realistic simulation. Retrospective validation study with post-operative scan of a real patient showed that there were considerable improvements in simulation accuracy by incorporating template-based facial muscle anatomy and sliding contact.

Keywords: soft-tissue simulation, mass-tensor model, cranio-maxillofacial surgery.

1 Introduction

There has been a comprehensive amount of studies on facial soft-tissue modeling, which can be classified into three categories depending on the computational method: mass-spring model (MSM), finite-element model (FEM), and mass-tensor model (MTM). MSM was widely accepted for real-time application due to its computational efficiency and geometric simplicity [1][2][3], whereas FEM has strong bio-mechanical relevance at the cost of long calculation time [4][5][6]. MTM was originally developed to find a golden mean between speed and accuracy [7], later it was extended to non-linear, anisotropic elasticity [8]. Recently, linear MTM was successfully applied to cranio-maxillofacial (CMF) surgeries, and validated with 10 clinical cases both in quantitative and qualitative ways [9]. However, most of the previous works have failed to be seamlessly integrated into the clinical workflow in two aspects. First, the simulation accuracy was not enough for surgeons to realize delicate soft-tissue variation around the error-sensitive regions. Second, it requires laborious pre-processing work and long computation time which cannot be acceptable in daily clinical practice.

In this study, we propose a method for accurate and computationally efficient soft-tissue simulation after CMF intervention, while respecting the current clinical workflow. To overcome the limitation in accuracy, we hypothesize that patient-specific muscle characterization is necessary. We propose an efficient way to incorporate patient-specific facial muscles by morphing a template model to patient data. Transversely isotropic MTM was adopted to consider the directional property of facial muscles, and achieve reasonable computation times. Two different template muscle models were evaluated to measure the sensitivity of muscle template on the simulation result. Additionally, modeling sliding contact around teeth and mucosa was introduced to realize the delicate soft-tissue variation on this area. The simulation accuracy was retrospectively validated with a real clinical case.

2 Methods

The overview of our simulation pipeline is depicted in Fig. 1. Detailed explanations on featured steps will be given in the following sub-sections.

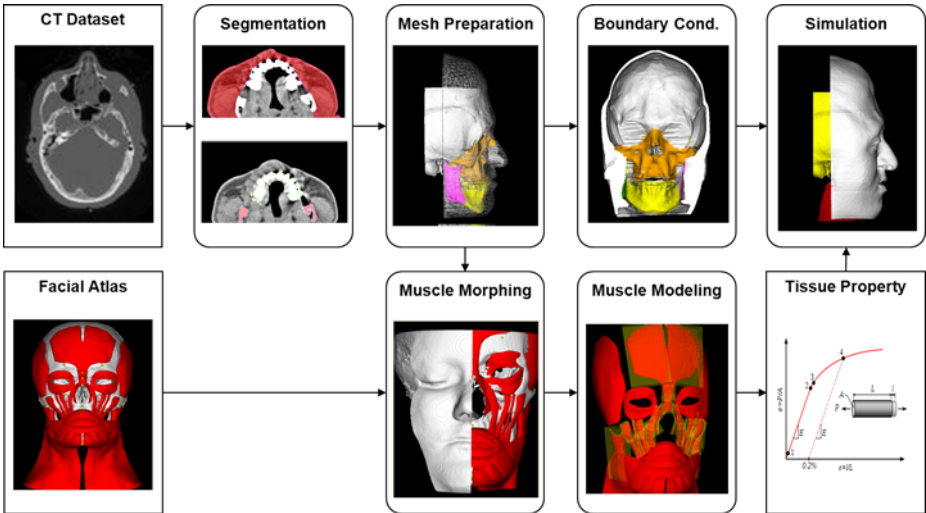


Fig. 1. The pipeline of the anatomically-driven facial soft-tissue simulation method using muscle template model

2.1 Patient-Specific Muscle Modeling

Since it is almost impossible to identify individual muscles from clinical Computed Tomographic (CT) scans, we propose to construct patient-specific facial muscles by morphing muscles from facial template models, as shown in Fig. 2.

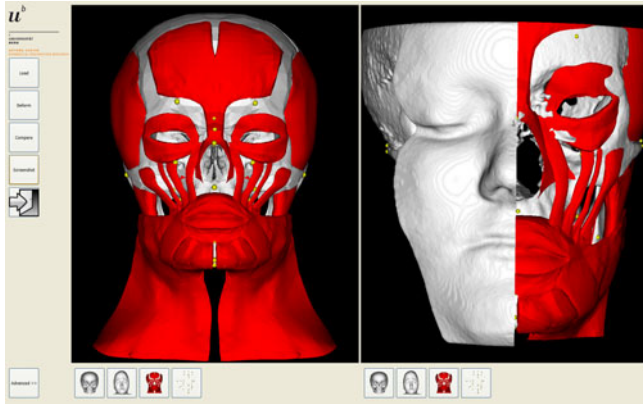


Fig. 2. Morphing of facial muscles from template model (*left*) to patient-specific anatomy (*right*), based on skull anatomical landmarks (*yellow dots*)

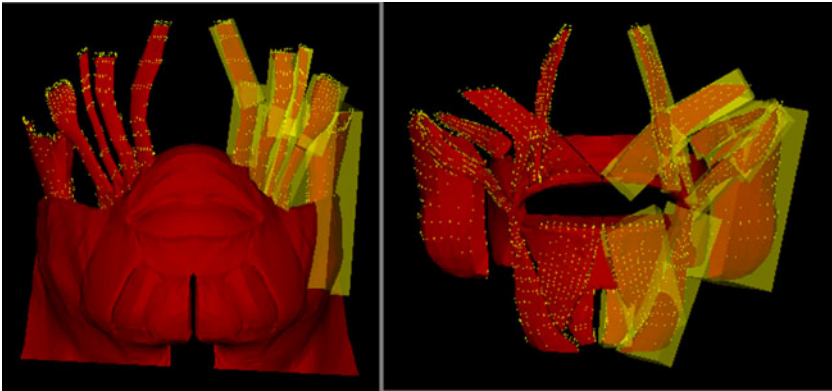


Fig. 3. Extraction of muscle direction using OBB for different muscle templates: muscle template I (*left*) and muscle template II (*right*)

For comparison, we evaluated two muscle template models, which consist of different representation of muscles, as shown in Fig. 3. The first template model is a synthetic three-dimensional craniofacial atlas built for educational purposes, and freely available online [10]. The second one has more detailed description of individual muscles, since it was obtained by manual segmentation from a high resolution Magnetic Resonance Imaging (MRI) data [11]. The morphing procedure was driven by landmark-based thin-plate-spline (TPS) algorithm [12]. We adopted thirty-two anatomical landmarks, which are commonly used for measuring skin depth in forensic science [13].

In order to obtain the direction of muscles, oriented-bounding boxes (OBB) were extracted for linear-type muscles, then the longitudinal direction of each bounding box was regarded as the direction of muscle for each segment.

2.2 Material Property Assignment

To date, there is no consensus in the literature on mechanical properties of facial soft-tissue, since different material parameters can be obtained depending on chosen characterization method and material model. We adopted Young’s modulus of muscle tissue along and across fiber from [14]. We also refer to the studies of estimating optimal facial tissue properties by comparing the simulated results with real post-operative data, which provides values in ranges [15] [16]. The material properties that we adopted for our simulation are shown in Table 1. Since there was no information available for directional Poisson ratio of muscle in the literature review, we assumed isotropic Poisson ratio for both directions.

In order to assign appropriate material properties, the intersecting portion of muscle in each tetrahedron needs to be determined. We adopted a random point sampling method inside of each tetrahedron [17], and simple geometrical tests on directional vectors to muscle surface.

2.3 Boundary Condition Assignment

We classified vertices on the volumetric mesh into four categories: fixed, joined, sliding and free node, as shown in Fig. 4. Fixed nodes are defined on the boundary area, such as the most posterior plane of volumetric soft-tissue. Joined nodes are defined on the interface between movable bony parts and surrounding soft-tissues. Sliding nodes are specially defined on the area around teeth. Sliding contact has already been tried to simulate the behavior of soft-tissues in breast augmentation [18].

In MTM, the internal force on the j -th node of the volumetric mesh can be given by

$$\mathbf{f}_j = \mathbf{K}_{jj}\mathbf{u}_j + \sum_{\forall k \in \Psi_j} \mathbf{K}_{jk}\mathbf{u}_k \quad (1)$$

where Ψ_j is the set of vertices connected to the j -th node, \mathbf{u}_j is the displacement vector on j -th node. \mathbf{K}_{jj} and \mathbf{K}_{jk} are the stiffness tensors for vertex and edge respectively, which are only dependent on material properties and the initial mesh configuration.

Table 1. Material properties used for simulation

	Young’s modulus (MPa)	Poisson Ratio
Fat	0.003	0.46
Muscle across fiber	0.79	0.43
Muscle along fiber	0.5	0.43

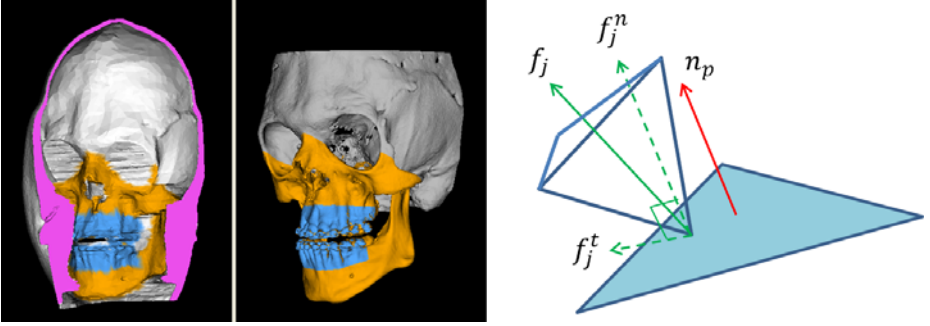


Fig. 4. Classification of volumetric nodes: fixed(pink), joined(orange), sliding(blue) and free(white) (*left*), force diagram on sliding contact surface (see Eq.2) (*right*).

Especially for sliding nodes, only the tangential component to closest contacting triangle will affect the movement along the contacting surface. No frictional forces are assumed. The tangential force \mathbf{f}_j^t can be calculated by

$$\mathbf{f}_j^t = \mathbf{f}_j - (\mathbf{f}_j \cdot \mathbf{n}_p)\mathbf{n}_p \quad (2)$$

where \mathbf{n}_p is the normal vector of nearest contacting triangle.

3 Results

The validation study was retrospectively performed with one clinical case called Pfeiffer-Syndrome. The patient was planned to be treated by Le-Fort III osteotomy with osteodistraction. The amount of maxillary advancement was about 12mm, while the rotational movement was around 3 degrees. Post-operative CT scan of the patient (in-plane resolution: 0.43mm, slice distance: 0.6mm) was regarded as a ground-truth to compare simulation accuracy using different tissue models: homogeneous and transversely isotropic elasticity. In addition, we evaluated the simulation results employing two different muscle templates, and sliding contact around teeth and mucosa area. The distance between each simulation and post-operative result was measured based on corresponding points using landmark-based TPS deformation and closest point matching. Fourteen facial anatomical feature points were manually selected for the TPS deformation.

As shown in Fig.5, the median of error was decreased by incorporating transversely isotropy of muscles: homogeneous (2.36mm), muscle template I (2.12mm). A further improvement was observed using the more detailed muscle template II (2.01mm). Higher simulation accuracy was obtained by incorporating sliding contact as additional boundary condition: homogeneous with sliding contact (1.95mm), muscle template I with sliding contact (1.07mm) and muscle template II with sliding contact (1.06mm). The improvement in simulation accuracy was around 1mm average, which corresponds to approximately 8% of

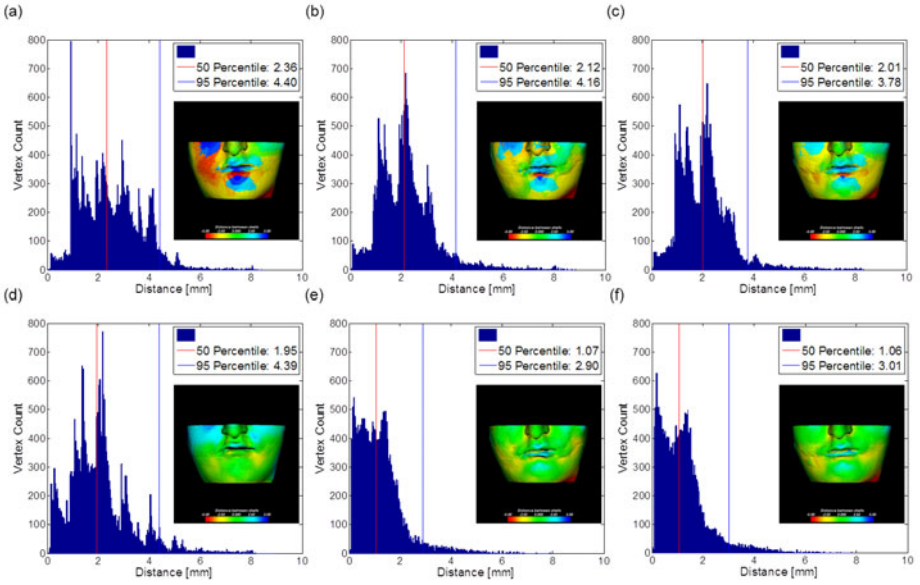


Fig. 5. Comparison of distance errors between simulations and post-operative result: histogram and color maps of (a) homogeneous, (b) muscle template I, (c) muscle template II, (d) homogeneous + sliding contact, (e) muscle template I + sliding contact, (f) muscle template II + sliding contact.

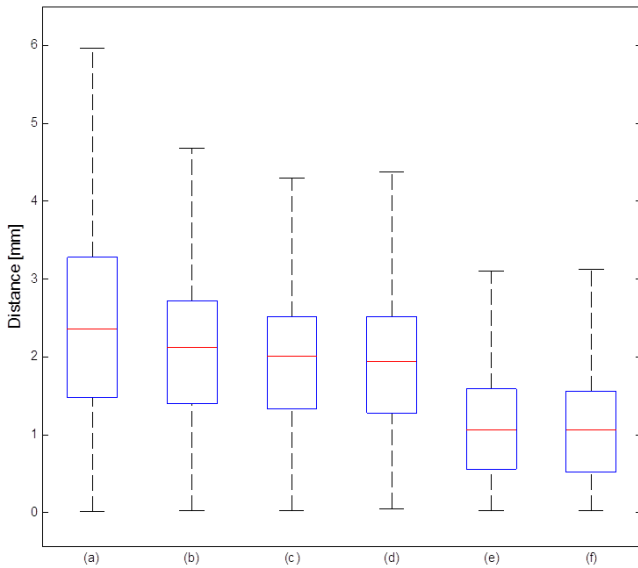


Fig. 6. Box plots of distance errors for the cases presented in Fig. 5; box encloses 25th and 75th quartiles, center bar indicates median, whiskers stand for max./min. values excluding outliers

overall translational movement. This amount of improvements make a difference in the surgeon's perspective, since the improvements occur around nose and lip areas, which are the error-sensitive regions for CMF surgeons. The distribution of the improvements in accuracy can be qualitative assessed by the color-coded distance maps in Fig 5. Statistical relevance was evaluated by Wilcoxon test ($p < 0.05$), which confirmed that there were statistically significant differences in all the pairs of simulated results.

4 Conclusion

In this study, we developed an efficient patient-specific soft-tissue simulation method for CMF surgery, with emphasis on its integration into the clinical workflow. A template-based facial muscle morphing technique was proposed to minimize the efforts of segmenting muscle structures without requiring additional imaging modality. A transversely isotropic MTM was applied to realize the directional property of muscle in reasonable calculation time. The retrospective validation study with a real post-operative CT scan confirmed the hypothesis that improvements in accuracy are obtained by introducing template-based facial muscles. Additionally, the results showed that sliding contact was quite effective in increasing simulation accuracy, and so it should be considered in CMF soft-tissue simulations.

As future work, the accuracy of morphing template models to patient-specific anatomy needs to be validated with manual segmentation of the morphology, since it is known that considerable amount of individuality of facial muscle's configuration is present. Additionally, we plan to follow up on the experimental studies on facial tissue characterization to compensate the assumptions of directional properties of muscles in this model. Finally, the extension to non-linear MTM is anticipated to overcome the limited range of deformation provided by linear models, and moreover, the need for non-linear model has yet to be justified for CMF soft-tissue simulations.

Acknowledgments

This work was funded by the Swiss National Center of Competence in Research "Computer Aided and Image Guided Medical Interventions (Co-Me)", and the AO/ASIF Foundation, Davos, Switzerland.

References

1. Lee, Y., Terzopoulos, D., Walters, K.: Realistic modeling for facial animation. In: Proceedings of the 22nd annual conference on Computer graphics and interactive techniques, pp. 55–62. ACM, New York (1995)
2. Keeve, E., Girod, S., Kikinis, R., Girod, B.: Deformable modeling of facial tissue for craniofacial surgery simulation. *Computer Aided Surgery* 3(5) (1998)

3. Teschner, M., Girod, S., Girod, B.: Direct computation of nonlinear soft-tissue deformation. In: Proc. Vision, Modeling, Visualization, VMV 2000 (2000)
4. Koch, R.M., Gross, M.H., Carls, F.R., von Büren, D.F., Fankhauser, G., Parish, Y.I.H.: Simulating facial surgery using finite element models. In: Proceedings of ACM SIGGRAPH, pp. 421–428. ACM Press, New York (1996)
5. Gladilin, E.: Biomechanical modeling of soft tissue and facial expressions for craniofacial surgery planning. PhD thesis, Free University Berlin (2003)
6. Chabanas, M., Luboz, V., Payan, Y.: Patient specific finite element model of the face soft tissues for computer-assisted maxillofacial surgery. *Medical Image Analysis* 7(2), 131–151 (2003)
7. Cotin, S., Delingette, H., Ayache, N.: A hybrid elastic model for real-time cutting, deformations, and force feedback for surgery training and simulation. *The Visual Computer* 16(8), 437–452 (2000)
8. Picinbono, G., Delingette, H., Ayache, N.: Non-linear anisotropic elasticity for real-time surgery simulation. *Graphical Models* 65(5) (September 2003)
9. Mollemans, W., Schutyser, F., Nadjmi, N., Maes, F., Suetens, P.: Predicting soft tissue deformations for a maxillofacial surgery planning system: from computational strategies to a complete clinical validation. *Medical image analysis* 11(3), 282–301 (2007)
10. Smith, D.M., Olikier, A., Carter, C.R., Kirov, M., McCarthy, J.G., Cutting, C.B.: A virtual reality atlas of craniofacial anatomy. *Plastic and reconstructive surgery* 120(6), 1641–1646 (2007)
11. Barbarino, G., Jabarene, M., Trzewik, J., Mazza, E.: Physically Based Finite Element Model of the Face. In: Bello, F., Edwards, E. (eds.) ISBMS 2008. LNCS, vol. 5104, pp. 1–10. Springer, Heidelberg (2008)
12. Bookstein, F.: Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on pattern analysis and machine intelligence* 11(6) (1989)
13. Phillips, V.M., Smuts, N.A.: Facial reconstruction: utilization of computerized tomography to measure facial tissue thickness in a mixed racial population. *Forensic Science International* 83(1), 51–59 (1996)
14. Duck, F.A.: Physical properties of tissue: a comprehensive reference book. Academic Press, London (1990)
15. Zachow, S., Hierl, T., Erdmann, B.: A quantitative evaluation of 3D soft tissue prediction in maxillofacial surgery planning. In: Proc. of CURAC, pp. 75–79 (2004)
16. Mollemans, W., Schutyser, F., Nadjmi, N., Maes, F., Suetens, P.: Parameter Optimisation of a Linear Tetrahedral Mass Tensor Model for a Maxillofacial Soft Tissue Simulator. In: Harders, M., Székely, G. (eds.) ISBMS 2006. LNCS, vol. 4072, pp. 159–168. Springer, Heidelberg (2006)
17. Uesu, D., Bavoil, L., Fleishman, S., Shepherd, J., Silva, C.: Simplification of unstructured tetrahedral meshes by point sampling. In: Fourth International Workshop on Volume Graphics (2005)
18. Roose, L., De Maerteleire, W., Mollemans, W., Maes, F., Suetens, P.: Simulation of soft-tissue deformations for breast augmentation planning. In: Harders, M., Székely, G. (eds.) ISBMS 2006. LNCS, vol. 4072, pp. 197–205. Springer, Heidelberg (2006)

Dynamic Shape Instantiation for Intra-operative Guidance

Su-Lin Lee¹, Adrian Chung¹, Mirna Lerotic¹, Maria A. Hawkins²,
Diana Tait², and Guang-Zhong Yang¹

¹ Institute of Biomedical Engineering, Imperial College London, London, United Kingdom

² Royal Marsden Hospital, London, United Kingdom

{su-lin.lee,ajchung,m.lerotic,g.z.yang}@imperial.ac.uk

Abstract. Primary liver cancer and oligometastatic liver disease are one of the major causes of mortality worldwide and its treatment ranges from surgery to more minimally invasive ablative procedures. With the increasing availability of minimally invasive hepatic approaches, a real-time method of determining the 3D structure of the liver and its location during the respiratory cycle is clinically important. However, during treatment, it is difficult to acquire images spanning the entire 3D volume rapidly. In this paper, a dynamic 3D shape instantiation scheme is developed for providing subject-specific optimal scan planning. Using only limited planar information, it is possible to instantiate the entire 3D geometry of the organ of interest. The efficacy of the proposed method is demonstrated with both detailed numerical simulation and a liver phantom with known ground-truth data. Preliminary clinical application of the technique is evaluated on a patient group with metastatic liver tumours.

Keywords: shape modeling, shape instantiation, intra-operative guidance, patient-specific deformation analysis, regression analysis.

1 Introduction

Dynamic shape instantiation can aid in the compensation of tissue deformation during surgery. With many intra-operative tasks involving complex anatomical pathways, only through the use of a detailed 3D deformation model, one can be certain to adjust and optimize the interventional approach. For liver tumour treatment where the tissue undergoes major translation and deformation due to respiratory motion, through the entire volume of the liver, localization of the tumours within becomes possible. Liver tumours or metastases are one of the major causes of death from cancer worldwide [1] and there is an increasing need for its early detection and minimally invasive treatment. For minimally invasive treatment and targeted therapy, such as those involving radio frequency (RF), focused ultrasound, cryogenic ablation and chemoembolisation, patient-specific tissue motion is an important factor to consider. The use of RF or focused ultrasound has the advantage of being performed within the magnetic resonance (MR) scanner, without patient re-positioning taking place, but methods such as laser ablation, which involve the insertion of needles, require the

imaging of patients under different imaging modalities. Using ultrasound rather than computed tomography (CT) prevents a radiation dose to the patient but is difficult to use for needle guidance. Although the benefit of using ultrasound imaging for intra-operative guidance is fully appreciated, the acquisition of the entire 3D structure in real-time is not feasible for routine clinical use. The relatively low resolution for 3D ultrasound prevents the building of accurate models of the liver and therefore another rapid method of 3D shape instantiation is required.

The use of a sparse set of 2D images for re-creating 3D models has been attempted previously by using radial basis functions [2] and statistical shape models [3]. Statistical shape models have also been used to determine the femur surface from sparse ultrasound points [4]. These methods are promising but both require a number of ultrasound images and significant processing. The reduction of the number of X-ray imaging planes used along with a template model [5] and a 2D-3D registration method based on a nearest-neighbor mapping operator and 2D thin-plate splines based deformations [6] have been investigated to instantiate a femur shape. This limitation on the number of planes acquired reduces the radiation dose to the patient. To release the restriction of the planes to be orthogonal to each other, a method for optimal scan planning was developed for MR scanning of soft tissue [7]. While the technique allows for any number of planes to be acquired, the planes found may not be suitable for ultrasound image acquisition due to their orientation. Instead of scan planes, the use of a surface trace to predict cardiac motion has also been investigated [8]. The current limitations are unable to generalize to any imaging modality, especially considering the limited viewing depth of ultrasound. A technique is required whereby only a very limited set of points is able to instantiate the geometry of the organ of interest rapidly for intra-operative guidance.

In this paper, a dynamic shape instantiation method is proposed utilizing an *a priori* model of the organ of interest and image acquisition using only limited imaging planes. An image collected along the optimal scan plane is pre-segmented and is used as input to a previously trained partial least squares regression (PLSR) to determine the output of an entire liver surface at arbitrary respiratory positions. The method was validated on both a phantom liver simulating realistic respiratory deformation and in one healthy subject *in vivo*. Initial clinical application is demonstrated on four patients with metastatic liver tumours. The geometry of their livers differs from normal subjects and highlights the need for subject-specific modeling.

2 Methods

2.1 Dynamic Shape Instantiation

For the proposed shape instantiation scheme, a statistical shape model was first built to examine the principal modes of variation of the training set. To find a single plane from the model, for example, the first mode of variation was examined and a subset of the points derived [7]. This subset included points that most affected the first mode of variation. In the liver coordinates, a plane was fitted to this point subset, with the constraint that the plane covers the motion of those points as defined by the first mode of variation. This constraint can be enforced by gathering the points of the subset at

each instance of the mode of variation and fitting the plane. For phantom and *in vivo* validation as described below, the plane is accessed by an ultrasound probe. A single optimal scan plane was determined for each of the datasets described below. The method can be extended to any number of optimal planes; we examine the use of multiple planes versus a single plane with the patient dataset to demonstrate the basic principle of the proposed shape instantiation scheme.

By using the optimal scan plane(s), the few points that most affect the shape of the liver can be used to instantiate the shape. Partial least squares regression (PLSR) is a statistical method that relies on a training set of input data X and its corresponding output data Y ; these variables are projected into a new space (finding new axes, similar to PCA) and the method finds the relationship between them:

$$\begin{bmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_M \end{bmatrix} = TDQ^T = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} B_{PLS} \quad (1)$$

In Eq. 1 and Fig. 1, T and U are the extracted score vectors, Q is the loading matrix on Y , and D is a matrix with the regression weights stored on the diagonal. B_{PLS} is the resultant matrix that can be calculated and when multiplied by X , estimates the predicted output \hat{Y} .

In this application, the dependents are the variables defining the surface of the liver (a 3D mesh) while the predictors are a very small subset of points on the surface as defined by the real-time imaging method (here validated using ultrasound).

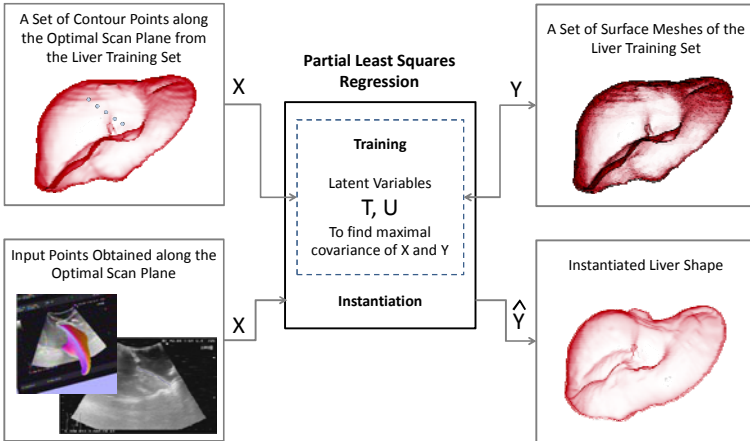


Fig. 1. An overview of the proposed dynamic shape instantiation scheme. The input X consists of contour points along the optimal scan plane from output Y , the liver meshes at various time points used in the training set. To determine the shape of the liver at a particular time point, input in the form of points obtained along the optimal scan plane using any imaging modality will result in instantiation of the liver geometry.

The method is described in Fig 1. With this method, a very limited amount of data acquired *in vivo* can be used to instantiate the entire organ surface for use during patient treatment.

2.2 Phantom Validation

For validating the accuracy of the proposed shape instantiation method, a silicone model of the internal organs was built by casting the plastic organs from a three-quarter sized teaching model. For the main body, a hollow plastic mannequin was used. As the liver undergoes respiratory motion and deformation, the lungs were sliced horizontally to accommodate inserts that would imitate its expansion during respiration. Seven 5mm thick foam board inserts were put into each lung, providing for 8 respiratory positions, and thus the corresponding liver deformation induced.

A CT scan of the phantom in all its respiratory positions was acquired using a Siemens 64 slice SOMATOM™ Sensation CT Scanner with the volume consisting of images with $0.77 \text{ mm} \times 0.77 \text{ mm}$ in-plane resolution and 1mm slice separation. Segmentation and 3D model building were performed in Analyze (AnalyzeDirect, Inc, Overland Park, KS, USA).

For validation, ultrasound images were acquired using the ALOKA prosound $\alpha 10$ system (Aloka Co. Ltd, Tokyo, Japan). A passive infrared tracker was affixed to the handheld ultrasound transducer to enable recording of spatial position and orientation of the imaging plane via the NDI Polaris (Northern Digital, Inc, Waterloo, ON, Canada). The rigid transformation between the CT coordinate frame of the phantom and the Polaris tool tracking coordinates was established by locating three known landmarks on the phantom liver in both frames of reference. The ultrasound image was captured using a PC video capture card connected to the S-video output feed from the ALOKA.

The silicone phantom was filled with water and the optimal plane was targeted for ultrasound imaging at all respiratory gates. The surface of the phantom liver in the ultrasound images was delineated using a semi-automated method using dynamic snakes [9] to find the path of maximal ultrasound response between two manually chosen points. The resulting points were converted to the CT coordinate frame to be used for prediction, shape instantiation, and validation.

2.3 *In Vivo* Validation and Patient Studies

For *in vivo* validation of the shape instantiation scheme with known ground-truth, one normal subject was first used. The method was then applied to a group of patients for further assessment of what may be used *in vivo*. One normal female subject was scanned in a GE 3T Discovery MR750 MR scanner. A LAVA (Liver Acquisition with Volume Acquisition) imaging sequence of the liver (TR = 4.26 ms, TE = 1.94 ms, $0.664 \text{ mm} \times 0.664 \text{ mm}$ in-plane resolution, 4mm slice thickness) was acquired at five breathhold positions. Markers (oil capsules) were fixed on the skin surface over the liver during the scan; the positions were used later for registration during the ultrasound image acquisition.

Four patients (3 male, 1 female, mean age 66 ± 8) were scanned twice, a fortnight apart, on a Philips Intera 1.5T MR scanner using a T1 weighted free-breathing sequence (TR = 7.83 ms, TE = 2.24 ms, $3.5\text{mm} \times 3.5\text{mm}$ in-plane resolution, 4.5mm slice thickness). This sequence took approximately 1.2s per volume of 45 slices and 60 sequences were collected. The livers from the initial scan were segmented and modeled at full inhalation and full exhalation.

Image constrained finite element modeling [10] was used to model the motion of the liver throughout the full breathing cycle and construct a series of liver meshes at different time points as only end-inspiration and end-expiration surface meshes were available. An end-expiration surface mesh was used to construct a 3D tetrahedral mesh of the liver in Gmsh [11]. A non-linear finite element model (FEM) was implemented in the Open Source SOFA framework [12], chosen for its emphasis on real-time medical simulations. The material was assumed to be isotropic and elastic, and was described by the generalized Hooke's law

$$\varepsilon_{x,y,z} = \frac{1}{E} \left[\sigma_{x,y,z} - \nu (\sigma_{y,z,x} + \sigma_{z,x,y}) \right] \quad (2)$$

where ε_i , and σ_i denote the strain and stress in the x, y and z direction. Young's modulus and Poisson's ratio are given by E and ν . The values used in the paper were $E = 640\text{Pa}$ and $\nu = 0.3$ [13]. Correspondent end-expiration and end-inspiration surface meshes extracted from the MRI data were used as constraints to build a non-linear elastic motion model of liver. The movement of the liver tissue was represented by a time series of 3D tetrahedral meshes given at different time points throughout the respiratory cycle. The resulting series of livers were used to build a statistical shape model and find the optimal plane. The livers from the second scans were used to obtain the optimal scan data for shape instantiation.

The optimal plane that could be used for real-time shape instantiation was also established for the human subject. By attaching a passive infrared marker to one of the oil capsule locations, respiratory gating information was also recorded during image acquisition that could be synchronized with the MRI-derived motion model for subsequent validation. Surface features on the liver were located manually for select frames which were converted into the MRI reference frame using the NDI Polaris tracking information.

For the patient data, the plane and model were found and built using the MR data resulting from the initial scan while the optimal surface contours were determined from the second scan. The patients all had metastatic liver tumours and two had undergone a liver resection, highlighting the need for patient-specific models and shape instantiation. Both one and two optimal scan planes were investigated for each patient data set. All the acquired data was used along with the trained regression models to instantiate the entire liver at the corresponding point in the respiratory cycle. Validation was performed on the phantom and single female *in vivo* livers with the patient livers showing the application of the proposed method. Errors were

calculated as a vertex to closest point on the surface measure and the ground truth was the meshes created from the scans during deformation.

3 Results

Fig. 2a shows a CT image of the silicone phantom along with the inserts in the lung used to simulate respiratory motion along with the model of the silicone liver. Ultrasound imaging of the silicone phantom effectively mimicked the properties of human tissue. The ultrasound images along the optimal scan plane and the surface meshes are also shown for the phantom (Figs. 2b and 2c) and the *in vivo* human liver (Figs. 2d and 2e).

For the silicone liver phantom, Fig. 2f displays the mean and standard deviation of the errors between the instantiated shapes from the ultrasound points to the original meshes found using CT imaging. Results are shown for each of the respiratory positions simulated with the lung inserts. For the *in vivo* liver from the normal female subject, at end-inspiration, the mean error was found to be $3.11\text{mm} \pm 2.12\text{mm}$ while at end-expiration, the mean error was $3.15\text{mm} \pm 2.30\text{mm}$.

As the patient liver scans do not cover the same regions of the liver, the instantiation was mainly used to validate the method for localization of the liver. Fig. 3 shows the results from all of the patients, where the error is shown between the instantiated shapes to the reference shapes. It is evident that the method is able to adapt to the location of the liver as well as the general shape of it as it changes through the respiratory cycle. The use of both one and two scan planes was investigated and with the exception of one liver set, the use of more input data results in better results. The outlier may be explained by the MR imaging not covering the same region of liver between scans.

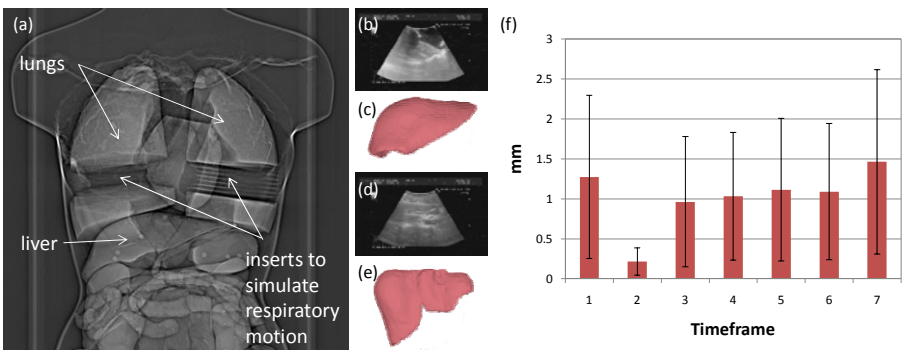


Fig. 2. (a) 3D reconstruction of the silicone phantom from CT. (b) An ultrasound image along the optimal scan plane of the silicone liver phantom and (c) the corresponding mesh. (d) An ultrasound image along the optimal scan plane of the *in vivo* study and (e) the corresponding liver mesh. (f) The mean and standard deviation of the errors of the phantom liver instantiation.

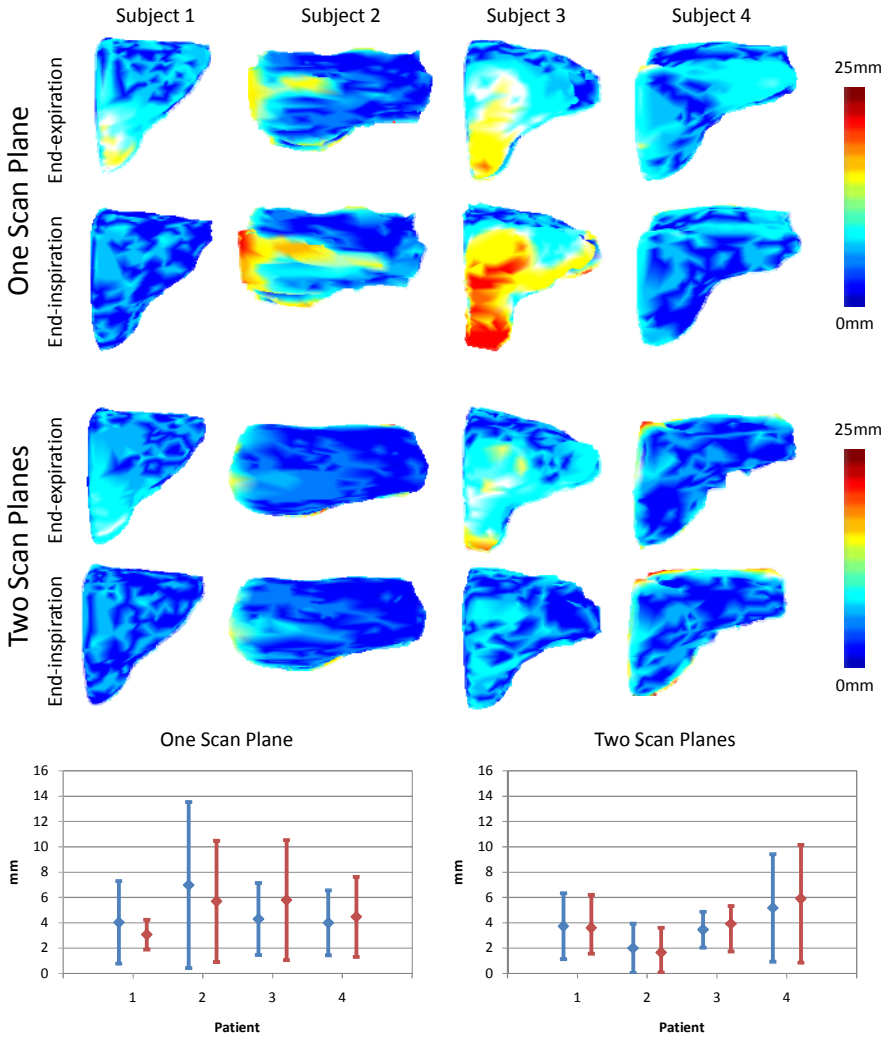


Fig. 3. Results of the proposed instantiation on patient data with one scan plane versus two scan planes. The colours indicate the error from the instantiated shape points to the surface of the original reference shape. Underneath, the mean and standard deviation for end-expiration (blue) and end-inspiration (red) are shown.

4 Conclusion

In conclusion, a rapid method for liver surface instantiation and localization has been introduced in this paper. Using a prior model of the liver derived from CT or MR imaging, and partial least squares regression for shape instantiation, the input can be only a few points on the liver surface as found via an optimal scan plane. It was also

shown that using data from multiple planes improves the instantiation. With a 3D model of the liver available, tissue motion may be compensated for intra-operative guidance. It should be noted that the proposed method is not limited to real-time data acquisition via ultrasound. If the treatment were to take place in an MR scanner, for example, a single MR slice may be used as the input to the instantiation.

Acknowledgments. The authors wish to thank Trupti Patel at the Royal Brompton Hospital for providing the CT scan of the phantom and Wady Gedroyc and his team at St Mary's Hospital for the normal subject MR images.

References

1. Bosch, F.X., Ribes, J., Díaz, M., Cléries, R.: Primary liver cancer: Worldwide incidence and trends. *Gastroenterology* 127, S5–S16 (2004)
2. Cool, D., Downey, D., Izawa, J., Chin, J., Fenster, A.: 3D prostate model formation from non-parallel 2D ultrasound biopsy images. *Medical Image Analysis* 10, 875–887 (2006)
3. Barratt, D.C., Chan, C.S.K., Edwards, P.J., Penney, G.P., Slomczykowski, M., Carter, T.J., Hawkes, D.J.: Instantiation and registration of statistical shape models of the femur and pelvis using 3D ultrasound imaging. *Medical Image Analysis* 12, 358–374 (2008)
4. Rajamani, K.T., Styner, M.A., Talib, H., Zheng, G., Nolte, L.P., Ballester, M.A.G.: Statistical deformable bone models for robust 3D surface extrapolation from sparse data. *Medical Image Analysis* 11, 99–109 (2007)
5. Filippi, S., Motyl, B., Bandera, C.: Analysis of existing methods for 3D modelling of femurs starting from two orthogonal images and development of a script for a commercial software package. *Computer Methods and Programs in Biomedicine* 89, 76–82 (2008)
6. Zheng, G., Gollmer, S., Schumann, S., Dong, X., Feilkas, T., González Ballester, M.A.: A 2D/3D correspondence building method for reconstruction of a patient-specific 3D bone surface model using point distribution models and calibrated X-ray images. *Medical Image Analysis* 13, 883–899 (2009)
7. Lee, S.-L., Horkaew, P., Caspersz, W., Darzi, A., Yang, G.-Z.: Assessment of Shape Variation of the Levator Ani With Optimal Scan Planning and Statistical Shape Modeling. *Journal of Computer Assisted Tomography* 29, 154–162 (2005)
8. Ablitt, N.A., Jianxin, G., Keegan, J., Stegger, L., Firmin, D.N., Guang-Zhong, Y.: Predictive cardiac motion modeling and correction with partial least squares regression. *IEEE Transactions on Medical Imaging* 23, 1315–1324 (2004)
9. Geiger, D., Gupta, A., Costa, L.A., Vlontzos, J.: Dynamic Programming for Detecting, Tracking, and Matching Deformable Contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17, 294–302 (1995)
10. Lerotic, M., Lee, S.-L., Keegan, J., Yang, G.-Z.: Image constrained finite element modelling for real-time surgical simulation and guidance. In: *Proceedings of the Sixth IEEE international Conference on Symposium on Biomedical Imaging: From Nano To Macro*, Boston, Massachusetts, USA, pp. 1063–1066 (2009)
11. Geuzaine, C., Remacle, J.-F.: Gmsh, <http://geuz.org/gmsh/>
12. Allard, J., Cotin, S., Faure, F., Bensoussan, P.-J., Poyer, F., Duriez, C., Delingette, H., Grisoni, L.: SOFA - an Open Source Framework for Medical Simulation. In: *Medicine Meets Virtual Reality (MMVR'15)*, pp. 13–18 (2007)
13. Yeh, W.-C., Li, P.-C., Jeng, Y.-M., Hsu, H.-C., Kuo, P.-L., Li, M.-L., Yang, P.-M., Lee, P.H.: Elastic modulus measurements of human liver and correlation with pathology. *Ultrasound in medicine & biology* 28, 467–474 (2002)

Dynamic Guidance for Robotic Surgery Using Image-Constrained Biomechanical Models

Philip Pratt, Danail Stoyanov,
Marco Visentini-Scarzanella, and Guang-Zhong Yang

Institute of Biomedical Engineering
Imperial College of Science, Technology and Medicine, London SW7 2AZ, UK
{p.pratt,danail.stoyanov,mv402,g.z.yang}@imperial.ac.uk

Abstract. The use of physically-based models combined with image constraints for intraoperative guidance is important for surgical procedures that involve large-scale tissue deformation. A biomechanical model of tissue deformation is described in which surface positional constraints and internally generated forces are derived from endoscopic images and preoperative 4D CT data, respectively. Considering cardiac motion, a novel technique is presented which minimises the average registration error over one or more complete cycles. Features tracked in the stereo video stream provide surface constraints, and an inverse finite element simulation is presented which allows internal forces to be recovered from known preoperative displacements. The accuracy of surface texture, segmented mesh and volumetrically rendered overlays is evaluated with detailed phantom experiments. Results indicate that by combining preoperative and intraoperative images in this manner, accurate intraoperative tissue deformation modelling can be achieved.

1 Introduction

Image guidance plays an important role in the deployment of minimally invasive surgery for complex procedures. It needs to be capable of seamlessly integrating preoperative and intraoperative images of dynamic anatomical structure, thereby enhancing the surgeon's view with detailed navigational cues, attaining a high level of spatiotemporal accuracy. To this end, physically-based modelling has attracted significant interest since during surgery, the exposed tissue surface for co-registration with preoperative data is typically small, and the approach allows for the effective incorporation of biomechanical constraints. The primary target application for this work is vessel identification and harvest prior to TECAB, but it is also applicable to image-guided liver resection and partial nephrectomy for the removal of metastases, in the presence of respiratory motion.

Dynamic guidance systems present significant technological challenges, but several notable advances have already been made. Sun *et al.* [1] describe a method for modelling tissue retraction during image-guided neurosurgery. A poroelastic brain model is driven by the stereoscopically-measured motion of a retractor to produce a full volume displacement field, which is used to update the preoperative MR images. Szpala *et al.* [2] derive a sequence of linear free-form

deformations from preoperative CT images, which are applied subsequently to a high-quality heart mesh model generated at end-diastole. Temporal synchronisation is driven by an electrocardiographic trigger signal, whereas spatial registration of optically-tracked phantom casing and endoscope coordinate systems is achieved using linear least squares error minimisation. Figl *et al.* [3] build a 4D B-spline motion model of the beating heart from multiple coronary CT phases. Temporal alignment with endoscopic video is obtained using an ECG signal. Two methods are proposed for spatial alignment: the first considers model registration to 3D feature positions tracked in the endoscopic feed, and the second employs photo-consistency as a similarity metric.

Su *et al.* [4] propose a near real-time 3D-CT to stereo endoscopic video registration method for robot-assisted partial nephrectomy. Selected features on the kidney surface are tracked from one video frame to the next, and are registered to a preoperatively-generated model of the anatomy using a modified iterative closest point technique. Lerotic *et al.* [5] introduce an image-constrained linear finite element simulation of the lung parenchyma capable of real-time performance. A model for the motion of points on the lung surface, incorporating amplitude, period and phase parameters, is fitted to landmarks acquired from preoperative MR scan data at full inhale and exhale. The motion model parameters are optimised using golden section search, which is then used to provide surface constraints for the finite element simulation which, in turn, determines the overlay deformation.

However, existing methods fail to combine preoperative and intraoperative images in a way that is simultaneously accurate for surface texture, segmented mesh and volumetrically rendered overlay regimes. Furthermore, initial spatiotemporal registrations typically do not exploit positional data acquired over an extended period of time. This paper proposes a method which achieves these goals, and promises to enhance the surgeon’s sensory experience, reduced initially as a direct consequence of minimally invasive robotic access. The method is also applicable to the wider context that includes preoperative rehearsal and intraoperative simulation of tool-tissue interaction, where the surface constraints are derived or enhanced, respectively, by known instrument positions. The method has also been implemented to exploit GPU acceleration and is capable of real-time operation, thus facilitating practical clinical use.

2 Material and Methods

2.1 Stereo Feature Tracking

In this study, the exposed tissue surface geometry is captured by the *da Vinci* stereo endoscope, with camera calibration following Zhang [6]. The motion tracking technique described by Stoyanov *et al.* [7] is used to recover tissue surface motion. Robust performance is ensured by using a combination of landmarks,

including maximally stable extremal regions (MSER) and traditional gradient-based image features.

2.2 Spatiotemporal Registration

The problem is formulated as a least-squares distance optimisation problem over projected 3D fiducial positions segmented from scan data and corresponding points recovered from the stereo video sequence. Rather than perform the optimisation at a single point in time, squares of the residual distances are summed over one or more complete phases of cardiac motion. It then remains to find the minimal overall sum by varying the initially unknown temporal phase difference between the first 4D scan phase and the first video frame. Symbolically, if fiducials are indexed by i and video frames by j , and λ is the temporal phase shift, and $\theta_x, \theta_y, \theta_z, t_x, t_y, t_z$ are the rotations and translations, respectively, that parameterise the transformation \mathbf{R} from the scan coordinate system to the camera coordinate system, then the quantity to be minimised is

$$\sum_{i,j} \left[\|\mathbf{p}^L(\mathbf{R}\mathbf{s}_i(\tau_j + \lambda)) - \mathbf{t}^L(\tau_j)\|^2 + \|\mathbf{p}^R(\mathbf{R}\mathbf{s}_i(\tau_j + \lambda)) - \mathbf{t}^R(\tau_j)\|^2 \right] \quad (1)$$

where \mathbf{p}^L and \mathbf{p}^R are the left and right 2D vector-valued projection functions, including lens distortion, determined by the stereo camera calibration stage. The \mathbf{s}_i represent the 3D fiducials coordinates derived from scan data, and the \mathbf{t}^L and \mathbf{t}^R are the 2D coordinates of the corresponding left and right points in the video sequence. They are evaluated at the discrete points in time τ_j corresponding to each video frame. Scan phases and video frames are assumed to extend through time in a cyclic manner, and linear interpolation in time is used to evaluate fiducials coordinates lying between scan phases.

2.3 Finite Element Simulation

In previous work, Miller *et al.* [8] describe the total Lagrangian explicit dynamics (TLED) algorithm. An efficient GPU implementation is used here to model deformation of the heart resulting from surface positional constraints and internal forces implied by the original 4D scan motion. In contrast to the *updated* Lagrangian formulation, this method expresses stress and strain measures in terms of the reference configuration. As such, many quantities can be either completely or partially precomputed, and together with an explicit integration scheme, the algorithm enables nonlinear finite element simulation using relatively complex meshes at interactive rates.

Using the notation of Bathe [9], the equations of motion for a mesh comprising N nodes, written in semi-discrete form and expressed in terms of the displacements $\mathbf{U} = [u^0 u^1 \dots u^{3N-1}]^\top$ from the initial configuration, are

$$\mathbf{M} \ddot{\mathbf{U}} + \mathbf{C} \dot{\mathbf{U}} + {}_0^t \mathbf{F} = {}^t \mathbf{R} \quad (2)$$

where ${}^t\dot{\mathbf{U}}$ and ${}^t\ddot{\mathbf{U}}$ are the velocity and acceleration vectors, respectively, and \mathbf{M} is the constant lumped mass matrix. The damping matrix $\mathbf{C} = \alpha\mathbf{M}$ is assumed to be proportional to the mass matrix, where α is the Rayleigh damping coefficient. The vectors ${}^t_0\mathbf{F}$ and ${}^t\mathbf{R}$ represent the nodal reaction forces equivalent to element stresses, and externally applied forces, respectively.

Subsequent time discretisation, with step size Δt , and application of an explicit central difference integration scheme gives the following update rule

$${}^{t+\Delta t}\mathbf{U} = \frac{2}{2 + \alpha\Delta t} \left[\frac{\Delta t^2}{\mathbf{M}} ({}^t\mathbf{R} - {}^t_0\mathbf{F}) + 2{}^t\mathbf{U} + \left(\frac{\alpha\Delta t}{2} - 1\right) {}^{t-\Delta t}\mathbf{U} \right] \quad (3)$$

The constitutive law employed is that of a compressible neo-Hookean material. The following constants are used in the simulation: Young's modulus $E = 3.0\text{E} + 03\text{ Pa}$; Poisson's ratio $\nu = 0.45$; material density $\rho = 1.0\text{E} + 03\text{ kg/m}^3$; and mass damping coefficient $\alpha = 7.5\text{E} + 01$. In this instance, the equations of motion are integrated using the conservatively small time step of $\Delta t = 0.0002$ seconds.

2.4 Mesh Generation and Temporal Smoothing

The heart phantom is segmented from the initial scan phase using edge-based snakes, as made available in the *ITK-SNAP* tool (see Yushkevich *et al.* [10]). Additional manual segmentation follows to make the simplification whereby the phantom is assumed to have a solid spherical topology. Subsequently, the high-resolution surface mesh, generated at the voxel level, is further simplified using quadric edge collapse decimation to a resolution suitable for interactive finite element simulation. Finally, tetrahedralisation is applied using *Gmsh* [11], and the resulting mesh is optimised for element quality.

The Image Registration Toolkit (see Rueckert *et al.* [12] and Schnabel *et al.* [13]) is used to determine the sequence of 3D tensor product B-spline deformations which maps the initial phase on to each subsequent phase in turn. These deformations are applied to the initial tetrahedral mesh to create a sequence of vertex displacements at discrete points in time. Finally, cyclic 1D B-splines over the time dimension are fitted through each vertex, both to smooth the motion implied from the scan sequence and to provide a means for obtaining phantom geometry at arbitrary positions in the cardiac cycle.

2.5 Application of Surface Constraints

Once a spatial and temporal registration has been established, tracked features in the initial video frame are associated with interpolated mesh node positions at the corresponding point in time. For each feature, the nearest node is identified in the collection of mesh surface nodes. These associations are maintained throughout the duration of the simulation. Each such node will have one or more feature associations, and as the latter displace over time, the average of their displacements is assigned to the node displacement, thereby enforcing a dynamic positional constraint. Typically, the simulation time step is smaller than the period of each video frame, so linear interpolation is applied to all positional

constraints over the duration of the frame. In addition, static regions in the original scan, comprising the heart phantom mounting, are segmented and meshed, and then used to locate surface mesh nodes that are to be held fixed with zero displacement during the simulation.

2.6 Force Recovery and Resolution

The method of force recovery and resolution described by Pratt [14] is used to incorporate cyclic organ motion derived from 4D scan data. Firstly, the forward update rule given by equation 3 is inverted, such that the externally applied forces can be expressed in terms of known displacements, as follows

$${}^t\mathbf{R} = \frac{\mathbf{M}}{\Delta t^2} \left[\left(1 + \frac{\alpha \Delta t}{2}\right)^{t+\Delta t} \mathbf{U} - 2{}^t\mathbf{U} - \left(\frac{\alpha \Delta t}{2} - 1\right)^{t-\Delta t} \mathbf{U} \right] + {}_0^t\mathbf{F} \quad (4)$$

Subsequently, in the absence of positional constraints, if the recovered forces were to be applied at the appropriate instants during forward simulation, then the original motion would result by construction. In the presence of positional constraints, however, the recovered forces must be resolved from the global coordinate systems into systems local to each mesh element. At a given node, the recovered force \mathbf{r} at that node is shared equally amongst the M elements to which it belongs. Symbolically, the contribution from each node in each element is written barycentrically as the weighted sum of adjacent element edge vectors \mathbf{b}_0 , \mathbf{b}_1 and \mathbf{b}_2 , as follows

$$\frac{1}{M} \mathbf{r} = w_0 \mathbf{b}_0 + w_1 \mathbf{b}_1 + w_2 \mathbf{b}_2 \quad (5)$$

By expressing these forces in terms of local geometry in this manner, they are made to act in the directions consistent with external constraint-induced deformation. In practice, the inverse and forward simulations are performed simultaneously to avoid storage of intermediate results. Forces are always recovered from the original displacements, and are expressed in terms of the original geometry. The forward simulation is then propagated using forces implied by the current geometry.

2.7 Overlay Deformation

As the current geometry is updated in time, the resulting new node positions are used to deform high-resolution overlay information in one of three ways. Firstly, high-resolution meshes segmented and built from the original CT scan data are embedded in the finite element mesh by associating each vertex in the former with its containing element, and precomputing their barycentric coordinates. Vertex update is then performed by transforming back to the global coordinate system in the context of the current geometry. Secondly, volumetrically rendered scan data is deformed by slicing each tetrahedral element individually, from back to front in planes perpendicular to the z-axis, and alpha-compositing

the results. The resulting overlays are presented using the *Inverse Realism* technique described by Lerotic *et al.* [15]. Finally, preoperatively determined surface landmarks and annotations are rendered as textures covering the finite element mesh, and are alpha-blended directly with the output video feed.

2.8 Validation

In order to assess 2D overlay accuracy when positional surface constraints are combined with recovered and resolved internal forces, the results of manual frame-by-frame video feature tracking are compared, considering the left stereo channel, with the positions of equivalent textured features rendered using the deforming geometry. In addition to the combined case, results are also evaluated for simulations where surface constraints and internal forces are applied individually. A similar approach is adopted for 3D validation, where simulation results are compared over multiple cardiac cycles against the spatial locations of volumetric features identified in the original scan data.

3 Results and Discussion

For quantitative assessment of the proposed method, detailed phantom experiments were conducted using a Chamberlain Group CABG beating heart phantom. In the first instance, the phantom was scanned at 90 bpm with a Philips 64-slice CT scanner, producing 20 uniformly spaced phases. Fiducial markers embedded in the outer layer of the phantom were manually segmented in each of the phases in order to recover the sequence of spatial locations. Using the corresponding left and right 2D stereo locations of five such fiducials in the captured *da Vinci* endoscope video stream, the optimal spatiotemporal registration produced by the procedure described in section 2.2 had average left and right errors of 0.4712 mm and 0.4667 mm, respectively. The output frames in figure 1 shows an original image captured from the *da Vinci* endoscope, the same frame blended with the spatiotemporally registered FEM mesh, and surface texture and internal structure overlay regimes.

The results shown on the left of table 1 compare 2D surface overlay accuracies against a typical feature manually tracked in the left video stream, considering the surface-only, force-only (i.e. volume) and combined constraint modes. Average and maximum errors, and root-mean-square deviations are determined over

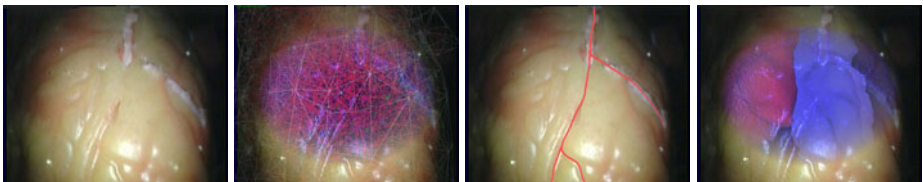
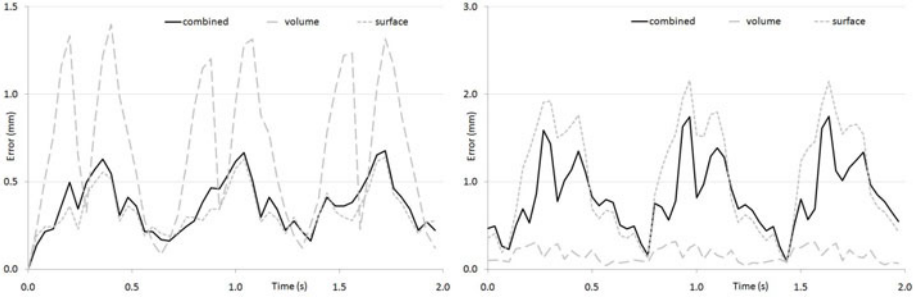


Fig. 1. Single *da Vinci* endoscope video frame, registered FEM mesh and overlays

Table 1. Projected surface texture and internal structure overlay errors

<i>error (mm)</i>	<i>surface overlay</i>			<i>internal overlay</i>		
	surface	volume	combined	surface	volume	combined
average	0.33	0.69	0.36	1.05	0.16	0.83
maximum	0.64	1.40	0.68	2.16	0.32	1.75
RMSD	0.36	0.80	0.40	1.21	0.18	0.92

**Fig. 2.** Overlay errors for surface texture (2D) and internal structure (3D) regimes**Table 2.** Relative analysis of fiducial deformation and overlay errors

<i>error (mm)</i>	fiducial #1	fiducial #2	fiducial #3	fiducial #4	fiducial #5	average
deformed	0.23	0.31	0.20	0.23	0.23	0.24
volume	0.25	0.30	0.20	0.23	0.33	0.26

a period covering three complete cardiac cycles. It can be seen that the application of surface constraints alone gives rise to the most accurate overlay, but that the inclusion of recovered and resolved forces degrades the performance only marginally. The left graph in figure 2 illustrates the manner in which overlay accuracy changes during this period.

Similarly, the results shown on the right of table 1 compare the location of a typical deforming feature identified in the original CT scan against the location of the same feature produced by barycentric interpolation of the point embedded in the FEM tetrahedral mesh, undergoing deformation implied by each of the constraint modes. As expect, simulation using only recovered and resolved forces produces the most accurate overlay. However, the combined constraint mode performs better than the application of surface constraints alone. The right graph in figure 2 illustrates changes in overlay accuracy during three consecutive cardiac cycles. Thus it can be seen that the combined constraint mode offers a good compromise over the surface-only and internal force-only alternatives, and results in a level of accuracy acceptable for dynamic guidance during interventions.

In-depth analysis of the different factors affecting the total overlay error reveals that a significant proportion can be attributed to the recovery of 3D B-spline deformations, as described in section 2.4. Table 2 compares the positions of fiducials manually segmented from each scan phase against the positions resulting from 3D B-spline deformation of the first phase, and also the full simulation results using the force-only constraint mode. Again, results are averaged over three complete cardiac cycles. They differ marginally, indicating that temporal smoothing and the simulation itself contribute relatively little to the total error.

4 Conclusion

This work proposes a new modelling framework for fusing preoperative tomographic and intraoperative endoscopic data in a dynamic environment, such that accurate overlays can be rendered in both surface texture and internal structure regimes. A physically-based, inverse finite element simulation transforms preoperative motion into a representation where it can be readily combined with tissue surface constraints. Furthermore, the method is cast in a setting that fulfils the key requirement of real-time operation, necessary to guarantee minimal latency in the surgeon's visual-motor feedback loop. Finally, a novel spatiotemporal registration technique is described which minimises errors over one or more complete cardiac cycles.

References

1. Sun, H., Kennedy, F.E., Carlson, E.J., Hartov, A., Roberts, D.W., Paulsen, K.D.: Modeling of brain tissue retraction using intraoperative data. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) MICCAI 2004, Part II. LNCS, vol. 3217, pp. 225–233. Springer, Heidelberg (2004)
2. Szpala, S., Wierzbicki, M., Guiraudon, G., Peters, T.M.: Real-time fusion of endoscopic views with dynamic 3-D cardiac images: A phantom study. *IEEE Transactions on Medical Imaging* 24(9), 1207–1215 (2005)
3. Figl, M., Rueckert, D., Hawkes, D., Casula, R., Hu, M., Pedro, O., Zhang, D.P., Penney, G., Bello, F., Edwards, P.: Image guidance for robotic minimally invasive coronary artery bypass. In: Dohi, T., Sakuma, I., Liao, H. (eds.) MIAR 2008. LNCS, vol. 5128, pp. 202–209. Springer, Heidelberg (2008)
4. Su, L.M., Vagvolgyi, B.P., Agarwal, R., Reiley, C.E., Taylor, R.H., Hager, G.D.: Augmented reality during robot-assisted laparoscopic partial nephrectomy: Toward real-time 3D-CT to stereoscopic video registration. *Urology* 73(4), 896–900 (2009)
5. Lerotic, M., Lee, S.L., Keegan, J., Yang, G.Z.: Image constrained finite element modelling for real-time surgical simulation and guidance. In: Proceedings of the IEEE International Symposium on Biomedical Imaging, pp. 1063–1066 (2009)
6. Zhang, Z.: A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(11), 1330–1334 (2000)
7. Stoyanov, D., Mylonas, G.P., Deligianni, F., Darzi, A., Yang, G.Z.: Soft-tissue motion tracking and structure estimation for robotic assisted MIS procedures. In: Duncan, J.S., Gerig, G. (eds.) MICCAI 2005, Part II. LNCS, vol. 3750, pp. 139–146. Springer, Heidelberg (2005)

8. Miller, K., Joldes, G., Lance, D., Wittek, A.: Total lagrangian explicit dynamics finite element algorithm for computing soft tissue deformation. *Communications in Numerical Methods in Engineering* 23, 121–134 (2007)
9. Bathe, K.J.: *Finite Element Procedures*. Prentice-Hall, Inc., Englewood Cliffs (1996)
10. Yushkevich, P.A., Piven, J., Cody Hazlett, H., Gimpel Smith, R., Ho, S., Gee, J.C., Gerig, G.: User-guided 3D active contour segmentation of anatomical structures. *Neuroimage* 31(3), 1116–1128 (2006)
11. Geuzaine, C., Remacle, J.F.: Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering* 79(11), 1309–1331 (2009)
12. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Non-rigid registration using free-form deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging* 18(8), 712–721 (1999)
13. Schnabel, J.A., Rueckert, D., Quist, M., Blackall, J.M., Castellano-Smith, A.D., Hartkens, T., Penney, G.P., Hall, W.A., Liu, H., Truwit, C.L., Gerritsen, F.A., Hill, D.L.G., Hawkes, D.J.: A generic framework for non-rigid registration based on non-uniform multi-level free-form deformations. In: Niessen, W.J., Viergever, M.A. (eds.) *MICCAI 2001*. LNCS, vol. 2208, pp. 573–581. Springer, Heidelberg (2001)
14. Pratt, P.: Image guidance and surgery simulation using inverse nonlinear finite element methods. In: Bello, F., Edwards, E. (eds.) *ISBMS 2008*. LNCS, vol. 5104, pp. 185–190. Springer, Heidelberg (2008)
15. Lerotic, M., Chung, A.J., Mylonas, G., Yang, G.Z.: pq -space based non-photorealistic rendering for augmented reality. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part II*. LNCS, vol. 4792, pp. 102–109. Springer, Heidelberg (2007)

Actin Filament Segmentation Using Spatiotemporal Active-Surface and Active-Contour Models

Hongsheng Li¹, Tian Shen¹, Dimitrios Vavylonis², and Xiaolei Huang¹

¹ Department of Computer Science & Engineering, Lehigh University, USA

² Department of Physics, Lehigh University, USA

Abstract. We introduce a novel algorithm for actin filament segmentation in a 2D TIRFM image sequence. We treat the 2D time-lapse sequence as a 3D image volume and propose an over-grown active surface model to segment the body of a filament on all slices simultaneously. In order to locate the two ends of the filament on the over-grown surface, a novel 2D spatiotemporal domain is created based on the resulting surface. Two 2D active contour models deform in this domain to locate the two filament ends accurately. Evaluation on TIRFM image sequences with very low SNRs and comparison with a previous method demonstrate the accuracy and robustness of this approach.

1 Introduction

Actin proteins spontaneously assemble into long polymers to build networks and bundles of filaments that are used by cells to move and change shape. An experimental method to study the kinetics of single actin filament growth *in vitro* is total internal reflection microscopy (TIRFM) [1], [2]. The information on growth kinetics is used to develop mathematical models that describe the dependence of actin filament elongation rate and fluctuations on the concentration of actin and cofactors [1], and the mechanisms that cells use to regulate the dynamics of the actin cytoskeleton [2]. In order to study kinetics based on image information, an essential step is to segment actin filaments in TIRFM image sequences.

There has been previous research on segmentation of actin filaments or microtubules (MT). Li *et al.* [3] proposed a slice-by-slice Stretch Open Active Contour model to segment and to track actin filaments in a time-lapse sequence. Hadjidemetriou *et al.* [4] minimized an image-based energy function to segment MTs using consecutive level sets methods. Saban *et al.* [5] automatically detected MT tips in the first frame and tracked tips by searching for the closest match in subsequent frames. Sargin *et al.* [6] located tip locations using second order derivative of Gaussian filtering and iteratively calculated geodesic paths to segment MT bodies. However, the above methods process an image sequence slice by slice and do not take advantage of the time-lapse sequence's temporal information.

Other methods of tracking filament tips employ Particle Filters (PF) [7], [8]. However, the two methods ignore the supporting information from filament bodies and may fail to locate a tip away from the filament body. To solve this problem, Li *et al.* [9] used an over-grown active contour model to segment the filament

body slice by slice and proposed a 1D PF to track filament tips across time. This method performs well in real time scenarios. Occasionally, this method generates some errors, such as segmenting two filaments as one, because it does not fully utilize all available temporal information: (i) it segments the filament body slice by slice and thus does not take into consideration of the location of filament bodies across time slices; and (ii) when tracking the tip in time t , only information prior to t ($1, \dots, t-2, t-1$) is considered. This is due to the PF's Bayesian tracking framework.

Since TIRFM image sequences are taken prior to data analysis, accuracy is more important than speed. In this paper, we treat a 2D time-lapse image sequence as a 3D spatiotemporal (2D spatial and 1D temporal) image volume. In such 3D domain, we use an over-grown active surface model to segment filament bodies. Given a reference point on resulting curves of all time slices, we can calculate the distances of all points along the resulting curves with respect to that reference point. An additional 2D spatiotemporal (1D spatial and 1D temporal) domain can be created by combining this distance-to-reference-point dimension with the time dimension. An active contour model in this 2D domain is then used to locate filament ends. In this proposed method, segmentation of the filament at time t depends on the information of the whole sequence ($1, \dots, t-1, t, t+1, \dots, n$). In evaluation on experimental images, this method shows more accurate segmentation performance over a previous method [3]. The overall framework of this method is illustrated in Fig. 1.

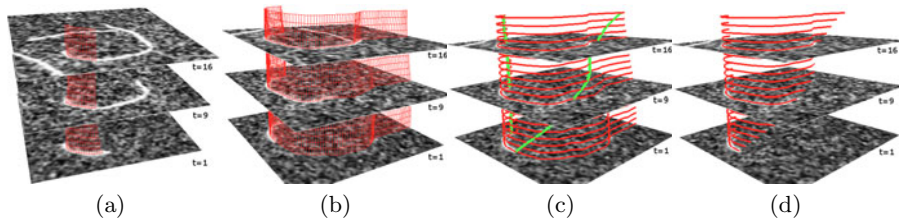


Fig. 1. An overview of the proposed method. (a) Initialization of the active surface. (b) The resulting over-grown active surface of (a) (from a different viewpoint). (c) The resulting over-grown curves (red) on all slices and two active contours (green) across temporal domain for locating filament ends. (d) Final segmentation results on all slices.

2 Spatiotemporal Active-Surface for Filament Body Segmentation

We treat the time-lapse 2D image sequence as a 3D spatiotemporal (2D spatial and 1D temporal) image volume, where each point has coordinates (x, y, t) , where x and y are the spatial coordinates, and t is the discrete time coordinate. One important observation of filament behavior is that filament bodies tend to remain static across time. Thus, if we connect points on the same filament across time, they form a smooth surface in the (x, y, t) domain. Based on this

observation, we create an open active surface model [10] not in a pure 3D spatial domain but a 3D spatiotemporal domain to segment the filament in all slices simultaneously (Fig. 1(b)).

2.1 Surface Initialization and Registration

Human interaction is necessary in this application since there is a limited number of filaments of interest in a sequence. Users specify a short line (2-3 clicks) close to the filament of interest on the first slice and use a 2D open active contour model [3] to fit only to a part of the filament. Then this initial short curve is used to fit a part of the same filament in all slices. The resulting curves on all slices are registered using the ICP method [11] to compensate for image drift [2] between contiguous frames. Since the drift between two contiguous frames is usually small, the ICP method generates satisfactory registration performance. We then create a grid mesh as the initial surface by connecting points on curves in all slices (Fig. 1(a)).

2.2 The Over-Grown Spatiotemporal Active-Surface Model

After being initialized, an active surface can be defined by a mapping from parameters p and q to the 3D spatiotemporal domain, $\mathbf{r} : \Omega = [0, 1] \times [0, 1] \rightarrow \mathbb{R}^3$,

$$(p, q) \rightarrow \mathbf{r}(p, q) = (x(p, q), y(p, q), t(p, q)). \quad (1)$$

We seek to minimize its overall energy E_{3D} :

$$E_{3D}(\mathbf{r}) = \int_{\Omega} \left[w_{10} \left\| \frac{\partial \mathbf{r}}{\partial p} \right\|^2 + w_{01} \left\| \frac{\partial \mathbf{r}}{\partial q} \right\|^2 + 2w_{11} \left\| \frac{\partial^2 \mathbf{r}}{\partial p \partial q} \right\|^2 + w_{20} \left\| \frac{\partial^2 \mathbf{r}}{\partial p^2} \right\|^2 + w_{02} \left\| \frac{\partial^2 \mathbf{r}}{\partial q^2} \right\|^2 + E_{ext}^{3D}(\mathbf{r}(p, q)) \right] dp dq, \quad (2)$$

where the internal energy is represented by the first 5 derivative terms where the elasticity is weighted by (w_{10}, w_{01}) , rigidity by (w_{20}, w_{02}) , and the resistance to twist by w_{11} . These terms help to keep the active surface smooth. $E_{ext}^{3D}(\mathbf{r}(p, q))$ is the external energy, which is derived from the image and other external constraints. In our case, the external energy generates forces to attract the active surface to fit the central line of the filament.

A local minimum of E_{3D} can be obtained by iteratively solving its associated Euler-Lagrange equations [10],

$$\bar{\mathbf{r}}^k = (I + \tau A^{3D})^{-1} (\bar{\mathbf{r}}^{k-1} + \tau \nabla E_{ext}^{3D}), \quad (3)$$

where $\bar{\mathbf{r}}^k$ is the vector consisting of values of $\mathbf{r}(p, q)$ for all p and q at iteration k , I is the identity matrix, τ is the step size parameter, and A^{3D} is a sparse matrix encoding the structure of the surface mesh and weights of the internal energy (details on A^{3D} can be found in [10]). The initial vector $\bar{\mathbf{r}}^0$ is given by the initial

surface (Section 2.1). With this model and properly defined external energy, the resulting surface would locate the central line of a filament in all slices.

However, the computation time for the 3D surface increases significantly compared with 2D active contour models. The intersections between the active surface and slices are also difficult to calculate. A solution is to use a simplified 3D surface model similar to that proposed in [10]. The degrees of freedom of the deformation are limited to two components instead of three by constraining the time component t , which corresponds to the discrete time coordinate, to depend linearly only on one parameter. The external force along the time dimension then becomes zero. Therefore, the surface we seek is now represented by a sequence of parametrized curves. We have $\mathbf{r}(p, q) = (x(p, q), y(p, q), q)$ so that for a given slice index q , there corresponds an open curve parametrized by p lying on the q th slice of the 3D spatiotemporal volume.

Compared to slice-by-slice approaches in [9], this new active surface model better utilizes temporal information to constrain the segmentation results. For instance, if a filament is missing or hardly observed in a slice, the surface bridges the neighboring curves, creating a smooth surface. In contrast, the previous methods may fail or generate large errors in this scenario.

2.3 The External Energy

The external energy consists of an image term and a stretching term, *i.e.*, $E_{ext}^{3D} = E_{img} + E_{str}$. As we mentioned above, the external force of the constrained surface is always zero along time dimension. Therefore, we can calculate it slice by slice. The image term represents the image constraints. Since we want to locate the bright ridges in images, we directly use the Gaussian-filtered or Anisotropic Diffusion filtered image of the q th slice as the image term of the q th slice, *i.e.*, $E_{img}(\mathbf{r}(p, q)) = G_\sigma * I(\mathbf{r}(p, q))$, $p \in [0, 1]$, for any q . The gradient vectors of such an image term, ∇E_{img} , point toward the center of filaments and attract the active surface to fit the filament.

On the one hand, with the image term only, the active surface would only fit a part of the filament because there is no force attracting the active surface to the ends of filaments. On the other hand, filament ends usually have very low SNR and might be difficult to distinguish during the segmentation process. One important observation is that an over-grown surface almost always covers tips of the filament [9]. Based on this observation, we let the active surface over-grow in both directions and distinguish the ends on the resulting over-grown surface in a subsequent step (Section 3). To make the surface over-grow, we add stretching forces to the boundaries of the active surface:

$$\nabla E_{str}(\mathbf{r}(p, q)) = \begin{cases} -k_{str} \cdot \frac{\mathbf{r}_p(p, q)}{|\mathbf{r}_p(p, q)|} & p = 0, \\ k_{str} \cdot \frac{\mathbf{r}_p(p, q)}{|\mathbf{r}_p(p, q)|} & p = 1, \\ 0 & otherwise, \end{cases} \quad \forall q \in [0, 1] \quad (4)$$

where \mathbf{r}_p is the derivative of \mathbf{r} with respect to p , $\mathbf{r}_p(p, q)/|\mathbf{r}_p(p, q)|$ is the direction pointing outwards along the surface, and k_{str} is the stretching weight and only needs to be large enough to let the surface over-grow. Fig. 2(a) illustrates the

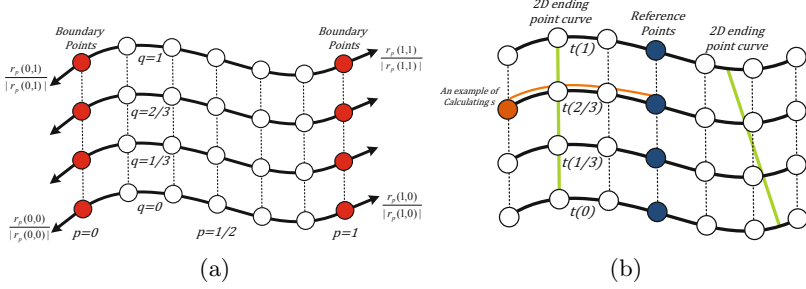


Fig. 2. Illustration of the 3D surface mesh. Solid and dashed lines represent spatial and temporal connections between sampled points, respectively. (a) Stretching forces (black arrows) are added on the boundary points (red). (b) Illustration of the reference points (blue) and ending-point active contours (green). The brown curve gives an example of calculating the brown point’s length s with respect to the reference point.

stretching forces. After every 3-4 iterations, the stretching force is removed for one iteration to let the active surface better fit to the filament. We deform the surface until intensity values corresponding to all boundary points, $\mathbf{r}(0, q)$ and $\mathbf{r}(1, q), \forall q \in [0, 1]$, are lower than the foreground intensity value, which is estimated in the surface initialization step (Section 2.1). An over-grown active surface is shown in Fig. 1(b).

3 Spatiotemporal Active-Contour Model for Locating Filament Ends

3.1 The 2D Spatiotemporal Domain

After we obtained the over-grown active surface, we can constrain the search of filament ends along the resulting curve on every slice (red curves in Fig. 1(c)). One simple way of locating ends might be to perform 1D Sobel filtering along each curve and to locate ends as locations with highest filtering response. But this operation ignores the relations between filament ends on different time slices: filament ends usually elongate at a consistent speed. Based on this observation, we create a 2D spatiotemporal active contour model to locate filament ends at all slices simultaneously (green curves in Fig. 1(c)).

Given a reference point on the first slice’s curve, we can obtain its nearest corresponding reference points on all other slices, such that every point on every curve is of length s away from the reference point along the curve (Fig. 2(b)). The length s dimension, together with the time t dimension, create a 2D spatiotemporal domain. Every point (s, t) in this 2D domain implicitly specifies a possible 2D ending point $(x(s), y(s))$ of the filament on the t th slice. Therefore,

if we connect the true ending points of a filament on different slices and map them to the (s, t) domain, these mappings form a smooth curve in the (s, t) domain.

3.2 The 2D Active Contour Model for Locating Filament Ends

The active contour is then defined by a mapping of parameter u to the 2D (s, t) domain, $\mathbf{v} : \Omega = [0, 1] \rightarrow \mathbb{R}^2$,

$$u \rightarrow \mathbf{v}(u) = (s(u), t(u)).$$

Similarly to the above constrained 3D model, we constrain the time t dimension. The point on the t th slice can only deform along the resulting over-grown curve on the t th slice. Therefore, the internal force along t dimension should not be included in the model, and the external force ∇E_{ext}^{2D} along the t dimension is zero. Based on the constrained model and the above consistent growing speed assumption on filament ends, we define the 2D active contour's energy E_{2D} as

$$E_{2D}(\mathbf{v}) = \int_{\Omega} w_1 \left\| \frac{\partial s}{\partial u} \right\|^2 + w_2 \left\| \frac{\partial^2 s}{\partial u^2} \right\|^2 + E_{ext}^{2D}(\mathbf{v}(u)) du, \quad (5)$$

where w_1 and w_2 weight the first-order and second-order derivative terms of the internal energy. For the external force ∇E_{ext}^{2D} , because it is zero along the t dimension, we use the 1D Sobel filtering response along each over-grown curve as the external energy on that slice.

A local minimum of E_{2D} can be solved iteratively:

$$\bar{\mathbf{v}}^k = (I + \tau A^{2D})^{-1} (\bar{\mathbf{v}}^{k-1} + \tau \partial E_{ext}^{2D} / \partial s), \quad (6)$$

where notations are similarly defined. Iteratively solving Eqn. (6) would deform the active contour to fit the desired filament end locations on all slices simultaneously (Fig. 1(c)). We then cut the over-grown curves using those estimated filament end locations to generate the final segmentation results (Fig. 1(d)).

3.3 Ending Point Initialization

Eqn. (6) only converges to a local minimum of E_{2D} . An ending point curve needs to be initialized in the (s, t) domain. One straightforward strategy is to initialize ending points on the t th slice at locations on the over-grown curves with highest Sobel filtering response.

Prior knowledge, if available, can also be taken into consideration when initializing the ending point curve. Given an expected end growing speed m , we can initialize a straight line, $(s_1, 1), (s_2, 2), \dots, (s_n, n)$ in the (s, t) domain, where $s_{i+1} - s_i = m, \forall i \in 2, 3, \dots, n$. We shift the straight line along the s dimension to find the best initialization location with the highest Sobel filtering response.

4 Application to Experimental Images

We used two TIRFM image sequences from [1]. In these experiments, polymerization of muscle Mg-ADP-actin was monitored in the presence of varying concentrations of inorganic phosphate (Pi) and actin monomers. The pixel size was $0.17 \mu\text{m}$. There were 20 images in sequence I and 34 images in sequence II. The time interval between frames was 30 sec in sequence I and 10 sec in II.

4.1 Evaluation and Comparison with the Previous Method [9]

For both sequences, we set $w_{01} = w_{10} = 0.1$, $w_{11} = 0$, $w_{02} = w_{20} = 0.02$, $w_1 = 0.01$, $w_2 = 0.05$, and $k_{str} = 1$. The average growing speed of filaments is used as the prior knowledge for ending point curve initialization (Section 3.3) and was set according to the average elongation rate estimated by a manual method [2].

The filament bodies were usually tracked accurately by our algorithm. Most of the errors are among two types: (i) intersected filaments may be mis-segmented as one filament (Fig. 3(1)), and (ii) tip locations were wrongly estimated (Fig. 3(3)). Therefore, we evaluated our algorithm by counting the number of mis-segmentation of filament bodies and by measuring errors on tip location estimation. We selected 15 and 5 actin filaments from image sequence I and II respectively to evaluate our algorithm. For all selected filaments, we manually labeled their two tips in each frame as ground truth and calculated L_2 distances between the ground truth and our algorithm’s results.

We also compared with a previous method [9], which only utilizes partial temporal information on filament tips. We show two examples of comparison with the previous method in Fig. 3. The first example shows the advantages of our spatiotemporal surface model. In Fig. 3(1), the filament of interest intersects with another one. The previous method segments this filament slice by slice and mis-stretches to the other filament. Our proposed method utilizes information about the filament body at all slices and thus can better handle filament intersection (Fig. 3(2)). The second example illustrates the advantages of the 2D spatiotemporal ending point curve. In Fig. 3(3), the filament grows much faster than other filaments in the sequence. The prior knowledge on the growing speed is thus inaccurate. The previous method [9] generates large tip estimation errors because of the wrong prior knowledge (red arrows in Fig. 3(3)). As shown in

Table 1. Tip tracking error statistics of selected filaments in both image sequences I and II. Smaller errors are shown in bold. (Unit: pixel)

	Sequence	Mean	Maximum	Number of Mis-segmentation
Proposed Method	I	0.93	2.60	1 out of 225
	II	1.47	6.01	3 out of 170
Previous Method [9]	I	1.29	3.91	4 out of 225
	II	1.73	5.83	4 out of 170

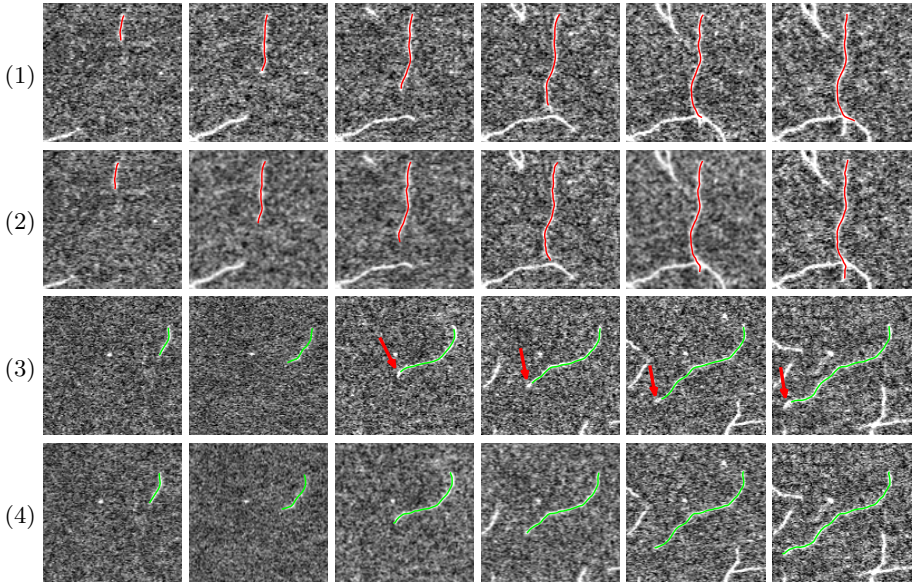


Fig. 3. Two examples of comparison on filament segmentation in a time-lapse sequence. (Odd rows) segmentation results by the previous method [9]. (Even rows) segmentation results by our proposed method.

Fig. 3 (4), our proposed ending point curve better utilizes data from observation and is less sensitive to inaccurate prior knowledge, and thus it accurately segments the filament. Table 1 shows tracking error statistics of our algorithm in comparison to the previous method; our new algorithm clearly outperforms the previous method.

5 Conclusions

In this paper, we present novel active surface and contour models for the segmentation of actin filaments in a time-lapse sequence. We solved the 2D tracking problem as a segmentation problem with smoothness constraints in a higher spatiotemporal 3D domain by introducing an active surface model. A novel (s, t) domain is created in which a 2D active contour deforms to locate filament ends accurately. Because it utilizes temporal constraints on filament body and tips from the entire sequence, it compares favorably to previous methods that use partial temporal information.

Acknowledgments. This work was supported by NIH grant R21GM083928.

References

1. Fujiwara, I., Vavylonis, D., Pollard, T.D.: Polymerization kinetics of ADP- and ADP-Pi-actin determined by fluorescence microscopy. *Proc. Natl. Acad. Sci. USA* 104, 8827–8832 (2007)
2. Kuhn, J.R., Pollard, T.D.: Real-time measurements of actin filament polymerization by total internal reflection fluorescence microscopy. *Biophys. J.* 88, 1387–1402 (2005)
3. Li, H., Shen, T., Smith, M., Fujiwara, I., Vavylonis, D., Huang, X.: Automated actin filament segmentation, tracking and tip elongation measurements based on open active contour models. In: *Proc. ISBI* (2009)
4. Hadjidemetriou, S., Toomre, D., Duncan, J.: Motion tracking of the outer tips of microtubules. *Medical Image Analysis* 12, 689–702 (2008)
5. Saban, M., Altinok, A., Peck, A., Kenney, C., Feinstein, S., Wilson, L., Rose, K., Manjunath, B.: Automated tracking and modeling of microtubule dynamics. In: *Proc. ISBI*, vol. 1, pp. 1032–1035 (2006)
6. Sargin, M.E., Altinok, A., Kiris, E., Feinstein, S.C., Wilson, L., Rose, K., Manjunath, B.S.: Tracing microtubules in live cell images. In: *Proc. ISBI* (2007)
7. Smal, I., Draegestein, K., Galjart, N., Niessen, W., Meijering, E.: Particle filtering for multiple object tracking in dynamic fluorescence microscopy images: Application to microtubule growth analysis. *IEEE Trans. on Medical Imaging* 27, 789–804 (2008)
8. Kong, K., Marcus, A., Giannakakou, P., Wang, M.: Using particle filter to track and model microtubule dynamics. In: *Proc. ICIP*, vol. 5, pp. 517–520 (2007)
9. Li, H., Shen, T., Vavylonis, D., Huang, X.: Actin filament tracking based on particle filters and stretching open active contour models. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009, Part II. LNCS*, vol. 5762, pp. 673–681. Springer, Heidelberg (2009)
10. Cohen, L.D., Cohen, I.: Finite element methods for active contour models and balloons for 2d and 3d images. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 15, 1131–1147 (1991)
11. Besl, P., McKay, H.: A method for registration of 3-D shapes. *IEEE Trans. Pattern Analysis and Machine Intelligence* 14, 239–256 (1992)

Model-Based Esophagus Segmentation from CT Scans Using a Spatial Probability Map

Johannes Feulner^{1,3}, S. Kevin Zhou², Martin Huber³, Alexander Cavallaro⁴,
Joachim Hornegger¹, and Dorin Comaniciu²

¹ Pattern Recognition Lab, University of Erlangen-Nuremberg, Germany

² Siemens Corporate Research, Princeton, NJ, USA

³ Siemens Corporate Technology, Erlangen, Germany

⁴ Radiology Institute, University Hospital Erlangen, Germany

Abstract. Automatic segmentation of the esophagus from CT data is a challenging problem. Its wall consists of muscle tissue, which has low contrast in CT. Sometimes it is filled with air or remains of orally given contrast agent. While air holes are a clear hint to a human when searching for the esophagus, we found that they are rather distracting to discriminative models of the appearance because of their similarity to the trachea and to lung tissue. However, air inside the respiratory organs can be segmented easily. In this paper, we propose to combine a model based segmentation algorithm of the esophagus with a spatial probability map generated from detected air. Threefold cross-validation on 144 datasets showed that this probability map, combined with a technique that puts more focus on hard cases, increases accuracy by 22%. In contrast to prior work, our method is not only automatic on a manually selected region of interest, but on a whole thoracic CT scan, while our mean segmentation error of 1.80mm is even better.

1 Introduction

Atrial fibrillation is a major cause of stroke. It can be treated with a catheter ablation therapy in the heart. However, this intervention imposes the risk of an atrio-esophageal fistula. The air from the esophagus can enter the left atrium, which normally causes the death of the patient [1]. Here, a segmentation of the esophagus can be helpful during intervention planning. A segmentation can also help to find lymph nodes in CT images of the chest area. The attenuation coefficients are so similar that it is often impossible even for a human to separate them given a single slice. Here, a segmentation can provide valuable overview to a physician, and also automatic detectors benefit because the esophagus can be excluded from search.

Automatic segmentation of the esophagus is challenging because it easily can be confused with vessels, muscles or lymph nodes. Both shape and appearance can vary a lot. It may appear solid, but it can also contain air or remains of contrast agent given orally to the patient.

Prior work on the topic is limited. In [2], the esophagus is segmented with a non-parametric shape prior of the centerline and an appearance model.

The method is semi-automated: The user has to specify two points on the centerline. Moreover, segmentations of the left atrium and the aorta are required as anchors for the shape prior. Another semiautomatic segmentation method which also relies on a shape prior that is relative to surrounding structures is described in [3]. In [4], contours that were manually drawn into axial slices are interpolated without using the image itself.

In this work, we follow the approach of [5]. There, a four step method is proposed. First, a box detector based on a discriminative model is run for each axial slice and a set of candidates of the approximate esophagus contour is generated. Then, the candidates are clustered and merged. Now a Markov chain model is used to infer the path through the axial slices that has the highest probability. Finally, a surface is generated and refined, again with a discriminative model. However, the method is only automatic on a manually cropped region of interest (ROI) and not on a whole CT volume.

In this work, the method of [5] is extended in three ways. First, it is made fully automatic on uncropped CT volumes as acquired by the scanner. The region of interest is determined by detecting a salient landmark in the chest. Second, the accuracy was improved by incorporating a new intermediate detection step. Finally, we explicitly handle air in the esophagus and in the respiratory organs in order to further increase the robustness.

The remainder of the paper is structured as follows: In section 2.1, we shortly summarize the approach of [5]. Section 2.2 describes our approach for automatically detecting the region of interest. In section 2.3, the detector which finds esophagus contour candidates is described along with the new intermediate detection step. Section 2.4 explains how the distribution of air was incorporated. Section 3 presents experiments and results, and section 4 concludes the paper.

2 Method

2.1 Model Based Segmentation

In [5], the esophagus contour in axial slices is first approximated with ellipses. They are parameterized as

$$\mathbf{e} = (\mathbf{t}, \theta, \mathbf{s}), \quad (1)$$

where $\mathbf{t} = (x, y)$ is the center within the slice, θ is the rotation angle and $\mathbf{s} = (a, b)$ contains the semi-axes of the ellipse.

Ellipses are detected using a technique called marginal space learning [6]. Instead of directly searching the five dimensional search space, three classifiers are trained. As classifiers, we use probabilistic boosting-trees (PBT). These are binary decision trees with a strong AdaBoost classifier at each node. The first PBT classifier learns the probability $p(m = 1 | \mathbf{H}(\mathbf{t}))$ of whether there is a model instance in the image at a certain location \mathbf{t} . Here, $\mathbf{H}(\mathbf{t})$ denote Haar-like features extracted at position \mathbf{t} . These are simple cuboid filters similar to the rectangle filters described in [7]. They can be computed very efficiently and thus even allow to search the volume exhaustively. The second one is trained to learn

the probability $p(m = 1 | \mathbf{S}(\mathbf{t}, \theta))$ of whether there is a model instance at a given position with a given rotation angle θ . It uses steerable features \mathbf{S} as proposed in [6], which are simple point features sampled on a regular grid, and the sampling pattern is rotated according to θ . The third classifier is trained on the whole search space to learn $p(m = 1 | \mathbf{S}(\mathbf{t}, \theta, \mathbf{s}))$. It uses steerable features as well, but now the sampling pattern is also scaled. By first using the detectors trained on marginal spaces, large portions of the search space can be pruned at early stages. The result is a set of ellipse candidates $\{\mathbf{e}^{(1)} \dots \mathbf{e}^{(N)}\}$ per axial slice. These candidates are spatially clustered and merged. Each cluster center $\mathbf{c}^{(k)}$ is associated with a weight $\sigma^{(k)}$ which is the sum of detection scores $p(m = 1 | \mathbf{e}^{(i)})$ of candidates belonging to cluster k .

Now the most likely path through the axial slices is inferred using a Markov chain model. Each slice i is associated with a random variable \mathbf{s}_i . Possible states of the random variable \mathbf{s}_i are the cluster centers $\mathbf{c}^{(k)}$ of slice i . The transition probability $p(\mathbf{s}_{i+1} | \mathbf{s}_i)$ from one slice to the next is modelled by normal distributions whose parameters are estimated from manual annotations, and the probability $p(\mathbf{s}_i | \mathbf{v})$ of a variable given an observed image slice \mathbf{v} is set to weight $\sigma^{(i)}$. The maximum a posteriori estimate of all state variables given all observed image slices $p(\mathbf{s}_{1:T} | \mathbf{v}_{1:T})$ can be computed efficiently using dynamic programming.

After the best path has been computed, it is converted into a triangulated surface representation and refined to better fit the boundary of the organ. The vertices of the surface are individually deformed according to the output of another PBT classifier which was trained with steerable features. Then a mesh smoothing algorithm is applied. Deformation and smoothing is repeated once.

2.2 Automatic ROI Detection

The method described in section 2.1 only works on a region of interest showing the esophagus. In order to make it fully automatic also on uncropped CT scans showing for instance the thoracic and abdominal region, we detect the ROI automatically. A salient point close to the esophagus that can be detected very robustly is the bifurcation of the trachea. The landmark detection method used here is described in [8]. An axis-aligned cuboidal region is rigidly attached to that landmark such that the esophagus is always inside in 144 datasets with a minimum margin of 3cm in x and y direction. The resulting region is of size $13.3 \times 15.6 \text{cm}^2$. In vertical (z) direction, the size is set to 26cm. This ROI is fairly large, which makes the detection harder because it contains more structures that may be confused with the esophagus.

2.3 Discriminative Model

Within the detected ROI, we run for each axial slice a classifier that was trained to learn $p(m = 1 | \mathbf{H}(\mathbf{t}))$ as described in section 2.1. A set of candidates $C_{T_1} = \{\mathbf{t}_1 \dots \mathbf{t}_{N_{T_1}}\}$ is generated which contains the N_{T_1} positions with highest detection score. We now propose to train another classifier of the same type and the same features. But now, the negative training examples are generated by

scanning images with the first classifier and collecting false alarms instead of randomly drawing samples from the images. Thus, the second classifier gets specialized on the difficult cases. It only considers the candidates from C_{T_1} and generates a set C_{T_2} which contains the N_{T_2} candidates with highest scores of the second detector. The remaining steps in the detection pipeline are similar to [5]: A classifier trained to learn the probability $p(m = 1 | \mathcal{S}(\mathbf{t}, \theta))$ considers the candidate set C_{T_2} and generates a set C_{TR} of position and orientation candidates of size N_{TR} . Finally, a classifier that was trained on the full search space of translation, rotation and scale is run on C_{TR} to generate the ellipse candidates of a slice. Clustering, merging, path inference and surface generation are adopted from [5].

In Fig. 1, example output is displayed for each step of the detection pipeline.

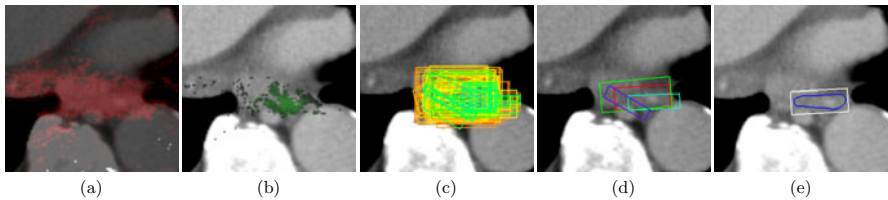


Fig. 1. The proposed detection pipeline. (a) shows the detection scores generated by the first translation detector. The output of the second translation detector is visualized in (b). The bounding boxes of the ellipse candidates are shown in (c). The confidence is color coded in HSV color space: Violet means lowest score, red is highest. The candidate set is clustered and merged (d). The best path through the axial slices is inferred, converted into a surface and further refined (e). The blue contour is the final segmentation.

2.4 Including the Distribution of Air

When a human tries to find the esophagus in a CT dataset, s/he also looks for air holes as they are clearly visible and a strong indicator for the esophagus. One might expect the classifiers to learn this correlation, but we found that air holes rather distract them. The reason is that the classifiers only rely on local features. Then, air holes can be confused easily with the trachea or lung tissue, and both are a priori much more likely because they cover a larger volume. A human, however, recognizes and excludes the respiratory organs effortlessly.

Separating esophageal from respiratory air. Therefore, we propose to support the detector by adding the knowledge that air belonging to the respiratory organs cannot belong to the esophagus, and air elsewhere most likely is inside the esophagus. This is modelled with a binary mask $B(\mathbf{t})$ that is zero if \mathbf{t} belongs to a respiratory organ and one otherwise, and a probability map $S(\mathbf{t})$ of the esophagus based on detected air holes.

Segmenting the air of the respiratory organs in CT is straightforward because it is one connected region. First, voxels with an attenuation coefficient below -625HU are labeled as air. Vessels and airways in the lung are labeled as air as well by identifying small connected components in axial, sagittal and coronal slices. Now all regions marked as air which touch the left, right, front or back border of the ROI are removed. The result is stored in B . Remaining regions marked as air probably belong to the esophagus. They are labeled with 1 in a mask E . Elsewhere, E is zero. A similar method to find esophageal air holes is described in [9]. Now for each axial slice, it is checked whether E contains exactly one connected region labeled as esophageal air. If so, we set the current slice of the probability map S to $g(|\mathbf{t} - \mathbf{p}|)$, where \mathbf{p} is the point of gravity of the region within the slice and g is a Gaussian with standard deviation s that is deformed and trimmed to have a maximum of 1 and limited support in $[-w, w]$. We selected a value of 7mm as s and 10mm as w .

Integration into detector. We now define a combined probability map $C(\mathbf{t})$ as

$$C(\mathbf{t}) = \frac{B(\mathbf{t}) + S(\mathbf{t})}{2} \quad (2)$$

and model the probability $p(m = 1|C(\mathbf{t}))$ of observing the esophagus at position \mathbf{t} given the global distribution of air as being proportional to $C(\mathbf{t})$:

$$p(m = 1|C(\mathbf{t})) \propto C(\mathbf{t}). \quad (3)$$

During position detection, we are finally interested in the probability $p(m = 1|\mathbf{H}(\mathbf{t}), C(\mathbf{t}))$ of observing the esophagus at a certain location \mathbf{t} given the Haar-like feature response $\mathbf{H}(\mathbf{t})$ and the information from the global distribution of air $C(\mathbf{t})$. In order to simplify the notation, we will omit the argument \mathbf{t} in the remainder of this section. Using Bayes' rule, this can be rewritten as

$$p(m = 1|\mathbf{H}, C) = \frac{p(\mathbf{H}, C|m = 1)p(m = 1)}{p(\mathbf{H}, C)} \quad (4)$$

Now we assume that the feature vector \mathbf{H} is statistically independent from the distribution of air C . This is of course a simplifying assumption. \mathbf{H} and C are to some extent statistically dependent. The assumption is justified by the fact that the map C does improve the performance as we will see, which means that \mathbf{H} does not contain much information about C . With this assumption, (4) can be transformed into

$$p(m = 1|\mathbf{H}, C) = \frac{p(\mathbf{H}|m = 1)p(C|m = 1)p(m = 1)}{p(\mathbf{H})p(C)} \quad (5)$$

$$= \frac{p(m = 1|\mathbf{H})p(m = 1|C)}{p(m = 1)}, \quad (6)$$

which is proportional to the product $p(m = 1|\mathbf{H})C(\mathbf{t})$ of the classifier output and the probability map C . This means we can integrate C into the translation detector simply by multiplying it with the detection score. In Fig. 2, the probability map C is visualized for two axial CT slices.

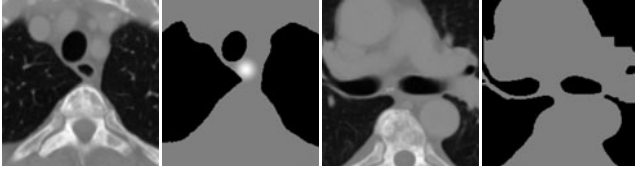


Fig. 2. Two examples of CT slices along with their combined probability map $C(t)$ generated from the distribution of air inside the volume. Left: The air hole is a clear hint for the esophagus. Right: No air hole is present, but respiratory air can be excluded.

Regions filled with respiratory air are not considered by the detector. Therefore, we also do not generate negative training examples from these regions. This makes the learning problem easier because now air is a priori more likely to be part of the esophagus.

3 Results

The method has been evaluated on 144 CT scans of the thoracic or the thoracic and abdominal region. No patient was included twice. The voxel spacing in x and y direction was in the range of 0.7mm to 0.8mm. The spacing in (longitudinal) z direction was 5mm. After ROI detection, the volumes were resampled to a voxel spacing of $0.7 \times 0.7 \times 5\text{mm}^3$. Manual segmentations were available for all datasets. The segmentations typically ranged from the thyroid gland down to a level below the left atrium.

The accuracy was measured using threefold cross-validation. For each fold, all five classifiers for translation ($2\times$), orientation, scale and surface were trained on the training data, and the parameters of the Markov model were estimated from the same training data. The remaining data was used for testing. There was no overlap between training and testing data. For evaluation, the detector was run in z direction on the same interval covered by the manual annotation in order not to introduce artificial errors because of different lengths of the segmentations.

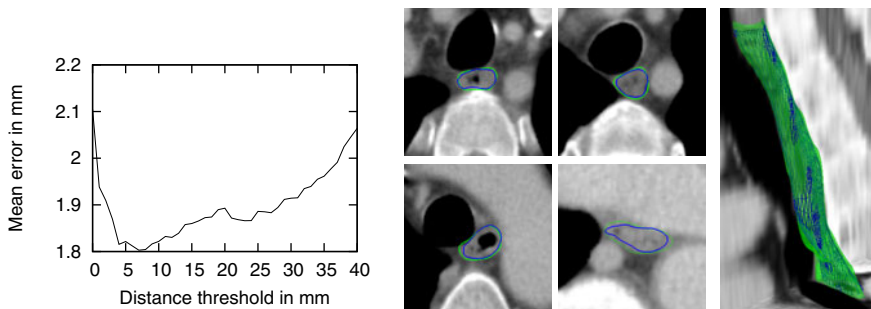
ROI detection succeeded in all of the 144 datasets. Due to the large ROI, the segmentation method can tolerate normal anatomical variations and detection errors of the bifurcation of the trachea.

Table 1 shows the results of accuracy evaluation. As error measures, we used the mean segmentation error and the Hausdorff distance, averaged over all test datasets. The mean error measures the mean distance between the ground truth and the detected surface, while the Hausdorff distance measures the maximum. The number of candidates N_{T1} generated by the first translation detector was set to 400, N_{T2} to 120, N_{TR} to 50, and the number of final candidates per slice to 200. The distance threshold in the clustering step was set to 8mm.

The proposed method segmented the esophagus with a mean error of 1.80mm. If only the binary mask B is used instead of the combined probability map C , the error is 1.88mm. If the air model as described in section 2.4 is omitted, the error is 1.94mm, meaning that explicitly modelling the air significantly improved the

Table 1. Results of performance evaluation. Shown is the mean error and the mean Hausdorff distance along with the corresponding standard deviations.

Method	mean error in mm	Hausdorff dist. in mm
Proposed method	1.80 ± 1.17	12.62 ± 7.01
Only binary air model $B(t)$	1.88 ± 1.24	13.00 ± 7.88
No air model	1.94 ± 1.39	13.06 ± 7.21
Single translation class.	2.07 ± 1.47	14.50 ± 8.92
No air model, single translation class.	2.32 ± 1.87	15.02 ± 9.83
No Markov model	2.30 ± 1.49	17.29 ± 11.42
Proposed method, best 80%	1.34 ± 0.31	9.65 ± 3.07
Feulner et al. [5]	2.28 ± 1.58	14.5
Inter observer variability	0.78 ± 0.17	7.29 ± 2.22

**Fig. 3.** Left: The mean segmentation error for different distance thresholds used for clustering. Right: Examples of automatic segmentations on unseen data (blue) along with ground truth (green).

accuracy. If the air probability map is used, but the second translation detector as described in section 2.3 is omitted, we get an error of 2.07mm. Without both the air model and the additional detector, performance is 2.32mm, meaning that using both decreases the error by 22%. When the Markov model is turned off and for each slice, the ellipse candidates are simply averaged, the error becomes 2.30mm. This is different from what is described in [5]: There, the use of the Markov chain did hardly influence the numeric results. Here, it clearly improves the accuracy. The reason is that the Markov model is especially useful to resolve ambiguity that occurs much more frequently in a larger ROI. The data used for evaluation does also contain extreme cases which in principle can be handled by our method if enough training examples are available. This was not always the case. If the most difficult cases are excluded from the test set, the mean error becomes 1.34mm. For comparison, the results of [5] are shown. We furthermore did an experiment to measure the inter observer variability: Ten datasets were manually segmented a second time by another observer. The second segmentations were evaluated like automatic ones. The result is shown in the last row.

Table 2. Computation time in seconds for different steps of the method

ROI detec.	prob. map gen.	ellipse detec.	path inference	refinement	total
6.96	1.13	7.40	$0.40 \cdot 10^{-3}$	0.34	15.83

Table 2 shows the computational requirements of the different steps of the method, measured on a 2.2GHz dual core PC. Though the ROI is larger, ellipse detection is slightly faster compared to [5] due to the spatial probability map and the second translation detector which rejects most candidates of the first one. In total, the method takes less than 16s.

4 Discussion

In contrast to prior work, our method is able to segment the esophagus from uncropped CT volume images without any user interaction. Segmentation on an automatically detected ROI is harder because the ROI has to be made large. Still, our results are better than what has been reported in [5]. A mean error of 1.8mm, which is only 1.0mm above the inter observer variability, can be considered as a good result for this problem.

References

1. Pappone, C., Oral, H., Santinelli, V., Vicedomini, G., Lang, C.C., Manguso, F., Torracca, L., Benussi, S., Alfieri, O., Hong, R., Lau, W., Hirata, K., Shikuma, N., Hall, B., Morady, F.: Atrio-Esophageal Fistula as a Complication of Percutaneous Transcatheter Ablation of Atrial Fibrillation. *Circulation* 109(22), 2724–2726 (2004)
2. Rousson, M., Bai, Y., Xu, C., Sauer, F.: Probabilistic minimal path for automated esophagus segmentation. In: *Proceedings of the SPIE*, vol. 6144, pp. 1361–1369 (2006)
3. Kurugol, S., Sharp, G., Dy, J., Brooks, D.: Esophagus Segmentation in Thoracic CT Images for Radiotherapy Planning. In: *AAPM*, vol. 36, pp. 2454–2454 (2009)
4. Fieselmann, A., Lautenschläger, S., Deinzer, F., John, M., Poppe, B.: Esophagus Segmentation by Spatially-Constrained Shape Interpolation. In: *BVM*, p. 247–251 (2008)
5. Feulner, J., Zhou, S.K., Cavallaro, A., Seifert, S., Hornegger, J., Comaniciu, D.: Fast Automatic Segmentation of the Esophagus from 3D CT Data Using a Probabilistic Model. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009, Part I. LNCS*, vol. 5761, pp. 255–262. Springer, Heidelberg (2009)
6. Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., Comaniciu, D.: Fast Automatic Heart Chamber Segmentation from 3D CT Data Using Marginal Space Learning and Steerable Features. In: *ICCV*, pp. 1–8 (2007)
7. Viola, P., Jones, M.: Rapid Object Detection using a Boosted Cascade of Simple Features. In: *CVPR*, vol. 1, p. 511 (2001)
8. Seifert, S., Barbu, A., Zhou, S.K., Liu, D., Feulner, J., Huber, M., Suehling, M., Cavallaro, A., Comaniciu, D.: Hierarchical parsing and semantic navigation of full body CT data. In: *Medical Imaging*, vol. 7259, p. 725902. *SPIE* (2009)
9. Fieselmann, A., Lautenschläger, S., Deinzer, F., Poppe, B.: Automatic Detection of Air Holes Inside the Esophagus in CT Images. In: *BVM*, p. 397–401 (2008)

Automatic Segmentation and Components Classification of Optic Pathway Gliomas in MRI

Lior Weizman¹, Liat Ben-Sira², Leo Joskowicz¹, Ronit Precel²,
Shlomi Constantini², and Dafna Ben-Bashat²

¹ School of Eng. and Computer Science, Hebrew University of Jerusalem, Israel

² Sourasky Medical Center, Tel-Aviv, Israel

`lweizm45@cs.huji.ac.il`

Abstract. We present a new method for the automatic segmentation and components classification of brain Optic Pathway Gliomas (OPGs) from multi-spectral MRI datasets. Our method accurately identifies the sharp OPG boundaries and consistently delineates the missing contours by effectively incorporating prior location, shape, and intensity information. It then classifies the segmented OPG volume into its three main components – solid, enhancing, and cyst – with a probabilistic tumor tissue model generated from training datasets that accounts for the datasets grey-level differences. Experimental results on 25 datasets yield a mean OPG boundary surface distance error of 0.73mm and mean volume overlap difference of 30.6% as compared to manual segmentation by an expert radiologist. A follow-up patient study shows high correlation between the clinical tumor progression evaluation and the component classification results. To the best of our knowledge, ours is the first method for automatic OPG segmentation and component classification that may support quantitative disease progression and treatment efficacy evaluation.

1 Introduction

Optic Pathway Gliomas (OPGs) are the most common brain tumors of the central nervous system in patients with Neurofibromatosis (NF) [1]. OPGs are low-grade pilocytic astrocytomas that arise in the optic nerve and chiasm and may involve the hypothalamus and post-chiasmatal regions. OPGs may be asymptomatic, but may become very aggressive and cause severe complications depending on their location [2]. Patients with known OPGs are typically screened serially for progressive visual loss and for changes on MR images. Precise follow-up of an OPG requires the quantification of the tumor volume and the classification of its components into solid, enhancing, and cyst regions. Evolution or changes in the tumor volume and its components may serve as markers for disease progression and may be used to determine the proper treatment and to evaluate its efficacy. Therefore, the accurate quantification of the tumor volume and identification of its components is crucial [3]. Currently, OPG volume is coarsely estimated manually by the physician with a few measurements on axial, coronal, and sagittal slices. This is inaccurate, time consuming, error prone, user dependent, and may compromise the follow-up of the disease progression and its treatment.

Brain tumor detection, characterization, and follow-up based on CT and MR images is currently the standard of care in radiology. The ample spectrum of tumor types and locations has given rise to a plethora of methods for tissue classification and quantification. Most studies focus on the automatic detection of Glioblastoma Multiforme (GBM) tumors [3,4,5], as they account for 40% of all primary malignant brain tumors in adults [6]. Additional studies address other brain lesions, e.g. astrocytoma [7] and low-grade glioma [8].

While effective, most methods do not take into account the anatomic location of the tumor, which is key for the detection and segmentation of OPGs. A common problem of OPGs and other tumors is the delineation of their boundaries due to the tumor inhomogeneity, the surrounding tissues with overlapping image intensity values, the uneven tumor ingrowth into nearby structures, and the imaging partial volume effect. In addition, most existing automatic tumor components classification methods are based on learning the grey-level range of every component from a training set [4]. Therefore, they might suffer from sensitivity to grey-level differences between the learning and the testing sets.

In this paper we describe a new automatic method for the segmentation and components classification of OPG from multi-spectral MRI datasets. Our method effectively incorporates prior location, shape, and intensity information to accurately identify the sharp OPG boundaries and to consistently delineate the OPG contours that cannot be clearly identified on standard MR images. It then classifies the segmented OPG volume into its solid, enhancing, and cyst components based on a probabilistic tumor tissue model generated from training datasets that overcomes the grey-level differences between the learning and the test datasets. Our experimental study on 25 datasets yields a mean surface distance error of 0.73mm and a mean volume overlap difference of 30.6% as compared to manual segmentation by an expert radiologist. A follow-up study shows high correlation between the clinical tumor progression evaluation and the component classification results. The advantages of our method are that it is automatic, accurate, consistent, and that it may support quantitative disease progression, treatment decision-making, and treatment efficacy evaluation.

2 OPG Segmentation and Classification

Our method inputs the patient multi-spectral MRI datasets, which include T1-weighted, T2-weighted, and Fluid Attenuated Inversion Recovery (FLAIR) pulse sequences, and a prior OPG spatial location. The OPG prior spatial location consists of the OPG Region Of Interest (ROI) M , and the chiasm core O , both defined by an expert radiologist on an anatomy atlas. The output is the OPG boundary and OPG voxel classification into solid, enhancing, and cyst components.

The method proceeds in four steps. First, the multi-spectral MR images are coregistered, normalized for intensity, and registered to the anatomy atlas to detect prior OPG ROI and chiasm core. Next, the OPG sharp boundaries are found. In the third step, the missing OPG boundary segments are computed from a probabilistic tumor tissue model generated from training datasets. Finally, the OPG voxels are classified into solid, enhancing, and cyst components.

2.1 MR Images Coregistration and Normalization

Since the patient may move during image acquisition, we first coregister the MR images with the SPM affine registration method [9]. We then standardize the patient MRI intensity values and the probabilistic OPG intensity model with the Dynamic Histogram Warping method [10]. The OPG ROI and the chiasm core are identified in the resulting intensity normalized and aligned patient MR images by registering them to the labeled anatomy atlas with the SPM normalization method [9]. The OPG ROI $M = \{m_1, \dots, m_{n_M}\}$ and chiasm core $O = \{o_1, \dots, o_{n_O}\}$ point sets are then mapped back from the prior atlas space to the patient image space. The resulting sets $\tilde{M} = \{\tilde{m}_1, \dots, \tilde{m}_{n_{\tilde{M}}}\}$ and $\tilde{O} = \{\tilde{o}_1, \dots, \tilde{o}_{n_{\tilde{O}}}\}$ represent the chiasm core and the OPG ROI in the patient image space.

2.2 OPG Sharp Boundaries Detection

The OPG is mostly surrounded by the Cerebral Spine Fluid (CSF), whose intensity value in the FLAIR pulse sequence is very low. Thus, the OPG sharp boundaries are clearly distinguishable where the CSF surrounds the OPG. The CSF voxels are identified in FLAIR by fixed-value thresholding. The sharp OPG boundary voxels are identified as follows. For every voxel $\tilde{m}_i \in \tilde{M}$, we find the shortest Euclidean distance path to \tilde{O} and label it as $P_i = \{p_1, \dots, p_l\}$. If at least one of the voxels in P_i is a CSF voxel, then \tilde{m}_i is removed from \tilde{M} . The resulting \tilde{M} does not contain the voxels in the OPG ROI that lie beyond the CSF borders surrounding the OPG. This step enforces a convex shape, which is mostly the case in OPG. Fig. 1 illustrates this step.

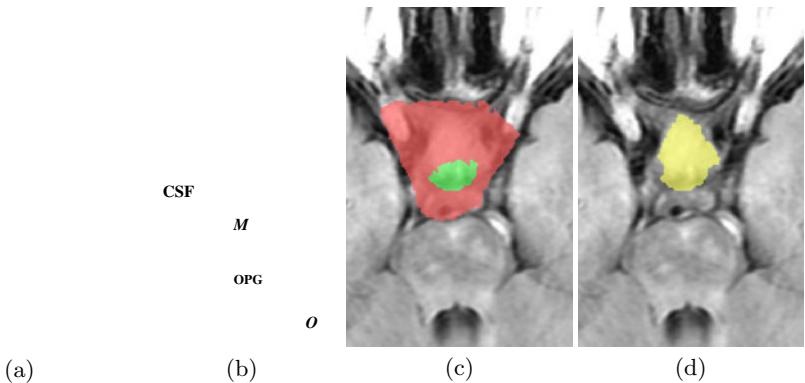


Fig. 1. (a) OPG location in the brain; (b) OPG ROI \tilde{M} (red), chiasm core \tilde{O} (green), OPG (yellow), CSF (blue) areas; (c) example of the OPG ROI (red) and chiasm core (green) on a sample slice; (d) sharp boundary detection result (yellow)

2.3 OPG Boundary Completion

To find the missing OPG boundary segments where a clear border with CSF does not exist, we use the Generalized Likelihood Ratio Test (GLRT) [11]. We define two complementary hypotheses – healthy tissue and OPG tissue – and choose between them based a probabilistic measure computed from an estimate of their unknown model parameters. We describe these two steps in detail next.

Probabilistic tissue model. We represent the multi-spectral MRI dataset consisting of k pulse sequences, each with n voxels, as a set $V = \{v_1(r), \dots, v_n(r)\}$ where $v_i(r)$ is a k -dimensional vector, and $v_i(r) = (v_i^1, v_i^2, \dots, v_i^j, \dots, v_i^k)$, where v_i^j represents the intensity value of the voxel v_i in the j -th pulse sequence. The parameter r denotes the spatial location of the voxel $v_i(r)$.

We postulate two hypotheses for voxel $v_j(r)$:

H_0 : voxel $v_j(r)$ corresponds to healthy tissue.

H_1 : voxel $v_j(r)$ corresponds to OPG tissue.

The probability of $v_j(r)$ to be OPG tissue depends on its spatial location and on the voxel intensity values in the MR images.

Since the every voxel in the image can have any intensity level, the spatial location of a voxel can be assumed to be independent of its intensity level. Therefore, the Probability Density Function (PDF) of $v_j(r)$ for a given hypothesis is:

$$f(v_j(r), r|H_i) = f_I(v_j(r)|H_i) \cdot f_S(r|H_i), \quad i = 0, 1$$

where $f_I(v_j(r)|H_i)$ and $f_S(r|H_i)$ are the respective intensity and spatial location contributions to $f(v_j(r), r|H_i)$.

Since the OPG spreads from the center of the core to the margins of the chiasm, we model $f_S(r|H_1)$ as a Gaussian, with mean r_S and covariance matrix \mathbf{C}_S . Since H_0 is the complementary hypothesis of H_1 , we obtain:

$$f_S(r|H_0) = 1 - f_S(r|H_1)$$

We model the intensity value of healthy/OPG voxels as a mixture of Gaussians:

$$f_I(v_j(r)|H_i) = \sum_{q=1}^3 a_{iq} \cdot \frac{1}{(2\pi)^{k/2} |\mathbf{C}_{iq}|^{1/2}} \exp\left\{-\frac{1}{2}(v_j(r) - \mu_{iq})^T \mathbf{C}_{iq}^{-1} (v_j(r) - \mu_{iq})\right\}$$

where the superscript T denotes the matrix transpose. The parameters $\{\mu_{0q}\}_{q=1}^3$ and $\{\mathbf{C}_{0q}\}_{q=1}^3$ denote the mean vector and covariance matrix of the healthy tissue component: air, CSF, and non-enhancing healthy tissue, respectively. The parameters $\{\mu_{1q}\}_{q=1}^3$ and $\{\mathbf{C}_{1q}\}_{q=1}^3$ denote the mean vector and covariance matrix of solid, enhancing, and cyst OPG components, respectively. Since we do not have the prior probabilities for these components for either the healthy or the OPG hypothesis, we set them to have equal prior probability, i.e. $\forall i, q \quad a_{iq} = \frac{1}{3}$.

Unknown parameters estimation. The Maximum Likelihood Estimators (MLEs) of the unknown model parameters, given the training data, are as follows [11]. The parameters $\{\hat{\mu}_{0q}\}_{q=1}^3$ and $\{\hat{\mathbf{C}}_{0q}\}_{q=1}^3$ are the sample mean and covariance matrix of the CSF, air, and healthy non-enhancing tissue components of healthy tissues, respectively. Similarly, $\{\hat{\mu}_{1q}\}_{q=1}^3$ and $\{\hat{\mathbf{C}}_{1q}\}_{q=1}^3$ are the sample mean and covariance of solid, enhancing, and cystic components of OPG, respectively. The parameters \hat{r}_S and $\hat{\mathbf{C}}_s$ are the center of mass and the spatial sample covariance matrix of \tilde{O} . The GLRT is thus:

$$A(v_j(r)) = \frac{f(v_j(r), r | \hat{\theta}_1, \hat{\theta}_2; H_1)}{f(v_j(r), r | \hat{\theta}_0, \hat{\theta}_2; H_0)} \stackrel{H_1}{\underset{H_0}{\gtrless}} \gamma \quad (1)$$

where γ is a predetermined threshold that reflects the trade-off between false and missed detections. The notation $\stackrel{H_1}{\underset{H_0}{\gtrless}}$ means that if $A(s_i(r))$ is greater than γ , H_1 is chosen for voxel $s_i(r)$, otherwise, H_0 . The final segmentation result is the intersection between the GLRT result and \tilde{M} . The set of voxels $S = \{s_i(r)\}$ that are detected as OPG is thus:

$$S = \{s_i(r) : A(s_i(r)) > \gamma \text{ and } s_i(r) \in \tilde{M}\}$$

2.4 OPG Internal Classification

A common problem of the state-of-the-art supervised classification methods is that the classification results are affected by different acquisition parameters of the training and testing datasets. We propose to use a classification technique that overcomes this phenomenon when the training and the testing datasets intensities differ by a multiplicative factor, as is a common case in OPG datasets.

To determine if a given OPG voxel is solid, enhancing, or cyst, we use the Spectral Angle Mapper (SAM) method [12]. SAM classification is based on the angle measured between the given vector of pulse sequences grey-levels and a training vector previously computed for every OPG component. To classify a given set of OPG voxels, $S = \{s_i(r)\}_{i=1}^{N_S}$, we use the estimations of the solid, enhancing, and cystic components, $\hat{\mu}_{11}, \hat{\mu}_{12}, \hat{\mu}_{13}$, which were previously calculated. Following the SAM approach, the angle between $s_i(r)$ and $\hat{\mu}_{1q}$ is: $\varphi_q = \text{acos}(s_i(r) \cdot \hat{\mu}_{1q})$, where \cdot denotes the vector dot product. Consequently, $s_i(r)$ is assigned to the component represented by $\hat{\mu}_{1q}$ that yields the lowest φ_q for $q = 1, 2, 3$.

3 Experimental Results

We conducted a quantitative evaluation of our method with clinical multi-spectral MRI datasets of 7 pediatric patients, 3-7 years old with OPGs. The patients were serially screened every several months to produce a total of 28 datasets. The MR images were acquired by General Electric Signa 1.5T HD. The study was approved by the local ethical research committee. Each scan consists of T1-weighted, T2-weighted, and FLAIR. Each dataset has $512 \times 512 \times 30$

voxels with voxel size $0.5 \times 0.5 \times 5.0 \text{mm}^3$. An expert radiologist defined the prior spatial inputs, O and M , on the Johns Hopkins University International Consortium of Brain Mapping T2 atlas [13], and manually produced ground-truth classified segmentations for each scan. A second expert radiologist reviewed and revised the segmentations. To separate the training and testing datasets and to provide robust performance of our methods, three data sets were used to estimate the unknown parameters of the model and to determine the CSF value in the FLAIR sequence to distinguish the OPG from CSF in their tangency region. The remaining 25 scans were used to evaluate the proposed method. All the results were obtained with an experimentally determined threshold of $\gamma = 1.2$.

In the first study, we applied the OPG segmentation algorithm (Secs 2.1-2.3) to each of the 25 cases. Fig. 2 shows the segmentation results in three common validation measures [16]. The average symmetric surface distance is 0.73mm , and the volumetric overlap error is 30.6%. These values are comparable to those of other automatic detection methods of brain tumors reported in the literature [4,5], and to the inter/intra observer variability of manual brain tumor segmentation [14,15].

In the second study, we evaluated the results of our OPG segmentation and component classification method with a follow-up study. MR images of three patients with OPG were serially acquired at subsequent time intervals. The OPG and its three components were then manually segmented by an expert radiologist. For the automatic processing, we defined the first scan of every patient as the reference scan and registered all subsequent scans to it. We then applied our method to each dataset, and computed the segmented OPG volume and that of its solid, enhancing, and cystic components (Sec. 2.4).

We computed the difference vector for every OPG component over time for both manual and automatic classification results. The difference vector consists of the volume differences between consecutive scans, and therefore represents the changes of the OPG component over the time for each patient.

Fig. 3 shows an illustrative example and the results the OPG automatic classification results as compared to the manual classification. We computed the correlation coefficients between the manual and automatic difference vectors.

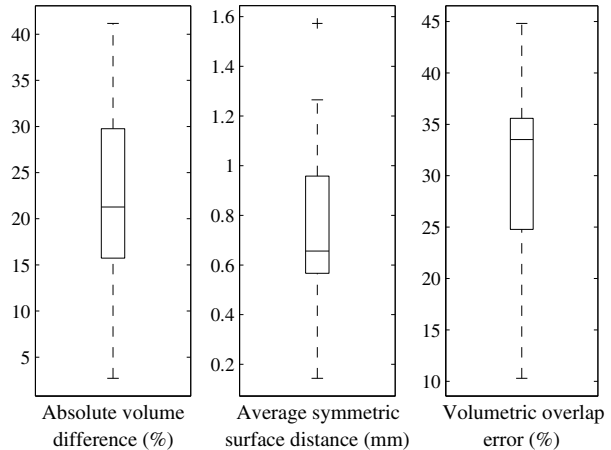


Fig. 2. Segmentation results summary for 25 cases

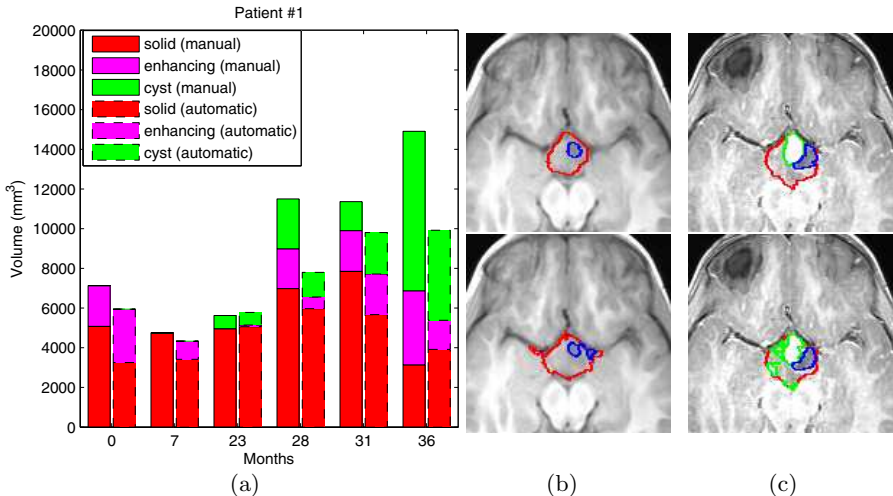


Fig. 3. Illustration of patient 1 follow-up study: (a) manual vs. automatic component classification chart; (b) and (c): ground truth (top) vs. our method (bottom) segmentation results on two patient 1 sample slices for months 23 (left) and 31 (right).

We also computed the same values for the standard Euclidean Distance (ED) classifier. Table 1 shows the results.

We conclude from Fig. 3 that our method successfully estimates the OPG volume progression. For example, the increase in the OPG volume of Patient 1, starting after 23 months, and the development of the

Table 1. SAM and ED correlation with ground truth

Component	Patient 1		Patient 2		Patient 3	
	SAM	ED	SAM	ED	SAM	ED
<i>Solid</i>	0.778	0.085	0.597	0.487	0.161	-0.468
<i>Enhancing</i>	0.503	0.319	0.905	0.875	0.869	0.426
<i>Cystic</i>	0.864	0.520	N/A	N/A	0.854	0.845

enhancing component after 28 months, can be observed in both the manual and the automatic segmentation. These findings are an indicator for positive tumor progression, which may require altering the current patient treatment. From Table 1, we conclude that our method successfully estimates the OPG components progression. In addition, we found that our classification method outperforms the ED classifier, which relies on absolute grey-level intensity values.

4 Conclusions

We have presented a method for the automatic segmentation and component classification of OPGs from multi-spectral MRI. The paper makes three main contributions. First, our segmentation method uses a spatial a priori anatomical atlas to find the initial location of the OPG tumor. This is usually done manually via seed selection or by other means in existing segmentation methods. Second, our method classifies voxels according to the learned ratio between the pulse sequences, rather

than by their absolute values. This yields a robust classification method that can handle gray-level intensity imaging variations. Third, we evaluated our method with a follow-up study on three patients, in addition to the standard measures of volume overlapping and surface distance. The study compares the relative volume progression of the OPG components at different times, and quantitatively supports the clinical findings. This constitutes a methodological improvement over the manual method currently used.

For future work, we are planning an extensive follow-up study. We plan to use the new ROI-based segmentation and SAM classification techniques for the automatic segmentation and classification of other types of brain tumors.

References

1. Huson, S.M., Hughes, R.A.: *The Neurofibromatoses: a Pathogenetic and Clinical Overview*. Chapman & Hall, London (1994)
2. Binning, M.J., Liu, J.K., Kestle, J.R.W., Brockmeyer, D.L., Walker, M.L.: Optic pathway gliomas: a review. *Neurosurgical Focus* 23(5) (2007)
3. Liu, J., et al.: A system for brain tumor volume estimation via MR imaging and fuzzy connectedness. *Comput. Med. Imag. Grap.* 29(1), 21–34 (2005)
4. Corso, J.J., et al.: Efficient multilevel brain tumor segmentation with integrated bayesian model classification. *IEEE T. Med. Imaging* 27(5), 629–640 (2008)
5. Prastawa, M., et al.: Automatic brain tumor segmentation by subject specific modification of atlas priors. *Acad. Radiol.* 10, 1341–1348 (2003)
6. Smirniotopoulos, J.G.: The new WHO classification of brain tumors. *Neuroimag. Clin. N. Am.* 9(4), 595–613 (1999)
7. Lee, C.H., et al.: Segmenting brain tumor with conditional random fields and support vector machines. In: *Proc. Int. Conf. Comput. Vision*, Beijing, China, pp. 469–478 (October 2005)
8. Kaus, M., et al.: Automated segmentation of MRI of brain tumors. *Radiology* 218, 586–591 (2001)
9. Friston, K.J., Holmes, A.P., Ashburner, J.: *Statistical Parametric Mapping (SPM)* (1999), <http://www.fil.ion.ucl.ac.uk/spm/>
10. Cox, I.J., Hingorani, S.L.: Dynamic histogram warping of image pairs for constant image brightness. In: *Int. Conf. on Image Proc.*, Washington, D.C, USA, vol. II, pp. 366–369. IEEE, Los Alamitos (October 1995)
11. Kay, S.: *Fundamentals of statistical signal processing: detection theory*. Prentice Hall, Englewood (1998)
12. Park, B., et al.: Classification of hyperspectral imagery for identifying fecal and ingesta contaminants. In: *Proc. of SPIE*, vol. 5271, pp. 118–127 (2003)
13. Laboratory of brain anatomical MRI. Johns Hopkins Medical Institute, <http://cmrm.med.jhmi.edu/>
14. Weltens, C., et al.: Interobserver variations in gross tumor volume delineation of brain tumors on computed tomography and impact of magnetic resonance imaging. *Radiother. Oncol.* 60, 49–59 (2001)
15. Wetzel, S.G., et al.: Relative cerebral blood volume measurements in intracranial mass lesions: interobserver and intraobserver reproducibility study. *Radiology* 224, 797–803 (2002)
16. Gerig, G., et al.: Valmet: A new tool for assessing and improving 3D object segmentation. In: Niessen, W.J., Viergever, M.A., et al. (eds.) *MICCAI 2001*. LNCS, vol. 2208, pp. 516–523. Springer, Heidelberg (2001)

Spatial Decision Forests for MS Lesion Segmentation in Multi-Channel MR Images

Ezequiel Geremia^{1,3}, Bjoern H. Menze^{1,2}, Olivier Clatz¹, Ender Konukoglu³, Antonio Criminisi³, and Nicholas Ayache¹

¹ Asclepios Research Project, INRIA Sophia-Antipolis, France

² Computer Science and Artificial Intelligence Laboratory, MIT, USA

³ Machine Learning and Perception Group, Microsoft Research Cambridge, UK

Abstract. A new algorithm is presented for the automatic segmentation of Multiple Sclerosis (MS) lesions in 3D MR images. It builds on the discriminative random decision forest framework to provide a voxel-wise probabilistic classification of the volume. Our method uses multi-channel MR intensities (T1, T2, Flair), spatial prior and long-range comparisons with 3D regions to discriminate lesions. A symmetry feature is introduced accounting for the fact that some MS lesions tend to develop in an asymmetric way. Quantitative evaluation of the data is carried out on publicly available labeled cases from the MS Lesion Segmentation Challenge 2008 dataset and demonstrates improved results over the state of the art.

1 Introduction

Multiple Sclerosis (MS) is a chronic, inflammatory and demyelinating disease that primarily affects the white matter of the central nervous system. Automatic detection and segmentation of MS lesions can help diagnosis and patient follow-up. It offers an attractive alternative to manual segmentation which remains a time-consuming task and suffers from intra- and inter-expert variability. MS lesions, however, show high appearance variability which makes automatic segmentation a challenging task. Indeed, MS lesions lack common intensity and texture characteristics, their shapes are variable and their location within the white matter varies across patients.

A variety of methods have been proposed for the automatic segmentation of MS lesions. Generative methods were proposed consisting in a tissue classification by means of an expectation maximization (EM) algorithm. The EM algorithm can be modified to be robust against lesion affected regions, its outcome is then parsed in order to detect outliers which, in this case, coincide with MS lesions [1]. Another approach consists in adding to the EM a partial volume effect model between tissue classes and combining it with a Mahalanobis thresholding which highlights the lesions [2]. Morphological postprocessing on resulting regions of interest was shown to improve the classification performance [3]. In [4], a constrained Gaussian mixture model is proposed, with no spatial prior, to capture the tissue spatial layout. MS lesions are detected as outliers

and then grouped in an additional tissue class. Final delineation is performed using probability-based curve evolution. Multi-scale segmentation can be combined with discriminative classification to take into account regional properties [5]. Beyond the information introduced via the spatial prior atlases, these methods are limited in their ability to take advantage of long-range spatial context in the classification task.

To overcome this shortcoming, we propose the use of an ensemble of discriminative classifiers. It builds on the random decision forest framework which has multiple applications in bioinformatics [6], but more recently also in the image processing community [7]. Adding spatial and multi-channel features to this classifier proved effective in object recognition [8], brain tissue segmentation in MR images [9], myocardium delineation in 3D echocardiography [10] and organ localization in CT volumes [11]. Applying multi-channel and context-rich random forest classification to the MS lesion segmentation problem is novel, to our knowledge. The presented classifier also exploits a specific discriminative symmetry feature assuming that the healthy brain is approximately symmetric with respect to the mid-sagittal plane and that MS lesions tend to develop in asymmetric ways.

2 Materials and Methods

This section describes our adaptation of the random decision forests to the segmentation of MS lesions and illustrates the visual features employed.

2.1 Dataset

Our dataset contains 20 labeled cases which are publicly available from the MS Lesion Segmentation Challenge 2008 website [12]. For each case, three MR channels are made available T1-, T2-weighted and Flair.

After being sub-sampled and cropped, all the images have the same size, $159 \times 207 \times 79$ voxels, and the same resolution, $1 \times 1 \times 2 \text{ mm}^3$. RF acquisition field inhomogeneities are corrected [13] and inter-subject intensity variations are normalized [14]. The images are then aligned on the mid-sagittal plane [15]. Spatial prior is added by registering the MNI atlas [16] to the anatomical images, each voxel of the atlas providing the probability of belonging to the white matter (WM), the grey matter (GM) and the cerebro-spinal fluid (CSF) (cf. Fig. 1).

We will adhere to the following notation: the data consists of a collection of voxel samples $\mathbf{v} = (\mathbf{x}, \mathbf{C})$, each characterized by a position $\mathbf{x} = (x, y, z)$ and associated with a list of signal channels \mathbf{C} . Signal channels $\mathbf{C} = (\mathbf{I}, \mathbf{P})$ include multi-sequence MR images $\mathbf{I} = (I_{T1}, I_{T2}, I_{Flair})$ and spatial priors $\mathbf{P} = (P_{WM}, P_{GM}, P_{CSF})$. Anatomical images and spatial priors, although having different semantics, can be treated under the unified term “signal channel”. We account for noise in MR images by averaging values over a 3^3 voxels box centered on \mathbf{x} , such an average is noted $C_c(\mathbf{x})$, e.g. $C_c = I_{Flair}$ or P_{GM} .

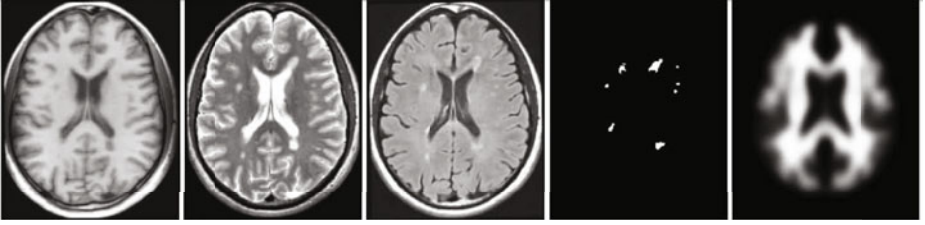


Fig. 1. Data. From left to right: preprocessed T1-weighted, T2-weighted and Flair MR images, the associated ground truth and the registered white matter atlas.

2.2 Context-Rich Decision Forest

Our detection and segmentation problem can be formalized as a binary classification of voxel samples into either background or lesions. This classification problem is addressed by a supervised method: discriminative random decision forest, an ensemble learner using decision trees as base learners. Decision trees are discriminative classifiers which are known to suffer from over-fitting. A random decision forest [17] achieves better generalization by growing an ensemble of many independent decision trees on a random subset of the training data and by randomizing the features made available to each node during training [18].

Forest training. The forest has T components with t indexing each tree. The training data consists in a set of labeled voxels $\mathcal{T} = \{\mathbf{v}_k, Y(\mathbf{v}_k)\}$ where the label $Y(\mathbf{v}_k)$ is given by an expert. When asked to classify a new image, the classifier aims to assign every voxel \mathbf{v} in the volume a label $y(\mathbf{v})$. In our case, $y(\mathbf{v}) \in \{0, 1\}$, 1 for lesion and 0 for background.

During training, all observations \mathbf{v}_k are pushed through each of the trees. Each internal node applies a binary test [8,9,10,11] as follows:

$$t^{\tau_{low}, \tau_{up}, \theta}(\mathbf{v}_k) = \begin{cases} true, & \text{if } \tau_{low} \leq \theta(\mathbf{v}_k) < \tau_{up} \\ false, & \text{otherwise} \end{cases}$$

where θ is a function identifying the visual feature extracted at position \mathbf{x}_k . There are several ways of defining θ , either as a local intensity-based average, local spatial prior or context-rich cue. These are investigated in more detail in the next section. The value of the extracted visual feature is thresholded by τ_{low} and τ_{up} . The voxel \mathbf{v}_k is then sent to one of the two child nodes based on the outcome of this test. Training the classifier means selecting the most discriminative binary test for each node by optimizing over $(\tau_{low}, \tau_{up}, \theta)$ in order to maximize the information gain on the input data partition [19], noted \mathcal{T}_p , defined as follows: $IG_{\tau_{low}, \tau_{up}, \theta}(\mathcal{T}_p) = H(\mathcal{T}_p) - H(\mathcal{T}_p | \{t^{\tau_{low}, \tau_{up}, \theta}(\mathbf{v}_k)\})$ where $\mathcal{T}_p \subset \mathcal{T}$, H stands for the entropy.

Only a randomly sampled subset Θ of the feature space is available for internal node optimization, while the threshold space is uniformly discretized. The optimal $(\tau_{low}^*, \tau_{up}^*, \theta^*)$ is selected by exhaustive search jointly over the feature and threshold space. Random sampling of the features leads to increased

inter-node and inter-tree variability which improves generalization. Nodes are grown to a maximum depth D . Another stopping criterion is to stop growing a node when too few training points reach it, i.e. when the information gain is below a minimal value IG_{min} .

As a result of the training process, each leaf node l of every tree t receives a partition \mathcal{T}_l of the training data. The following empirical posterior probability is then stored at the leaf $p_{l_t}(Y(\mathbf{v}) = b) = |\{(\mathbf{v}, Y(\mathbf{v})) \in \mathcal{T}_l | Y(\mathbf{v}) = b\}| / |\mathcal{T}_l|$ where $b \in \{0, 1\}$ denotes the background or lesion class, respectively.

Prediction. When applied to a new test data $\mathcal{T}_{test} = \{\mathbf{v}_k\}$, each voxel \mathbf{v}_k is propagated through all the trees by successive application of the relevant binary tests. When reaching the leaf node l_t in all trees $t \in [1..T]$, posteriors $p_{l_t}(Y(\mathbf{v}) = b)$ are gathered in order to compute the final posterior probability defined as follows: $p(y(\mathbf{v}) = b) = \frac{1}{T} \sum_{t=1}^T p_{l_t}(Y(\mathbf{v}) = b)$. This probability may be thresholded at a fixed value $T_{posterior}$ if a binary segmentation is required.

2.3 Visual Features

In this section, two kinds of visual features are computed: 1) local features: $\theta_c^{loc}(\mathbf{v}) = C_c(\mathbf{x})$ where c indexes an intensity or a prior channel; 2) context-rich features comparing the voxel of interest with distant regions. The first context-rich feature looks for relevant 3D boxes R_1 and R_2 to compare within an extended neighborhood: $\theta_{c_1, c_2, R_1, R_2}^{cont}(\mathbf{v}) = C_{c_1}(\mathbf{x}) - \frac{1}{vol(R_1 \cup R_2)} \sum_{\mathbf{x}' \in R_1 \cup R_2} C_{c_2}(\mathbf{x}')$ where c_1 and c_2 are two signal channels. The regions R_1 and R_2 are sampled randomly in a large neighborhood of the voxel \mathbf{v} (cf. Fig. 2). The sum over these regions is efficiently computed using integral volume processing [8]. The second context-rich feature compares the voxel of interest at \mathbf{x} with its symmetric counterpart with respect to the mid-sagittal plane, noted $S(\mathbf{x})$: $\theta_c^{sym}(\mathbf{v}) = C_c(\mathbf{x}) - C_c \circ S(\mathbf{x})$ where c is an intensity channel. Instead of comparing with the exact symmetric $S(\mathbf{x})$ of the voxel, we consider, respectively, its 6, 26 and 32 neighbors in a sphere

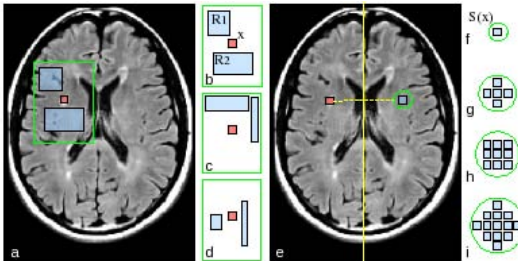


Fig. 2. 2D view of context-rich features. (a) A context-rich feature depicting two regions R_1 and R_2 with constant offset relatively to \mathbf{x} . (b-d) Three examples of randomly sampled features in an extended neighborhood. (e) The symmetric feature with respect to the mid-sagittal plane. (f) The hard symmetric constraint. (g-i) The soft symmetry feature considering neighboring voxels in a sphere of increasing radius. See text for details.

\mathcal{S} (cf. Fig. 2), centered on $S(\mathbf{x})$. We obtain a softer version of the symmetric feature which reads: $\theta_{c,S}^{sym}(\mathbf{v}) = \min_{\mathbf{x}' \in \mathcal{S}} \{C_c(\mathbf{x}) - C_c(\mathbf{x}')\}$.

3 Results

In our experiments, forest parameters are fixed to the following values; number of random regions $|\Theta| \simeq 950$, number of trees $T = 30$, tree depth $D = 20$, lower bound for the information gain $IG_{min} = 10^{-5}$, posterior threshold $T_{posterior} = 0.5$. These values were chosen based on prior parameter optimization on synthetic data and worked well for real data too.

For quantitative evaluation, the 20 available cases are classified and compared to one of the state of the art methods [3]. A three-fold cross-validation is carried out on this dataset: the forest is trained on $\frac{2}{3}$ of the cases and tested on the other $\frac{1}{3}$, this operation is repeated three times in order to collect test errors for each case. Note that the random forest is trained on the preprocessed data.

The binary classification is evaluated using two measures, true positive rate (TPR) and positive predictive value (PPV), both equal 1 for perfect segmentation. Formally, $TPR = \frac{TP}{TP+FN}$ and $PPV = \frac{TP}{TP+FP}$ where TP counts the number of true positive voxels in the classification compared to the ground truth, FP the false positives, FN the false negatives.

Random forest based segmentations are compared with an aligned and sub-sampled version of the ground truth (cf. Fig. 3), whereas segmentations from the winner algorithm were compared with original and sub-sampled segmentations for similar results. Our segmentation compares favorably to one of the state of the art algorithms (cf. Table 1) for both TPR and PPV .

The Grand Challenge 2008 website carried out a complementary and independent evaluation of the algorithm on their private dataset [12]. The results confirm a significant improvement over the winner algorithm of the challenge [3]. The presented spatial random forest achieves, in average, slightly larger true positive (TPR), which is beneficial, and comparable false positive (FPR) rates but significantly lower volume difference (VD), and surface distance (SD) values.

3.1 Discussion

Influence of preprocessing. Data normalization is critical. Indeed, features selected during training should be applied exactly in the same way to new data. For instance, context-rich features, θ^{cont} , are sensitive to rotation and thus require aligned images. Moreover, intensity based features require inter-image normalization to ensure consistency of threshold values in binary tests. This limitation is merely due to our supervised approach. On the contrary, image sub-sampling does not affect feature evaluation much as we are considering averages over rectangular regions.

Analysis of feature relevance. Decision trees are interesting because of ease of interpretability which is highly relevant in classification algorithms considering a large number of input features [6]. We exploit this to analyze the selected

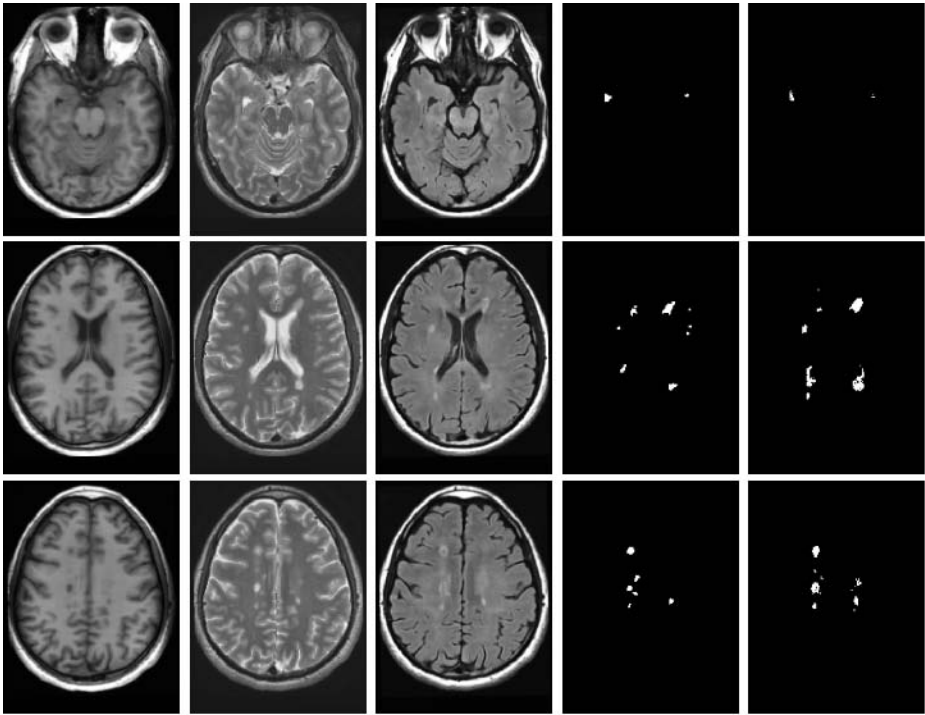


Fig. 3. Segmentation results on a multi-channel 3D MR image. Rows: Axial slices. Columns (from left to right): T1, T2, Flair MR images, ground truth and the output posterior after thresholding.

Table 1. Comparison of context-rich random forests with a state of the art method. In bold we indicate where we do better than the winner algorithm of the MS Segmentation Challenge 2008.

	Ch. winner [3]		Context-rich RF			Ch. winner [3]		Context-rich RF	
Patient	TPR	PPV	TPR	PPV	Patient	TPR	PPV	TPR	PPV
CHB01	0.22	0.41	0.49	0.64	UNC01	0.01	0.01	0.02	0.01
CHB02	0.18	0.29	0.44	0.63	UNC02	0.37	0.39	0.48	0.36
CHB03	0.17	0.21	0.22	0.57	UNC03	0.12	0.16	0.24	0.35
CHB04	0.12	0.55	0.31	0.78	UNC04	0.38	0.54	0.54	0.38
CHB05	0.22	0.42	0.40	0.52	UNC05	0.38	0.08	0.56	0.19
CHB06	0.13	0.46	0.32	0.52	UNC06	0.09	0.09	0.15	0.08
CHB07	0.13	0.39	0.40	0.54	UNC07	0.57	0.18	0.76	0.16
CHB08	0.13	0.55	0.46	0.65	UNC08	0.27	0.20	0.52	0.32
CHB09	0.03	0.18	0.23	0.28	UNC09	0.16	0.43	0.67	0.36
CHB10	0.05	0.18	0.23	0.39	UNC10	0.22	0.28	0.53	0.34

features and understand what are the most discriminative channels for MS lesion segmentation. For this analysis, we consider one of the random forest classifiers which generated the results in Table II. For every tree in the forest, the root node always applies a test on the Flair sequence (θ_{Flair}^{loc}). It means that out of all available features, containing local, context-rich and symmetry multi-channel features, θ_{Flair}^{loc} was found to be the most discriminative. This automated guess coincides with the first step in [3]. At the second level of the tree, a context-rich feature on prior information ($\theta_{WM,GM}^{cont}$) appears to be the most discriminative over all trees in the forest. The associated test discards all voxels which do not belong to the white matter. Again, our algorithm automatically reproduced the second step in [3]. In deeper levels of the tree, local, context-rich and symmetry features adjust the segmentation by combining spatial and multi-channel information. Contribution of each feature to the forest can be quantified by counting the nodes in which they were selected. This indicates a feature discrimination power for the task of MS lesion classification. Local features were selected in 24% of the nodes, context-rich features were selected in 71% of the nodes whereas symmetry features were selected in only 5%. Successive decisions based on local features may learn a non-parametric multi-channel appearance model with spatial prior. Context-rich features exhibit high variability (900 of them are randomly sampled at every node). This variability combined with their ability to highlight regions which differ from their neighborhood explains the high selection frequency. In addition, this kind of features may learn a spatial layout for lesion patterns in peri-ventricular regions (cf. second row in Fig. 3). Symmetry features are under-represented in the forest and thus prove to be the least discriminative ones. Nevertheless, they appear in top levels of the tree (up to third level) which indicates that, they provide an alternative to local and context-rich features when these two fail.

4 Conclusion

We introduce a new algorithm for the segmentation of MS lesions in multi-channel MR images. We present three kinds of 3D features based on multi-channel intensity, prior and context-rich information. Those features are part of a spatial random decision forest classifier which demonstrates improved results on one of the state of the art algorithms on the public MS challenge dataset.

Acknowledgment. The authors thank wholeheartedly the Grand Challenge 2008 organizers and J.C. Souplet and G. Malandain who kindly provided their pre-processed images of the MICCAI Challenge to allow a fair comparison of results. This work was partially funded by the Microsoft Prize, the CompuTumor INRIA funding and the German Academy of Sciences Leopoldina (Fellowship Programme LPDS 2009-10).

References

1. Leemput, K.V., Maes, F., Vandermeulen, D., Colchester, A.C.F., Suetens, P.: Automated segmentation of multiple sclerosis lesions by model outlier detection. *IEEE Trans. Med. Imaging* 20(8), 677–688 (2001)

2. Dugas-Phocion, G., Ballester, M.Á.G., Malandain, G., Ayache, N., Lebrun, C., Chanalet, S., Bensa, C.: Hierarchical segmentation of multiple sclerosis lesions in multi-sequence MRI. In: ISBI, pp. 157–160. IEEE, Los Alamitos (2004)
3. Souplet, J.C., Lebrun, C., Ayache, N., Malandain, G.: An automatic segmentation of T2-FLAIR multiple sclerosis lesions. The MIDAS Journal - MS Lesion Segmentation (MICCAI 2008 Workshop) (2008)
4. Freifeld, O., Greenspan, H., Goldberger, J.: Multiple sclerosis lesion detection using constrained GMM and curve evolution. *J. of Biomed. Imaging* 2009, 1–13 (2009)
5. Akselrod-Ballin, A., Galun, M., Basri, R., Brandt, A., Gomori, M.J., Filippi, M., Valzasina, P.: An integrated segmentation and classification approach applied to multiple sclerosis analysis. In: CVPR 2006, pp. 1122–1129. IEEE, Los Alamitos (2006)
6. Menze, B.H., Kelm, B.M., Masuch, R., Himmelreich, U., Petrich, W., Hamprecht, F.A.: A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinformatics* 10, 213 (2009)
7. Andres, B., Köthe, U., Helmstaedter, M., Denk, W., Hamprecht, F.A.: Segmentation of SBFSEM volume data of neural tissue by hierarchical classification. In: Rigoll, G. (ed.) DAGM 2008. LNCS, vol. 5096, pp. 142–152. Springer, Heidelberg (2008)
8. Shotton, J., Winn, J.M., Rother, C., Criminisi, A.: Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. *Int. J. Comp. Vision* 81(1), 2–23 (2009)
9. Yi, Z., Criminisi, A., Shotton, J., Blake, A.: Discriminative, semantic segmentation of brain tissue in MR images. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5762, pp. 558–565. Springer, Heidelberg (2009)
10. Lempitsky, V.S., Verhoek, M., Noble, J.A., Blake, A.: Random forest classification for automatic delineation of myocardium in real-time 3D echocardiography. In: Ayache, N., Delingette, H., Sermesant, M. (eds.) FIMH 2009. LNCS, vol. 5528, pp. 447–456. Springer, Heidelberg (2009)
11. Criminisi, A., Shotton, J., Bucciarelli, S.: Decision forests with long-range spatial context for organ localization in CT volumes. In: MICCAI workshop on Probabilistic Models for Medical Image Analysis, MICCAI-PMMA (2009)
12. Styner, M., Lee, J., Chin, B., Chin, M., Commowick, O., Tran, H., Markovic-Plese, S., Jewells, V., Warfield, S.: 3D segmentation in the clinic: A grand challenge II: MS lesion segmentation. *MIDAS Journal*, 1–5 (September 2008)
13. Prima, S., Ayache, N., Barrick, T., Roberts, N.: Maximum likelihood estimation of the bias field in MR brain images: Investigating different modelings of the imaging process. In: Niessen, W.J., Viergever, M.A. (eds.) MICCAI 2001. LNCS, vol. 2208, pp. 811–819. Springer, Heidelberg (2001)
14. Rey, D.: Détection et quantification de processus évolutifs dans des images médicales tridimensionnelles: application à la sclérose en plaques. Thèse de sciences, Université de Nice Sophia-Antipolis (October 2002) (in French)
15. Prima, S., Ourselin, S., Ayache, N.: Computation of the mid-sagittal plane in 3d brain images. *IEEE Trans. Med. Imaging* 21(2), 122–138 (2002)
16. Evans, A.C., Collins, D.L., Mills, S.R., Brown, E.D., Kelly, R.L., Peters, T.M.: 3D statistical neuroanatomical models from 305 MRI volumes. In: IEEE-Nuclear Science Symposium and Medical Imaging Conference, pp. 1813–1817 (1993)
17. Amit, Y., Geman, D.: Shape quantization and recognition with randomized trees. *Neural Computation* 9(7), 1545–1588 (1997)
18. Breiman, L.: Random forests. *Machine Learning* 45(1), 5–32 (2001)
19. Quinlan, J.R.: C4.5: Programs for Machine Learning. Morgan Kaufmann, San Francisco (1993)

Segmentation Subject to Stitching Constraints: Finding Many Small Structures in a Large Image

Elena Bernardis¹ and Stella X. Yu²

¹ University of Pennsylvania, Philadelphia, PA 19104, USA

² Boston College, Chestnut Hill, MA 02467, USA

Abstract. Extracting numerous cells in a large microscopic image is often required in medical research. The challenge is to reduce the segmentation complexity on a large image without losing the fine segmentation granularity of small structures. We propose a constrained spectral graph partitioning approach where the segmentation of the entire image is obtained from a set of patch segmentations, independently derived but subject to stitching constraints between neighboring patches. The constraints come from mutual agreement analysis on patch segmentations from a previous round. Our experimental results demonstrate that the constrained segmentation not only stitches solutions seamlessly along overlapping patch borders but also refines the segmentation in the patch interiors.

1 Introduction

There is often a need in medical research to count, measure, and compare numerous tiny cells in a large image. Segmentation methods such as watershed and k -means clustering [1,2] are efficient but unable to deal with large intensity variation, whereas spectral graph partitioning methods [3,4] are robust but unable to efficiently find small structures in a large image. Our goal is to make the latter approach scale effectively with the image size yet without losing the fine granularity of small segments (Fig. 1).

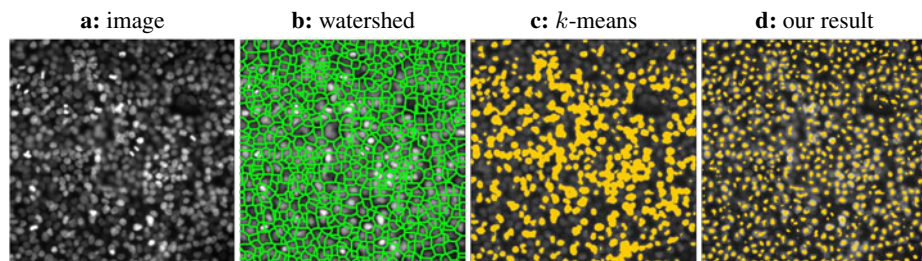


Fig. 1. Segmenting numerous small regions in a large image. **a)** Epithelial cells, which stained with Hoechst fluoresce blue normally and red when transformed by virus, are small structures of varying intensity in this microscopic image. They must be segmented, counted, and measured. **b,c,d)** Two-way segmentations by watershed, k -means, and our method. With large intensity variation for the cells, watershed fails with oversegmentation and k -means fails with conjoined cells of similar intensities, whereas our method correctly pops out all the individual cells.

Spectral graph partitioning methods [3,4] are prized for their ability to grasp the large structural organization of an image from the global integration of local cues. While this property is desired for understanding a real-scene image, it not only unnecessarily handles a huge number of pixels in a large image (since segmenting cells in one region really should not be influenced by cells far from them), but also prevents small structures from being segmented all at once (since a larger image size leads to larger regions instead of numerous small ones given a fixed number of segments). Therefore, finding many small regions in a large image faces two challenges: segmentation complexity from dealing with the large and segmentation granularity from dealing with the small.

The two main approaches to reduce complexity, coarse-to-fine and multiresolution segmentations [4,5,6,7,8,9,10], are not suitable for this task. A coarse-to-fine approach speeds up the segmentation by initializing a finer segmentation with the results of a coarser one, whereas a multiresolution approach integrates features at multiple scales to yield a better segmentation. Since small structures are not present in either coarser-scale segmentations or coarser-scale features, there is no help to be gained from either approach.

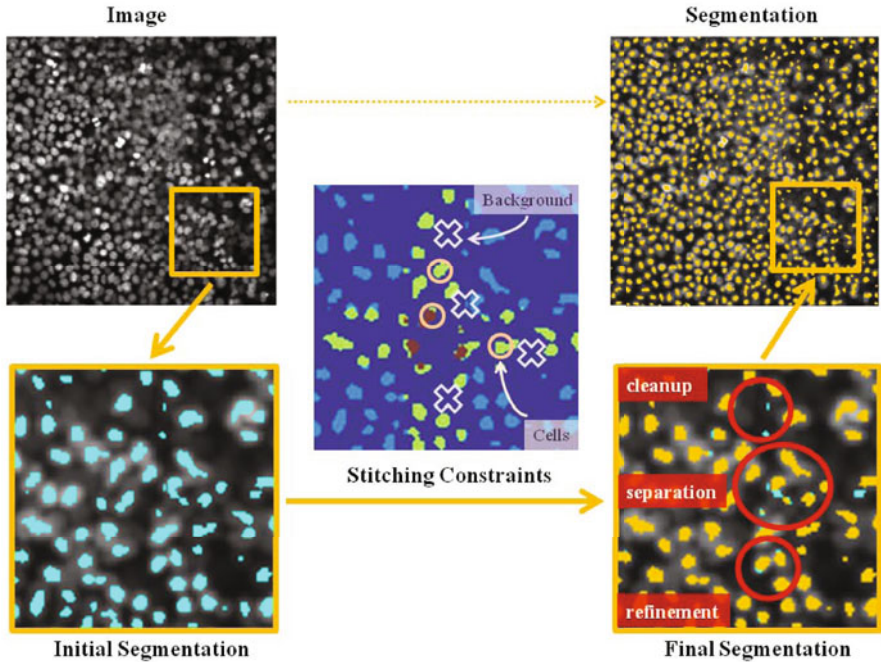


Fig. 2. Algorithm overview. Segmenting numerous small structures in a large image can be performed as a series of independent patch segmentation subject to stitching constraints between neighbouring patches. The constraints are derived from mutual agreement analysis on adjacent patch segmentations from a previous round. Segmentations between neighbouring patches are marked in blue, green, or maroon, if 1, 2, or more than 3 patches agree. Constrained segmentation (gold) improves the initial segmentation (cyan) in three different ways: clean-up of spurious small regions, separation of conjoined cells, and refinement of cell boundaries.

We propose performing segmentation on smaller patches while subject to stitching constraints in their overlapping areas (Fig. 2). We first break down the image into patches and segment each patch independently, based on local pairwise cues that distinguish cells from their background [11]. Their individual segmentations are then used to establish the agreement between the patches, which provide pairwise long-range stitching constraints to be respected by each patch. We run the independent segmentation again on each patch, but now subject to these pairwise constraints on its pixels. The segmentation can be solved efficiently as a constrained eigenvalue problem [12]. Since these segmentations have mutual agreement in the overlapping areas, their individual solutions can be collapsed into one segmentation on the entire large image.

Segmentation subject to stitching constraints does more than stitching solutions together at patch borders. The constraints in the overlapping regions are propagated through local cues in the optimization process to improve the interior segmentation of a patch. Constrained segmentation in individual patches achieves reduced complexity without losing global consistency, refining segmentations both inside and between patches. We detail our model and experiments in Sections 2 and 3 respectively.

2 Spectral Graph Partitioning Subject to Stitching Constraints

We formulate our cell segmentation task as a constrained graph partitioning problem on a set of overlapping patches. Each patch is represented by a weighted graph, where nodes denote pixels and weights attached to edges connecting two nodes encode grouping cues between the pixels. Segmenting small structures in the image becomes a two-way node partitioning problem: pixels inside cells form a foreground node set, and those outside form the other background node set.

We need to address: **1)** What features and grouping cues to use to facilitate this foreground-background segregation [11]; **2)** How to set up constraints between neighboring patches; **3)** How to integrate these constraints into the segmentation [12][13].

2.1 Features F and Grouping Cues W

We characterize cells of small convex bright regions as the sinks of local gradient fields. Each pixel is associated with a peak direction vector p that indicates where pixels of higher intensity are located in its convex vicinity. Two pixels are attracted to the same region if their pixel-centric local gradient fields F 's are similar, and repelled into different regions if their F 's are of opposite types (e.g. sources and sinks).

Consider pixel i and its neighbourhood $N(i)$. If neighbour $a \in N(i)$ can be reached in a straight line from i with non-decreasing intensity, a is a higher intensity pixel in the same convex region. Let $p(i)$ be the average direction from its a neighbours, weighted by the total non-decreasing intensity $T(i, a)$ along the straight line from i to a :

$$p(i) \propto \sum_{a \in N(i)} T(i, a)(L(a) - L(i)), |p(i)| = 1 \quad (1)$$

$$T(i, a) = \sum_{\substack{I(m_1) \leq I(m_t) \leq \dots \leq I(m_k) \\ m_1 m_2 \dots m_k = \text{line}(i, a)}} I(m_t) \quad (2)$$

where $L(i)$ denotes the 2D location of pixel i in the image, $I(i)$ the intensity of pixel i , and $|\cdot|$ the L_2 norm of a vector. Peak direction vector $p(i)$ thus points from i towards the core of the cell that i belongs to, i.e., the highest intensity of its local convex region. It measures the direction and distance from pixel i to the center of the cell.

We define $F(i, a)$ as the inner product of $p(i)$ and $p(a)$, measuring how much a 's cell center estimate agrees with i 's. The ensemble of $\{F(i, a): a \in N(i)\}$ is a pixel-centric vector field (i.e. with the absolute direction of $p(i)$ factored out) that characterizes where pixel i is in the shape of a convex region, and we can use the feature similarity S to establish pairwise pixel grouping cues:

$$F(i, a) = \langle p(i), p(a) \rangle, \quad a \in N(i) \quad (3)$$

$$S(i, j) = \frac{\langle F(i, :), F(j, :)\rangle}{|F(i, :)| \cdot |F(j, :)|}, \quad j \in N(i) \quad (4)$$

$S(i, j)$ is more likely to be positive for nearby pixels inside the same dot, and negative for distant pixels between different dots, giving rise to two kinds of grouping cues [11]: The short-range attraction A is proportional to similarity S and the long-range repulsion R is proportional to dissimilarity $1 - S$. The total effective weight W is $A - R$.

Unlike real-scene image segmentation [34], we do not use single edge features (e.g. large intensity gradients along region boundaries) to delineate regions. We use distributive local gradient fields to characterize geometrical distinction between region cores in the foreground and region peripheries in the background. Similar ideas about such features can be found in [14] on detecting critical points in images with topological numbers. While the individual pairwise grouping cues have poor precisions for localizing region boundaries, they taken together in global integration result in segmentations that are sensitive to geometrical variation yet robust to intensity variations.

2.2 Stitching Constraints U

A two-way node partitioning can be described by a $n \times 2$ binary partition matrix X , where n is the total number of pixels, $X(i, 1)$ and $X(i, 2)$ indicating whether pixel i belongs to the inside or outside of a cell.

Our stitching constraints are imposed on the partition indicator X that is to be solved in the optimization. If pixels a and b are known to belong in the same region, we have the constraint $X(a, :) = X(b, :)$, or $X(a, :) - X(b, :) = 0$. All these equations can be described in a linear constraint $U^T X = 0$, where $U(a, k) = 1$, $U(b, k) = -1$ is the k -th constraint that a and b belong to the same region.

The initial first-round patch segmentation does not require any constraints U , although simple intensity thresholding or initial seeds can be introduced. In the second-round patch segmentation, where each patch has been segmented, U comes from a mutual agreement analysis of X in the overlapping regions between neighbouring patches: those pixels that two patches agree on the segmentation become either foreground or background pixels. Only a sparse set of pairwise constraints (usually between distant pixels) are needed to ensure that two neighbouring patches will have consistent segmentations in their overlapping areas.

2.3 Segmentation with Stitching: Constrained Graph Partitioning X

For each image patch I , after having computed its pairwise grouping cues W and stitching constraints U , we obtain a two-way segmentation using the constrained normalized cuts criterion [12]. Formally, this criterion can be written in the following matrix form:

$$\text{maximize} \quad \varepsilon(X) = \sum_{g=1}^2 \frac{X_g^T W X_g}{X_g^T D X_g} \quad (5)$$

$$\text{subject to} \quad X \in \{0, 1\}^{n \times 2}, X 1_2 = 1_n \quad (6)$$

$$U^T X = 0 \quad (7)$$

where 1_n denote an $n \times 1$ vector of 1's and D is the diagonal *degree* matrix for a $n \times n$ weight matrix W . Note that W could have both positive and negative weights, and the negative ones are essential for popping out disconnected regions [15][11].

The near-global optimal solution is given by the eigenvectors of QPQ , where

$$P = D^{-1}W \quad (8)$$

$$Q = I - D^{-1}U(U^T D^{-1}U)^{-1}U^T \quad (9)$$

While the eigensolution of QPQ takes a longer time than that of P (unconstrained version) to compute at each iteration, it often requires fewer iterations and could be in fact faster. We follow the eigensolution and its discretization procedures developed in [12][13] and their code online to obtain a binary segmentation.

The space and time complexity is much reduced using patch segmentation with stitching constraints, as the image is broken down to smaller patches and finding numerous small regions becomes possible in a single two-way segmentation.

3 Experiments

We implement our algorithm in MATLAB and apply it to sets of 512×512 microscopic images. The cells in these images are 15 pixels in diameter on average. We use the same set of parameters as in [11] for deriving the weights W on each patch of size 256×256 . The pixel neighbourhood radius is 12 pixels, and the overlap radius between patches is 20 pixels. Since the computational complexity mainly depends on the patch size, the entire image can be arbitrarily large.

Fig. 3 shows our results on 3 appearance types of medical images, each representative of a large class of images in its own domain. The cells have a large range of intensity, and fainter ones could be darker than some background pixels elsewhere in the image. Worse still, cells are not always isolated, but rather packed closely next to each other, making the separation even hard for the human eye. In the last rows of Fig. 3, fainter cells are so overwhelmed by those brightest cells that they can only be seen with a close inspection of what appears to be smudged backgrounds.

These images are very challenging to segment and most existing approaches fail. Our method, however, is capable of finding these cells, including faint ones and conjoined ones, all at once in a two-way segmentation, without any need for post-processing.

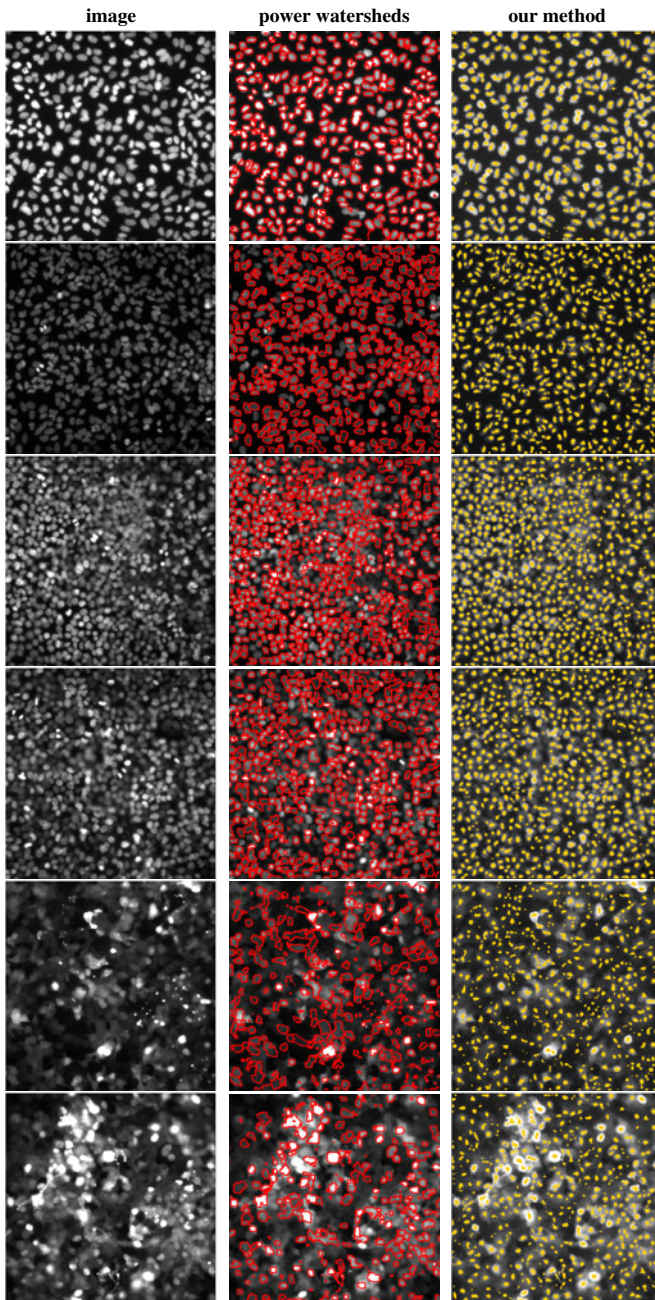


Fig. 3. Results by Power watersheds and our method on human alveolar basal epithelial A549 cells (rows 1-4) and embryonic kidney HEK293T cells (rows 5-6). While the quality of segmentation degrades for power watersheds when the cells have a larger intensity variation, our method pops out all the cells in these images with the same parameters and no post-processing.

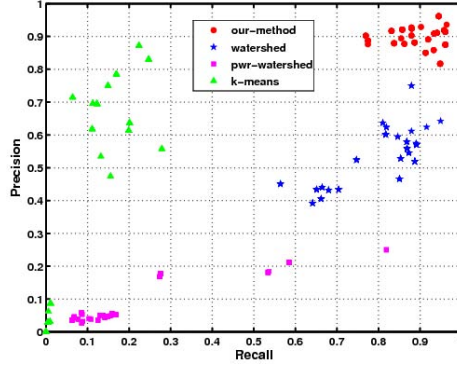


Fig. 4. Precision-recall statistics for k -means, watershed, power watersheds, and our method. Our method (red round dots, upper right corner) has better precision and recall overall.

We benchmark our method against human labeled dot centers, in comparison with 3 other approaches: k -means, standard watershed (MATLAB built-in implementation), and power watersheds [2] provided by its authors.

Given m ground-truth dot centers and n segment centers for an image, let D_{ij} be the Euclidean distance between dot i and segment j . If it is less than a certain radius threshold ρ , we consider (i, j) a matched detection.

$$\text{precision} = \frac{\#\{j : \min_{i=1}^m D_{ij} \leq \rho\}}{n} = \frac{\# \text{ nearest dots within radius } \rho}{\# \text{ segments}} \quad (10)$$

$$\text{recall} = \frac{\#\{i : \min_{j=1}^n D_{ij} \leq \rho\}}{m} = \frac{\# \text{ nearest segments within radius } \rho}{\# \text{ dots}} \quad (11)$$

The precision measures the proportion of true dots among all the segments, and the recall measures the proportion of segments among all the true dots.

Fig. 4 shows that our method performs much better than these other methods in terms of both precision and recall. k -means, clustering pixels based on their intensity values, particularly has trouble separating conjoined like-intensity cells, while increasing k only leads to clustering instability. While power watersheds has lower precision, it does noticeably improve boundary shapes of *segmented cells* over standard watershed, which is not properly measured in the precision-recall statistics. However, it tends to miss faint cells and the segmentation degrades with larger intensity variation (Fig. 3).

The quality of our segmentations depends on our patch segmenter and stitching constraints. The precision-recall statistics in Fig. 4 shows that our spectral graph partitioning approach works better than others at popping out small regions. Our stitching constraints can be appreciated by comparing the quality of segmentation without and with constraints: While there is no significant improvement in the recall ($p = 0.46$, t test), there is an average improvement of 0.04 in the precision ($p = 0.007$, t test).

To summarize, we present an efficient method for segmenting many small regions in a large image by constructing a set of patch segmentations which are independently derived but subject to stitching constraints between them. The quality of entire

segmentation depends on how to segment individual patches and how to stitch them together. Although our method works with any patch segmenter, it naturally integrates the stitching problem into each patch's spectral graph partitioning formulation. Our results demonstrate that stitching as partitioning constraints not only reduces segmentation complexity, but also corrects segmentation incongruences and imperfections.

Segmenting small structures in a large image faces a scale dilemma between the image size and the segment size, and our approach resolves the dilemma by decoupling the two sizes in constrained patch segmentations. We can expand it to a broader variety of shapes, e.g. thin structures such as blood vessels, by tailoring the weights to allow directional extension in the evaluation of pairwise pixel similarity.

Acknowledgements. This research is funded by NSF CAREER IIS-0644204 and a Clare Boothe Luce Professorship to Stella X. Yu. We would like to thank Nisha Sosale at the University of Pennsylvania for providing images and Camille Couprie at Université Paris Est and ESIEE for providing power watersheds results.

References

1. Meyer, F.: Topographic distance and watershed lines. *Signal Process* 38(1), 113–125 (1994)
2. Couprie, C., Grady, L., Najman, L., Talbot, H.: Power watersheds: a new image segmentation framework extending graph cuts, random walker and optimal spanning forest. In: *ICCV* (2009)
3. Shi, J., Malik, J.: Normalized cuts and image segmentation. *PAMI* 22(8), 888–905 (2000)
4. Yu, S.X.: Segmentation induced by scale invariance. In: *CVPR* (2005)
5. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient graph-based image segmentation. *IJCV* 59(2), 167–181 (2004)
6. Cour, T., Benezit, F., Shi, J.: Spectral segmentation with multiscale graph decomposition. In: *CVPR 2005*, pp. 1124–1131. IEEE Computer Society, Washington (2005)
7. Barbu, A., Zhu, S.C.: Graph partition by swendsen-wang cuts. In: *ICCV* (2003)
8. Galun, M., Sharon, E., Basri, R., Brandt, A.: Texture segmentation by multiscale aggregation of filter responses and shape elements. In: *ICCV* (2003)
9. Hofmann, T., Puzicha, J., Buhmann, J.M.: An optimization approach to unsupervised hierarchical texture segmentation. In: *ICIP* (1997)
10. Yu, S.X.: Segmentation using multiscale cues. In: *CVPR* (2004)
11. Bernardis, E., Yu, S.X.: Finding dots: Segmentation as popping out regions from boundaries. In: *CVPR* (2010)
12. Yu, S.X., Shi, J.: Segmentation given partial grouping constraints. *PAMI* 26(2) (2004)
13. Yu, S.X., Shi, J.: Multiclass spectral clustering. In: *ICCV*, October 11-17 (2003)
14. Staal, J., Kalitzin, S., Romeny, B.T.H., Viergever, M.: Detection of critical structures in scale space. In: Nielsen, M., Johansen, P., Fogh Olsen, O., Weickert, J. (eds.) *Scale-Space 1999*. LNCS, vol. 1682, p. 105. Springer, Heidelberg (1999)
15. Yu, S.X., Shi, J.: Understanding popout through repulsion. In: *CVPR* (2001)

Measurement Selection in Untracked Freehand 3D Ultrasound

Catherine Laporte^{1,2} and Tal Arbel²

¹ Dept. of Electrical Engineering, École de Technologie Supérieure, Montréal, Canada
² Centre for Intelligent Machines, McGill University, Montréal, Canada
catherine.laporte@etsmtl.ca, arbel@cim.mcgill.ca

Abstract. In freehand 3D ultrasound, out-of-plane transducer motion can be estimated via speckle decorrelation instead of using a position tracking device. This approach was recently adapted to arbitrary media by predicting elevational decorrelation curves from local image statistics. However, such adaptive models tend to yield biased measurements in the presence of spatially persistent structures. To account for such failures, this paper introduces a new iterative algorithm for probabilistic fusion and selection of correlation measurements. In experiments with imagery of animal tissue, the approach yields significant accuracy improvements over alternatives which do not apply principled measurement selection.

1 Introduction

Freehand 3D ultrasound (US) involves the reconstruction of a volume from 2D US images [15]. The relative positions of the 2D images are generally obtained from an external position tracking device, at the expense of clinical convenience. Alternatively, in-plane motion can be recovered using image registration techniques, and out-of-plane motion can be estimated from elevational speckle decorrelation [2,16,14]. By scanning a speckle phantom at regular elevational intervals and pooling the correlation coefficients computed between pairs of images, a transducer specific decorrelation curve can be constructed so as to estimate out-of-plane motion between pairs of correlated images in new data sets [2,16].

The above approach only works under Rayleigh scattering conditions, which seldom occur in real tissue. Recent solutions to this problem predict the shape of the local decorrelation curve from signal statistics within the image plane [3,8]. While this dramatically reduces error, experiments suggest that accuracy deteriorates for larger displacements [6], as shown in Fig. 1. The problem is that these adaptive models predict complete image decorrelation as displacement increases. Yet, in the presence of spatially persistent structures like the bright layer of fat in the pork flank sample of Fig. 1(b), correlation remains significant even for large displacements. When computing transducer trajectory, distance estimates arising from inaccurate parts of the adapted model should be discarded. Moreover, the displacement range over which the adaptive model is accurate and the extent of the bias outside this range are subject dependent. Therefore, the measurement selection strategy must dynamically adapt to the data.

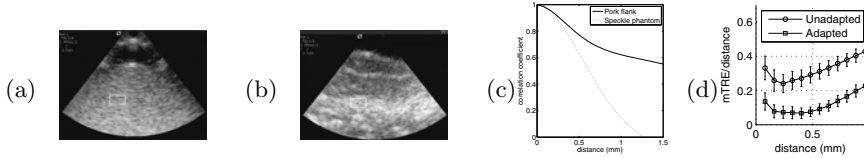


Fig. 1. (a) Speckle phantom image; (b) Pork flank image; (c) Decorrelation curves associated with the data in the white rectangles; (d) Relative motion estimation error obtained for the pork flank data set with and without medium specific adaptation [8]

Measurement selection for image-based transducer tracking is discussed in [4], which proposes to compute a robust average of reconstructions obtained from interleaved subsets of the image frames. Optimal assignment of frames to independent reconstructions requires consecutive frames of a given subset to be correlated, but not so much that they lie on the flat, error-prone part of the decorrelation curve [13]. However, this does not explicitly account for the kind of significant but uninformative image correlations shown in Fig. 1.

This paper proposes a new iterative measurement fusion and selection algorithm based on statistical tests to account for persistent structures in the medium. The method extends the probabilistic approach of [7] for use with imagery of real tissue using the technique in [8] and prevents errors arising from situations like Fig. 1. The structure of this paper is as follows. Section 2 reviews the probabilistic speckle decorrelation model and its extension to arbitrary tissue. Section 3 describes the new iterative measurement selection and probabilistic fusion scheme based on this model. Section 4 discusses experiments on animal tissue data, showing that the new approach offers improved accuracy over approaches that do not use principled measurement selection.

2 Probabilistic Speckle Decorrelation Model

In this work, US images are divided into Q patches (e.g. Fig. 2, left) to account for the spatial variation of elevational beam width. Each patch has its own nominal speckle decorrelation model. During a calibration step, a speckle phantom is scanned at regular elevational intervals. Gaussians are fitted to the patch-wise axial, lateral and elevational decorrelation curves derived from this imagery; their standard deviations are termed the nominal axial, lateral and elevational *correlation lengths* of the transducer, and labelled \hat{w}_a , \hat{w}_l and \hat{w}_e . A refined nominal probabilistic elevational speckle decorrelation model is also computed for each patch by maximum entropy analysis [11, 7], yielding a log-likelihood function $\hat{L}(\rho|\delta, q) = \log \hat{p}(\rho|\delta, q)$ that captures the stochasticity of correlation measurements ρ in relation to displacement δ for image location q .

In real tissue, US correlates over larger distances than in a speckle phantom. Here, this phenomenon is modelled as a scaling of the nominal elevational decorrelation model along the δ axis by a factor r_e , determined from four statistical

features measured locally within the image plane: the axial and lateral correlation lengths, w_a and w_l , and the squared signal to noise ratio $R^2 = \frac{E\{I\}^2}{VAR\{I\}}$ and skewness $S = \frac{E\{(I-E\{I\})^3\}}{VAR\{I\}^{\frac{3}{2}}}$, where I is the squared US envelope.

As described in [8], the relationship between these image statistics and the local elevational correlation length w_e is learned from a pool of synthetic US data depicting varied media, using sparse Gaussian process regression [14]. To enforce transducer independence, the regressive relationship is established using normalised features $\{r_a = \frac{w_l}{w_a}, r_l = \frac{w_l}{w_l}, r_{R^2} = \frac{R^2}{R^2}, r_S = \frac{S}{S}\}$ and $r_e = \frac{w_e}{w_e}$. In spite of its simplicity, this model was shown [8] to have accuracy comparable to that of [3]. Both models tend to fail for large elevational separations. This paper shows how to detect such failures.

In new, real imagery, r_e is estimated locally from in-plane image features using the learned regressor. The probabilistic nominal speckle decorrelation model is then scaled, yielding a locally adapted probabilistic estimator of elevational separation, $L(\rho_{ij}|\delta, q) = \hat{L}(\rho_{ij}|r_e\delta, q)$. In practice, δ is reasonably approximated by a Gaussian random variable with mean $\bar{\delta}$ and variance σ^2 given by

$$\bar{\delta} = \underset{\delta}{\operatorname{argmax}} L(\rho|\delta, q) \quad \text{and} \quad \sigma^2 = - \left(\frac{d^2}{d\delta^2} L(\rho|\bar{\delta}, q) \right)^{-1}. \quad (1)$$

3 Recovering Transducer Motion

In order to eliminate sources of error outside the control of the proposed techniques from our experiments, it is assumed throughout the paper that the transducer does not move in-plane (or else the images have been correctly re-aligned by an image registration technique) and that any out-of-plane rotations are small. Note that the measurement selection technique introduced in this paper would apply equally to more complete motion models.

Subdividing the image frames in the US data set into Q patches corresponding to those used to define the speckle decorrelation model breaks the data set into Q individual smaller US data sets consisting of “mini-frames” the size of an individual image patch. The relative positions of the full US image frames are then recovered as follows: (1) For every pair of corresponding mini-frames i and j , a correlation measurement is obtained and $r_{e_{ij}}$ is estimated using the learning-based approach of [8]; (2) For each image location $q \in Q$, estimates of elevational mini-frame positions are derived from redundant correlation-based distance estimates deemed shorter than $w_{e_{ij}} = r_{e_{ij}}\hat{w}_e$. This is done using the measurement selection and fusion algorithm of section 3.1.1 (3) The resulting data are smoothed across each frame using a median filter (as suggested in [3]); (4) Rigid transformations relating every frame to the first are computed using a least-median-of-squares approach [12,7].

¹ Although the data sets used in the experiments involved only monotonic probe motion, the combinatorial frame ordering algorithm of [7] was also applied at this stage and may have influenced the outcome of the experiments.

3.1 Iterative Measurement Selection and Fusion

Let $z_i, i = 0, \dots, n$ be the position of the i th mini-frame corresponding to patch q along the elevational direction relative to $z_0 = 0$ and let \mathbf{z} be the vector formed by the $z_i, i = 1, \dots, n$. The correlation coefficient ρ_{ij} between mini-frames i and j provides an estimate of their elevational separation $\delta_{ij} = |z_j - z_i|$ via (11). Assuming that the distance estimates obtained for different mini-frame pairs are mutually independent given the configuration of the frames, an estimate of \mathbf{z} and its covariance matrix \mathbf{C} are obtained by solving (9.7):

$$\mathbf{G}\mathbf{z}^* = \mathbf{b} \quad \text{and} \quad \mathbf{C} = \mathbf{G}^{-1}, \quad (2)$$

$$\text{with } G_{ij} = \begin{cases} \sum_{k=0}^n \frac{1}{\sigma_{ik}^2} & i = j \\ -\frac{1}{\sigma_{ij}^2} & \text{otherwise} \end{cases} \quad \text{and} \quad b_i = \sum_{j=0, j \neq i}^n \frac{\bar{\delta}_{ij}}{\sigma_{ij}^2}. \quad (3)$$

In principle, this approach could be used to combine all available correlation measurements into a trajectory estimate for image location q . In practice, this is highly error prone because, as illustrated in Fig. 1, the (adapted) speckle decorrelation model exhibits bias in certain parts of its range. Biased measurements can be excluded from the data fusion process through an iterative algorithm based on statistical hypothesis testing. Assuming an unbiased initial trajectory estimate $\mathbf{z}_{[0]}$, iteration t proceeds as follows:

1. A new subset of the correlation measurements are tested against the current estimate of the trajectory. For each candidate δ_{ij} derived via (11), let H_0 denote the hypothesis that it is consistent with $\mathbf{z}_{[t-1]}$, and let H_a denote the hypothesis that it is biased. H_0 is rejected in favour of H_a at the 1% significance level (and δ_{ij} is discarded) if

$$|\bar{\delta}_{ij} - |z_{[t-1]_j}^* - z_{[t-1]_i}^*|| \geq \Phi^{-1}(0.995, \sigma_{ij}^2 + C_{[t-1]_{jj}} + C_{[t-1]_{ii}} - 2C_{[t-1]_{ij}}), \quad (4)$$

where $\Phi^{-1}(\beta, \nu^2)$ is the inverse cumulative distribution function of a Gaussian random variable with variance ν^2 evaluated at probability β .

2. The measurements accepted by the above hypothesis tests are fused with all other previously accepted measurements to obtain $\mathbf{z}_{[t]}^*$ and $\mathbf{C}_{[t]}$ by solving (2).

The effectiveness of this approach depends on the initial trajectory estimate (see section 3.2) and the order in which the available measurements are processed (or *testing schedule*). This paper adopts a very simple testing schedule: the first iteration considers only correlation measurements between consecutively acquired mini-frames (unless they were already used to compute $\mathbf{z}_{[0]}$). The second iteration considers correlation measurements between frames which are two acquisition steps apart, and so on until no new measurements are accepted.

3.2 Initial Trajectory Estimate

The proposed scheme rests on the assumption of an unbiased $\mathbf{z}_{[0]}$. In the absence of other data, $\mathbf{z}_{[0]}$ is obtained from n correlation measurements. These must form

a spanning tree of the graph \mathcal{G} whose vertices and edges represent the mini-frames and the correlation measurements between them. A *minimal* spanning tree (MST) of \mathcal{G} is hence an optimal initial measurement set with respect to a cost function κ over pairs of mini-frames. One possible cost function is

$$\kappa_1(i, j) = |i - j|. \quad (5)$$

The associated MST consists of the n correlation measurements between consecutive mini-frames. Because the log-likelihood function is flat for large ρ , the uncertainty of the associated $\mathbf{z}_{[0]}$ is typically high so that initially, H_0 is rarely rejected by (4). The initial estimate is thus diluted by new, more certain measurements early on in the data fusion process. As long as modelling inaccuracies occur mostly for widely separated images (e.g. Fig. 1), this can lead to good results. However, empirical evidence suggests that model failures can occur for shorter displacements when the overall US signal is low, due to the increased contribution of unmodelled noise in the image [6]. In such cases, inaccurate estimates of r_e are obtained and short displacements are over-estimated. The following two cost functions are designed to alleviate this problem:

$$\kappa_2(i, j) = \text{sgn}(r_{e_{ij}} - 1)\bar{\delta}_{ij} \quad \text{and} \quad \kappa_3(i, j) = \text{sgn}(r_{e_{ij}} - 1)\bar{\delta}_{ij}/\sigma_{ij}^2. \quad (6)$$

Where the estimated $r_e > 1$, κ_2 and κ_3 are similar to κ_1 , preferring measurements with large relative variance. This changes in areas where the estimated $r_e < 1$, which is taken to be indicative of model failure.² In these error-prone areas, κ_2 and κ_3 select measurements with smaller relative variance, making the hypothesis test of (4) more discriminative.

4 Experiments

Imagery was acquired using a video frame-grabber connected to an Acuson Cypress US scanner with a 7V3C transducer at a depth setting of 2.7 cm. Images were divided into 26 patches of 50×30 pixels and back-converted to polar coordinates. Motion was purely elevational and controlled to 0.01 mm precision using a manually operated linear stage. Fifty-one 161 frame long scans of pork tenderloin, turkey breast and beef brisket samples were acquired at a range of different regular spacings. Fig. 2 shows sample images. A nominal probabilistic speckle decorrelation model was estimated from speckle phantom imagery acquired at regular 0.05 mm spacing. The algorithm of [10] was applied to phantom images to reverse log-compression in all data sets. The mapping from US image statistics to elevational correlation length was learned from synthetic image sequences depicting varied media, as described in [8]. For each animal tissue scan, the probe trajectory was estimated as described in Section 3.

² Indeed, true $r_e < 1$ is a very rare occurrence in the noiseless training set because correlation length is theoretically minimal under Rayleigh scattering conditions ($r_e = 1$). Thus, small values of r_e , which arise in the presence of noisy real data, while often estimated as small (< 1), are unlikely to be estimated *accurately* by the learning-based algorithm.

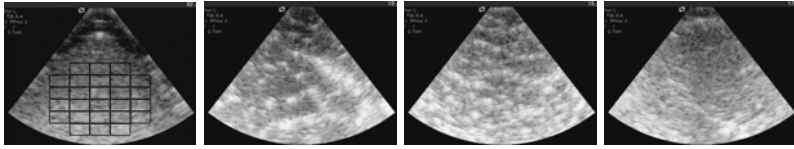


Fig. 2. From left to right: sample speckle phantom, pork, beef and turkey images. The phantom image shows the subdivision of imagery into 26 patches.

Three versions of the proposed approach, HYP-K1, HYP-K2 and HYP-K3, were considered, respectively corresponding to the initialisation cost functions of (5) and (6). Two baseline approaches were also tested: (a) a nearest-neighbour (NN) approach, which uses only deterministic distance estimates from consecutive pairs of mini-frames and (b) solving (2) once using all δ_{ij} deemed shorter than $w_{e_{ij}} = r_{e_{ij}}^* \hat{w}_e$, an approach labelled NMS (no measurement selection). The latter is a direct adaptation of [7] (shown to work well under Rayleigh scattering conditions) for use in imagery of real tissue via [8].

For statistical analysis, accuracy was summarised by the mean target registration error (mTRE) measured at the mini-frame centres, averaged over all frames and normalised by the length of the trajectory. A mixed within/between-subject factor ANOVA was carried out on the results of the 51 experiments to evaluate the effect of algorithm choice and tissue type.³ The HYP-K2 and HYP-K3 variants of the proposed approach significantly outperformed both baselines ($p < 0.001$ in both cases). HYP-K1 significantly outperformed NMS ($p < 0.001$), but was not significantly more accurate than NN ($p = 0.251$).

Fig. 3 illustrates these results with examples. In most cases, HYP-K2 and HYP-K3 gave excellent results (e.g. Fig. 3(a)-(c)). Though NN and HYP-K1 may exhibit similar mTRE, they differ in qualitative behaviour. NN underestimates displacement because the deterministic decorrelation model is biased for short displacements [13]. HYP-K1 drifts because it easily accepts displacement over-estimates arising from unmodelled image noise (e.g. Fig. 3(a), (c)), as discussed in Section 3.2. This likely explains the significantly lower accuracy obtained with turkey data (which has lower echogenicity) than with beef data ($p = 0.042$). HYP-K2 and HYP-K3 were usually able to compensate for this modelling limitation (see Fig. 3(c)).

Though they generally gave good results, HYP-K2 and HYP-K3 were not always able to overcome failures of the adaptive speckle decorrelation model. For example, in Fig. 3(d), distances were over-estimated by all approaches except, uncharacteristically, NMS. Here, the κ_2 and κ_3 cost functions probably failed to protect the trajectory estimate against distance over-estimates because there were few areas where $r_e < 1$. While $r_e < 1$ is a good indicator of model failure, $r_e > 1$ does not always indicate success.

³ The Bonferroni adjustment was used when carrying out multiple comparisons and statistical significance was evaluated at the 5% level.

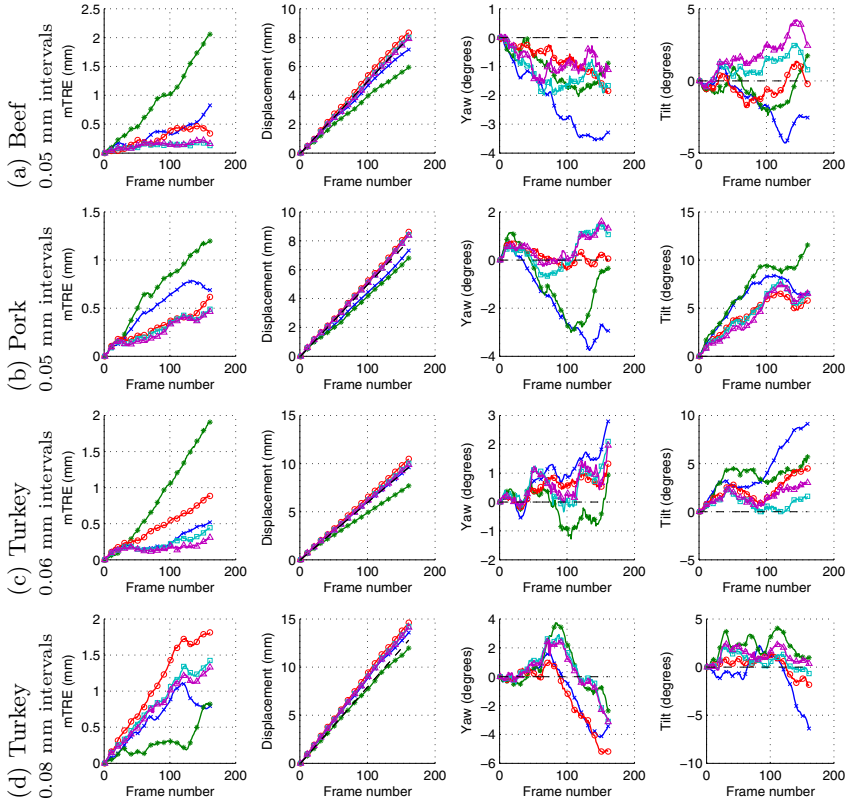


Fig. 3. Sample results, markers placed every 10 data points for greyscale legibility. From left to right: mTRE, estimated displacement, yaw and tilt evaluated at the centroid of the grid of image patches. Dashed: ground truth; blue crosses: NN; green stars: NMS; red circles: HYP-K1; cyan squares: HYP-K2; magenta triangles: HYP-K3.

5 Conclusions

This paper considered a new approach for image-based recovery of out-of-plane US transducer trajectories which generalises the probabilistic data fusion of [7] to arbitrary media using the learning-based approach of [8]. In order to discard significant, but uninformative image correlation measurements arising from spatially persistent structures in the medium, a novel iterative algorithm was presented which alternates measurement fusion steps with measurement selection steps based on statistical hypothesis testing. Despite some imperfections, experiments show that, if well initialised, the algorithm is statistically significantly more accurate than a deterministic nearest-neighbour approach or probabilistic fusion without measurement selection. It would be interesting to compare the new approach to that of [4] and possibly combine them by applying the probabilistic techniques in this paper to the frame subsets selected and averaged in [4].

Future research could also exploit the measurement selection approach in a context where image-based measurements complement an external tracking device providing mostly unbiased but highly imprecise positional data [5]. The latter, along with a valid noise model, could initialise the iterative algorithm, thereby improving its robustness.

References

1. Baker, R.: Probability estimation and information principles. *Struct. Saf.* 9, 97–116 (1990)
2. Chen, J.-F., Fowlkes, J.B., et al.: Determination of scan-plane motion using speckle decorrelation: theoretical considerations and initial test. *Int. J. Imaging Syst. Technol.* 8(1), 38–44 (1997)
3. Gee, A.H., Housden, R.J., et al.: Sensorless freehand 3D ultrasound in real tissue: speckle decorrelation without fully developed speckle. *Med. Image Anal.* 10(2), 137–149 (2006)
4. Housden, R.J., Gee, A.H., et al.: Sensorless reconstruction of unconstrained freehand 3D ultrasound data. *Ultrasound Med. Biol.* 33(3), 408–419 (2007)
5. Lang, A., Mousavi, P., et al.: Fusion of electromagnetic tracking with speckle-tracked 3D freehand ultrasound using an unscented Kalman filter. In: *Proc. SPIE Med. Imaging*, pp. 72651A-1–72651A-12 (2009)
6. Laporte, C.: Statistical methods for out-of-plane ultrasound transducer motion estimation. PhD thesis, McGill University (2009)
7. Laporte, C., Arbel, T.: Combinatorial and probabilistic fusion of noisy correlation measurements for untracked freehand 3D ultrasound. *IEEE TMI* 27(7), 984–994 (2008)
8. Laporte, C., Arbel, T.: Learning a tissue invariant ultrasound speckle decorrelation model. In: *Proc. ISBI*, pp. 995–998 (2009)
9. Lu, F., Milios, E.: Globally consistent range scan alignment for environment mapping. *Auton. Robots* 4(4), 333–349 (1997)
10. Prager, R.W., Gee, A.H., et al.: Decompression and speckle detection for ultrasound images using the homodyned k-distribution. *Patt. Recognit. Lett.* 24, 705–713 (2003)
11. Prager, R.W., Gee, A.H., et al.: Sensorless freehand 3-D ultrasound using regression of the echo intensity. *Ultrasound Med. Biol.* 29(3), 437–446 (2003)
12. Rousseeuw, P.J., Leroy, A.M.: *Robust Regression & Outlier Detection*. Wiley, Chichester (1987)
13. Smith, W., Fenster, A.: Statistical analysis of decorrelation-based transducer tracking for three-dimensional ultrasound. *Med. Phys.* 30(7), 1580–1591 (2003)
14. Snelson, E., Ghahramani, Z.: Sparse Gaussian processes using pseudo-inputs. In: *NIPS*, vol. 18, pp. 1257–1264 (2006)
15. Solberg, O.V., Lindseth, F., et al.: Freehand 3D ultrasound reconstruction algorithms — a review. *Ultrasound Med. Biol.* 33(7), 991–1009 (2007)
16. Tuthill, T.A., Krücker, J.F., et al.: Automated three-dimensional US frame positioning computed from elevational speckle decorrelation. *Radiology* 209(2), 575–582 (1998)

Efficient MR Image Reconstruction for Compressed MR Imaging

Junzhou Huang, Shaoting Zhang, and Dimitris Metaxas

Division of Computer and Information Sciences,
Rutgers University,
NJ, USA 08854

Abstract. In this paper, we propose an efficient algorithm for MR image reconstruction. The algorithm minimizes a linear combination of three terms corresponding to a least square data fitting, total variation (TV) and $L1$ norm regularization. This has been shown to be very powerful for the MR image reconstruction. First, we decompose the original problem into $L1$ and TV norm regularization subproblems respectively. Then, these two subproblems are efficiently solved by existing techniques. Finally, the reconstructed image is obtained from the weighted average of solutions from two subproblems in an iterative framework. We compare the proposed algorithm with previous methods in term of the reconstruction accuracy and computation complexity. Numerous experiments demonstrate the superior performance of the proposed algorithm for compressed MR image reconstruction.

1 Introduction

Magnetic Resonance (MR) Imaging has been widely used in medical diagnosis because of its non-invasive manner and excellent depiction of soft tissue changes. Recent developments in compressive sensing theory [1][2] show that it is possible to accurately reconstruct the Magnetic Resonance (MR) images from highly undersampled K-space data and therefore significantly reduce the scanning duration.

Suppose x is a MR image and R is a partial Fourier transform, the sampling measurement b of x in K-space is defined as $b = Rx$. The compressed MR image reconstruction problem is to reconstruct x giving the measurement b and the sampling matrix R . Motivated by the compressive sensing theory, Lustig et al. [3] proposed their pioneering work for the MR image reconstruction. Their method can effectively reconstruct MR images with only 20% sampling. The improved results were obtained by having both a wavelet transform and a discrete gradient in the objective, which is formulated as follows:

$$\hat{x} = \arg \min_x \left\{ \frac{1}{2} \|Rx - b\|^2 + \alpha \|x\|_{TV} + \beta \|\Phi x\|_1 \right\} \quad (1)$$

where α and β are two positive parameters, b is the undersampled measurements of K-space data, R is a partial Fourier transform and Φ is a wavelet transform. It is based on the fact that the piecewise smooth MR images of organs can be sparsely represented by the wavelet basis and should have small total variations. The TV

was defined discretely as $\|x\|_{TV} = \sum_i \sum_j ((\nabla_1 x_{ij})^2 + (\nabla_2 x_{ij})^2)$ where ∇_1 and ∇_2 denote the forward finite difference operators on the first and second coordinates, respectively. Since both $L1$ and TV norm regularization terms are nonsmooth, this problem is very difficult to solve. The conjugate gradient (CG) [3] and PDE [4] methods were used to attack it. However, they are very slow and impractical for real MR images. Computation became the bottleneck that prevented this good model [1] from being used in practical MR image reconstruction.

Other methods tried to reconstruct compressed MR images by performing L_p -quasinorm ($p < 1$) regularization optimization [5][6][7]. These nonconvex methods do not always give global minima and are also relatively slow. Trzasko et al. [8] used the homotopic nonconvex L_0 -minimization to reconstruct MR images. It was faster than those L_p -quasinorm regularization methods. However, it still needed 1–3 minutes to obtain reconstructions of 256×256 images in MATLAB on a 3 GHz desktop computer. Recently, two fast methods were proposed to directly solve [1]. In [9], Ma et al. proposed an operator-splitting algorithm (TVCMRI) to solve the MR image reconstruction problem. In [10], a variable splitting method (RecPF) was proposed to solve the MR image reconstruction problem. Both of them can replace iterative linear solvers with Fourier domain computations, which can gain substantial time savings. In MATLAB on a 3 GHz desktop computer, they can be used to obtain good reconstructions of 256×256 images in ten seconds or less. They are two of the fastest MR image reconstruction methods so far.

In this paper, we propose a new MR image reconstruction method based on the combination of variable and operator splitting techniques. We decompose the hard composite regularization problem [1] into two simpler regularization subproblems by: 1) splitting variable x into two variables $\{x_i\}_{i=1,2}$; 2) performing operator splitting to minimize total variation regularization and L1 norm regularization subproblems over $\{x_i\}_{i=1,2}$ respectively and 3) obtaining the solution x by linear combination of $\{x_i\}_{i=1,2}$. This includes both variable splitting and operator splitting. We call it the Composite Splitting Algorithm (CSA). Motivated by the effective acceleration scheme in Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) [11], the proposed CSA is further accelerated with an additional acceleration step. Numerous experiments have been conducted on real MR images to compare the proposed algorithm with previous methods. Experimental results show that it impressively outperforms previous methods for the MR image reconstruction in terms of both reconstruction accuracy and computation complexity.

2 Related Acceleration Algorithm

In this section, we briefly review the FISTA in [11], since our methods are motivated by it. FISTA consider to minimize the following problem:

$$\min\{F(x) \equiv f(x) + g(x), x \in \mathbf{R}^n\} \quad (2)$$

where f is a smooth convex function with Lipschitz constant L_f , and g is a convex function which may be nonsmooth.

Algorithm 1. FISTA [11]

Input: $\rho = 1/L_f$, $r^1 = x^0$, $t^1 = 1$
for $k = 1$ **to** K **do**
 $x_g = r^k - \rho \nabla f(r^k)$
 $x^k = \text{prox}_\rho(g)(x_g)$
 $t^{k+1} = \frac{1 + \sqrt{1 + 4(t^k)^2}}{2}$
 $r^{k+1} = x^k + \frac{t^k - 1}{t^{k+1}}(x^k - x^{k-1})$
end for

Algorithm 2. CSD

Input: $\rho = 1/L$, α , β , $z_1^0 = z_2^0 = x_g$
for $j = 1$ **to** J **do**
 $x_1 = \text{prox}_\rho(2\alpha \|x\|_{TV})(z_1^{j-1})$
 $x_2 = \text{prox}_\rho(2\beta \|\Phi x\|_1)(z_2^{j-1})$
 $x^j = (x_1 + x_2)/2$
 $z_1^j = z_1^{j-1} + x^j - x_1$
 $z_2^j = z_2^{j-1} + x^j - x_2$
end for

ϵ -optimal Solution: Suppose x^* is an optimal solution to (2). $x \in \mathbf{R}^p$ is called an ϵ -optimal solution to (2) if $F(x) - F(x^*) \leq \epsilon$ holds.

Gradient: $\nabla f(x)$ denotes the gradient of the function f at the point x .

The proximal map: given a continuous convex function $g(x)$ and any scalar $\rho > 0$, the proximal map associated to function g is defined as follows [11] [12]:

$$\text{prox}_\rho(g)(x) := \arg \min_u \{g(u) + \frac{1}{2\rho} \|u - x\|^2\} \quad (3)$$

Algorithm 1 outlines the FISTA. It can obtain an ϵ -optimal solution in $\mathcal{O}(1/\sqrt{\epsilon})$ iterations.

Theorem 1. (Theorem 4.1 in [11]): Suppose $\{x^k\}$ and $\{r^k\}$ are iteratively obtained by the FISTA, then, we have

$$F(x^k) - F(x^*) \leq \frac{2L_f \|x^0 - x^*\|^2}{(k+1)^2}, \forall x^* \in X_*$$

The efficiency of the FISTA highly depends on being able to quickly solve its second step $x^k = \text{prox}_\rho(g)(x_g)$. For simpler regularization problems, it is possible, i.e., the FISTA can rapidly solve the $L1$ regularization problem with cost $\mathcal{O}(n \log(n))$ [11] (where n is the dimension of x), since the second step $x^k = \text{prox}_\rho(\beta \|\Phi x\|_1)(x_g)$ has a close form solution; It can also quickly solve the TV regularization problem, since the step $x^k = \text{prox}_\rho(\alpha \|x\|_{TV})(x_g)$ can be compute with cost $\mathcal{O}(n)$ [12]. However, the FISTA can not efficiently solve the composite $L1$ and TV regularization problem (1), since no efficient algorithm exists to solve the step

$$x^k = \text{prox}_\rho(\alpha \|x\|_{TV} + \beta \|\Phi x\|_1)(x_g). \quad (4)$$

To solve the problem (1), the key problem is thus to develop an efficient algorithm to solve problem (4). In the following section, we will show that a scheme based on composite splitting techniques can be used to do this.

3 Algorithm

From the above introduction, we know that, if we can develop a fast algorithm to solve problem (4), the MR image reconstruction problem can then be efficiently

solved by the FISTA, which obtains an ϵ -optimal solution in $\mathcal{O}(1/\sqrt{\epsilon})$ iterations. Actually, problem (4) can be considered as a denoising problem:

$$x^k = \arg \min_x \left\{ \frac{1}{2} \|x - x_g\|^2 + \rho\alpha \|x\|_{TV} + \rho\beta \|\Phi x\|_1 \right\}. \quad (5)$$

We use composite splitting techniques to solve this problem: 1) splitting variable x into two variables $\{x_i\}_{i=1,2}$; 2) performing operator splitting over each of $\{x_i\}_{i=1,2}$ independently and 3) obtaining the solution x by linear combination of $\{x_i\}_{i=1,2}$. We call it Composite Splitting Denoising (CSD) method, which is outlined in Algorithm 2. Its validity is guaranteed by the following theorem:

Theorem 2. *Suppose $\{x^j\}$ the sequence generated by the CSD. Then, x^j will converge to $\text{prox}_\rho(\alpha\|x\|_{TV} + \beta\|\Phi x\|_1)(x_g)$, which means that we have $x^j \rightarrow \text{prox}_\rho(\alpha\|x\|_{TV} + \beta\|\Phi x\|_1)(x_g)$.*

Due to page limitations, the proof for this theorem is given in the supplemental material.

Combining the CSD with FISTA, a new algorithm, FCSA, is proposed for MR image reconstruction problem (1). In practice, we found that a small iteration number J in the CSD is enough for the FCSA to obtain good reconstruction results. Especially, it is set as 1 in our algorithm. Numerous experimental results in the next section will show that it is good enough for real MR image reconstruction.

Algorithm 4 outlines the proposed FCSA. In this algorithm, if we remove the acceleration step by setting $t^{k+1} \equiv 1$ in each iteration, we will obtain the Composite Splitting Algorithm (CSA), which is outlined in Algorithm 3. A key feature of the FCSA is its fast convergence performance borrowed from the FISTA. From Theorem 1, we know that the FISTA can obtain an ϵ -optimal solution in $\mathcal{O}(1/\sqrt{\epsilon})$ iterations.

Another key feature of the FCSA is that the cost of each iteration is $\mathcal{O}(n \log(n))$, as confirmed by the following observations. The step 4, 6 and 7 only involve adding vectors or scalars, thus cost only $\mathcal{O}(n)$ or $\mathcal{O}(1)$. In step 1, $\nabla f(r^k = R^T(Rr^k - b)$ since $f(r^k) = \frac{1}{2}\|Rr^k - b\|^2$ in this case. Thus, this step only costs $\mathcal{O}(n \log(n))$. As introduced above, the step $x^k = \text{prox}_\rho(2\alpha\|x\|_{TV})(x_g)$ can be computed quickly with cost $\mathcal{O}(n)$ [12]; The step $x^k = \text{prox}_\rho(2\beta\|\Phi x\|_1)(x_g)$ has a close form solution and can be computed with cost $\mathcal{O}(n \log(n))$. In the step $x^k = \text{project}(x^k, [l, u])$, the function $x = \text{project}(x, [l, u])$ is defined as: 1) $x = x$ if $l \leq x \leq u$; 2) $x = l$ if $x < l$; and 3) $x = u$ if $x > u$, where $[l, u]$ is the range of x . For example, in the case of MR image reconstruction, we can let $l = 0$ and $u = 255$ for 8-bit gray MR images. This step costs $\mathcal{O}(n)$. Thus, the total cost of each iteration in the FCSA is $\mathcal{O}(n \log(n))$.

With these two key features, the FCSA efficiently solves the MR image reconstruction problem (1) and obtains better reconstruction results in terms of both the reconstruction accuracy and computation complexity. The experimental results in the next section demonstrate its superior performance compared with all previous methods for compressed MR image reconstruction.

Algorithm 3. CSA

Input: $\rho = 1/L$, α , β , $t^1 = 1$ $x^0 = r^1$
for $k = 1$ **to** K **do**
 $x_g = r^k - \rho \nabla f(r^k)$
 $x_1 = \text{prox}_\rho(2\alpha \|x\|_{TV})(x_g)$
 $x_2 = \text{prox}_\rho(2\beta \|\Phi x\|_1)(x_g)$
 $x^k = (x_1 + x_2)/2$
 $x^k = \text{project}(x^k, [l, u])$
 $r^{k+1} = x^k$
end for

Algorithm 4. FCSA

Input: $\rho = 1/L$, α , β , $t^1 = 1$ $x^0 = r^1$
for $k = 1$ **to** K **do**
 $x_g = r^k - \rho \nabla f(r^k)$
 $x_1 = \text{prox}_\rho(2\alpha \|x\|_{TV})(x_g)$
 $x_2 = \text{prox}_\rho(2\beta \|\Phi x\|_1)(x_g)$
 $x^k = (x_1 + x_2)/2$; $x^k = \text{project}(x^k, [l, u])$
 $t^{k+1} = (1 + \sqrt{1 + 4(t^k)^2})/2$
 $r^{k+1} = x^k + ((t^k - 1)/t^{k+1})(x^k - x^{k-1})$
end for

4 Experiments

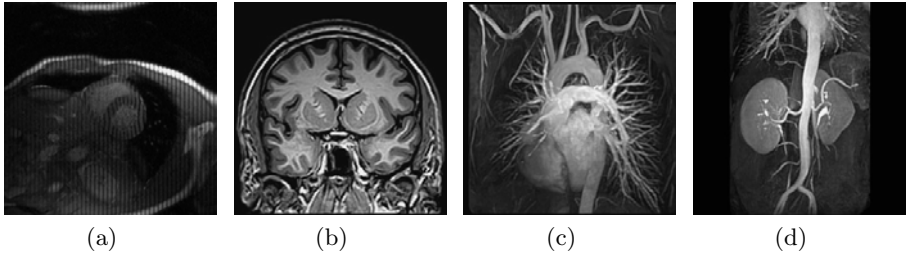


Fig. 1. MR images: (a) Cardiac; (b) Brain; (c) Chest and (d) Artery

4.1 Experiment Setup

Suppose a MR image x has n pixels, the partial Fourier transform R in problem (1) consists of m rows of a $n \times n$ matrix corresponding to the full 2D discrete Fourier transform. The m selected rows correspond to the acquired b . The sampling ratio is defined as m/n . The scanning duration is shorter if the sampling ratio is smaller. In MR imaging, we have certain freedom to select the rows, which correspond to certain frequencies. In the k-space, we randomly obtain more samples in low frequencies and less samples in higher frequencies. This sampling scheme is the same of those in [3, 9, 10] and has been widely used for compressed MR image reconstruction. Following these guidelines, we randomly created sampling matrices. Practically, the sampling scheme and speed in MR imaging also depend on the physical and physiological limitations [3].

We implement our CSA and FCSA for problem (1) and apply them on 2D real MR images. All experiments are conducted on a 2.4GHz PC in Matlab environment. We compare the CSA and FCSA with the classic MR image reconstruction method based on the CG [3]. We also compare them with two of the fastest MR image reconstruction methods, TVCMRI [9] and RecPF [10]. For fair comparisons, we download the codes from their websites and carefully follow

their experiment setup. For example, the observation measurement b is synthesized as $b = Rx + \mathbf{n}$, where \mathbf{n} is Gaussian white noise with standard deviation $\sigma = 0.01$. The regularization parameter α and β are set as 0.001 and 0.035. R and b are given as inputs, and x is the unknown target. For quantitative evaluation, we compute the Signal-to-Noise Ratio (SNR) for each reconstruction result.

4.2 Numerical Results

We apply all methods on four 2D MR images: cardiac, brain, chest and artery respectively. Figure 1 shows these images. For convenience, they have the same size of 256×256 . The sample ratio is set to be approximately 25%. To perform fair comparisons, all methods run 50 iterations except that the CG runs only 8 iterations due to its higher complexity. The CPU time and SNR are traced in each iteration for each of methods. To reduce the randomness, we run each experiments 100 times for each parameter setting of each method.

Figure 2 shows the visual comparisons of the reconstructed results by different methods. Figure 3 gives the performance comparisons between different methods in terms of the CPU time over SNR. The FCSA always obtains the best reconstruction results on all MR images by achieving the highest SNR in less CPU time. The CSA is always inferior to the FCSA, which shows the effectiveness of acceleration steps in the FCSA for the MR image reconstruction. While the classical CG [3] is far worse than others because of its higher cost in each iteration, the RecPF sound be slightly better than the TVCMRI, which is consistent to the observations in [9] and [10].

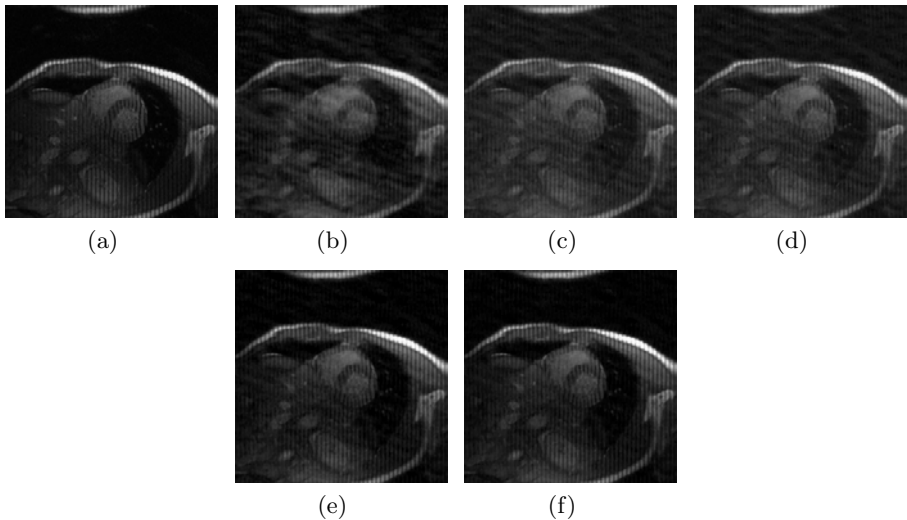


Fig. 2. Cardiac MR image reconstruction (a) Original image; (b), (c), (d) (e) and (f) are the reconstructed images by the CG [3], TVCMRI [9], RecPF [10], CSA and FCSA. Their SNR are 9.86, 14.43, 15.20, 16.46 and 17.57 (db). Their CPU time are 2.87, 3.14, 3.07, 2.22 and 2.29 (s).

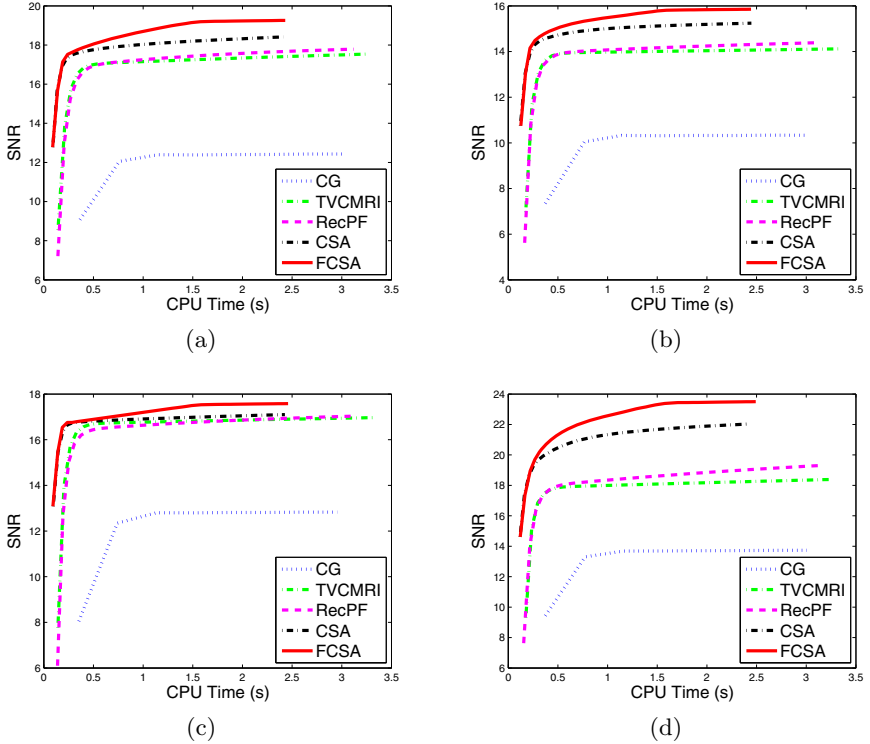


Fig. 3. Performance comparisons (CPU-Time vs. SNR) on different MR images: a) Cardiac image; b) Brain image; c) Chest image and d) Artery image

The reconstructed results produced by the FCSA are far better than those produced by the CG, TVCMRI and RecPF. The reconstruction performance of the FCSA is always the best in terms of both the reconstruction accuracy and the computational complexity, which clearly demonstrate the effective and efficiency of the FCSA for the compressed MR image construction.

5 Conclusion

We have proposed an efficient algorithm for the compressed MR image reconstruction. Our work has the following contributions. First, the proposed FCSA can efficiently solve a composite regularization problem including both TV term and L1 norm term, which can be easily extended to other medical image applications. Second, the computational complexity of the FCSA is only $\mathcal{O}(n \log(n))$ in each iteration. It also has strong convergence properties. These properties make the real compressed MR image reconstruction much more feasible than before. Finally, we conduct numerous experiments to compare different reconstruction methods. Our method is shown to impressively outperform the classic methods and two of fastest methods so far in terms of both accuracy and complexity.

References

1. Candes, E.J., Romberg, J., Tao, T.: Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory* 52, 489–509 (2006)
2. Donoho, D.: Compressed sensing. *IEEE Transactions on Information Theory* 52(4), 1289–1306 (2006)
3. Lustig, M., Donoho, D., Pauly, J.: Sparse mri: The application of compressed sensing for rapid mr imaging. *Magnetic Resonance in Medicine* 58, 1182–1195 (2007)
4. He, L., Chang, T.C., Osher, S., Fang, T., Speier, P.: Mr image reconstruction by using the iterative refinement method and nonlinear inverse scale space methods. Technical report, UCLA CAM 06-35 (2006)
5. Ye, J., Tak, S., Han, Y., Park, H.: Projection reconstruction mr imaging using focuss. *Magnetic Resonance in Medicine* 57, 764–775 (2007)
6. Chartrand, R.: Exact reconstruction of sparse signals via nonconvex minimization. *IEEE Signal Processing Letters* 14, 707–710 (2007)
7. Chartrand, R.: Fast algorithms for nonconvex compressive sensing: Mri reconstruction from very few data. In: *Proceedings of ISBI* (2009)
8. Trzasko, J., Manduca, A., Borisch, E.: Highly undersampled magnetic resonance image reconstruction via homotopic l_0 -minimization. *IEEE Transactions on Medical Imaging* 28, 106–121 (2009)
9. Ma, S., Yin, W., Zhang, Y., Chakraborty, A.: An efficient algorithm for compressed mr imaging using total variation and wavelets. In: *Proceedings of CVPR* (2008)
10. Yang, J., Zhang, Y., Yin, W.: A fast alternating direction method for tvl1-l2 signal reconstruction from partial fourier data. *IEEE Journal of Selected Topics in Signal Processing*, Special Issue on Compressive Sensing 4(2) (2010)
11. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences* 2(1), 183–202 (2009)
12. Beck, A., Teboulle, M.: Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Transaction on Image Processing* 18(113), 2419–2434 (2009)

4D Computed Tomography Reconstruction from Few-Projection Data via Temporal Non-local Regularization

Xun Jia¹, Yifei Lou², Bin Dong³, Zhen Tian^{1,4}, and Steve Jiang¹

¹ Department of Radiation Oncology University of California, San Diego,
La Jolla, CA 92037-0843, USA

² Department of Mathematics, University of California, Los Angeles,
Los Angeles, CA 90095-1555, USA

³ Department of Mathematics, University of California, San Diego,
La Jolla, CA 92093-0112, USA

⁴ Department of Biomedical Engineering, Graduate School at
Tsinghua University, Shenzhen, Guangdong 518055, China

Abstract. 4D computed tomography (4D-CT) is an important modality in medical imaging due to its ability to resolve patient anatomy motion in each respiratory phase. Conventionally 4D-CT is accomplished by performing the reconstruction for each phase independently as in a CT reconstruction problem. We propose a new 4D-CT reconstruction algorithm that explicitly takes into account the temporal regularization in a non-local fashion. By imposing a regularization of a temporal non-local means (TNLM) form, 4D-CT images at all phases can be reconstructed simultaneously based on extremely under-sampled x-ray projections. Our algorithm is validated in one digital NCAT thorax phantom and two real patient cases. It is found that our TNLM algorithm is capable of reconstructing the 4D-CT images with great accuracy. The experiments also show that our approach outperforms standard 4D-CT reconstruction methods with spatial regularization of total variation or tight frames.

1 Introduction

Four-dimensional Computed Tomography (4D-CT) is one of the most important topics in medical imaging field that attract tremendous interests nowadays. In addition to providing three dimensional volumetric anatomical information as in conventional CT, 4D-CT is capable of resolving organ motions due to, for example, patient respiration by reconstructing a set of CT images corresponding to different respiratory phases in a breathing cycle. Such an imaging modality is particularly of use in many clinical applications regarding thorax or upper abdomen area, where a considerable amount of blurring artifacts would appear, if conventional CT is used instead.

In 4D-CT acquisition, x-ray projection data are usually extensively over-sampled. Those projections are then grouped according to their associated respiratory phase information and 4D-CT are accomplished by reconstructing

tomography images corresponding to different phases independently [1] as in a conventional CT reconstruction problem. Since the current commercial standard Filtered Back Projection (FBP) algorithm [2] used to reconstruct each individual phase usually requires hundreds of projections to achieve decent quality, the 4D-CT reconstruction scheme demands an extremely large number of x-ray projections. The consequent long acquisition process therefore leads to an excess amount of radiation dose to the patient, potentially elevating cancer risks.

One obvious way of reducing the imaging dose in 4D-CT is to reconstruct the CT images of each breathing phase from fewer x-ray projections. However, the images reconstructed by conventional FBP-type algorithms would be severely degraded due to insufficient sampling. Recently, a bloom of vast and exciting research in compressed sensing [3,4] has demonstrated the feasibility of recovering signals from incomplete measurements through optimization methods, providing us new perspectives of solving the CT reconstruction problem. Though this approach allows us to retrieve CT images in each phase with only a few number of projections, the total number of projections used for an entire 4D-CT reconstruction is still large due to many breathing phases to be considered. Therefore, it is highly desirable to develop new techniques to reconstruct 4D-CT with a greatly reduced number of projections, while image quality can still be well maintained.

One idea deeply buried in all approaches currently applied to the 4D-CT reconstruction is that images at different respiratory phases are reconstructed individually. Nevertheless, 4D-CT images in a breathing cycle are never independent of each other due to the smooth breathing pattern. Taking this temporal correlation into account can in principle facilitate the 4D-CT reconstruction process and potentially achieve the goal of even lowering projection number. In fact, it is reasonable to believe that there are usually common anatomical features within successive CT images, though the precise locations of those features may slightly vary. Inspired by this fact, we propose in this work a new 4D-CT reconstruction approach by imposing regularization among neighboring phases via a *Temporal Non-local Means* (TNLM) method. Specifically, each feature in a CT image is searched in nearby area in images of neighboring breathing phases and similar features are grouped together to constructively enhance each other. Such an approach is found to be capable of solving the few-view 4D-CT reconstruction problem, as will be seen in the rest of this paper.

2 Our Method

2.1 Conventional CT Reconstruction

Before presenting our method, we first describe how a conventional CT image is reconstructed from highly under-sampled projections. In fact, the conventional CT reconstruction problem, challenging by itself, is a subproblem of the 4D-CT reconstruction, if each phase is reconstructed independently. Let us denote one horizontal slice of patient anatomical information by a vector f . An x-ray projection matrix P maps f into another vector Y on x-ray detectors in a fan-beam geometry, such that $Pf = Y$. A CT reconstruction problem is formulated

as the retrieval of the vector f based on the observation Y given the projection matrix P .

The so-called few-projection CT reconstruction problem is well known to be highly under-determined in that there are infinitely many solution vectors f satisfying the equation $Pf = Y$. In order to single out an ideal CT image f , additional information needs to be imposed properly. For this purpose, regularization models are usually used to reconstruct a desirable CT image in this highly under-sampling situation. As such, one considers the optimization problem

$$f = \operatorname{argmin}_f \|Pf - Y\|_2^2 + \frac{\mu}{2} J[f], \tag{2.1}$$

where the first term ensures the consistency between the reconstructed CT image f and the observation Y . The second term $J[f]$, known as a regularization term, is imposed a priori to guarantee that the reconstructed image from (2.1) satisfies some desirable properties, *e.g.* sharp in edges and smooth in homogeneous regions. Examples of $J[f]$ include *Total Variation* (TV) [5] and *Tight Frames* (TF) [6], to name a few. The parameter μ in (2.1) balances the fidelity of the solution f to the data Y and the imposed regularization.

It is straightforward to utilize this reconstruction approach to solve the 4D-CT reconstruction problem by applying it to each phase independently. The results produced by the TV-based or TF-based methods will be used to benchmark our 4D-CT reconstruction algorithm in our experiments.

2.2 4D-CT Reconstruction

Let us divide a respiratory cycle into N phases labeled by $\alpha = 1, \dots, N$. Denote the CT image at phase α by a vector f_α . A projection matrix P_α at phase α maps the image into an observation vector Y_α , *i.e.* $P_\alpha f_\alpha = Y_\alpha$. As opposed to reconstructing images at each phase independently, we propose the following 4D-CT reconstruction model

$$\{f_\alpha\} = \operatorname{argmin}_{\{f_\alpha\}} \sum_{\alpha=1}^N \|P_\alpha f_\alpha - Y_\alpha\|_2^2 + \frac{\mu}{2} [J(f_\alpha, f_{\alpha-1}) + J(f_\alpha, f_{\alpha+1})], \tag{2.2}$$

where $J(\cdot, \cdot)$ is a TNLM functional imposed as a temporal regularization term between successive phases. Specifically, for two images f_α and f_β , $J(f_\alpha, f_\beta)$ is defined as:

$$J(f_\alpha, f_\beta) = \iint [f_\alpha(x) - f_\beta(y)]^2 w_{f_\alpha^*, f_\beta^*}(x, y) dx dy. \tag{2.3}$$

The weighting factors $w_{f_\alpha^*, f_\beta^*}(x, y)$ are ideally independent of f_α, f_β but defined according to the ground truth images f_α^* and f_β^* as

$$w_{f_\alpha^*, f_\beta^*}(x, y) = \frac{1}{Z} \exp^{-\|p_{f_\alpha^*}(x) - p_{f_\beta^*}(y)\|_2^2 / h^2}, \tag{2.4}$$

where $p_{f_\alpha^*}(x)$ denotes a small patch in image f_α^* centering at the coordinate x and Z is a normalization factor. The scale parameter h controls to what extent similarity between patches is enforced.

The above definition of TNLM functional $J(\cdot, \cdot)$ resembles, and yet is fundamentally different from, the original *Non-local Means* (NLM) idea that have been widely used for restoring natural images [7,8]. A key assumption in their approach is that a natural image, such as a photograph of buildings, usually contains some repetitive features. Therefore, the similarity between features at different locations can be utilized to constructively enhance each other for image restoration. In contrast, similar structures are hardly found in a single medical image. The efficacy of a direct application of the NLM to restore a single medical image, for example in CT reconstruction context, is thus limited. Nonetheless, the crucial similarity assumption naturally exists along a temporal direction, namely among images of adjacent phases. Therefore, in (2.2), we consider a TNLM regularization $J(f_\alpha, f_{\alpha\pm 1})$ as opposed to $J(f_\alpha, f_\alpha)$ as in the NLM approach. Another reason why the TNLM is superior to the NLM method in this 4D-CT reconstruction problem is the capability of removing streaking artifacts (straight lines along CT projection directions), which usually present in highly under-sampled CT reconstruction problem (see the conventional FBP reconstruction results in Fig. 1). If NLM were used in the reconstruction, the streaking artifacts would be in fact strengthened rather than removed, since NLM method tends to locate those straight lines in a single image and enhance them. On the other hand, since the projection directions in two breathing phases are different, the streaking artifacts do not repeat in different phases. TNLM method is therefore able to remove them in all phases simultaneously.

To solve (2.2) efficiently, we adopt a forward-backward splitting algorithm [9,10], where the solution to (2.2) can be obtained by alternatively performing the following two steps

$$v_\alpha^k = f_\alpha^{k-1} - \frac{1}{\lambda} P_\alpha^T (P_\alpha f_\alpha^{k-1} - Y_\alpha), \quad \forall \alpha, \tag{2.5}$$

$$\{f_\alpha^k\} = \operatorname{argmin}_{\{f_\alpha\}} \sum_\alpha \lambda \|f_\alpha - v_\alpha^k\|_2^2 + \frac{\mu}{2} [J(f_\alpha, f_{\alpha-1}) + J(f_\alpha, f_{\alpha+1})]. \tag{2.6}$$

Here v_α are auxiliary vectors and λ is a constant introduced by the splitting algorithm. Note that (2.5) is simply one step of gradient descent algorithm towards minimizing an energy functional $E[f_\alpha] = \sum_\alpha \|P_\alpha f_\alpha - Y_\alpha\|_2^2$ with a step size $1/2\lambda$. In order to speed up the convergence, we substitute this step by a conjugate gradient minimization for the energy functional $E[f_\alpha]$ with an initial solution f_α^{k-1} , denoted as $v_\alpha^k = \operatorname{CG}_{E[f_\alpha]}(f_\alpha^{k-1})$. As for (2.6), we perform one step of gradient descent for each exterior iteration as

$$f_\alpha^k(x) = v_\alpha^k(x)(1 - 2\mu) + \mu \left[\int v_{\alpha+1}^k(y) w_{f_\alpha^*, f_{\alpha+1}^*}(x, y) dy + \int v_{\alpha-1}^k(y) w_{f_\alpha^*, f_{\alpha-1}^*}(x, y) dy \right]. \tag{2.7}$$

Note that, in the above equation, the images at phases $(\alpha \pm 1)$ are naturally imposed to enhance the image quality of f_α . Another issue worth mentioning here is the weighting factors $w_{f_\alpha^*, f_\beta^*}(x, y)$. In (2.4) they are defined with respect

to the ground truth images f_{α}^* , which are not available during the reconstruction. In practice, for each iteration k , the weights $w_{f_{\alpha}^*, f_{\beta}^*}(x, y)$ are replaced by $w_{f_{\alpha}^{k-1}, f_{\beta}^{k-1}}(x, y)$. In summary, the algorithm solving (2.2) is to perform the following two steps alternatively:

$$\begin{aligned} v_{\alpha}^k &= \text{CG}_{E[f_{\alpha}]}(f_{\alpha}^{k-1}), \quad \forall \alpha, \\ f_{\alpha}^k(x) &= v_{\alpha}^k(x)(1 - 2\mu) + \mu \left[\int v_{\alpha+1}^k(y) w_{f_{\alpha}^{k-1}, f_{\alpha+1}^{k-1}}(x, y) dy \right. \\ &\quad \left. + \int v_{\alpha-1}^k(y) w_{f_{\alpha}^{k-1}, f_{\alpha-1}^{k-1}}(x, y) dy \right]. \end{aligned} \quad (2.8)$$

3 Experiments

We first test our 4D-CT reconstruction algorithm on a digital NCAT thorax phantom [11]. We consider a simple breathing model with only two phases in a respiratory cycle. The image resolution is 256×256 . Due to its simple anatomical structure in this digital phantom, only 20 x-ray projections per respiratory phase are used. These projections are generated in a fan beam geometry and are equally spaced in an entire 360° rotation. In Fig. 1, we show the ground truth images and the reconstruction results from the conventional FBP algorithms in the first and the second columns, respectively. Clearly, FBP algorithm produces severe streaking artifacts in this context of extremely under-sampling, making these images clinically unacceptable. The reconstruction results shown in the columns 3 and 4 in Fig. 1 correspond to TV and TF methods. Despite a great improvement over the FBP algorithm, the images obtained from these two methods are smeared out with reduced contrast and still contain obvious

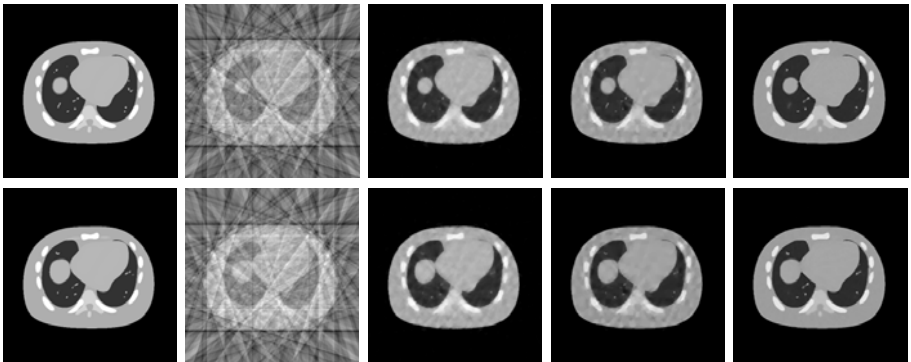


Fig. 1. A Digital NCAT phantom case. Images at two breathing phases are shown in two rows. Columns from left to right: ground truth images, reconstruction results from FBP, TV, TF, and our TNLM method, respectively.

fluctuating artifacts. Finally, the images reconstructed from our TNLM regularization method are shown in the last column, where the image quality is considerably enhanced by imposing the temporal regularization between these two phases. Notice that for these iterative methods (TV, TF and TNLM), the results are presented with parameters, such as μ in Eq. (2.1), carefully tuned for the optimal results.

To further validate our 4D-CT reconstruction algorithm, we study two patient cases obtained in the real clinic. There are 10 breathing phases in both cases. Since the real clinical 4D-CT images are full of detailed structures, 30 fan beam x-ray projections per breathing phase are used. Other parameters are same as those in the NCAT phantom case. We exclude presenting the FBP results since it is apparently not able to provide clinically acceptable 4D-CT images. Due to the space limitation, only phase 1, 4, and 7 out of the 10 phases in an entire breathing cycle are illustrated in Fig. 2 and 3 respectively for the two real clinical cases. Again, our method is able to reconstruct the 4D-CT images with great quality, while low contrast and streaking artifacts are found to some extent in those images obtained from TV or TF methods.

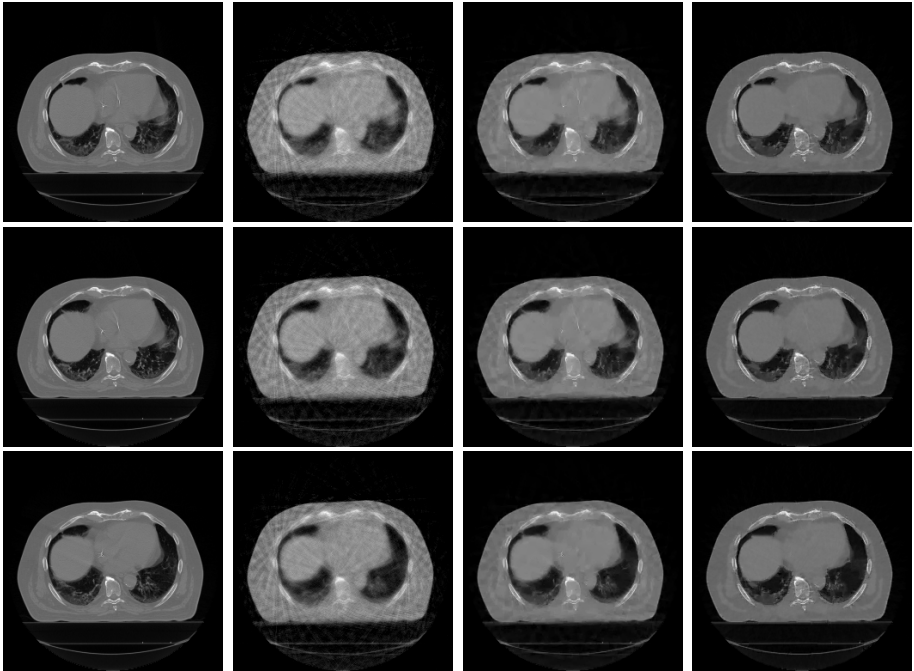


Fig. 2. Patient case A. Rows from top to bottom: phase 1, 4, and 7 in a 10-phase respiratory cycle. Columns from left to right: ground truth images, reconstruction results from TV, TF, and our TNLM method.

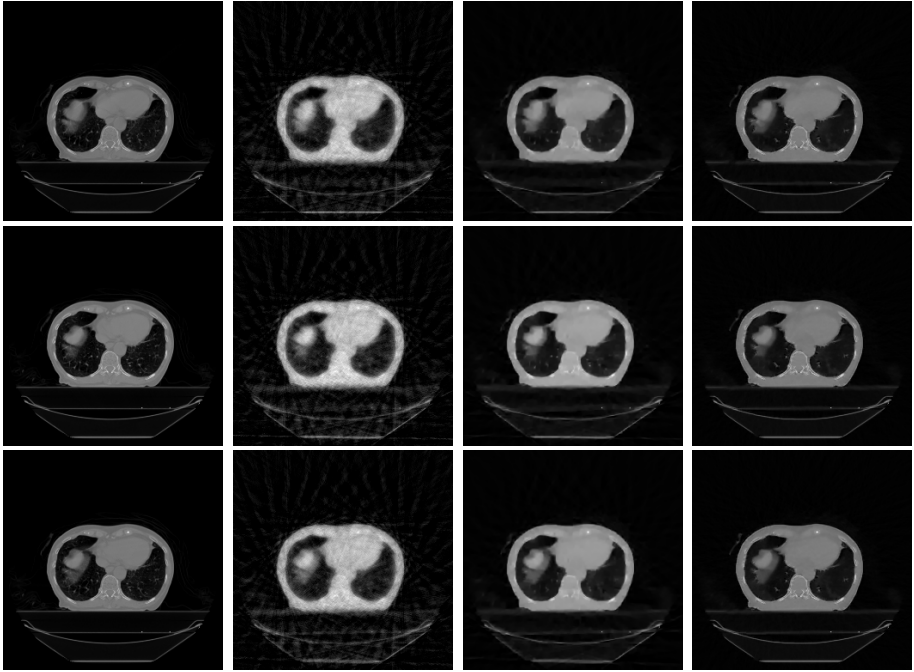


Fig. 3. Patient case B. Rows from top to bottom: phase 1, 4, and 7 in a 10-phase respiratory cycle. Columns from left to right: ground truth images, reconstruction results from TV, TF, and our TNLM method.

In order to quantitatively evaluate the reconstruction results, we take signal-to-noise ratio (SNR) as a metric defined as

$$\text{SNR}(f^*, f) = 20 \log_{10} \left\{ \frac{\|f - \bar{f}\|_{L^2}}{\|f - f^*\|_{L^2}} \right\}, \quad (3.1)$$

where f^* is the ground truth image and \bar{f} denote the mean value of the image f . The SNRs for images obtained from different methods are summarized in Tab. [1](#). Our method yield the highest SNR values in all cases, undoubtedly outperforms the other two methods.

Table 1. SNR for different methods

	NCAT		patient A			patient B		
	Phase 1	Phase 2	Phase 1	Phase 4	Phase 7	Phase 1	Phase 4	Phase 7
TV	20.35	20.14	15.04	14.97	15.47	11.66	11.39	11.40
TF	21.06	20.72	18.71	18.66	19.46	15.85	15.44	15.14
our method	25.68	25.59	20.63	20.29	20.44	18.99	19.22	19.13

4 Conclusion

In this paper, we have presented a new 4D-CT reconstruction algorithm that exploits the recurrence of the anatomical structures at different locations between adjacent breathing phases. TNLM regularization is specifically designed to take this fact into the reconstruction process. We have tested our reconstruction results on one digital NCAT phantom and two patient cases. The reconstruction results indicate that our TNLM algorithm outperforms the conventional FBP-type reconstruction algorithm and the TV-based or TF-based spatial regularization methods. One concern about our TNLM method is the speed. Currently it takes much longer time for the TNLM method to reconstruct a set of 4D-CT images due to its inherent complicated mathematical structure. However, our preliminary work on the speed-up via advanced GPU technology indicates promising perspectives.

References

1. Low, D.A., Nystrom, M., Kalinin, E., et al.: A method for the reconstruction of four-dimensional synchronized ct scans acquired during free breathing. *Med. Phys.* 30(1254), 1254–1263 (2003)
2. Deans, S.: *The Radon Transform and Some of Its Applications*. Wiley, New York (1983)
3. Candes, E.J., Romberg, J., Tao, T.: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory* 52(2), 489–509 (2006)
4. Donoho, D.L.: Compressed sensing. *IEEE Transactions on Information Theory* 52, 1289–1306 (2006)
5. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Physics D* 60, 259–268 (1992)
6. Daubechies, I., Han, B., Ron, A., Shen, Z.: *Applied and Computational Harmonic Analysis*, pp. 1–46 (2003)
7. Buades, A., Coll, B., Morel, J.M.: A review of image denoising algorithms, with a new one. *Multiscale Modeling and Simulation* 4(2), 490–530 (2005)
8. Lou, Y., Zhang, X., Osher, S., Bertozzi, A.: Image recovery via nonlocal operators. *Journal of Scientific Computing* 42(2), 185–197 (2010)
9. Combettes, P., Wajs, V.: Signal recovery by proximal forward-backward splitting. *Multiscale Modeling and Simulation* 4(4), 1168–1200 (2006)
10. Hale, E., Yin, W., Zhang, Y.: A fixed-point continuation method for ℓ_1 -regularization with application to compressed sensing. CAAM Technical Report TR, Rice University, Houston, TX, p. 07 (2007)
11. Segars, W.P.: Development and application of the new dynamic NURBS-based cardiac-torso (NCAT) phantom. PhD thesis, University of North Carolina. Chapel Hill (2001)

ECG-Gated Interventional Cardiac Reconstruction for Non-periodic Motion

Christopher Rohkohl^{1,2}, Günter Lauritsch²,
Lisa Biller², and Joachim Hornegger¹

¹ Pattern Recognition Lab, Department of Computer Science,
Friedrich-Alexander University Erlangen-Nuremberg,
91058 Erlangen, Germany

² Siemens AG, Healthcare Sector, 91301 Forchheim, Germany

Abstract. The 3-D reconstruction of cardiac vasculature using C-arm CT is an active and challenging field of research. In interventional environments patients often do not have arrhythmic heart signals or cannot hold breath during the complete data acquisition. This important group of patients cannot be reconstructed with current approaches that do strongly depend on a high degree of cardiac motion periodicity for working properly. In a last year's MICCAI contribution a first algorithm was presented that is able to estimate non-periodic 4-D motion patterns. However, to some degree that algorithm still depends on periodicity, as it requires a prior image which is obtained using a simple ECG-gated reconstruction. In this work we aim to provide a solution to this problem by developing a motion compensated ECG-gating algorithm. It is built upon a 4-D time-continuous affine motion model which is capable of compactly describing highly non-periodic motion patterns. A stochastic optimization scheme is derived which minimizes the error between the measured projection data and the forward projection of the motion compensated reconstruction. For evaluation, the algorithm is applied to 5 datasets of the left coronary arteries of patients that have ignored the breath hold command and/or had arrhythmic heart signals during the data acquisition. By applying the developed algorithm the average visibility of the vessel segments could be increased by 27%. The results show that the proposed algorithm provides excellent reconstruction quality in cases where classical approaches fail. The algorithm is highly parallelizable and a clinically feasible runtime of under 4 minutes is achieved using modern graphics card hardware.

1 Introduction

1.1 Purpose of This Work

One key component of image guidance in the field of interventional cardiology is three-dimensional image information before, during and after interventional procedures. Three-dimensional image data can support complex interventional procedures, such as transcatheter valve replacement, interventional therapy of

atrial fibrillation, implantation of biventricular pacemakers and the assessment of myocardial perfusion.

With the technology of C-arm CT it is possible to reconstruct intraprocedural 3-D images from angiographic projection data [1]. Currently, the major limitation of this technology is its insufficient temporal resolution which limits the visualization of fast moving parts of the heart. Due to the long acquisition time of several seconds, at which a couple of heart beats and breathing motion can occur, motion related image artifacts, e.g. blurring or streaks are observed. Therefore it is essential to develop algorithms that can cope with both cardiac and respiratory motion.

1.2 State-of-the-Art

Based on the electrocardiogram (ECG) a relative cardiac phase is assigned to each projection image assuming a periodic heart motion [2]. The phase information is used for a phase-correlated reconstruction by gating or motion estimation and compensation. A gated reconstruction takes only those images into account that lie inside a defined phase window, that is centered at the targeting cardiac phase to be reconstructed [2, 3]. This is, however, not ideal in terms of missing data for single-run acquisitions. The incomplete data leads to streak artifacts and a poor signal-to-noise ratio. To increase the data usage, the cardiac motion is estimated and motion compensated reconstruction algorithms are applied [4, 2, 5, 6, 7]. Reasonable results are provided in the presence of regular heart rates without breathing or other patient motion. However, in the field of intraprocedural cardiac reconstruction, the patients suffer from heart diseases and cannot completely hold breath, stay still or have irregular heart beats. Those aspects do conflict with the periodicity assumption of ECG-based methods.

In literature these problems were addressed by approximate 2-D corrections in the projection image [4, 8]. In a last year's MICCAI contribution [9], a first algorithm was presented that is able to estimate true 4-D non-periodic motion patterns using a time-continuous B-spline motion model (NOPMEC). However, this algorithm requires an initial 3-D reference image as a priori information. This reference is easy to obtain for almost periodic motion by ECG-gated reconstruction. However, strong respiratory motion and severe arrhythmia can make it impossible to obtain a sufficient initial image. This weak point is to be solved in this paper.

2 Affine 4-D Motion Estimation and Reconstruction

2.1 General Idea

In previous investigations, e.g. by Shechter et al. [10] it has been shown that heart phase variations and respiratory motion of cardiac vasculature can be modelled using global transformations, i.e. rigid body transformations or affine transformations. In the following sections an algorithm for the estimation of such an affine motion between the different heart beats is presented.

2.2 Affine 4-D Time-Continuous Motion Model

We assume a time-continuous motion model function that maps a voxel $\mathbf{x} = (x_0, x_1, x_2)^T$ to a new voxel location \mathbf{x}' for each time when a projection image is acquired. It is conceptualized by a function $M : \mathbb{N} \times \mathbb{R}^3 \times \mathbb{S} \mapsto \mathbb{R}^3$ with $M(i, \mathbf{x}, \mathbf{s}) = \mathbf{x}'$ transforming the voxel coordinate \mathbf{x} at the time of the i -th projection image. The mapping is based on the motion model parameters $\mathbf{s} \in \mathbb{S}$. In this work a global affine motion model is used which is defined in the following.

A set of time points is created and each point in time is assigned 12 affine parameters that describe the affine transformation at that moment. The set of time points is determined from the ECG-signal. We are choosing two time points per heart beat closest to the reference heart phase h_r and to an additional phase $h_r + \Delta h$. For a heart phase h this set of projection images is denoted \mathcal{N}_h . The complete set L of temporal control points is then given by

$$L = \{1, N\} \cup \mathcal{N}_{h_r} \cup \mathcal{N}_{h_r + \Delta h}, \quad (1)$$

where the number of projections ranges from 1 to N . The first and last projection image is added such that no boundary problems will occur in the following.

The affine parameters $\mathbf{s}_l \in \mathbb{R}^{12}$ for a single time point $l \in L$ are collected in a vector $\mathbf{s}_l = (t_0, t_1, t_2, \alpha_0, \alpha_1, \alpha_2, a_0, a_1, a_2, b_0, b_1, b_2)^T$, where t_i represents the translation along, α_i the rotation around, a_i the scaling along and b_i the shearing of the i -th coordinate axis. The complete parameter vector $\mathbf{s} \in \mathbb{S}$, $\mathbb{S} = \mathbb{R}^{12|L|}$ is then given by

$$\mathbf{s} = (\mathbf{s}_{L_1}, \dots, \mathbf{s}_{L_{|L|}})^T \quad (2)$$

with L_i being the i -th smallest element of L . For an arbitrary projection image i the affine transform parameters $\tilde{\mathbf{s}}_i$ are then obtained by temporal interpolation of each component. In our work a cubic B-spline interpolation [11] has been used. The final motion model is then formally given by

$$M(i, \mathbf{x}, \mathbf{s}) = \mathbf{x}' \quad \text{with} \quad \mathbf{A}_{\tilde{\mathbf{s}}_i} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} \cong \begin{pmatrix} \mathbf{x}' \\ 1 \end{pmatrix} \quad (3)$$

where $\mathbf{A}_{\tilde{\mathbf{s}}_i}$ is the affine transformation matrix in homogenous coordinates for the affine parameters $\tilde{\mathbf{s}}_i$. The formation of the affine transformation matrix is straightforward and can be found e.g. in [12].

2.3 Motion Compensated ECG-Gated Reconstruction

Projection Image Preprocessing. For motion estimation we are only interested in the motion of the cardiac vasculature. Therefore we apply a background reduction technique previously proposed by Hansis et al. [3] using a morphological top-hat filter. In the following this preprocessed projection data will be referred to by the function $p : \mathbb{N} \times \mathbb{R}^2 \mapsto \mathbb{R}$ where $p(i, \mathbf{u})$ returns the value of the i -th preprocessed projection image at the pixel \mathbf{u} .

Reconstruction Algorithm. For motion estimation and correction a dynamic reconstruction $f(\mathbf{x}, \mathbf{s})$ is defined. The function f returns the reconstructed object value at a voxel \mathbf{x} based on the motion model parameters \mathbf{s} . In this paper, an ECG-gated motion compensated FDK reconstruction method is utilized [2, 5]. The ECG-gating is performed by applying a weighting factor λ to each image which is calculated from the distance to the reference heart phase. The dynamic ECG-gated FDK reconstruction $f_{h_r} : \mathbb{R}^3 \times \mathbb{S} \mapsto \mathbb{R}$ is then given by

$$f_{h_r}(\mathbf{x}, \mathbf{s}) = \sum_i \lambda(i, h_r) \cdot w(i, M(i, \mathbf{x}, \mathbf{s})) \cdot p_F(i, A(i, M(i, \mathbf{x}, \mathbf{s}))). \quad (4)$$

The function $w : \mathbb{N} \times \mathbb{R}^3 \mapsto \mathbb{R}$ is the distance weight of the FDK-formula. The pre-processed, filtered and redundancy weighted projection data is accessed by the function $p_F : \mathbb{N} \times \mathbb{R}^2 \mapsto \mathbb{R}$ where $p_F(i, \mathbf{u})$ returns the value of the i -th image at the pixel \mathbf{u} . The pixel location \mathbf{u} is determined by the perspective projection $A : \mathbb{N} \times \mathbb{R}^3 \mapsto \mathbb{R}^2$, where $A(i, \mathbf{x}) = \mathbf{u}$ maps a voxel \mathbf{x} to a pixel location \mathbf{u} in the i -th projection image. The function λ is a weighting function adapted from [5] that is used for obtaining an ECG-phase correlated reconstruction for the heart phase $h_r \in [0, 1]$. It is given by

$$\lambda(i, h_r) = \begin{cases} \cos^\beta \left(\frac{d(h(i), h_r)}{\omega} \pi \right) & \text{if } d(h(i), h_r) \leq \frac{\omega}{2} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where $h(i)$ is the heart phase of the i -th projection image and $\omega \in (0, 1]$ is the width of the non-zero support region of the weighting function. The parameter $\beta \in [0, \infty)$ controls the shape of the support region. The distance function $d(h_1, h_2) = \min_{c \in \{0, 1, -1\}} |h_1 - h_2 + c|$ returns the distance between two relative heart phases.

2.4 Objective Function for Motion Estimation

Motion estimation is formulated as a multi-dimensional optimization problem where the motion model parameters $\hat{\mathbf{s}} \in \mathbb{S}$ maximizing the objective function $\mathcal{L} : \mathbb{S} \mapsto \mathbb{R}$ need to be estimated. The objective function introduced in this paper is motivated by the basic relationship of the motion compensated reconstruction f with the measured projection data p . Maximum intensity projections of the motion state observed in a projection image can be created from a motion compensated reconstruction $f(\mathbf{x}, \mathbf{s})$ by dynamic forward projection:

$$r(i, \mathbf{u}, \mathbf{s}) = \max_{\mathbf{x} \in L_{i, \mathbf{u}}} f_{h_r}(M^{-1}(i, \mathbf{x}, \mathbf{s}), \mathbf{s}). \quad (6)$$

The function $r : \mathbb{N} \times \mathbb{R}^2 \times \mathbb{S} \rightarrow \mathbb{R}$ returns the dynamic maximum intensity forward projection of the ECG-gated and motion compensated reconstruction f_{h_r} . The voxels on the straight measurement ray $L_{i, \mathbf{u}}$ of the i -th image hitting the detector at pixel \mathbf{u} are transformed by the inverse motion model to consider the motion state observed at the projection image i . The matching of the measured

and preprocessed data p and the forward projected data r is assessed by computing the average normalized cross-correlation (NCC). Formally, the objective function to be maximized is then given by:

$$\mathcal{L}(\mathbf{s}) = \arg \max_{\mathbf{s} \in \mathbb{S}} \frac{1}{\nu} \sum_i^N \left(\lambda(i, h_r) \sum_{\mathbf{u}} \frac{(p(i, \mathbf{u}) - \mu_{p_i})(r(i, \mathbf{u}, \mathbf{s}) - \mu_{r_{i,s}})}{\sigma_{p_i} \sigma_{r_{i,s}}} \right) \quad (7)$$

with the normalizing factor $\nu = (I_p - 1) \sum_i^N \lambda(i, h_r)$, I_p being the number of image pixels \mathbf{u} of a projection and μ, σ being the mean and standard deviation of the subscripted image. The value of the objective function ranges $\mathcal{L}(\mathbf{s}) \in [-1, 1]$ with the maximum value representing a perfect linear relationship of the measured and forward projected data. The NCC of the i -th projection image pair is weighted by the gating function λ as it characterizes the influence on the dynamic ECG-gated reconstruction.

2.5 Optimization Strategy

For maximization of Eq. (7) a stochastic gradient ascent method is used. In each iteration one of the temporal points is selected with a certain probability. For initialization, all time points are assigned the same probability. The gradient is computed using finite differences by varying the 12 affine parameters for the selected time point. Next, one step is taken into the gradient direction with a fixed step size. The probability of the time point of being selected in the next iteration, is set proportional to the increase of cost function value. Optimization stops after a certain number of iterations or if the convergence ratio drops below a certain threshold. This kind of stochastic selection of parameters for the gradient-based optimization procedure has the benefit that it picks out the most beneficial parameters. In that way a fast convergence with as few as possible cost function evaluations was found.

2.6 Implementation Details

One evaluation of the objective function Eq. (7) comprises an ECG-gated reconstruction, an ECG-gated forward projection and the computation of the quality measure. Each step is very well parallelizable on the graphics card using CUDA [13]. The backprojection of the FDK-reconstruction and forward projection are based on projection matrices. The affine matrix $\mathbf{A}_{\tilde{s}_i}$ in Eq. (3) only depends on the projection geometry and is independent of the voxel location. This allows us to replace the voxel-wise computation of the motion transform M by a right-side multiplication of the projection matrix with the affine transformation matrix. Consequently, no additional overhead is introduced during forward or backward projection operations.

3 Experimental Setup

The presented algorithm is integrated into the non-periodic motion estimation and compensation scheme of [9] (NOPMEC). It requires as input a prior

image which is in the standard version a simple ECG-gated reconstruction. For our experiments the standard version is compared with a version that uses the proposed ECG-gating for non-periodic motion presented in this paper.

Datasets of left coronary arteries divided into two groups are reconstructed. The first group G_1 consists of five cases with good breath hold and a regular heart rate. The second group G_2 consists of five cases that have ignored the breath hold command and/or had arrhythmic heart signals during the data acquisition. For both groups the coronary segments S_5 to S_{15} , named according to the guidelines of the American Heart Association [14] are identified and the visibility is rated with a value from zero to three with the following meaning: 0 (not visible), 1 (partially visible, strong artifacts), 2 (completely visible, some artifacts), 3 (close to perfect).

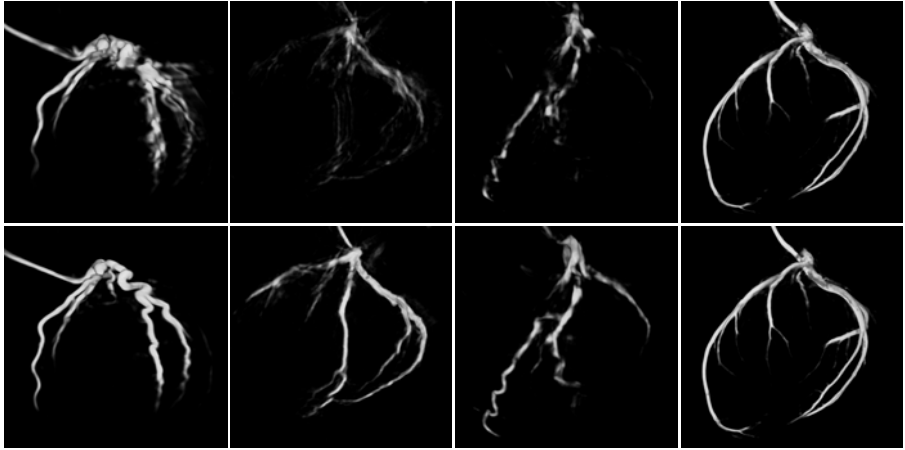
The datasets were acquired on an Artis Zeego C-arm system (Siemens AG, Healthcare Sector, Forchheim, Germany). They consist each of $N = 133$ projection images acquired in 5 seconds with a size of 960×1240 pixels at an isotropic resolution of 0.32 mm/pixel. The number of iterations was set to 300. The reconstruction parameters were set to $h_r = 0.7$, $h_r + \Delta h = 0.2$, $\omega = 0.2$, $\beta = 0$. Image reconstruction is performed on an image volume of 20^3 cm^3 distributed on a 256^3 voxel grid. The parameters of the NOPMEC algorithm [9] were set to 5 spatial and 35 temporal control points with 100 iterations.

4 Results and Discussion

The average rating for each of the segments S_5 to S_{15} is listed separately for each patient group and for both algorithms in Table 1. The performance of the proposed method does not significantly increase the visibility for group G_1 (1.3%). This is reasonable, as this group contains highly periodic data for which the simple ECG-gating already works well. The 3-D reconstruction of an example case from group G_1 is depicted in Fig. 1d.

Table 1. Results from the evaluation of the clinical data. For each segment the visibility for both patient groups is analyzed. Results are given for the standard algorithm and the proposed method.

	Group G_1			Group G_2		
	standard	proposed	increase	standard	proposed	increase
S_5	3.00	3.00	0.0%	3.00	3.00	0.0%
S_6	2.80	2.80	0.0%	2.60	2.80	7.7%
S_7	2.80	2.80	0.0%	2.00	2.20	10.0%
S_8	2.80	2.80	0.0%	2.40	2.80	16.7%
S_9	2.60	2.80	7.7%	2.00	2.60	30.0%
S_{10}	2.40	2.40	0.0%	1.00	1.75	75.0%
S_{11}	3.00	3.00	0.0%	2.20	2.60	18.2%
S_{12}	2.80	2.80	0.0%	0.75	1.50	100.0%
S_{13}	2.60	2.60	0.0%	1.80	2.60	44.4%
S_{14}	2.20	2.20	0.0%	1.00	1.50	50.0%
S_{15}	2.80	3.00	7.1%	1.75	2.75	57.1%
Avg.	2.71	2.75	1.3%	1.86	2.37	27.3%



(a) Ex. case of G_2 (b) Ex. case of G_2 (c) Ex. case of G_2 (d) Ex. case of G_1

Fig. 1. Example reconstruction results for the clinical datasets using the standard algorithm (top row) and the proposed approach (bottom row). The first three datasets belong to the patient group G_2 while the last one belongs to G_1 .

For patient group G_2 the situation changes. The average visibility can be increased significantly (27.3%) using the method proposed in this paper. Three of the cases are depicted in Fig. 1a to Fig. 1c to provide a visual impression of the quality increase. By applying the proposed method, the quality can be increased drastically (bottom row) in terms of intensity, connectedness, artifact level and smoothness of the vessels. In summary the proposed method never decreased the visibility of any of the vessel segments and provided superior image quality in cases of non-periodic motion (group G_2). The runtime of our algorithm was below 4 minutes in all cases.

5 Conclusions and Outlook

A major clinical challenge of C-arm based cardiac vasculature reconstruction is non-periodic motion. In this paper a framework for the ECG-gated 3-D reconstruction with affine motion correction for non-periodic motions has been introduced. It has been integrated into the NOPMEC algorithm for obtaining a prior image. It is demonstrated that for non-periodic motion, using simple ECG-gating as prior image fails, while the visibility of the vasculature could be increased in average by 27%, using the motion compensated ECG-gating presented in this paper. In summary, with the proposed method an important step towards the clinical usability has been made by addressing patient groups which cannot be reconstructed with previous methods.

Disclaimer: The concepts and information presented in this paper are based on research and are not commercially available.

References

1. Hetterich, H., Redel, T., Lauritsch, G., Rohkohl, C., Rieber, J.: New X-ray imaging modalities and their integration with intravascular imaging and interventions. In: *The International Journal of Cardiovascular Imaging* (2009) (online) (to be printed)
2. Schäfer, D., Borgert, J., Rasche, V., Grass, M.: Motion-compensated and gated cone beam filtered back-projection for 3-d rotational x-ray angiography. *IEEE Transactions on Medical Imaging* 25(7), 898–906 (2006)
3. Hansis, E., Schafer, D., Dossel, O., Grass, M.: Evaluation of iterative sparse object reconstruction from few projections for 3-d rotational coronary angiography. *IEEE Transactions on Medical Imaging* 27(11), 1548–1555 (2008)
4. Blondel, C., Malandain, G., Vaillant, R., Ayache, N.: Reconstruction of coronary arteries from a single rotational x-ray projection sequence. *IEEE Transactions on Medical Imaging* 25(5), 653–663 (2006)
5. Rohkohl, C., Lauritsch, G., Nöttling, A., Prümmer, M., Hornegger, J.: C-arm ct: Reconstruction of dynamic high contrast objects applied to the coronary sinus. In: *IEEE NSS-MIC*, Dresden, Germany, M10–M328 (October 2008)
6. Prümmer, M., Hornegger, J., Lauritsch, G., Wigström, L., Girard-Hughes, E., Fahrig, R.: Cardiac c-arm ct: A unified framework for motion estimation and dynamic ct. *IEEE Trans. Med. Imaging* 28(11), 1836–1849 (2009)
7. Keil, A., Vogel, J., Lauritsch, G., Navab, N.: Dynamic cone beam reconstruction using a new level set formulation. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009*. LNCS, vol. 5762, pp. 389–397. Springer, Heidelberg (2009)
8. Hansis, E., Schäfer, D., Dössel, O., Grass, M.: Projection-based motion compensation for gated coronary artery reconstruction from rotational x-ray angiograms. *Physics in Medicine and Biology* 53(14), 3807–3820 (2008)
9. Rohkohl, C., Lauritsch, G., Prümmer, M., Hornegger, J.: Interventional 4-d motion estimation and reconstruction of cardiac vasculature without motion periodicity assumption. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009*. LNCS, vol. 5761, pp. 132–139. Springer, Heidelberg (2009)
10. Shechter, G., Ozturk, C., Resar, J.R., McVeigh, E.R.: Respiratory motion of the heart from free breathing coronary angiograms. *IEEE Transactions on Medical Imaging* 23(8), 1046–1056 (2004)
11. Unser, M.: Splines: A perfect fit for signal and image processing. *IEEE Signal Processing Magazine* 16(6), 22–38 (1999); *IEEE Signal Processing Society's 2000 magazine award*
12. Hartley, R.I., Zisserman, A.: *Multiple View Geometry in Computer Vision*, 2nd edn. Cambridge University Press, Cambridge (2004) ISBN: 0521540518
13. Rohkohl, C., Keck, B., Hofmann, H., Hornegger, J.: RabbitCT - an open platform for benchmarking 3D cone-beam reconstruction algorithms. *Medical Physics* 36, 3940–3944 (2009)
14. Austen, W., Edwards, J., Frye, R., Gensini, G., Gott, V., Griffith, L., McGoon, D., Murphy, M., Roe, B.: A reporting system on patients evaluated for coronary artery disease. *Circulation* 51, 5–40 (1975)

Physiological Fusion of Functional and Structural Data for Cardiac Deformation Recovery

Ken C.L. Wong^{1,2}, Linwei Wang¹, Heye Zhang³, and Pengcheng Shi¹

¹Computational Biomedicine Laboratory, Rochester Institute of Technology, Rochester, USA
{linwei.wang, pengcheng.shi}@rit.edu

²ASCLEPIOS Research Project, INRIA Sophia Antipolis, Sophia Antipolis, France
kenclwong@sophia.inria.fr

³Bioengineering Institute, University of Auckland, Auckland, New Zealand
hey.e.zhang@auckland.ac.nz

Abstract. The advancement in meaningful constraining models has resulted in increasingly useful quantitative information recovered from cardiac images. Nevertheless, single-source data used by most of these algorithms have put certain limits on the clinical completeness and relevance of the analysis results, especially for pathological cases where data fusion of multiple complementary sources is essential. As traditional image fusion strategies are typically performed at pixel level by fusing commensurate information of registered images through various mathematical operators, such approaches are not necessarily based on meaningful biological bases, particularly when the data are dissimilar in physical nature and spatiotemporal quantity. In this work, we present a physiological fusion framework for integrating information from different yet complementary sources. Using a cardiac physiome model as the central link, structural and functional data are naturally fused together for a more complete subject-specific information recovery. Experiments were performed on synthetic and real data to show the benefits and potential clinical applicability of our framework.

1 Introduction

Cardiac measurements can be categorized as structural and functional. Structural data, such as magnetic resonance images (MRI), noninvasively provide cardiac structures and motions in various spatiotemporal resolutions. Functional data, such as body surface potential maps (BSPM), reveal physiological information of electrical propagation.

To provide physiologically plausible constraints to recover subject-specific information from measurements, cardiac physiome models have been utilized to describe the general macroscopic electromechanical physiology of the heart, which comprise an electrical propagation model (E model) describing electricity propagation within the myocardium, an electromechanical coupling model (EM model) converting the electrical stimulations into contraction stresses, and a biomechanical model (BM model) relating the contraction stresses with the cardiac deformation. In [1], an E model was used to recover volumetric myocardial electrical activities from BSPM. In [2,3], cardiac physiome models were coupled with MRI to estimate ventricular deformation. These applications show that cardiac physiome models are valuable for revealing subject-specific

cardiac functions. Nevertheless, regardless of the general physiological meaningfulness, these single-source algorithms could not benefit from the complementary information of the other measuring modalities, especially in pathological cases.

To enhance subject-specificity and reliability of the recovery results, measurements from different sources should be fused together to include complementary information. As MRI and BSPM are different manifestations of cardiac functions, fusion of these data aids improving the clinical relevance of the recovered information. In practice, the fusion techniques widely used in medical image analysis are pixel-level fusions, by which the spatial and spectral data of all layers of the registered commensurate images are combined through mathematical operators such as the Boolean or fuzzy logic operators. Nevertheless, the acquisition spaces (myocardium vs. body surface) and physical natures (structure vs. electricity) of MRI and BSPM are incommensurate, the direct use of these operators is physiologically meaningless, if not impossible. In view of the ability to couple subject-specific data from electrical and structural measurements, we propose to adopt a cardiac physiological model as the physiological link for fusing information from these data. Specifically, transmembrane potentials (TMP) recovered from the subject's BSPM through the E model are converted into active stresses through the EM model, which are subsequently related to the myocardial deformation through the cardiac electromechanical dynamics. By coupling this subject-specific *a priori* deformation with the motion information extracted from the MRI sequence of the same subject under a state-space filtering framework, the cardiac deformation is recovered. In this manner, electrical and structural data are fused together and tightly coupled through the cardiac physiome model. Experiments were performed on synthetic and real data to show the benefits and potential clinical applicability.

2 Cardiac Physiome Model

The cardiac physiome model acts as the physiological link between the electrical and structural information. Only the E model and the cardiac electromechanical dynamics are presented, as they directly associate with the data fusion.

2.1 Electric Wave Propagation Model

To model the TMP propagation of the myocardium, the E model described by a two-variable diffusion-reaction system is used, which matrix representation is given as [11]:

$$\frac{\partial \mathbf{U}_E}{\partial t} = -\mathbf{M}_E^{-1} \mathbf{K}_E \mathbf{U}_E + f_1(\mathbf{U}_E, \mathbf{V}); \quad \frac{\partial \mathbf{V}}{\partial t} = f_2(\mathbf{U}_E, \mathbf{V}) \quad (1)$$

where \mathbf{U}_E and \mathbf{V} are vectors of nodal TMP and repolarization variables respectively. Matrices \mathbf{M}_E and \mathbf{K}_E account for the intercellular electrical propagation. f_1 and f_2 are general descriptors of cellular TMP dynamics.

2.2 Cardiac Electromechanical Dynamics

The TMP can be converted into active forces through the EM model, which are further related to the cardiac deformation through the cardiac electromechanical dynamics.



Fig. 1. Synthetic data. (a) Heart represented as nodes bounded by surfaces, with tissue structures indicated by local coordinate systems (fiber, sheet, sheet normal: blue, yellow, cyan). (b) Infarcted regions shown in red color.

The displacement-based total-Lagrangian system dynamics is utilized, which is given as [3]:

$$\mathbf{M}_{\text{BM}} \ddot{\mathbf{U}}_{\text{BM}} + \mathbf{C} \dot{\mathbf{U}}_{\text{BM}} + \mathbf{K}_{\text{BM}} \Delta \mathbf{U}_{\text{BM}} = \mathbf{F}_c - \mathbf{F}_i \quad (2)$$

where \mathbf{M}_{BM} , \mathbf{C} , and \mathbf{K}_{BM} are the mass, damping, and stiffness matrices respectively. The hyperelastic and orthotropic material properties of the BM model are embedded in \mathbf{K}_{BM} . \mathbf{F}_c is the force vector containing the active forces from the E and EM models. \mathbf{F}_i is the force vector related to internal stresses. $\ddot{\mathbf{U}}_{\text{BM}}$, $\dot{\mathbf{U}}_{\text{BM}}$ and $\Delta \mathbf{U}_{\text{BM}}$ are the respective acceleration, velocity and incremental displacement vectors respectively.

3 Physiological Data Fusion

To fuse structural and functional image information, subject-specific volumetric TMP are first recovered from BSPM through the E model, which become the force inputs to the cardiac electromechanical dynamics through the EM model. This BSPM-embedded cardiac electromechanical dynamics is then used to recover cardiac deformation from MRI with fused information from both functional and structural images.

3.1 Functional Data and Volumetric Electrophysiological Recovery

To relate BSPM to the cardiac physiome model, the framework in [1] is used to image volumetric TMP from BSPM. Noninvasive observations of TMP are provided by potential maps on the body surface, with their relation follows the quasi-static electromagnetism on the personalized heart-torso model obtained from subject's MRI. The resulted linear relation between volumetric TMP and BSPM is given as:

$$\mathbf{y}_E = \mathbf{H}_E \mathbf{U}_E \quad (3)$$

where \mathbf{y}_E consists of the body surface potentials (BSP). This model combines the advantage of FEM-based and BEM-based approaches to TMP-to-BSP modeling.

As \mathbf{U}_E is of much higher dimension than \mathbf{y}_E , the E model (1) is necessary to provide the physiological constraint for obtaining a unique \mathbf{U}_E from \mathbf{y}_E . To account for the uncertainties of the model and the measurements, (1) and (3) are reformulated into state-space representations to utilize optimal estimation [4]. By letting the state vector

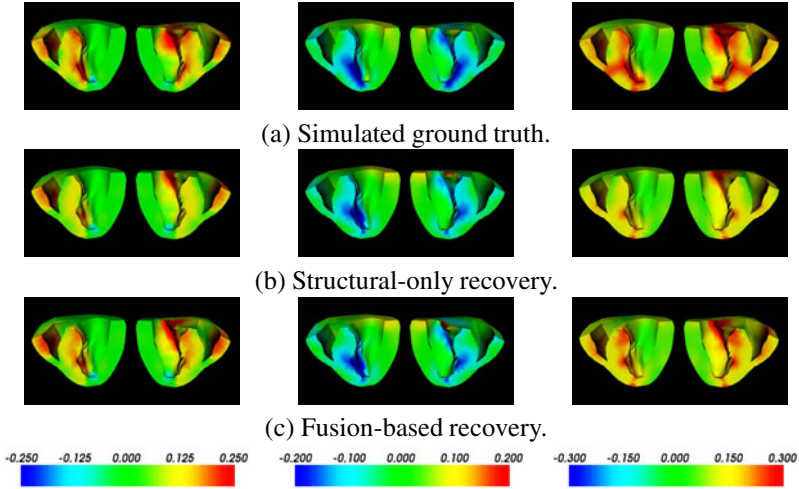


Fig. 2. Synthetic data. Strain maps at the end of systole (250 ms), with red, blue, and green colors representing extension, contraction, and no deformation respectively. Left to right: radial strain, circumferential strain, and first principal strain.

$\mathbf{x}_E^T = [\mathbf{U}_E^T, \mathbf{V}^T]$, and the model uncertainties and measurement errors as $\boldsymbol{\omega}_E$ and $\boldsymbol{\nu}_E$ respectively, the state-space system is given as:

$$\mathbf{x}_E(k) = f_E(\mathbf{x}_E(k-1)) + \boldsymbol{\omega}_E(k-1) \quad (4)$$

$$\mathbf{y}_E(k) = \tilde{\mathbf{H}}_E \mathbf{x}_E(k) + \boldsymbol{\nu}_E(k) \quad (5)$$

with f_E rearranged from (1) and discretized in time, and $\tilde{\mathbf{H}}_E = [\mathbf{H}_E, \mathbf{0}]$. With (4,5), given a subject's BSPM sequence, volumetric TMP can be estimated through the state-space filtering in Section 3.3, which can then be transformed into contraction forces through the EM model, and embedded to the cardiac electromechanical dynamics.

3.2 Structural Data and Electrophysiology-Guided Deformation Recovery

To relate the structural images with the cardiac deformation, the relation between the nodal displacements \mathbf{U}_{BM} and the image-extracted motions \mathbf{y}_{BM} is given as:

$$\mathbf{y}_{BM} = \mathbf{H}_{BM} \mathbf{U}_{BM} \quad (6)$$

where \mathbf{H}_{BM} is the measurement matrix with interpolation functions relating \mathbf{U}_{BM} to \mathbf{y}_{BM} . For a unique solution, \mathbf{U}_{BM} is related to the BSPM-embedded cardiac dynamics through (2), and this establishes the relation between the structural images and the BSPM. As (2) is already discretized in time, by letting the state vector \mathbf{x}_{BM} as \mathbf{U}_{BM} and the model uncertainties as $\boldsymbol{\omega}_{BM}$, the state-update equation is given as:

$$\begin{aligned} \mathbf{U}_{BM}(k) &= \mathbf{U}_{BM}(k-1) + \Delta \mathbf{U}_{BM} + \boldsymbol{\omega}_{BM}(k-1) \\ \Rightarrow \mathbf{x}_{BM}(k) &= f_{BM}(\mathbf{x}_{BM}(k-1)) + \boldsymbol{\omega}_{BM}(k-1) \end{aligned} \quad (7)$$

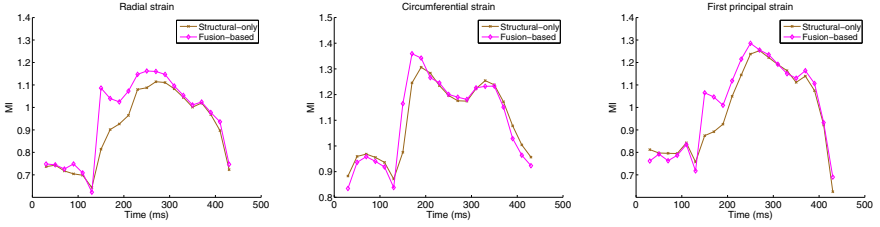


Fig. 3. Synthetic data. Mutual information between strain patterns of the ground truth and the recoveries, for the whole cardiac cycle.

with ΔU_{BM} obtained by solving (2), thus f_{BM} is nonlinear with respect to \mathbf{x}_{BM} . The corresponding stochastic measurement equation is obtained from (6) as:

$$\mathbf{y}_{\text{BM}}(k) = \mathbf{H}_{\text{BM}} \mathbf{x}_{\text{BM}}(k) + \boldsymbol{\nu}_{\text{BM}}(k) \quad (8)$$

where $\boldsymbol{\nu}_{\text{BM}}$ accounts for the measurement errors. With this system, the state-space filtering framework in Section 3.3 can be used to estimate subject-specific cardiac deformation with physiologically fused information from both BSPM and structural images.

3.3 Nonlinear State-Space Filtering for Model-Measurement Coupling

System (4.5) and system (7.8) are in the same form:

$$\mathbf{x}(k) = f(\mathbf{x}(k-1)) + \boldsymbol{\omega}(k-1) \quad (9)$$

$$\mathbf{y}(k) = \mathbf{H} \mathbf{x}(k) + \boldsymbol{\nu}(k) \quad (10)$$

In consequence, the same filtering framework can be applied. To preserve model non-linearity and maintain computational feasibility, we utilize the unscented Kalman filter (UKF) which comprises the advantages of Monte Carlo methods and Kalman filter updates [4]. In UKF, the prediction step is done by the unscented transformation, which calculates the statistics of a random variable undergoing a nonlinear transformation. A set of minimal sample points called sigma points is carefully chosen from the state statistical distribution, which completely captures the true mean and covariance. Each sigma point is projected with intact nonlinearity through (9), and all projected points are combined together for the projected mean and covariance. Afterwards, as (10) is linear, the correction procedures of Kalman filtering can be applied to obtain the estimation.

4 Experiments

4.1 Synthetic Data

The heart architecture from the University of Auckland was used to provide the anatomical cardiac geometry and tissue structure for the experiments [5] (Fig. 1(a)). To show that our framework can aid verifying possible diseased areas, part of the heart was assumed to be infarcted (Fig. 1(b)), where the mechanical stiffness was three times as that

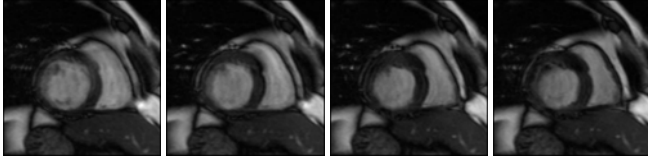


Fig. 4. Human data, case 2. MRI sequence during systole (frame #1, #3, #5, #7).

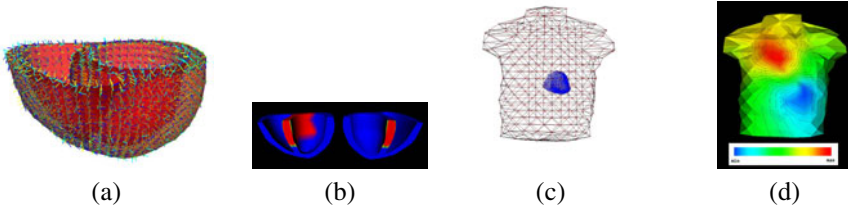


Fig. 5. Human data, case 2. (a) Cardiac geometry segmented from MRI, with mapped tissue structures. (b) Infarcted segments (3, 4, 9, 10) shown in red color. (c) Patient's heart-torso representation. (d) Exemplary input BSPM of ventricular activation.

of the normal tissue, and the cardiac electricity could not propagate through. A cardiac cycle of 450 ms was simulated through the cardiac physiome model as the ground truth. The nodal displacements on the heart boundaries were extracted, and added with noises of 5dB SNR as the noisy structural measurement inputs of the experiments.

Two frameworks were tested. The first framework has no active force \mathbf{F}_c , so only structural measurements were used (structural-only recovery). The second framework is our fusion-based recovery with both subject-specific (abnormal) TMP and structural measurements. In the recovery process, the infarcted regions were unknown.

Fig. 2 shows the strain maps at the end of systole. For both frameworks, the recovered strains show almost no deformation in the infarcted regions. Comparing between the frameworks, the radial and first principal strains recovered by the fusion-based framework are closer to the ground truth than those recovered by the structural-only framework, while the circumferential strains of both frameworks are similar. To quantify the similarities between strain patterns of the ground truth and the recoveries, mutual information (MI) comparison is utilized. Fig. 3 shows that similar to the observations of the strain maps, the recovered strains using the fusion-based framework are closer to the ground truth in general, especially during systole (120 ms to 250 ms).

4.2 Human Data

To show the potential clinical applicability of our framework, experiments were done on human data sets comprising both BSPM and MRI, which are case 1 and case 2 available at [6], from two patients with acute myocardial infarction. Case 1 contains a human short-axis MRI sequence of 19 frames (52.5 ms/frame), with 12 slices/frame. Case 2 contains a human short-axis MRI sequence of 16 frames (50 ms/frame), with

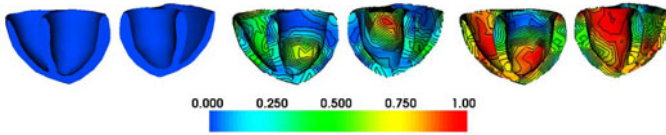


Fig. 6. Human data, case 2. Normalized volumetric myocardial TMP recovered from patient's BSPM, shown with isochrones. Left to right: 0 ms, 5 ms, and 12 ms.

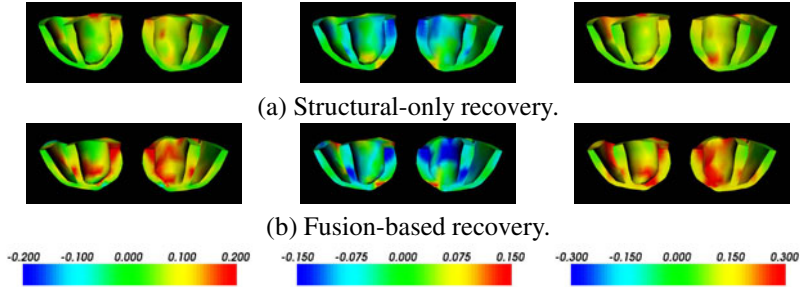


Fig. 7. Human data, case 2. Strain maps at the end of systole (300 ms). Left to right: radial strain, circumferential strain, and first principal strain.

13 slices/frame. Both sequences have 8 mm inter-slice spacing, and in-plane resolution 1.32 mm/pixel (Fig. 4). Each case has the patient's BSPM recorded from 123 electrodes and interpolated to 370 nodes on the torso surface. Segmentations were performed to obtain the initial heart geometries, with the fibrous-sheet structures mapped from the Auckland heart architecture using nonrigid registration (Fig. 5(a)). The expert-identified infarcted segments of case 1 are 1, 2, 3, 8, 9, 13, 14, 15, and of case 2 are 3, 4, 9, 10, with respect to the myocardial nomenclature in [7] (Fig. 5(b)). Fig. 5(c) illustrates the combined heart-torso model of the patient, where the torso was obtained through segmentation of the patient's torso MRI. Fig. 5(d) exemplifies an input BSPM.

In the experiments, the volumetric TMP were first recovered from the BSPM using the system described in Section 3.1. The recovered TMP show limited propagation in the infarcted regions, which is different from the normal regions (Fig. 6).

Motions of the heart boundaries were extracted from MRI using free-form image registration between consecutive image frames [8]. Experiments were performed using the structural-only framework and the fusion-based framework, on both case 1 and case 2. Fig. 7 shows that although both frameworks show relatively small deformation in the infarcted regions, the fusion-based framework has a better contrast between the normal and the infarcted regions, and the shapes of the regions with negligible deformation are more consistent with the shapes of the infarcted regions.

For clinical applications, the capabilities of frameworks to aid locating diseased regions are of great interest, thus the frameworks are compared in this aspect. As there is a strong inverse relationship between the first principal strain and the extent of myocardial infarction [9], we sorted the segments at different levels according to the mean

Table 1. Human data. Segment numbers at different levels sorted according to the mean of the first principal strains of each segment at the end of systole, in ascending order. The numbers of the infarcted segments are highlighted in red color.

Case 1	Basal	Mid	Apical	Case 2	Basal	Mid	Apical
Structural-only	3 4 2 5 1 6	10 9 8 7 11 12	14 15 13 16	Structural-only	3 5 2 1 4 6	10 9 11 12 8 7	14 15 16 13
Fusion-based	3 2 4 1 5 6	10 9 8 7 11 12	14 13 15 16	Fusion-based	3 4 1 2 5 6	10 9 11 7 8 12	14 15 16 13

Table 2. Human data. Optimal sensitivities (SEN) and specificities (SPEC) of the frameworks for identifying infarcted segments.

Case 1	Basal	Mid	Apical	Case 2	Basal	Mid
	SEN / SPEC	SEN / SPEC	SEN / SPEC		SEN / SPEC	SEN / SPEC
Structural-only	67% / 67%	100% / 75%	100% / 100%	Structural-only	50% / 100%	100% / 100%
Fusion-based	100% / 67%	100% / 75%	100% / 100%	Fusion-based	100% / 100%	100% / 100%

of the first principal strains of each segment at the end of systole, in ascending order (Table 1). From Table 1, we can calculate the various sensitivities and specificities with different cutoff segments, where sensitivities are the ratios between the correctly identified positives (infarcted segments) and the actual positives, and the specificities are the ratios between the correctly identified negatives (normal segments) and the actual negatives. Table 2 shows that for both cases, the fusion-based framework achieves the same or higher optimal sensitivities compared with the structural-only framework.

References

1. Wang, L., Zhang, H., Wong, K.C.L., Liu, H., Shi, P.: Noninvasive volumetric imaging of cardiac electrophysiology. In: IEEE CVPR, pp. 2176–2183 (2009)
2. Sermesant, M., Delingette, H., Ayache, N.: An electromechanical model of the heart for image analysis and simulation. IEEE Transactions on Medical Imaging 25(5), 612–625 (2006)
3. Wong, K.C.L., Wang, L., Zhang, H., Shi, P.: Nonlinear cardiac deformation recovery from medical images. In: IEEE ICIP, pp. 2513–2516 (2009)
4. Simon, D.: Optimal State Estimation: Kalman, H_∞ , and Nonlinear Approaches. John Wiley & Sons, Inc., Chichester (2006)
5. Nash, M.: Mechanics and Material Properties of the Heart using an Anatomically Accurate Mathematical Model. PhD thesis, The University of Auckland (1998)
6. PhysioNet/Computers in Cardiology challenge 2007: electrocardiographic imaging of myocardial infarction (2007), <http://www.physionet.org/challenge/2007/>
7. Cerqueira, M.D., Weissman, N.J., Dilsizian, V., Jacobs, A.K., Kaul, S., Laskey, W.K., Pennell, D.J., Rumberger, J.A., Ryan, T., Verani, M.S.: Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: a statement for healthcare professionals from the cardiac imaging committee of the council on clinical cardiology of the American Heart Association. Circulation 105, 539–542 (2002)
8. Perperidis, D., Mohiaddin, R.H., Rueckert, D.: Spatio-temporal free-form registration of cardiac MR image sequences. Medical Image Analysis 9(5), 441–456 (2005)
9. Tim Marcus, J., Götte, M.J.W., van Rossum, A.C., Kuijper, J.P.A., Heethaar, R.M., Axel, L., Visser, C.A.: Myocardial function in infarcted and remote regions early after infarction in man: assessment by magnetic resonance tagging and strain analysis. Magnetic Resonance in Medicine 38(5), 803–810 (1997)

Adaptive Noise Filtering for Accurate and Precise Diffusion Estimation in Fiber Crossings

Matthan W.A. Caan¹, Ganesh Khedoe², Dirk Poot², Arjan den Dekker²,
Silvia Olabarriga¹, Kees Grimbergen¹, Lucas van Vliet², and Frans Vos^{1,2}

¹ Academic Medical Center, University of Amsterdam, NL

² Delft University of Technology, NL

m.w.a.caan@amc.uva.nl

Abstract. Measuring the diffusion properties of crossing fibers is very challenging due to the high number of model parameters involved and the intrinsically low SNR of Diffusion Weighted MR Images. Noise filtering aims at suppressing the noise while pertaining the data distribution. We propose an adaptive version of the Linear Minimum Mean Square Error (LMMSE) estimator to achieve this. Our filter applies an adaptive filtering kernel that is based on a space-variant estimate of the noise level and a weight consisting of the product of a Gaussian kernel and the diffusion similarity with respect to the central voxel. The experiments show that the data distribution after filtering is still Rician and that the diffusivity values are estimated with a higher precision while pertaining an equal accuracy. We demonstrate on brain data that our adaptive approach performs better than the initial LMMSE estimator.

1 Introduction

Diffusion Weighted Magnetic Resonance Imaging (DW-MRI) is hampered by a low signal to noise ratio (SNR) for high values for the diffusion weighting parameter b . As a consequence, diffusion model parameters may be estimated with low precision. An incorrect representation of the noise properties, particularly assuming a Gaussian instead of a Rician noise distribution in the DWIs, may render a biased signal model [1]. Noise filtering aims to increase the accuracy and precision of the estimated diffusivity. A sufficiently high spatial resolution permits the identification of locally homogeneous tissue regions in which averaging reduces the noise, but pertains the signal.

In previous work, the Linear Minimum Mean Square Error (LMMSE) estimator [2] was proposed to reduce Rician distorted MR data, including DW-MRI. In DW-MRI data, anisotropic smoothing was applied to reduce noise [3]. An estimate of the noise level was obtained by extending the single tensor description with a Rician noise distribution and estimating the noise level by a Maximum Likelihood (ML) framework [4,5]. Rician noise reduction by spatial regularization [6,7] was used to limit the bias in estimates of a single tensor model. Alternatively, sequential anisotropic multichannel Wiener filtering allows the correction of the bias in the diffusivity estimates due to the Rician distribution of Diffusion

Weighted Images (DWIs) [8]. By modeling a Markov Random tensor Field, a joint optimization of the likelihood and a prior could be performed [9]. These methods have in common that they are applied to a single tensor model.

To date, relatively little attention is paid to filtering noise in fiber crossings. Crossing fiber orientations were more precisely reconstructed after filtering [8]. Estimating diffusion properties independently per crossing fiber bundle involves more model parameters and hence is inherently more susceptible to noise [10].

We propose a method for noise suppression in fiber crossings, which is also valid in voxels with a single fiber configuration. The technique is inspired by the LMMSE estimator introduced by Aja-Fernandez [2], but it contains two important improvements. First, the original Aja-Fernandez approach used an isotropic, uniform kernel, which mixes neighbouring tissues around transitions. We use an adaptive Gaussian kernel by only including voxels whose single tensors are similarly shaped. Consequently, the bias induced by the mixing is reduced. Second, the Aja-Fernandez filtering approach uses a global estimate of the noise level σ , while the noise level is known to be smoothly varying due to parallel imaging [5]. We perform an ML estimation of σ per voxel by fitting a constrained dual tensor model.

The distribution of the data after LMMSE-filtering should be known in order to allow accurate parameter estimation. In our experiments we will empirically show that our model, assuming a Rician distribution, still adequately describes the data. It is then demonstrated that the variance in the estimated diffusion parameters is decreased while the bias remains low. Finally, we will compare our adaptive noise suppression scheme with the original Aja-Fernandez approach on brain data.

2 Method

Linear Minimum Mean Square Estimation (LMMSE)-filtering [2] has been proposed to reduce noise in Rician distributed MRI data. This closed-form solution considers the DWIs as realizations of stochastic processes. The local signal variance is reduced, steered by the estimated noise level. The estimate of the squared signal \hat{A}^2 equals

$$\hat{A}^2 = \langle M^2 \rangle - 2\sigma^2 + \left(1 - \frac{4\sigma^2 (\langle M^2 \rangle - \sigma^2)}{\langle M^4 \rangle - \langle M^2 \rangle^2} \right) (M^2 - \langle M^2 \rangle), \quad (1)$$

in which M represents the measured signal and $\langle \cdot \rangle$ the estimator of a sample's expectation value. The LMMSE-filter is applied in 3-D to each DWI independently. The variance of the noise, σ^2 , was originally assumed constant in the entire field of view and estimated as the mode of the locally estimated signal variances. The expectation value of the sample is approximated by a weighted sum in a local neighbourhood η , $\langle I \rangle = \frac{1}{\sum d_p} \sum_{p \in \eta} d_p I_p$. Initially, uniform weighting, i.e. $d_p = 1 \forall p$, was proposed.

In the following Section 2.1 we will introduce an ML approach involving a dual tensor model for improved estimation of the noise level σ . Section 2.2 will discuss an adaptive filtering kernel which is valid at fiber crossings.

2.1 Maximum Likelihood Noise Level Estimation

DW-MRI records the magnitude of the complex MR signal. The real and imaginary channels of the complex MR image are assumed to be independently affected by Gaussian noise, such that the DWIs are Rician distributed. Hence, the probability density function of a measured signal $S_{m,j} \geq 0$ is given by

$$p(S_{m,j} | S_{\theta,j}, \sigma) = \frac{S_{m,j}}{\sigma^2} \exp\left(-\frac{S_{m,j}^2 + S_{\theta,j}^2}{2\sigma^2}\right) I_0\left(\frac{S_{m,j} S_{\theta,j}}{\sigma^2}\right). \quad (2)$$

Here $S_{\theta,j}$ denotes the true underlying value given the parameter vector θ , σ the standard deviation of the noise, and I_0 the zeroth order modified Bessel function of the first kind. Now the log likelihood function $\ln L(\theta, \sigma | \mathbf{S}_m)$ becomes

$$\ln L(\theta, \sigma | \mathbf{S}_m) = -2 \ln(\sigma) + \sum_{j=1}^{N_g} \left(\ln(S_{m,j}) - \frac{S_{m,j}^2 + S_{\theta,j}^2(\theta)}{2\sigma^2} + \ln I_0\left(\frac{S_{m,j} S_{\theta,j}(\theta)}{\sigma^2}\right) \right). \quad (3)$$

where the DWIs are assumed independent, such that the joint probability density function of the signal profile \mathbf{S}_m in a voxel is given by the product of the marginal distributions for the measured signal $S_{m,j}$ in each of the N_g diffusion weighted directions \mathbf{g}_j . ML estimation of the parameters θ is obtained by maximizing the log likelihood function: $\hat{\theta}_{ML} = \arg \{\max_{\theta} (\ln L)\}$.

We propose to estimate the noise level in a dual tensor model to avoid a bias in the estimated diffusivity of crossing fibers. We assert a model for the diffusion weighted signal $S_{\theta,j}$ consisting of two tensors and an isotropic compartment:

$$S_{\theta,j} = S_0 \left\{ \sum_{i=1,2} f_i \exp(-b_j \mathbf{g}_j^T \mathbf{D}_i \mathbf{g}_j) + f_{iso} \exp(-b_j D_{iso}) \right\}, \quad (4)$$

where S_0 is the signal measured without diffusion weighting and f_{\dots} the volume fractions, with $f_2 = 1 - f_1 - f_{iso}$. We choose $D_{iso} = 3.0 \cdot 10^{-3} \text{ mm}^2 \text{ s}^{-1}$ as an isotropic diffusion constant (to approximate diffusion values reported in CSF), because f_{iso} and D_{iso} are dependent. The signal is measured with a diffusion weighting b_j in gradient direction \mathbf{g}_j . The diffusion tensors are spectrally decomposed, $\mathbf{D}_i = \mathbf{R}_i \mathbf{E}_i \mathbf{R}_i$, with $\mathbf{E}_i = \text{diag}(\lambda_{\parallel} \lambda_{i\perp} \lambda_{i\perp})$ being the eigenvalue matrix with axial and planar diffusion values (assuming axially symmetric tensors). The two rotation matrices \mathbf{R}_i are parametrized using Euler angles.

The estimation is done per voxel in two steps, by optimizing the log-likelihood function (Eq. 3). First, σ is estimated in which we constrain the \mathbf{E}_i to be equal while pertaining constant f_{\dots} and D_{iso} , to reduce the number of parameters and hence the variance in the estimated σ . Second, all parameters are estimated, with the estimated σ substituted in Eq. 3.

2.2 Adaptive Filtering

Instead of a uniform filter we propose a filter d_p that adapts its shape automatically to the local structure (p indexes the filter components). The filter coefficients are the product of an isotropic Gaussian kernel $G_{p|k}$ defined at the central pixel k having a width σ_G , and a normalized weighting based on the single tensor similarity at position p ($D_{s,p}$) compared to the tensor in the central pixel k ($D_{s,k}$). We represent the six unique single tensor components by a vector D_s . The filter d_p yields

$$d_p = \frac{D_{s,p}^T D_{s,k}}{|D_{s,p}| \cdot |D_{s,k}|} \cdot G_{p|k}(\sigma_G), \quad (5)$$

in which $|\cdot|$ is the vector norm. Note that in a region of crossing fibers, the DWIs are filtered within the plane of the crossing, spanned by an oblate single tensor. We chose a Gaussian kernel width $\sigma_G = 4$ mm, coarsely corresponding to the uniform filtering neighbourhood of size 1×1 cm² in [2].

The adaptive filter was implemented in Matlab (The MathWorks, Natick, MA) and is illustrated in figure 1(a) in which the filter kernel is overlaid in red over an image containing the corpus callosum (CC). The original approach involved an isotropic kernel, due to which the filtering mixes white matter and cerebrospinal fluid (left image). The adaptation to the local structure (such as in [11]) of the corpus callosum reduces this effect (right). In crossings, the kernel gets an oblate shape (bottom).

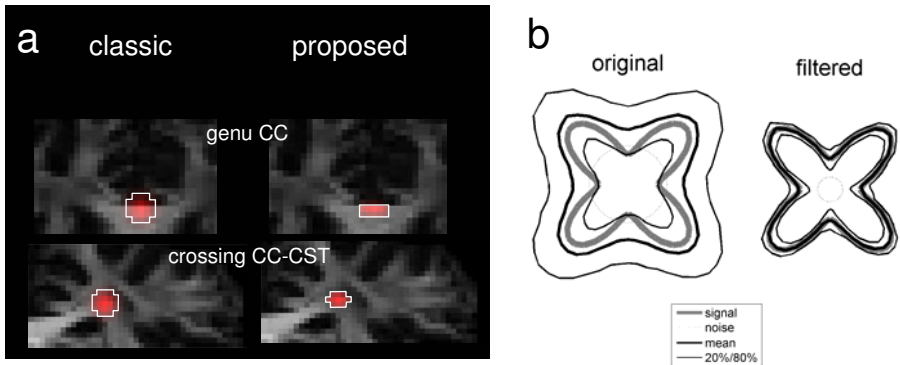


Fig. 1. (a) The original isotropic Aja-Fernandez kernel (left) and the proposed adaptive kernel (right) are shown in red (and manually outlined in white), in the genu of the CC (top) and crossing of CC and CST (bottom). (b) Original and filtered profile of a simulated measurement.

3 Experiments

3.1 Brain Data Acquisition

Diffusion Weighted Image acquisition of 20 axial slices (mid-brain) of a healthy volunteer was performed on a Philips Intera 3.0 Tesla MRI scanner by means of a spin echo EPI-sequence, on an eight-channel head coil with SENSE reconstruction. An imaging matrix of 112×110 was used, from which 128×128 sized slices were reconstructed, with a voxel size of $1.7 \times 1.7 \times 2.2 \text{ mm}^3$. The diffusion weighting was along 92 three-fold tessellated icosahedric gradient directions, with two b -values: $\mathbf{b} = [1.0 \ 3.0] \cdot 10^3 \text{ mm}^{-2}\text{s}$. Per b -value, one non-diffusion weighted image S_0 was acquired. 20 axial slices were acquired, resulting in a total scanning time of 30 minutes. The deformations induced by eddy currents were corrected with an affine registration in the phase encoding direction [12]. In addition, a rigid registration of the S_0 -images and coregistration of the DWIs corrected for patient head motion (up to 2 voxels).

3.2 Filtered Data Distribution

The distribution of the data after filtering must be known to accurately estimate the model parameters. Since it was stated that the distribution of the data after filtering may no longer be Rician distributed [2], we empirically studied two potential effects on the distribution. First, the mixing of Rician distributed data might induce a multivariate Rician distribution. Second, the Rician distribution might converge to a Gaussian due to the lower noise level and as result of averaging noisy realizations, in accordance to the Central Limit Theorem. Clearly, if the post-filtering distribution is multivariate Rician distributed, the ML estimation with the given Rician noise model is no longer valid.

The filter's effect on synthetic Rician distorted data was studied with the following procedure. Different FA-values for both tensors $\text{FA}_1 = 0.66$ and $\text{FA}_2 = 0.75$ were selected by constraining axial diffusivities $\lambda_{1\parallel} = \lambda_{2\parallel} = 1.4 \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$ and planar diffusivities $\lambda_{1\perp} = 0.4 \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$ and $\lambda_{2\perp} = 0.3 \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$. Furthermore, we choose $f_1 = 0.4$ and $f_{iso} = 0.15$ with $D_{iso} = 3.0 \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$ while the opening angle between the long axis of both tensors was $\frac{2\pi}{5} \text{ rad} = 72^\circ$. The magnitudes of the DWI simulated with these parameters were replicated on a 2D-grid. The signals on each grid point were independently distorted by Rician distributed noise ($\text{SNR} = \frac{S_0}{\sigma} = 1$), such that a significant bias due to the noise was introduced. The experiment was repeated 100 times. The aggregate signal values before and after filtering using the proposed method are shown in figure 1(b). It turns out that the filter removes the bias from the data.

In brain data, we focused on a set of highly attenuated DW-signal values, in which a Rician distribution is expected. We included voxels with an anisotropic diffusion profile, by thresholding on $\text{FA} \geq 0.5$ and $\text{MD} > 1.0 \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$. Per voxel the DW-signal value S_j was selected from the signal profile \mathbf{S}_m whose corresponding gradient direction \mathbf{g}_j was most closely aligned with the principal diffusion direction of a single tensor fit \mathbf{v}_1 , i.e. $\{S_j \in \mathbf{S}_m \mid \arg \max_j \mathbf{g}_j \cdot \mathbf{v}_1\}$. Histograms

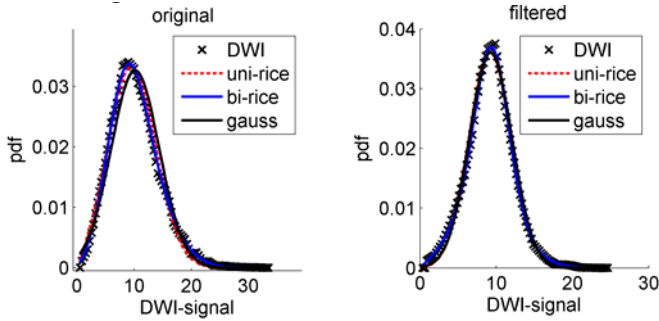


Fig. 2. Histograms of a set of highly attenuated DW-signal values, selected from brain data (see text). A Gaussian (black line), Rician (red line) and bivariate (weighting of two) Rician distributions (blue line) were fit to these points. The left image shows original data, the right image shows filtered data.

before and after applying the proposed noise reduction filter were determined to which uni- and bivariate Rician and Gaussian distributions were fitted: c.f. figure 2. The Kolmogorov-Smirnov test revealed that the Gaussian distribution significantly differed from the measured data, while both Rician distributions comply with the measured data. After filtering, no significant differences were found. This result shows that a univariate Rician is an appropriate noise model prior to filtering. Figure 2 shows that the filtering changes the original noise distribution to a Gaussian. An approximate Gaussian distribution also follows from the Central Limit Theorem. The Rician also fits well due to the higher signal level after filtering: for high SNR the Rician distribution approaches the Gaussian distribution. Practically, a Rician distribution was always fit, since it better generalizes to low signal levels.

3.3 Accurate and Precise Parameter Estimation

A Monte Carlo simulation was performed to assess the accuracy of the estimated parameters after noise filtering in the presence of a varying anisotropy in the neighbourhood (see figure 3(a)). A 2D-grid of measurements at crossings was simulated. By adjusting $\lambda_{i\perp}$ the FA-values of both tensors were varied along the horizontal and retained constant along the vertical axis (with a length of 500 voxels). Rician noise was added to all measurements such that the SNR=12. The grid was noise filtered after which we estimated SNR=25 ($\hat{\sigma} = 4$). Levenberg-Marquardt optimization of the log likelihood function (Eq. 3) was used in the parameter estimations.

Measurements were performed along the dashed line in figure 3(a). Figure 3(b) shows distributions of the estimated parameters and the mean errors, both before and after noise reduction. The distributions after filtering are narrower as a result of the noise reduction. A significant bias is only observed in f_{iso} , the estimated value is 13% lower than the modeled value. Simultaneously, the bias in the other

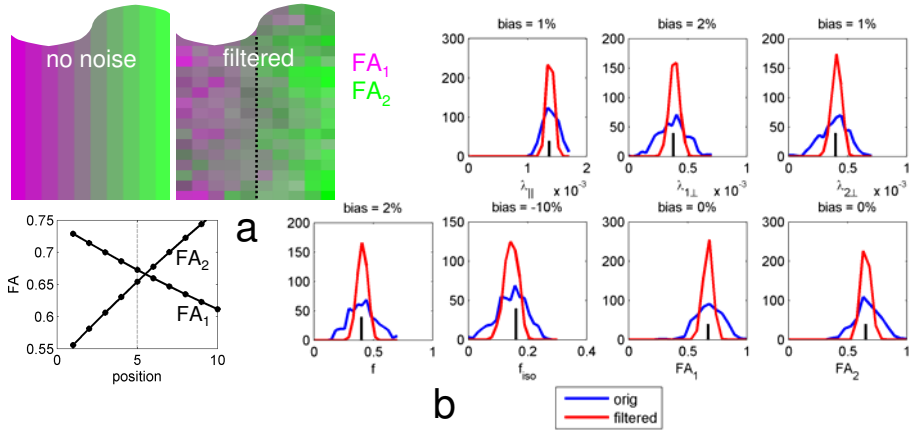


Fig. 3. (a) A 2D-grid of measurements at crossings was simulated. Generated and estimated FA-values are given. Measurements were performed along the dashed line. (b) Distributions of estimated parameters, prior to and after filtering. The true parameter values are indicated by a black bar.

parameters is negligible. We conclude that noise filtering can be safely applied to allow for accurate and precise estimation of diffusivity in crossing fibers.

3.4 Noise Reduction in Brain Data

The effect of our noise filtering applied to brain data is illustrated in figure 4. It demonstrates that the proposed method employing a space-variant adaptive kernel performs better than the original isotropic kernel. This is supported by a shift towards a higher SNR. It should also be noted from the smoothly varying SNR-field that the estimation of σ is stable. The average computation time for filtering and tensor estimation per voxel on a notebook computer equals 0.09s without filtering, and only 0.17s with the proposed method.

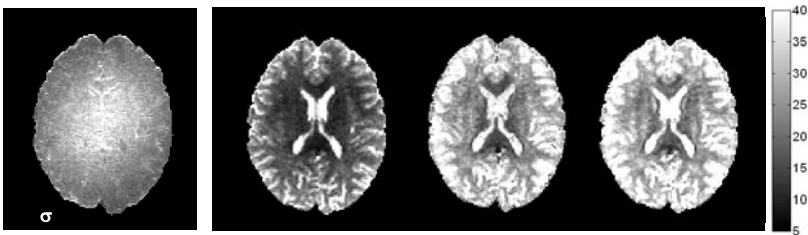


Fig. 4. Estimated SNR ($\frac{S_0}{\sigma}$) before filtering (left), after filtering with an isotropic kernel (middle), and filtering with the proposed method (right)

4 Discussion

An adaptive filter for suppressing noise in fiber crossings was proposed. The impact on clinical applications of our method is in comparative studies that aim to assess white matter defects in crossings. Future work will be to estimate a non-stationary noise-covariance for the different DWI directions [13]. The recently proposed Diffusion Type Based similarity measure [14] may even further reduce filtering to homogeneous tissue regions. A crossing fiber phantom whose diffusion properties are accurately known will aid in validating our method.

References

- den Dekker, A.J., Sijbers, J.: *Advanced Image Processing in Magnetic Resonance Imaging. Signal Processing and Communications*, vol. 27, pp. 85–143. CRC Press, Taylor & Francis Group, LLC, Boca Raton (2005)
- Aja-Fernandez, S., Niethammer, M., Kubicki, M., Shenton, M., Westin, C.F.: Restoration of DWI data using a rician LMMSE estimator. *IEEE Trans. Med. Im.* 27, 1389–1403 (2008)
- Ding, Z., Gore, J., Anderson, A.: Reduction of Noise in Diffusion Tensor Images Using Anisotropic Smoothing. *Magn. Reson. Med.* 53, 486–490 (2005)
- Andersson, J.L.: Maximum a posteriori estimation of diffusion tensor parameters using a rician noise model: Why, how and but. *Neuroimage* 42, 1340–1356 (2008)
- Landman, B., Bazin, P.L., Prince, J.: Diffusion tensor estimation by maximizing Rician likelihood. In: *IEEE 11th ICCV*, pp. 2433–2440 (2007)
- Fillard, P., Pennec, X., Arsigny, V., Ayache, N.: Clinical DT-MRI estimation, smoothing, and fiber tracking with log-euclidean metrics. *IEEE Transactions on Medical Imaging* 26, 1472–1482 (2007)
- Basu, S., Fletcher, T., Whitaker, R.: Rician noise removal in diffusion tensor MRI. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) *MICCAI 2006. LNCS*, vol. 4190, pp. 117–125. Springer, Heidelberg (2006)
- Martin-Fernandez, M., Munoz-Moreno, E., Cammoun, L., Thiran, J.P., Westin, C.F., Alberola-Lopez, C.: Sequential anisotropic multichannel Wiener filtering with Rician bias correction applied to 3D regularization of DWI data. *Med. Im. Anal.* 13, 19–35 (2009)
- Bansal, R., Staib, L., Xu, D., Laine, A., Liu, J., Peterson, B.: Using perturbation theory to reduce noise in diffusion tensor fields. *Med. Im. Anal.* 13, 580–597 (2009)
- Caan, M., Khedoe, H., Poot, D., den Dekker, A., Olabarriaga, S., Grimbergen, C., van Vliet, L., Vos, F.: Estimation of diffusion properties in crossing fiber bundles. *IEEE Trans. Med. Im.* (2010) (in press)
- van Kempen, G., van den Brink, N., van Vliet, L., van Ginkel, M., Verbeek, P., Blonk, H.: The application of a local dimensionality estimator to the analysis of 3d microscopic network structures. In: *Proc. 11th Scandinavian Conference on Image Analysis (SCIA)*, pp. 447–455 (1999)
- Mangin, J.F., Poupon, C., Clark, C., Bihan, D.L., Bloch, I.: Eddy-current distortion correction and robust tensor estimation for MR diffusion imaging. In: Niessen, W.J., Viergever, M.A. (eds.) *MICCAI 2001. LNCS*, vol. 2208, pp. 186–193. Springer, Heidelberg (2001)
- Tristán-Vega, A., Aja-Fernández, S.: DWI filtering using joint information for DTI and HARDI. *Medical Image Analysis* (2009)
- Moreno, E.M., Fernández, M.M.: Characterization of the similarity between diffusion tensors for image registration. *Computers in Biology and Medicine* 39, 251–265 (2009)

A Sobolev Norm Based Distance Measure for HARDI Clustering

A Feasibility Study on Phantom and Real Data

Ellen Brunenberg¹, Remco Duits^{2,1},
Bart ter Haar Romeny¹, and Bram Platel³

¹ Biomedical Engineering, Eindhoven University of Technology
² Mathematics and Computer Science, Eindhoven University of Technology
³ Biomedical Engineering, Maastricht University Medical Center
e.j.l.brunenberg@tue.nl

Abstract. Dissimilarity measures for DTI clustering are abundant. However, for HARDI, the \mathbb{L}_2 norm has up to now been one of only few practically feasible measures. In this paper we propose a new measure, that not only compares the amplitude of diffusion profiles, but also rewards coincidence of the extrema. We tested this on phantom and real brain data. In both cases, our measure significantly outperformed the \mathbb{L}_2 norm.

1 Introduction

Segmentation of gray matter nuclei in the brain can provide solutions for a multitude of clinical questions. A rich variety of clustering algorithms has been employed for this purpose, together with a large number of dissimilarity measures, mainly focusing on DTI. Wiegell et al. [1] clustered DTI in the thalamus, using k-means and the Frobenius norm. Ziyang et al. used graph cuts [2] on the same data to investigate the angular difference and K-L divergence.

For HARDI, the \mathbb{L}_2 norm has been used most often to compare ODFs, possibly represented by spherical harmonic (SH) coefficients. Grassi et al. [3] performed k-medoids clustering in the thalamus using the \mathbb{L}_2 norm on the ODFs. Similarly, Descoteaux [4] used the \mathbb{L}_2 norm on SH coefficients. Apart from clustering using SH coefficients, there are other ways to represent the ODF. Some studies have implemented mixture models like von Mises-Fisher distributions [5], while others have used a model-free representation of the ODF [6].

In this paper, we propose a new dissimilarity measure for ODFs represented by SH coefficients. Whereas the \mathbb{L}_2 norm only compares amplitudes of the diffusion profiles, our Sobolev norm also takes into account whether the extrema coincide.

2 Dissimilarity Measure Based on Sobolev Norm

Consider HARDI-image $U : \mathbb{R}^3 \times S^2 \rightarrow \mathbb{R}^+$ and assume it is square integrable, i.e. $U \in \mathbb{L}_2(\mathbb{R}^3 \times S^2)$. By restricting this HARDI-image to two fixed points, say $\mathbf{x}_1, \mathbf{x}_2$ in \mathbb{R}^3 we obtain two functions on the 2-sphere

$$S^2 \ni \mathbf{n} \mapsto U(\mathbf{x}_1, \mathbf{n}) \in \mathbb{R}^+ \text{ and } S^2 \ni \mathbf{n} \mapsto U(\mathbf{x}_2, \mathbf{n}) \in \mathbb{R}^+,$$

which we from now on denote by $U(\mathbf{x}_1, \cdot)$ and $U(\mathbf{x}_2, \cdot)$. These functions can be represented by a so-called glyph $\mathcal{S}_\mu(U)(\mathbf{x}_1)$ and $\mathcal{S}_\mu(U)(\mathbf{x}_2)$ as defined below.

Definition 1. A glyph of a distribution $U : \mathbb{R}^3 \times S^2 \rightarrow \mathbb{R}^+$ on positions and orientations is a surface $\mathcal{S}_\mu(U)(\mathbf{x}) = \{\mathbf{x} + \mu U(\mathbf{x}, \mathbf{n}) \mathbf{n} \mid \mathbf{n} \in S^2\} \subset \mathbb{R}^3$ for some $\mathbf{x} \in \mathbb{R}^3$, $\mu > 0$. A glyph visualization of distribution $U : \mathbb{R}^3 \times S^2 \rightarrow \mathbb{R}^+$ is a visualization of a field $\mathbf{x} \mapsto \mathcal{S}_\mu(U)(\mathbf{x})$ of glyphs, with $\mu > 0$ a suitable constant.

A common approach to compare two glyphs $\mathcal{S}_\mu(U)(\mathbf{x}_1)$ and $\mathcal{S}_\mu(U)(\mathbf{x}_2)$ is to compute the \mathbb{L}_2 distance between $U(\mathbf{x}_1, \cdot) : S^2 \rightarrow \mathbb{R}^+$ and $U(\mathbf{x}_2, \cdot) : S^2 \rightarrow \mathbb{R}^+$:

$$d(U(\mathbf{x}_1, \cdot), U(\mathbf{x}_2, \cdot)) = \sqrt{\int_{S^2} |U(\mathbf{x}_1, \mathbf{n}) - U(\mathbf{x}_2, \mathbf{n})|^2 d\sigma(\mathbf{n})}$$

where σ denotes the usual surface measure on S^2 . However, this distance only compares glyph amplitudes. It does not take into account robust regularization and more importantly, it does not consider whether the extrema of the glyphs coincide. Therefore we include a blob-detector [7] in our distance. We do not use higher-order derivatives because they hinder damping before the Nyquist frequency and are more ill-posed due to a higher operator norm.

We have defined the (squared) Sobolev distance between two glyphs as

$$\begin{aligned} & (d_{\alpha, \gamma, t}(U(\mathbf{x}_1, \cdot), U(\mathbf{x}_2, \cdot)))^2 \\ & := \int_{S^2} \left| (e^{-t|\Delta_{S^2}|^\alpha} U)(\mathbf{x}_1, \mathbf{n}) - (e^{-t|\Delta_{S^2}|^\alpha} U)(\mathbf{x}_2, \mathbf{n}) \right|^2 d\sigma(\mathbf{n}) \\ & + \gamma^{2\alpha} \int_{S^2} \left| |\Delta_{S^2}|^\alpha (e^{-t|\Delta_{S^2}|^\alpha} U)(\mathbf{x}_1, \mathbf{n}) - |\Delta_{S^2}|^\alpha (e^{-t|\Delta_{S^2}|^\alpha} U)(\mathbf{x}_2, \mathbf{n}) \right|^2 d\sigma(\mathbf{n}) \\ & =: \|U(\mathbf{x}_1, \cdot) - U(\mathbf{x}_2, \cdot)\|_{\mathbb{H}_t^{2\alpha}(S^2)}^2. \end{aligned} \tag{1}$$

This Sobolev distance basically is a sum of a standard (smoothed) \mathbb{L}_2 part (first integral) and a (smoothed) blob-sensitive second part (second integral). Next we provide a brief explanation on the involved parameters:

- The parameter $\alpha \in [\frac{1}{2}, 1]$ denotes the α -scale space regularization on a sphere [8], applied at time $t \geq 0$ (or scale $t^{\frac{1}{2\alpha}}$). Note that $e^{-t|\Delta_{S^2}|^\alpha}$ denotes a smoothing operator generated by a fractional power of the Laplace-Beltrami operator Δ_{S^2} , i.e. $W_U(\mathbf{x}, \mathbf{n}, t) = e^{-t|\Delta_{S^2}|^\alpha} U(\mathbf{x}, \mathbf{n})$ is the solution of

$$\begin{cases} \frac{\partial W_U}{\partial t}(\mathbf{x}, \mathbf{n}, t) = -|\Delta_{S^2}|^\alpha W_U(\mathbf{x}, \mathbf{n}, t), & \mathbf{x} \in \mathbb{R}^3, \mathbf{n} \in S^2, t \geq 0, \\ W_U(\mathbf{x}, \mathbf{n}, 0) = U(\mathbf{x}, \mathbf{n}). \end{cases} \tag{2}$$

- The parameter $t \geq 0$ determines the stopping time of the spherical α -scale space regularization.
- The parameter γ (physical dimension $[\text{Length}]^2$) determines the influence of the blob-sensitive and the intensity-sensitive \mathbb{L}_2 part of the Sobolev norm.

The blob-sensitive part is the same as the total difference of all scale space dynamics of the glyphs. This follows from the fact that (II) can be rewritten as

$$\int_{S^2} |W(\mathbf{x}_1, \mathbf{n}, t) - W(\mathbf{x}_2, \mathbf{n}, t)|^2 + \gamma^{2\alpha} \left| \frac{dW}{dt}(\mathbf{x}_1, \mathbf{n}, t) - \frac{dW}{dt}(\mathbf{x}_2, \mathbf{n}, t) \right|^2 d\sigma(\mathbf{n})$$

$$= \int_{S^2} |W(\mathbf{x}_1, \mathbf{n}, t) - W(\mathbf{x}_2, \mathbf{n}, t)|^2 + \gamma^{2\alpha} \left| |\Delta_{S^2}|^\alpha W(\mathbf{x}_1, \mathbf{n}, t) - |\Delta_{S^2}|^\alpha W(\mathbf{x}_2, \mathbf{n}, t) \right|^2 d\sigma(\mathbf{n}),$$

where we use $W = W_U$. Roughly speaking, $\gamma > 0$ balances the similarity of the extrema and the similarity of the amplitudes of $\mathcal{S}_\mu(U)(\mathbf{x}_1)$ and $\mathcal{S}_\mu(U)(\mathbf{x}_2)$.

Recall that the spherical harmonics $\{Y_{lm}\}_{l=0, \dots, \infty}^{m=-l, \dots, l}$ form an orthonormal basis of $\mathbb{L}_2(S^2)$, i.e.

$$(Y_{lm}, Y_{l'm'})_{\mathbb{L}_2(S^2)} := \int_{S^2} \overline{Y_{lm}(\mathbf{n})} Y_{l'm'}(\mathbf{n}) d\sigma(\mathbf{n}) = \delta_{ll'} \delta_{mm'}$$

and $\forall l=0, \dots, \infty \forall m=-l, \dots, l : (f, Y_{lm})_{\mathbb{L}_2(S^2)} = 0 \Rightarrow f = 0$.

with Y_{lm} eigenvalues of the negative semi-definite, fractional Laplacian operator

$$-|\Delta_{S^2}|^\alpha Y_{lm} = -l^\alpha(l+1)^\alpha Y_{lm}$$

and thereby the solution of (2) is given by

$$W(\mathbf{x}, \mathbf{n}(\theta, \phi), t) = \sum_{l=0}^\infty \sum_{m=-l}^l (Y_{lm}, U(\mathbf{x}, \cdot))_{\mathbb{L}_2(S^2)} e^{-t|\Delta_{S^2}|^\alpha} Y_{lm}(\theta, \phi)$$

$$= \sum_{l=0}^\infty \sum_{m=-l}^l (Y_{lm}, U(\mathbf{x}, \cdot))_{\mathbb{L}_2(S^2)} e^{-t(l(l+1))^\alpha} Y_{lm}(\theta, \phi). \tag{3}$$

In particular for $\alpha = \frac{1}{2}$ (Poisson scale space [8,9]) the solution of (2) is nearly equivalent to outward harmonic extension of the initial distribution $\mathbf{n} \mapsto U(\mathbf{n})$ on the 2-sphere, where one must set the radius $\rho = e^{-t}$, since

$$\rho \mathbf{n}(\theta, \phi) = (\rho \cos \phi \sin \theta, \rho \sin \phi \sin \theta, \rho \cos \theta) \mapsto \rho^l Y_{lm}(\theta, \phi)$$

is a harmonic function on \mathbb{R}^3 and $\rho^l = e^{-tl} \approx e^{-t\sqrt{(l+1)l}}$ for l sufficiently large. So intuitively one may consider a radial scale axis for scale spaces on glyphs. If we expand $U(\mathbf{x}_1, \cdot)$ and $U(\mathbf{x}_2, \cdot)$ into the orthonormal basis of spherical harmonics:

$$U(\mathbf{x}_i, \mathbf{n}(\theta, \phi)) = \sum_{l=0}^\infty \sum_{m=-l}^l c_i^{lm} Y_{lm}(\mathbf{n}(\theta, \phi)), \quad i = 1, 2,$$

with $Y_l^m(\phi, \theta) = \sqrt{\frac{(2l+1)(l-|m|)!}{4\pi(l+|m|)!}} P_l^m(\cos \theta) e^{im\phi}$,

with P_l^m the associated Legendre polynomial of degree l and order m , and with $c_i^{lm} = (Y_{lm}, U(\mathbf{x}_i, \cdot))_{\mathbb{L}_2(S^2)}$, $i = 1, 2$, the squared Sobolev distance (1) reads

$$(d_{\alpha, \gamma, t}(U(\mathbf{x}_1, \cdot), U(\mathbf{x}_2, \cdot)))^2 = \sum_{l=0}^\infty \sum_{m=-l}^l |c_1^{lm} - c_2^{lm}|^2 \tilde{m}_l^{\alpha, \gamma, t}.$$

This is a weighted ℓ_2 inner product on the coefficients, so we only have to study the multiplier

$$\tilde{m}_l^{\alpha, \gamma, t} = (1 + (\gamma l(l+1))^{2\alpha}) e^{-2t(l(l+1))^\alpha}$$

to see how the different spherical harmonics are weighted in the Sobolev distance. Note that if $\gamma = t = 0$, the Sobolev norm is equivalent to the \mathbb{L}_2 norm.

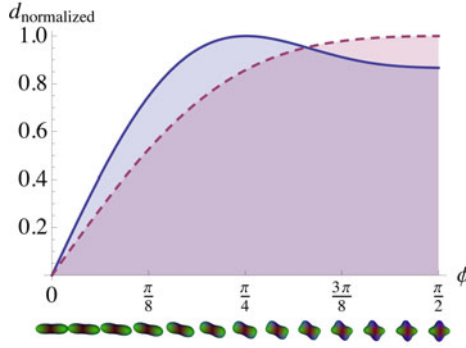


Fig. 1. Behavior of the Sobolev (blue, solid line) and the \mathbb{L}_2 norm (red, dashed line) for a combination of two single fiber profiles generated with angle ϕ in between

Analytical Example. To qualify the Sobolev and \mathbb{L}_2 norm, we investigated several diffusion profiles. The basis for each profile is a single fiber distribution

$$U(\mathbf{n}) = (\mathbf{n}^T M \mathbf{n})^4,$$

with M the diagonal diffusion matrix with $\{1, \epsilon, \epsilon\}$ as elements and $\mathbf{n} \in S^2$ describing the tessellation orientations. The fourth power is taken to sharpen the result. To this static part, a second single fiber profile is added, that is rotated over angle ϕ with respect to the first profile. The total profile then amounts to

$$U(\mathbf{n}) = (\mathbf{n}^T M \mathbf{n})^4 + (\mathbf{n}^T R_\phi M R_\phi^T \mathbf{n})^4,$$

with R_ϕ a transformation matrix resulting in a clockwise rotation over angle ϕ . We compare default diffusion profile f ($\phi = 0, f = U(\mathbf{x}_1, \cdot)$) with a set of profiles g with ϕ varying between 0 and $\pi/2$ ($g = U(\mathbf{x}_2, \cdot)$). Figure 1 shows the results for both the Sobolev and \mathbb{L}_2 norm, normalized by their maximum response:

$$d_{\text{normalized}}(f, g) = \frac{\|f - g\|}{\max_g \|f - g\|}.$$

It can be seen that for $\phi = \pi/4$, the diffusion profile has maxima that do not overlap with those at $\phi = 0$, while two of the four maxima at $\phi = \pi/2$ do coincide. The Sobolev norm clearly punishes the deviations in extrema. For small differences in angle ϕ , the Sobolev norm gives a larger response, but in the case of partly overlapping extrema, towards $\phi = \pi/2$, the Sobolev norm is again lower than its \mathbb{L}_2 counterpart. The latter yields its maximum at $\phi = \pi/2$, because the amplitudes of the diffusion profiles differ most at that point.

Implementation. In practice, discretization of the continuous spherical harmonics is done by the pseudo-inverse of the inverse spherical harmonic transform (DISHT) [10]. Similar to Descoteaux [4], we use only even orders of spherical harmonics to represent our HARDI data, i.e. order $l = 0, 2, 4, \dots, L_{max}$. As m still has range $-l, \dots, 0, \dots, l$, the total number of SH coefficients is defined as

$n_{SH} = \frac{1}{2}(L_{max} + 1)(L_{max} + 2)$. We define a single index j in terms of l and m such that $j(l, m) = (l^2 + l + 2)/2 + m$ and compute the spherical harmonic coefficients $\mathbf{s} \in \mathbb{C}^{n_{SH}}$ from the values $\mathbf{f} \in (\mathbb{R}^+)^{N_o}$ by means of

$$\mathbf{s} = (DISHT)^+[\mathbf{f}] = (\overline{\mathbf{M}}\mathbf{M}^T)^{-1}\overline{\mathbf{M}}\mathbf{f},$$

with $\mathbf{M} = [M_k^j] = [\frac{1}{\sqrt{C}}Y_{m(j)}^{l(j)}(\mathbf{n}_k)]$ and $C = \sum_{j=1}^{n_{SH}} |Y_{m(j)}^{l(j)}(0, 0)|^2$, such that $\mathbf{M}^\dagger\mathbf{M}$ has a diagonal of ones. If $\mathbf{f}_i \in (\mathbb{R}^+)^{N_o}$ denotes the discrete data on $U(\mathbf{x}_i, \cdot)$, $i = 1, 2$, i.e. $(\mathbf{f}_i)_k = U(\mathbf{x}_i, \mathbf{n}_k)$, $k = 1, \dots, N_o$ then the discrete Sobolev norm is

$$d_{\alpha, \gamma, t}^{\text{discrete}}(\mathbf{f}_1, \mathbf{f}_2) = \sqrt{\sum_{j=1}^{n_{SH}} |(DISHT)^+[\mathbf{f}_2][j] - (DISHT)^+[\mathbf{f}_1][j]|^2 \tilde{m}_{l[j]}^{\alpha, \gamma, t}}.$$

3 Phantom and Real Brain Data

We computer generated a HARDI phantom, employing a multitensor model [11], with 121 different gradient directions and b-value 3000 s/mm^2 . The phantom consists of 18 columns with different diffusion profiles, namely single fibers at angles $\{0, 1, 3, 6, 10, 15, 21, 28, 36, 45\}$ degrees, two fibers crossing in-plane at angles $\{40, 45, 55, 70, 90\}$ degrees, and three fibers crossing in-plane at angles $\{30, 40, 60\}$ degrees. Each column contains the original profile, followed by ten times this profile with Rician noise added, by applying the transformation

$$((\mathbf{y}, \mathbf{n}) \mapsto U(\mathbf{y}, \mathbf{n})) \mapsto ((\mathbf{y}, \mathbf{n}) \mapsto \sqrt{(U(\mathbf{y}, \mathbf{n}) \cos \eta_1 + \eta_2)^2 + (U(\mathbf{y}, \mathbf{n}) \sin \eta_1 + \eta_3)^2}),$$

where $\eta_2, \eta_3 \sim \mathcal{N}(0, \sigma)$ normally distributed and η_1 uniformly distributed over $[0, 2\pi)$. To obtain data with a realistic SNR of 30, σ was chosen to be 0.1. Using the Funk-Radon transform [4], the simulated signal was converted to spherical harmonic coefficients, with $L_{max} = 12$. Comparing the ℓ_2 norm of the SH coefficients, the noise was shown to lead to a disturbance of approximately 10%.

We also tested our algorithm on real data, analogous to for example Grassi et al. [3]. Human brain diffusion MRI data were acquired at 3T, measuring 52 slices of 128×128 2-mm isotropic voxels, using TE 85 ms, 128 unique gradient directions and b-value 2000 s/mm^2 . The data were registered onto the MNI152 template and the accompanying Talairach atlas, using an affine transformation within FSL [12]. Subsequently, an ROI of 30 pixels wide and 39 pixels high, containing the subject’s right thalamus, was selected on an axial slice. The data were again transformed to spherical harmonic coefficients ($L_{max} = 12$).

4 Parameter Tuning

The Sobolev norm contains three parameters – α, γ, t – that needed tuning, in order to yield the desired behavior. Equivalent to Lindeberg [7], α was set to 1, the value for Gaussian regularization. We chose t to be 0, as our data were sufficiently smooth. In addition, the \mathbb{L}_2 norm lacks regularization, so setting $t = 0$ enables a fair comparison. To assess the optimal value for γ , we performed

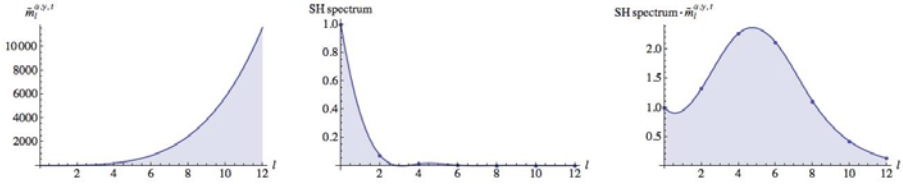


Fig. 2. Left: Behavior of the Sobolev multiplier $\tilde{m}_l^{\alpha,\gamma,t}$ for $\{\alpha = 1, \gamma = 0.69, t = 0\}$ and l from 0 to 12. **Middle:** Normalized mean spectrum of spherical harmonic coefficients of the phantom diffusion profiles. **Right:** Product of multiplier and spectrum.

a 1-nearest-neighbor classification, using the 18 original profiles of the phantom (i.e. without noise) as training and the whole phantom as test set, while γ was varied from 0 to 0.8 in steps of 0.01. The performance of the \mathbb{L}_2 norm for this phantom is 71.2%. However, using the Sobolev norm with $\gamma \geq 0.69$, the phantom classification reached a performance of 100%.

The behavior of the multiplier $\tilde{m}_l^{\alpha,\gamma,t}$ for $\{\alpha = 1, \gamma = 0.69, t = 0\}$ is plotted in Fig. 2 (left). It can be seen that the function exponentially rises towards higher values of l . However, the spherical harmonic coefficients of the used diffusion profiles at higher orders are quite small, as can be seen in Fig. 2 (middle). This means that the product of the multiplier and the spectrum is damped well enough to be truncated at $l = 12$ and avoid Gibbs artifacts (see Fig. 2 (right)).

5 K-Means Clustering

Following the example set by Wiegell et al. [1], we performed k-means clustering. For both the synthetic phantom and the real brain data, a set of seed points to serve as initial centroids was determined manually. The seed point placement for the phantom was straightforward, with a seed point in the middle of each column. The real brain ROI was masked using the atlas’ thalamus segmentation, while

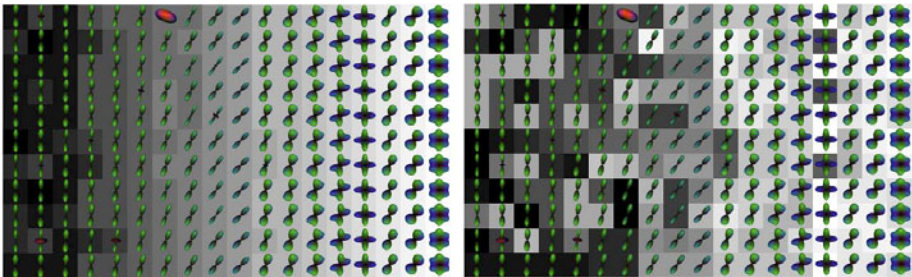


Fig. 3. Left: Result of \mathbb{L}_2 norm (20.7% correctly classified). **Right:** Result of Sobolev norm with $\{\alpha = 1, \gamma = 0.69, t = 0\}$ (73.7 % correctly classified). The diffusion profiles are visualized as min-max normalized Q-ball glyphs using SH coefficients up to $l = 8$.

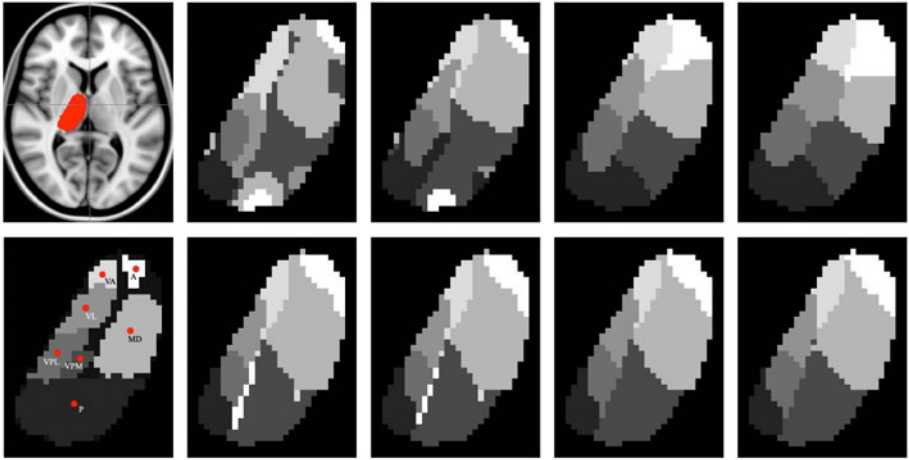


Fig. 4. Leftmost column: Thalamus ROI (top) and nuclei atlas labels (bottom). **Other columns:** Results using \mathbb{L}_2 (top) and Sobolev norm (bottom, $\{\alpha = 1, \gamma = 0.69, t = 0\}$). The ratio between diffusion and spatial information was varied, from left to right: ratio 1:1; $\frac{1}{5}$ -weight factor (as defined in [11]); 1-weight factor; 5-weight factor.

the seven seed points for this ROI were chosen to lie in the different nuclei of the thalamus, as defined by the atlas (see Fig. 4 (bottom left)). To associate each point of the data sets with a cluster, we used both the \mathbb{L}_2 norm and the Sobolev norm with the parameters calculated in Section 4 $\{\alpha = 1, \gamma = 0.69, t = 0\}$. The new centroids of each cluster were calculated as the mean of the voxel positions of all connected points. However, as our dissimilarity measures only concerned diffusion information, we calculated the distance between each data point and the mean diffusion profile of all points associated with each cluster.

The results of the k-means clustering for the phantom can be seen in Fig. 3. The \mathbb{L}_2 norm yields only a 20.7% correct classification (left). Clearly, the Sobolev norm has performed much better, obtaining 73.7% correctly classified diffusion profiles, as shown on the right-hand side. With respect to the thalamus ROI, in Fig. 4 we can see that the k-means clustering succeeds to classify the thalamus nuclei reasonably well. The pulvinar and ventral posterior medial nucleus are segmented but cannot be separated. The results of the Sobolev norm seem more stable, i.e. less dependent on the weight factor between diffusion and spatial information, as defined by Wiegell et al. [11], than the results of the \mathbb{L}_2 norm.

6 Discussion

In this paper, we introduced a new dissimilarity measure that can be employed for clustering of HARDI data. Instead of comparing only the amplitudes of the diffusion profiles, our Sobolev norm also takes into account whether the extrema of the profiles coincide. We illustrated the behavior of our norm for some simple synthetic glyphs and then built a more difficult phantom. The optimal

parameters for our norm were found using 1-nearest-neighbor clustering of this phantom, and these parameters were used for the subsequent k-means clustering. The Sobolev norm consistently performed better than the \mathbb{L}_2 norm, for both the phantom and the real brain data.

In order to improve upon this work, a clustering algorithm that is not biased by the number and placement of seed points could be employed, for example spectral clustering involving graph cuts [13]. It will remain a challenge to interpret the results of the real data, due to the absence of a proper ground truth.

Acknowledgments

Ellen Brunenberg is supported by the Netherlands Organization for Scientific Research (NWO). The authors would like to thank Bart Spronck, Vesna Prčkovska, and Markus van Almsick for their help.

References

1. Wiegell, M.R., Tuch, D.S., Larsson, H.B.W., Wedeen, V.J.: Automatic segmentation of thalamic nuclei from diffusion tensor magnetic resonance imaging. *Neuroimage* 19(2), 391–401 (2003)
2. Ziyang, U., Tuch, D., Westin, C.F.: Segmentation of thalamic nuclei from DTI using spectral clustering. In: Larsen, R., Nielsen, M., Sporrang, J. (eds.) MICCAI 2006. LNCS, vol. 4191, pp. 807–814. Springer, Heidelberg (2006)
3. Grassi, A., Cammoun, L., Pollo, C., Hagmann, P., Meuli, R., Thiran, J.P.: Thalamic nuclei clustering on high angular resolution diffusion images. *ISMRM 16*, 1777 (2008)
4. Descoteaux, M.: High Angular Resolution Diffusion MRI: from Local Estimation to Segmentation and Tractography. PhD thesis, INRIA Sophia Antipolis (2008)
5. McGraw, T., Vemuri, B., Yezierski, R., Mareci, T.: Segmentation of high angular resolution diffusion MRI modeled as a field of von Mises-Fisher mixtures. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3953, pp. 463–475. Springer, Heidelberg (2006)
6. Goh, A., Lenglet, C., Thompson, P.M., Vidal, R.: A nonparametric Riemannian framework for processing high angular diffusion images (HARDI). In: CVPR, pp. 2496–2503 (2009)
7. Lindeberg, T.: Scale-space behaviour of local extrema and blobs. *J. Math. Imaging Vision* 1(1), 65–99 (1992)
8. Duits, R., Felsberg, M., Florack, L.: α scale spaces on a bounded domain. In: Scale Space Conference, pp. 494–510 (June 2003)
9. Felsberg, M., Duits, R., Florack, L.: The monogenic scale space on a bounded domain and its applications. *Int. J. Comput. Vision* 2-3(64), 187–201 (2005)
10. Duits, R., Franken, E.: Left-invariant diffusions on the space of positions and orientations and their application to crossing-preserving smoothing of hardi images. *Int. J. Comput. Vision* (2010), <http://dx.doi.org/10.1007/s11263--010--0332--z>
11. Söderman, O., Jönsson, B.: Restricted diffusion in cylindrical geometry. *J. Magn. Reson. Series A* 117(1), 94–97 (1995)
12. Jenkinson, M., Smith, S.: A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* 5(2), 143–156 (2001)
13. Brunenberg, E., Pelgrim, E., ter Haar Romeny, B., Platel, B.: k-means and graph cuts clustering of diffusion MRI in rat STN. *ISMRM 18*, 4045 (2010)

MicroTrack: An Algorithm for Concurrent Projectome and Microstructure Estimation*

Anthony J. Sherbondy¹, Matthew C. Rowe², and Daniel C. Alexander²

¹ Psychology Department, Stanford University, USA
anthony.sherbondy@stanford.edu

² Computer Science Department, University College of London, UK

Abstract. This paper presents MicroTrack, an algorithm that combines global tractography and direct microstructure estimation using diffusion-weighted imaging data. Previous work recovers connectivity via tractography independently from estimating microstructure features, such as axon diameter distribution and density. However, the two estimates have great potential to inform one another given the common assumption that microstructural features remain consistent along fibers. Here we provide a preliminary examination of this hypothesis. We adapt a global tractography algorithm to associate axon diameter with each putative pathway and optimize both the set of pathways and their microstructural parameters to find the best fit of this holistic white-matter model to the MRI data. We demonstrate in simulation that, with a multi-shell HARDI acquisition, this approach not only improves estimates of microstructural parameters over voxel-by-voxel estimation, but provides a solution to long standing problems in tractography. In particular, a simple experiment demonstrates the resolution of the well known ambiguity between crossing and kissing fibers. The results strongly motivate further development of this kind of algorithm for brain connectivity mapping.

1 Introduction

Diffusion-weighted imaging (DWI) has been used to provide useful markers of brain tissue microstructure in a range of applications including stroke, normal development, aging, and highlighting diseased tissue. Most work uses simple indices, e.g., fractional anisotropy (FA), derived from the diffusion tensor model [7]. A major limitation of these simple indices is that they conflate multiple tissue properties. For example, FA correlates with axon density, axon radius distribution, coherence of axon orientations, myelin volume and permeability.

Other previous work [16, 5, 4, 2, 3] demonstrates the feasibility of measuring specific properties of the tissue directly. The techniques use more sophisticated

* AJS sponsored by Defense Advanced Research Projects Agency, Defense Sciences Office (DSO), Program: Systems of Neuromorphic Adaptive Plastic Scalable Electronics (SyNAPSE), Issued by DARPA/CMO under Contract No. HR0011-09-C-0002; DCA by EPSRC EP/E007748 and the EU CONNECT consortium www.brain-connect.eu

models that relate tissue properties like axon density and diameter to the diffusion-weighted MR signal. Early attempts of fitting the model in each voxel [16,4] have prohibitive data requirements preventing estimation of these parameters on live subjects within a reasonable imaging time. Recently, however, Barazany et al [6] successfully estimate the axon diameter distribution in the brain of a live rat with about 2h of imaging time. Moreover, Alexander et al. [2] demonstrate the potential for orientationally invariant estimation of axon density and diameter, even on human imaging systems with around 1h acquisition time, through a combination of simplification of the model and careful experiment design. The most recent work [3] shows promising results on data from both fixed monkey brains and live human brains. Orientational invariance [2,3] offers significant benefits over earlier work [16,4,6], which requires knowledge of the fiber orientation and thus limits parameter maps to structures with that specific orientation. Fiber orientation varies widely in the brain, so orientational invariance is essential for whole brain mapping and enables combination with tractography to study tractwise variation in microstructure. However, the maps of axon diameter and density in [3] are noisy, particularly for live subjects where hardware and acquisition time limitations weaken sensitivity.

Here, we introduce the MicroTrack algorithm that leverages the assumption that microstructure features, e.g., axon radius, myelin volume, and longitudinal diffusivity, remain constant along white matter fascicles. For example, one may estimate an axon diameter distribution that is fixed along the length of a tract such as the optic radiation. The assumption that the axon diameter distribution remains constant along the length of individual fascicles is common in histological studies that characterize tract properties using single cross-section measurements [1] and has been measured to be constant over a range of at least 1mm [14]. This model provides a powerful geometric constraint that allows us to collapse measurements across multiple voxels to estimate microstructure parameters shared along a single fascicle and potentially improve parameter estimation dramatically.

MicroTrack simultaneously solves for the projectome, i.e., the collection of long-range white matter fascicles projecting between all gray matter locations in the brain [11], and the microstructure tissue parameters of the fascicles within the projectome by fitting a holistic model of the white matter to all the DWI intensities. The white matter model is hierarchical in the sense that it describes both macroscopic features, i.e., pathways joining remote locations, and microscopic features, e.g., the density and diameter of axons in each pathway. This concurrent multiscale modeling is a fundamental augmentation of previous global tractography techniques [12,15,10] and we hypothesize that it offers major benefits over independent estimation of either quantity. First, as mentioned above, pooling data from voxels with similar microstructure improves estimates of axon diameter and density. However, equally important is the potential reduction in false positive connections in the projectome. Specifically, knowledge of the microstructure of individual fascicles resolves ambiguities that occur at, for example, fiber crossing and kissing configurations. At a kissing or crossing, where two

or more paths are equally likely from their local orientation, the algorithm can pick the path with microstructure most consistent to the rest of the path and reject paths along which significant changes in microstructure occur. This approach offers the first viable and biologically motivated solution to tractography ambiguities like crossing versus kissing fibers.

Section 2 outlines the white matter macro and microstructure models and introduces MicroTrack: the algorithm that simultaneously estimates the projectome and fascicle microstructure properties. Section 3 presents simulation experiments that confirm our central hypotheses of improved microstructure parameter estimation and reduction of false positive projections. We conclude in section 4 with a discussion of limitations and areas for further work.

2 Method

This section outlines the MicroTrack algorithm. We start with the forward model of white matter as a set of fascicles linking remote locations each with microstructure that determines the diffusion MR signal from the fascicles in individual voxels. Next, we describe the algorithm that solves the inverse problem of fitting the model to the DWI data. The section goes on to outline the cost function and optimization procedure that enable this fitting.

White Matter Tissue Model. The white matter model captures geometrical architecture of tissue across multiple length scales. Currently, our model considers two length scales: axon bundles or fascicles, with cross-sectional radius on the order of hundreds of microns, and axons, with cross-sectional radius on the order of a micron. The fascicles are simple tubular structures that are defined by a sequence of backbone points that extend between distant gray matter locations and a single cross-sectional radius (R) for the entire fascicle. At the sub-fascicle scale, the composition of the fascicle structure is characterized by several parameters: local orientation \mathbf{n} , average axon radius r , fraction of fascicle cross-section filled with axons f , intra-axonal diffusivity d , and extra-axonal diffusivity, which we assume is also d , but that tortuosity around the axons reduces the perpendicular apparent diffusivity to d_p . Orientation is the only parameter that varies along the fascicle and all other microstructural parameters are fixed within each.

The tissue model predicts MR attenuation within a voxel composed of a mixture of fascicles and non-fascicle volume by dividing each voxel into compartments that contain only one tissue type in a similar way to Close et al. [8]. Specifically, we divide the voxel grid into a much higher resolution subvoxel grid, typically 3x3x3 subvoxels per image voxel. Any subvoxels within fascicles are considered fascicle compartments and the rest contribute cerebral spinal fluid (CSF). Occasionally, fascicles overlap on a subvoxel; the closest fascicle determines the microstructure and orientation of these subvoxels' tissue.

Signal Model. The simplified CHARMED model in [23] provides the signal in subvoxels within fascicles. In this model, the normalized MR signal is

$$A(\mathbf{G}, \Delta, \delta; \mathbf{n}, r, f, d, d_p) = f A_r(\mathbf{G}, \Delta, \delta; \mathbf{n}, r, f, d) + (1 - f) A_h(\mathbf{G}, \Delta, \delta; \mathbf{n}, d, d_p), \quad (1)$$

where f is the volume fraction of the restricted water inside axons, which has normalized signal A_r , and A_h is the signal from the hindered water outside but amongst the axons; \mathbf{G} is the gradient vector, δ is the width of the gradient pulses in the pulsed-gradient spin-echo sequence, and Δ is their separation. The intracellular signal, A_r , is the product of attenuated signals from the components of \mathbf{G} parallel and perpendicular to the fiber direction [5]. We assume Gaussian displacements with diffusivity d parallel to the fiber and use the Gaussian Phase Distribution approximation [13] of the signal from spins with diffusivity d restricted within a cylinder of radius r [18] for the perpendicular signal. The signal, A_h , from the hindered compartment comes from the diffusion tensor model with principal eigenvector \mathbf{n} and eigenvalue d and minor eigenvalues d_p . The signal model for CSF subvoxels assumes isotropic Gaussian displacements with diffusivity d_{iso} . The final set of signals for each whole voxel is the average of each corresponding measurement over all the subvoxels within.

Algorithm Overview. The MicroTrack algorithm follows the general optimization procedure of a recent projectome identification algorithm [15]. MicroTrack takes as input a collection of N_C fascicle candidates (C) and outputs a vector of fascicle radii ($\bar{R} = [R_1, R_2, R_3, \dots, R_{N_C}]$), but it also outputs a vector of microstructure parameters ($\bar{r} = [r_1, r_2, r_3, \dots, r_{N_C}]$). The algorithm searches over configurations of \bar{r} and \bar{a} to minimize the error function $E(\bar{R}, \bar{r})$ that balances fitting the MRI data with volumetric constraints that penalize fascicle overlap. Similar to the previous projectome identification algorithm, \bar{R} is binarized so that each fascicle may either be included in the optimal projectome, $R_i > 0$, or excluded, $R_i = 0$. Unlike the previous algorithm, MicroTrack also searches over a discrete set of r_i values that best characterize the fascicle candidate.

Cost Function. MicroTrack searches over configurations of \bar{R} and \bar{r} in order to select the setting that minimizes the global error, $E = \lambda E_1 - (1 - \lambda) E_2$. E_1 measures the difference between the predicted and observed MRI signal attenuation, and, E_2 measures the amount of overlap between the fascicle structures. The parameter λ balances the effects of E_1 and E_2 .

The data fitting error, E_1 , is a straightforward square error metric summed over the number of voxels (V) and number of MR measurements per voxel (K),

$$E_1 = \sum_{v=1}^V \sum_{k=1}^K \frac{(\hat{A}_{k,v} - A_{k,v})^2}{\sigma^2}, \quad (2)$$

where $A_{k,v}$ is the predicted DWI signal, $\hat{A}_{k,v}$ is actual MRI and σ is the standard deviation of the measurements, which is fixed for the entire volume.

The fascicle volume overlap error, E_2 , simply reports the volume of fascicles that overlap, i.e., violate physical integrity of fascicles by occupying the same space. The units of E_2 are in fractions of the voxel volume. For example, if

a subvoxel is overlapped by two fascicles than a penalty of $1/N_s$, where N_s is the number of subvoxels, is added to E_2 for that voxel and if three fascicles overlap the subvoxel the penalty is $2/N_s$ and so on.

Fascicle Candidates. MicroTrack precomputes a large set, C , of candidate fascicles using either individual or combinations of standard local deterministic and probabilistic tracking algorithms (Camino, FSL, DTIStudio, MRTRIX, Track-Vis). For this work, the candidate set was created using the Camino software [9], which offers both deterministic and probabilistic tractography with HARDI measurement schemes. Other than the endpoint criteria, all default parameter values are used and seed points are randomly selected within the entire volume.

Optimization Strategy. The search space of all possible white matter configurations is very high-dimensional and the cost function has many local minima. Here we use a stochastic optimization algorithm, differential evolution (DE) [17], which we have found to outperform a variety of other off the shelf global optimization algorithms. DE fits in the class of genetic evolution algorithms that search over a parameter space by evolving individuals (parameter settings) amongst a community (several parameter settings) based on the individual's fitness (error function). The stochastic/genetic search is advantageous for our problem as multiple individuals may fit disparate parts of the white matter volume well and combine to provide the benefits of both. The complexity of the problem prevents guaranteed convergence on the global minimum, in practice we find that the following parameter settings: (generations = 1600, populations = 40, CR = 0.9, F = [0.5,0.8], resets = 8, $r_i \in [1, 5, 10]\mu m$, $R_i \in [0, 0.2]mm$) produce good, if suboptimal, solutions in a reasonable time (2 hours on 8 2.3 GHz CPUs).

3 Results

The intention of the experiments is to test the core hypothesis that simultaneous estimation of the projectome and microstructure is advantageous over estimation of either independently. We thus construct several simple synthetic data sets specifically to enable such a test. The synthetic data was produced using a modification of the Numerical Fiber Generator (NFG) software [8]. The NFG software offers an efficient way of producing crossing structures with complex intersections. For the synthetic data used here, we replace the tensor model in the original NFG software with the two compartment model in equation [1]. Each strand of the strand structure is assigned an individual axon radius, in our examples 1, 5 or 10 μm . The other parameters have fixed values of $d=1.7\times 10^{-9} m^2/s$, $d_p=1.2\times 10^{-9} m^2/s$, $d_{iso}=3.0\times 10^{-9} m^2/s$, $f = 0.7$. The system adds Rician noise so that the signal to noise ratio at $b = 0$ is 20.

The imaging protocol we simulate is a multi-shell HARDI acquisition obtained from the experiment-design optimization in [2], which optimizes sensitivity to axon diameter. The three shells have $b = 2242(|\mathbf{G}|=200 mT/m, \Delta=14.5 ms, \delta=8 ms)$, 3791 ($|\mathbf{G}|=121 mT/m, \Delta=23.4 ms, \delta=13.8 ms$), and 11329 s/mm^2

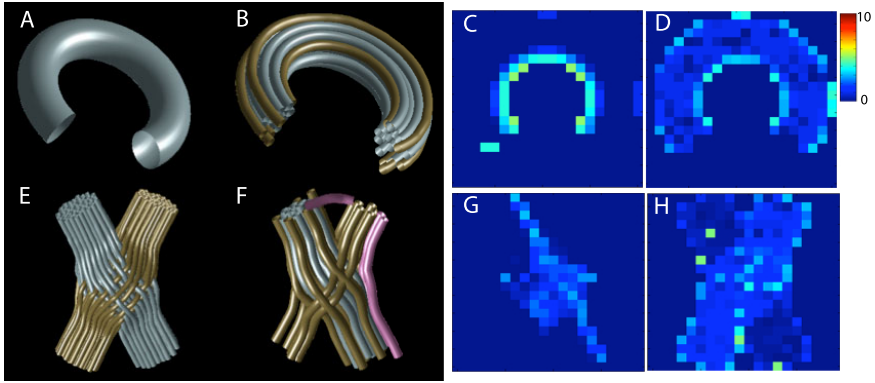


Fig. 1. MicroTrack improves microstructure estimation accuracy. Torus (A) and oblique crossing (E) data are simulated and reconstructed with MicroTrack (B,F). Error in mean axon radius estimation for a central slice (C,G) reduces error compared with the estimation using a typical voxel-by-voxel algorithm (D,H). Fibers are labeled based on mean axon radius as either 1 (blue), 5 (brown), or 10 μm (purple). Pink fibers represent false positive connections. Error bar is in units of μm .

($|\mathbf{G}|=200 \text{ mT/m}$, $\Delta=21.8 \text{ ms}$, $\delta=15.3 \text{ ms}$), and 100, 105 and 84 gradient directions, respectively. The protocol also includes 65 $b = 0$ measurements for a total of 364 images.

Improving Microstructure Estimation. We compare MicroTrack estimates of white matter fraction and average axon radius within each voxel to estimates from the grid search and maximum likelihood fit steps of the fitting algorithm in [3] a simple gradient-descent fit of the tissue parameters independently for each voxel. The precise model we fit is as in equation [1], but with an extra CSF compartment using the same CSF signal model as the simulation. The comparison is carried out on two synthetic data sets: cut torus (Figure [1]A) and oblique crossing (Figure [1]E). The MicroTrack estimates reduce the error over most voxels in both synthetic experiments and, in particular, are more accurate in the voxels where the fiber orientation distributions are multi-modal.

Reducing Tracking Ambiguities. In order to focus in on the hypothesis that concurrent microstructure and projectome estimation can improve the connectivity results over standard tracking techniques, we created two synthetic volumes that demonstrate the well-known kissing versus crossing ambiguity (Figure [2]A and D). The synthetic volume was carefully constructed so that the white matter in the oblique crossing configuration overlaps the white matter volume in the kissing configuration as much as possible. The result is that many voxels within the two volumes contain nearly identical fiber orientation distributions, despite a fundamental difference in connectivity structure. To further focus the experiment, the fascicle candidate set was derived from the gold standard fascicles used to create the synthetic data. This ensures that the candidate fascicle set,

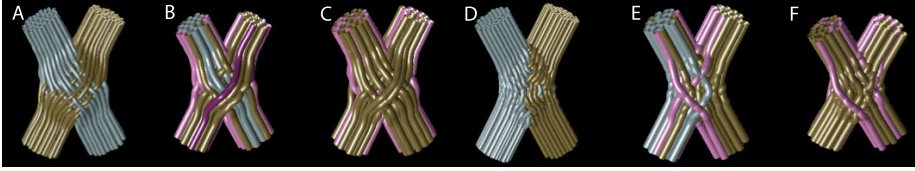


Fig. 2. MicroTrack addresses orientation ambiguity. Classic crossing (A-C) and kissing (D-F) data are simulated (A,D), reconstructed with MicroTrack (B,E) and reconstructed MicroTrack, but with concurrent microstructure estimation disabled (C,F). Fibers are labeled based on mean axon radius as either 1 (blue), 5 (brown), or 10 μm (purple). Pink fibers represent false positive connections.

C, contains equal numbers of candidate fibers that follow the crossing and kissing configurations, which maximizes the ambiguity for the algorithm to resolve. If we solve for the best fitting projectome without solving for microstructure parameters the false positive rate is $40 \pm 2\%$ (Figure 2C and F). When we simultaneously fit the microstructure parameters the false positive rate is reduced by nearly 3 fold to $14 \pm 1\%$ false positives (Figure 2B and E).

4 Conclusion

This paper demonstrates the first algorithm to constrain the solution of microstructural tissue properties with fascicle geometry estimates within an identified projectome. The algorithm is also the first to use microstructural tissue properties to determine the fascicle geometry. Despite the tremendous computational complexity of solving for parameters in this very high dimensional space, we demonstrate the feasibility of the algorithm. The simple simulation experiments demonstrate the advantages of employing such a detailed model of white matter tissue that not only incorporates global fascicular projection geometry, but also geometrical details of tissue microstructure. In particular, we demonstrate the potential to resolve crossing and kissing configurations. Further work will refine the model for use with real brain tissue and test the hypothesis on well-known anatomical crossings and kissings.

References

1. Aboitiz, F., Scheibel, A.B., Fisher, R.S., Zaidel, E.: Fiber composition of the human corpus callosum. *Brain Research* 598, 143–153 (1992)
2. Alexander, D.C.: A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features. *Magnetic Resonance in Medicine* 60, 439–448 (2008)
3. Alexander, D.C., Hubbard, P.L., Hall, M.G., Moore, E.A., Ptito, M., Parker, G.J.M., Dyrby, T.D.: Orientationally invariant indices of axon diameter and density from diffusion MRI. *NeuroImage* (2010) (in press)

4. Assaf, Y., Blumenfeld-Katzir, T., Yovel, Y., Basser, P.J.: AxCaliber: a method for measuring axon diameter distribution from diffusion MRI. *Magnetic Resonance in Medicine* 59, 1347–1354 (2008)
5. Assaf, Y., Freidlin, R.Z., Rohde, G.K., Basser, P.J.: New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magnetic Resonance in Medicine* 52, 965–978 (2004)
6. Barazany, D., Basser, P.J., Assaf, Y.: In-vivo measurement of the axon diameter distribution in the corpus callosum of a rat brain. *Brain* 132, 1210–1220 (2009)
7. Basser, P.J., Pierpaoli, C.: Microstructural and physiological features of tissues elucidated by quantitative diffusion tensor MRI. *Journal of Magnetic Resonance Series B* 111, 209–219 (1996)
8. Close, T.G., Tournier, J.-D., Calamante, F., Johnston, L.A., Mareels, I., Connelly, A.: A software tool to generate simulated white matter structures for the assessment of fibre-tracking algorithms. *NeuroImage* 47(4), 1288–1300 (2009)
9. Cook, P.A., Bai, Y., Nedjati-Gilani, S., Seunarine, K.K., Hall, M.G., Parker, G.J.M., Alexander, D.C.: Camino: Open-source diffusion-MRI reconstruction and processing. In: Proc. 14th Annual Meeting of the ISMRM, Berlin (2006)
10. Fillard, P., Poupon, C., Mangin, J.-F.: A novel global tractography algorithm based on an adaptive spin glass model. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5761, pp. 927–934. Springer, Heidelberg (2009)
11. Kasthuri, N., Lichtman, J.W.: The rise of the 'projectome'. *Nat. Methods* 4(4), 307–308 (2007)
12. Kreher, B.W., Mader, I., Kiselev, V.G.: Gibbs tracking: A novel approach for the reconstruction of neuronal pathways. *Magnetic Resonance in Medicine* 60, 953–963 (2008)
13. Murday, J.S., Cotts, R.M.: Self-diffusion coefficient of liquid lithium. *Journal of Chemical Physics* 48, 4938–4945 (1968)
14. Ong, H.H., Wright, A.C., Wehrli, S.L., Souza, A., Schwartz, E.D., Hwang, S.N., Wehrli, F.W.: Indirect measurement of regional axon diameter in excised mouse spinal cord with q-space imaging: Simulation and experimental studies. *NeuroImage* 40(4), 1619–1632 (2008)
15. Sherbondy, A.J., Dougherty, R.F., Ananthanarayanan, R., Modha, D.S., Wandell, B.A.: Think global, act local; projectome estimation with BlueMatter. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5761, pp. 861–868. Springer, Heidelberg (2009)
16. Stanisz, G.J., Szafer, A., Wright, G.A., Henkelman, M.: An analytical model of restricted diffusion in bovine optic nerve. *Magnetic Resonance in Medicine* 37, 103–111 (1997)
17. Storn, R., Price, K.: Differential evolution a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization* 11, 341–359 (1997)
18. Van Gelderen, P., DesPres, D., van Zijl, P.C.M., Moonen, C.T.W.: Evaluation of restricted diffusion in cylinders. Phosphocreatine in rabbit leg muscle. *Journal of Magnetic Resonance Series B* 103, 255–260 (1994)

Joint Generative Model for fMRI/DWI and Its Application to Population Studies

Archana Venkataraman¹, Yogesh Rathi², Marek Kubicki²,
Carl-Fredrik Westin^{3,1}, and Polina Golland¹

¹ MIT Computer Science and Artificial Intelligence Laboratory, Cambridge, MA

² Psychiatry Neuroimaging Laboratory, Harvard Medical School, Boston, MA

³ Laboratory for Mathematics Imaging, Harvard Medical School, Boston, MA

Abstract. We propose a novel probabilistic framework to merge information from DWI tractography and resting-state fMRI correlations. In particular, we model the interaction of *latent* anatomical and functional connectivity templates between brain regions and present an intuitive extension to population studies. We employ a mean-field approximation to fit the new model to the data. The resulting algorithm identifies differences in latent connectivity between the groups. We demonstrate our method on a study of normal controls and schizophrenia patients.

1 Introduction

The interaction between functional and anatomical connectivity provides a rich framework for understanding the brain. Functional connectivity is commonly measured via temporal correlations in resting-state fMRI data. These correlations are believed to reflect the intrinsic functional organization of the brain [1]. Anatomical connectivity is often measured using DWI tractography, which estimates the configuration of underlying white matter fibers [2]. In this work we propose and demonstrate a novel probabilistic framework to infer the relationship between these modalities. The model is based on *latent* connectivities between brain regions and makes intuitive assumptions about the data generation process. We present a natural extension of the model to population studies, which we use to identify widespread connectivity changes in schizophrenia.

To date, relatively little progress has been made in fusing information between the aforementioned anatomical and functional modalities. It has been shown that while a high degree of structural connectivity predicts higher functional correlations, the converse does not always hold [3,4]. For example, strong functional correlations can be found between spatially distributed locations in the brain. However, one is more likely to identify white matter tracts connecting nearby regions. Graph-theoretic models have previously been used to examine the correspondence between independently estimated structural hubs and functional networks [5,6]. In contrast, we infer a population template of connectivity using *both* resting-state fMRI correlations and DWI tractography.

We demonstrate the capability of our model to learn stable patterns on a population study of schizophrenia. Schizophrenia is a poorly-understood disorder

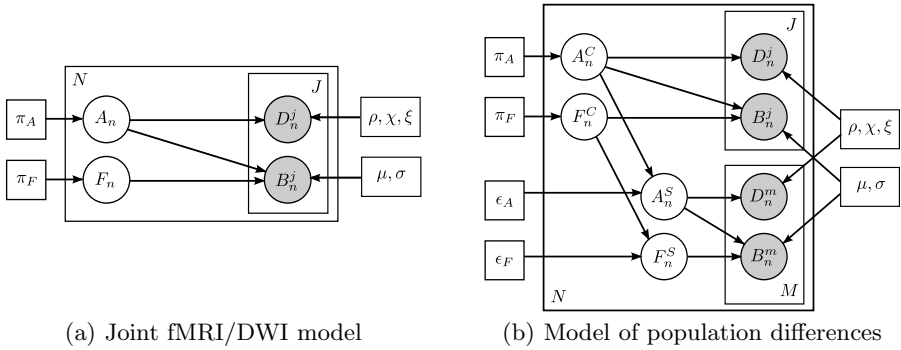


Fig. 1. (a) Joint connectivity model for a single population. (b) Joint model for the effects of schizophrenia. The pairwise connections are indexed by $n = 1, \dots, N$, the control subjects are indexed by $j = 1, \dots, J$, and the schizophrenia patients are indexed by $m = 1, \dots, M$. Squares indicate non-random parameters; circles indicate hidden random variables; all shaded variables are observed.

marked by widespread cognitive difficulties affecting intelligence, memory, and executive attention. These impairments are not localized to a particular cortical region; rather, they reflect abnormalities in widely-distributed functional and anatomical networks [748]. In accordance with these findings, our model identifies connectivity differences in spatially extensive networks.

Only a few studies to date have combined resting-state fMRI and DWI tractography to analyze schizophrenia [910]. Univariate statistical tests are commonly used to identify significant population differences in temporal correlations and in mean Fractional Anisotropy (FA) values. The relevant connections are then compared across modalities to draw conclusions. This approach treats functional and structural connections as *a priori* independent and ignores distributed patterns of connectivity. In contrast, our model jointly infers the entire pattern of functional and anatomical connectivity, as well as the group differences.

2 Generative Model and Inference

Unlike voxel- and ROI-based analysis, we model the behavior of *pairwise connections* between regions of the brain. Our observed variables are correlations in resting-state fMRI and average FA values along the white matter tracts.

Latent Connectivity. Fig. 1(a) shows our model for a single population. Let N be the total number of relevant connections in the brain. A_n and F_n are the latent anatomical and functional connectivity measures between the two regions associated with the n^{th} connection. A_n is a binary random variable with parameter π_A : $P(A_n; \pi_A) = \pi_A^{A_n} (1 - \pi_A)^{1 - A_n}$. It indicates the presence or absence of a *direct* anatomical pathway between the regions. In contrast, F_n

is a tri-state random variable drawn from a multinomial distribution π_F . These states represent little or no functional co-activation ($F_n = 0$), positive functional synchrony ($F_n = 1$), and negative functional synchrony ($F_n = -1$) between the regions. For notational convenience, we represent F_n as a length-three indicator vector with exactly one of its elements $\{F_{nk} : k = -1, 0, 1\}$ equal to one:

$$P(F_n; \pi_F) = \prod_{k=-1}^1 (\pi_{Fk})^{F_{nk}}. \quad (1)$$

Data Likelihood. Let J be the number of subjects. The DWI measurement D_n^j for the j^{th} subject is a noisy observation of the anatomical connectivity A_n :

$$P(D_n^j | A_n; \{\rho, \chi, \xi\}) = \left[\rho_0 \delta(D_n^j) + (1 - \rho_0) \mathcal{N}(D_n^j; \chi_0, \xi_0^2) \right]^{(1-A_n)} \cdot \left[\rho_1 \delta(D_n^j) + (1 - \rho_1) \mathcal{N}(D_n^j; \chi_1, \xi_1^2) \right]^{A_n}, \quad (2)$$

where $\delta(\cdot)$ is the Dirac delta function, $\mathcal{N}(\cdot; \chi, \xi^2)$ is a Gaussian distribution with mean χ and variance ξ^2 , and ρ is the probability of failing to find a tract between two regions. The value zero of D_n^j is arbitrarily chosen to represent ‘‘No Connection’’. D_n^j is strictly positive when a connection is present.

The BOLD fMRI correlation B_n^j for the j^{th} subject depends on both F_n and A_n since direct anatomical connections predict higher functional correlations:

$$P(B_n^j | A_n, F_n; \{\mu, \sigma\}) = \prod_{k=-1}^1 \left[\mathcal{N}(B_n^j; \mu_{0k}, \sigma_{0k}^2)^{(1-A_n)} \cdot \mathcal{N}(B_n^j; \mu_{1k}, \sigma_{1k}^2)^{A_n} \right]^{F_{nk}}. \quad (3)$$

Using histograms of the data, we verified that the Gaussian distributions in Eqs. (2)(3) provide reasonable approximations for the DWI and fMRI data. Pragmatically, they greatly simplify the learning/inference steps.

Population Differences. Fig. 1(b) depicts an extension of the model to a population study involving controls and schizophrenia patients. We model differences between the groups within the latent connectivities alone and share the data likelihood distributions between the two populations.

We treat the latent connectivity templates $\{A_n^S, F_n^S\}$ of the schizophrenia population as ‘‘corrupted’’ versions of the healthy template. In particular, with (small) probability ϵ , each connection can switch its state:

$$P(A_n^S | A_n^C; \epsilon_A) = \epsilon_A^{A_n^C(1-A_n^S) + (1-A_n^C)A_n^S} \cdot (1 - \epsilon_A)^{A_n^C A_n^S + (1-A_n^C)(1-A_n^S)}, \quad (4)$$

$$P(F_n^S | F_n^C; \epsilon_F) = \prod_{k=-1}^1 \frac{\epsilon_F}{2}^{(1-F_{nk}^C)F_{nk}^S} \cdot (1 - \epsilon_F)^{F_{nk}^C F_{nk}^S}. \quad (5)$$

For robustness, we rely on a single scalar to govern the probability of change within each modality. Additionally, in Eq. (5) we assume that functional connectivity switches to its other two states with equal probability.

Variational EM Solution. It is not difficult to show that the complete log-likelihood of all the random variables has multiplicative interactions among the hidden variables. For this reason, we employ the mean-field algorithm [11] to approximate the posterior probability distribution of the latent variables using a fully factorized distribution.

We let $\{\hat{p}_n^C, \hat{p}_n^S, \hat{q}_{nk}^C, \hat{q}_{nk}^S\}$ represent the posterior probability estimates for $\{A_n^C, A_n^S, F_n^C, F_n^S\}$. The variational EM algorithm alternates between updating the posterior estimates and the model parameters to minimize the variational free energy. Due to space constraints, we directly present the update rules.

Learning. We fix the posterior distributions and learn the model parameters. Let J_n^0 be the number of healthy subjects for which $D_n^j = 0$, and let M_n^0 be similarly defined for schizophrenia patients. The update rules are identical to those of the standard EM. The probability estimates are sums of the latent posteriors:

$$\begin{aligned} \pi_A &= \frac{1}{N} \sum_{n=1}^N \hat{p}_n^C & \pi_{Fk} &= \frac{1}{N} \sum_{n=1}^N \hat{q}_{nk}^C \\ \epsilon_A &= \frac{1}{N} \sum_{n=1}^N \hat{p}_n^C (1 - \hat{p}_n^S) + (1 - \hat{p}_n^C) \hat{p}_n^S & \epsilon_F &= 1 - \frac{1}{N} \sum_{n=1}^N \sum_{k=-1}^1 \hat{q}_{nk}^C \hat{q}_{nk}^S \end{aligned}$$

and the density parameters are equal to weighted statistics of the data:

$$\begin{aligned} \mu_{1k} &= \frac{\sum_{n=1}^N [\hat{p}_n^C \hat{q}_{nk}^C \sum_{j=1}^J B_{nk}^j + \hat{p}_n^S \hat{q}_{nk}^S \sum_{m=1}^M B_{nk}^m]}{\sum_{n=1}^N [\hat{p}_n^C \hat{q}_{nk}^C J + \hat{p}_n^S \hat{q}_{nk}^S M]} \\ \sigma_{1k}^2 &= \frac{\sum_{n=1}^N [\hat{p}_n^C \hat{q}_{nk}^C \sum_{j=1}^J (B_{nk}^j - \mu_{1k})^2 + \hat{p}_n^S \hat{q}_{nk}^S \sum_{m=1}^M (B_{nk}^m - \mu_{1k})^2]}{\sum_{n=1}^N [\hat{p}_n^C \hat{q}_{nk}^C J + \hat{p}_n^S \hat{q}_{nk}^S M]} \\ \rho_1 &= \frac{\sum_{n=1}^N [\hat{p}_n^C J_n^0 + \hat{p}_n^S M_n^0]}{\sum_{n=1}^N [\hat{p}_n^C J + \hat{p}_n^S M]} & \chi_1 &= \frac{\sum_{n=1}^N [\hat{p}_n^C \sum_{j: D_n^j > 0} D_n^j + \hat{p}_n^S \sum_{m: D_n^m > 0} D_n^m]}{\sum_{n=1}^N [\hat{p}_n^C (J - J_n^0) + \hat{p}_n^S (M - M_n^0)]} \\ \xi_1^2 &= \frac{\sum_{n=1}^N [\hat{p}_n^C \sum_{j: D_n^j > 0} (D_n^j - \chi_1)^2 + \hat{p}_n^S \sum_{m: D_n^m > 0} (D_n^m - \chi_1)^2]}{\sum_{n=1}^N [\hat{p}_n^C (J - J_n^0) + \hat{p}_n^S (M - M_n^0)]} \end{aligned}$$

The parameter updates for $\{\mu_{0k}, \sigma_{0k}^2, \rho_0, \chi_0, \xi_0^2\}$ are trivially obtained from these expressions by replacing \hat{p}_n^C with $(1 - \hat{p}_n^C)$ and \hat{p}_n^S with $(1 - \hat{p}_n^S)$.

Inference. We fix the model parameters and update the variational posteriors. We use $\mathcal{P}_i(\cdot)$ to denote the mixture distribution in Eq. (2) and $\mathcal{G}_{lk}(\cdot)$ to denote a normal distribution with parameters $\{\mu_{lk}, \sigma_{lk}\}$ in order to obtain:

$$\begin{aligned} \frac{\hat{p}_n^C}{1 - \hat{p}_n^C} &= \left(\frac{\pi_A}{1 - \pi_A} \right) \left(\frac{1 - \epsilon_A}{\epsilon_A} \right)^{(2\hat{p}_n^S - 1)} \prod_{j=1}^J \left(\frac{\mathcal{P}_1(D_n^j)}{\mathcal{P}_0(D_n^j)} \right) \prod_{k=-1}^1 \left(\frac{\mathcal{G}_{1k}(B_n^j)}{\mathcal{G}_{0k}(B_n^j)} \right)^{\hat{q}_{nk}^C} \\ \frac{\hat{p}_n^S}{1 - \hat{p}_n^S} &= \left(\frac{1 - \epsilon_A}{\epsilon_A} \right)^{(2\hat{p}_n^C - 1)} \prod_{m=1}^M \left(\frac{\mathcal{P}_1(D_n^m)}{\mathcal{P}_0(D_n^m)} \right) \prod_{k=-1}^1 \left(\frac{\mathcal{G}_{1k}(B_n^m)}{\mathcal{G}_{0k}(B_n^m)} \right)^{\hat{q}_{nk}^S} \end{aligned}$$

$$\hat{q}_{nk}^C \propto \pi_{Fk} \cdot \left(\frac{1 - \epsilon_F}{(\epsilon_F/2)} \right)^{\hat{q}_{nk}^S} \prod_{j=1}^J \left(\mathcal{G}_{1k}(B_n^j) \right)^{\hat{p}_n^C} \left(\mathcal{G}_{0k}(B_n^j) \right)^{(1 - \hat{p}_n^C)} \quad s.t. \quad \sum_{k=-1}^1 \hat{q}_{nk}^C = 1$$

$$\hat{q}_{nk}^S \propto \left(\frac{1 - \epsilon_F}{(\epsilon_F/2)} \right)^{\hat{q}_{nk}^C} \prod_{m=1}^M \left(\mathcal{G}_{1k}(B_n^m) \right)^{\hat{p}_n^S} \left(\mathcal{G}_{0k}(B_n^m) \right)^{(1 - \hat{p}_n^S)} \quad s.t. \quad \sum_{k=-1}^1 \hat{q}_{nk}^S = 1$$

As seen, the updates can be decomposed into a prior term (for normal subjects only), a term arising from the connectivity changes between the populations, and a data likelihood term involving the other modality.

Model Evaluation. Based on the latent posterior probabilities $\{\hat{p}_n^C, \hat{p}_n^S, \hat{q}_{nk}^C, \hat{q}_{nk}^S\}$, the empirical probability of change in the anatomical or functional connectivity of the n^{th} connection is

$$\hat{\epsilon}_A^n = \hat{p}_n^C(1 - \hat{p}_n^S) + (1 - \hat{p}_n^S)\hat{p}_n^S \quad \text{and} \quad \hat{\epsilon}_F^n = 1 - \sum_{k=-1}^1 \hat{q}_{nk}^C \hat{q}_{nk}^S \quad (6)$$

respectively. We evaluate the significance and robustness of our model through non-parametric permutation tests and cross-validation. To construct the distributions for $\hat{\epsilon}_A^n$ and $\hat{\epsilon}_F^n$ under the null hypothesis, we randomly permute the subject labels (NC vs. SZ) 10,000 times. For each permutation, we fit the model and compute the relevant statistics in Eq. (6). The significance (p-value) of each connection is equal to the proportion of permutations for which the computed statistic is greater than or equal to the value obtained under the true labeling.

We also quantify the model’s predictive power via ten-fold cross validation. The model is fit using the training subjects, and a likelihood ratio test is used to predict the diagnosis for the held-out group. The data is resampled 20 times to ensure stability of the results. For comparison, we perform the same ten-fold cross validation using support vector machine (SVM) classifiers trained on the fMRI correlations, the DWI FA values, and the combined fMRI and DWI data.

3 Results

Data. We demonstrate our model on a study of 18 male patients with chronic schizophrenia and 18 male healthy controls. The control participants were group matched to the patients on age, handedness, parental socioeconomic status, and an estimated premorbid IQ. For each subject, an anatomical scan (SPGR, $TR = 7.4s$, $TE = 3ms$, $FOV = 26cm^2$, $res = 1mm^3$), a diffusion-weighted scan (EPI, $TR = 17s$, $TE = 78ms$, $FOV = 24cm^2$, $res = 1.66 \times 1.66 \times 1.7mm$, 51 gradient directions with $b = 900s/mm^2$, 8 baseline scans with $b = 0s/mm^2$) and a resting-state functional scan (EPI-BOLD, $TR = 3s$, $TE = 30ms$, $FOV = 24cm^2$, $res = 1.875 \times 1.875 \times 3mm$) were acquired using a 3T GE Echospeed system.

Pre-Processing. We segmented the structural images into 77 anatomical regions with Freesurfer [12]. The DWI data was corrected for eddy-current distortions. Two-tensor tractography was used to estimate the white matter fibers [13]. We compute the DWI connectivity D_n^j by averaging FA along all fibers connecting the corresponding regions. If no tracts are found, D_n^j is set to zero.

Table 1. Parameters of the joint model in Fig. 1(a), estimated separately for control (NC) and schizophrenic (SZ) populations, and for the entire dataset (NC+SZ)¹

	π_A	$\pi_{F,-1}$	π_{F0}	π_{F1}	ρ_0	ρ_1	χ_0	χ_1	ξ_0	ξ_1
NC	0.37	0.43	0.40	0.17	0.66	0.10	0.42	0.34	0.005	0.003
SZ	0.37	0.43	0.41	0.16	0.67	0.11	0.41	0.34	0.005	0.003
NC+SZ	0.36	0.43	0.40	0.17	0.66	0.11	0.41	0.34	0.005	0.003

	$\mu_{0,-1}$	$\mu_{1,-1}$	μ_{00}	μ_{10}	μ_{01}	μ_{11}	$\sigma_{0,-1}^2$	$\sigma_{1,-1}^2$	σ_{00}^2	σ_{10}^2	σ_{01}^2	σ_{11}^2
NC	-0.063	-0.022	0.006	0.093	0.12	0.27	0.014	0.010	0.008	0.012	0.019	0.033
SZ	-0.083	-0.036	0.001	0.097	0.13	0.27	0.011	0.011	0.013	0.013	0.017	0.035
NC+SZ	-0.073	-0.027	0.003	0.099	0.12	0.28	0.012	0.011	0.012	0.012	0.017	0.035

We discarded the first five fMRI time points and performed motion correction by rigid body alignment and slice timing correction using FSL [14]. The data was spatially smoothed using a Gaussian filter, temporally low-pass filtered with $0.08Hz$ cutoff, and motion corrected via linear regression. Finally, we regressed out global contributions to the timecourses from the white matter, ventricles and the whole brain. We extract the fMRI connectivity B_n^j by computing Pearson correlation coefficients between every pair of voxels in the two regions of the n^{th} connection, applying the Fisher-r-to-z transform to each correlation (to enforce normality), and averaging these values. Since our anatomical regions are large, the correlation between the mean timecourses of two regions shows poor correspondence with the distribution of voxel-wise correlations between them. We believe our measure is more appropriate for assessing fMRI connectivity.

To inject prior clinical knowledge, we pre-selected 8 brain structures (corresponding to 16 regions) that are believed to play a role in schizophrenia: the superior temporal gyrus, rostral middle frontal gyrus, hippocampus, amygdala, posterior cingulate, rostral anterior cingulate, parahippocampal gyrus, and transverse temporal gyrus. We model only the 1096 ($16 \times 76 - \binom{16}{2}$) unique pairwise connections between these ROIs and all other regions in the brain.

Joint Connectivity Model. We first fit the joint model in Fig. 1(a) to each population separately as well as to the entire dataset. Table 1 reports the parameters of the three models. We observe that $\{\mu, \sigma, \rho, \chi, \xi\}$ are largely consistent across the three models. This supports our hypothesis that group differences appear in the latent connectivities rather than in the data likelihood parameters.

Population Study. Fig. 2 depicts the significantly different ($p < 0.05$, $\hat{\epsilon} > 0.5$) anatomical and functional connections identified by the algorithm. Table 2 lists the corresponding regions and p-values. Due to space limitations, we report just the 3 connections with the smallest p-values in each modality.

As seen from Fig. 2(b), schizophrenia patients exhibit increased functional connectivity between the parietal/posterior cingulate region and the frontal lobe and reduced functional connectivity between the parietal/posterior cingulate region and the temporal lobe. These results confirm the hypotheses of widespread

¹ $\chi_0 \chi_1$ implies that spurious DWI fibers arise due to artificially high anisotropy.

Table 3. Training and testing accuracy of ten-fold cross validation for the control (NC) and Schizophrenic (SZ) populations

	Training NC	Training SZ	Testing NC	Testing SZ
Joint fMRI/DWI Model	0.99 ± 0.005	0.88 ± 0.01	0.61 ± 0.06	0.55 ± 0.05
linear SVM fMRI	1.00 ± 0.00	1.00 ± 0.00	0.54 ± 0.05	0.61 ± 0.05
linear SVM DWI	1.00 ± 0.00	1.00 ± 0.00	0.59 ± 0.08	0.58 ± 0.06
linear SVM fMRI/DWI	1.00 ± 0.00	1.00 ± 0.00	0.67 ± 0.04	0.60 ± 0.05

4 Conclusion

We proposed a novel probabilistic framework that fuses information from resting-state fMRI data and DWI tractography. We further extended the basic approach to model connectivity differences between two populations. We show that our method captures changes in functional and anatomical connectivity induced by schizophrenia. In particular, we detect increased functional connectivity from the parietal lobe to the frontal lobe and decreased functional connectivity from the parietal lobe to the temporal lobe. We also find significant anatomical connectivity differences involving the superior temporal gyrus, the posterior cingulate and the hippocampus. Finally, we demonstrate the predictive power of our model through cross validation. These results establish the promise of our approach for combining multiple imaging modalities to better understand the brain.

Acknowledgments. This work was supported in part by the National Alliance for Medical Image Analysis (NIH NIBIB NIMIC U54-EB005149), the Neuroimaging Analysis Center (NIH NCRR NAC P41-RR13218), the NSF CAREER Grant 0642971 and NIH R01MH074794. A. Venkataraman is supported by the National Defense Science and Engineering Graduate Fellowship (NDSEG).

References

1. Buckner, R.L., Vincent, J.L.: Unrest at rest: Default activity and spontaneous network correlations. *NeuroImage* 37(4), 1091–1096 (2007)
2. Basser, P., Pierpaoli, C.: Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor mri. *Journal of Magnetic Resonance, Series B* 111, 209–219 (1996)
3. Honey, C., et al.: Predicting human resting-state functional connectivity from structural connectivity. *PNAS* 106(6), 2035–2040 (2009)
4. Koch, M.A., et al.: An investigation of functional and anatomical connectivity using magnetic resonance imaging. *NeuroImage* 16(1), 241–250 (2002)
5. Sporns, O., et al.: Theoretical neuroanatomy: Relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cerebral Cortex* 10(2), 127–141 (2000)
6. Honey, C., et al.: Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *PNAS* 104(24), 10240–10245 (2007)

7. Gabrieli-Whitfield, S., et al.: Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-degree relatives of persons with schizophrenia. *PNAS* 106(4), 1279–1284 (2009)
8. Burns, J., et al.: Structural disconnectivity in schizophrenia: A diffusion tensor magnetic resonance imaging study. *Br. J. Psychiatry* 182, 439–443 (2003)
9. Ke, M., et al.: Combined analysis for resting state fmri and dti data reveals abnormal development of function-structure in early-onset schizophrenia. In: Wang, G., Li, T., Grzymala-Busse, J.W., Miao, D., Skowron, A., Yao, Y. (eds.) *RSKT 2008. LNCS (LNAI)*, vol. 5009, pp. 628–635. Springer, Heidelberg (2008)
10. Zhou, Y., et al.: Altered resting-state functional connectivity and anatomical connectivity of hippocampus in schizophrenia. *Schiz. Res.* 100(1-3), 120–132 (2008)
11. Jordan, M., et al.: An introduction to variational methods for graphical models. *Machine Learning* 37(2), 183–233 (1999)
12. Fischl, B., et al.: Sequence-independent segmentation of magnetic resonance images. *NeuroImage* 23, 69–84 (2004)
13. Malcolm, J., Shenton, M., Rathi, Y.: Neural tractography using an unscented kalman filter. In: *IPMI*, pp. 126–138 (2009)
14. Smith, S.M., et al.: Advances in functional and structural mr image analysis and implementation as fsl. *NeuroImage* 23(51), 208–219 (2004)
15. Kubicki, M., et al.: A review of diffusion tensor imaging studies in schizophrenia. *Journal of Psychiatric Research* 41(1-2), 15–30 (2007)

Detection of Brain Functional-Connectivity Difference in Post-stroke Patients Using Group-Level Covariance Modeling

Gaël Varoquaux¹, Flore Baronnet², Andreas Kleinschmidt², Pierre Fillard¹,
and Bertrand Thirion¹

¹ Parietal project-team, INRIA Saclay-île de France

² INSERM, U562, CEA/Neurospin bât 145, 91191 Gif-Sur-Yvette, France

Abstract. Functional brain connectivity, as revealed through distant correlations in the signals measured by functional Magnetic Resonance Imaging (fMRI), is a promising source of biomarkers of brain pathologies. However, establishing and using diagnostic markers requires probabilistic inter-subject comparisons. Principled comparison of functional-connectivity structures is still a challenging issue. We give a new matrix-variate probabilistic model suitable for inter-subject comparison of functional connectivity matrices on the manifold of Symmetric Positive Definite (SPD) matrices. We show that this model leads to a new algorithm for principled comparison of connectivity coefficients between pairs of regions. We apply this model to comparing separately post-stroke patients to a group of healthy controls. We find neurologically-relevant connection differences and show that our model is more sensitive than the standard procedure. To the best of our knowledge, these results are the first report of functional connectivity differences between a single-patient and a group and thus establish an important step toward using functional connectivity as a diagnostic tool.

1 Introduction

The correlation structure of brain activity, measured via fMRI, reveals stable inter-subject networks of distant brain regions that can be the expression of cognitive function. In particular, some connectivity networks are present in the absence of stimuli. They can reveal intrinsic brain activity and are studied in the *resting-state* paradigm. These structures are of particular interest to study and diagnose brain diseases and disorders [1] as they can be used for deep probes of brain function on diminished subjects. Not only can they extract medically or cognitively relevant markers on subjects unconscious [2], or with limited cooperation [3], but they also give information on higher-level cognitive systems that are challenging to probe via medical imaging or behavioral clinical tests [4].

To use functional connectivity as a quantitative inference tool, principled probabilistic comparison of connectivity structures across subjects is needed. Unlike with stimuli-response studies used routinely in functional neuroimaging,

this comparison is challenging, as the underlying description of the signal is multivariate: each brain-activity time course is considered relative to others. Univariate group models, such as random effects or mixed effects, are in general not sound as they neglect the strong statistical dependence between parameters estimated from the data. Multivariate techniques have been successfully employed to single out outlying subjects [5], but have met little success: their results are difficult to interpret as they do not point to specific localized differences.

In this paper, we focus on the description of brain functional-connectivity using inter-regions correlation matrices. We first review the current practice in inter-subject functional covariance comparison and recall some results on the manifold of covariance matrices. Then, we introduce a probabilistic model at the group level for the different subjects' correlation matrices, and a corresponding algorithm to detect connectivity differences in a specific parametrization of the covariance matrices, as correlations are a form of covariance. We quantify on simulated data the performance of this detection. Finally, we apply the model to the individual comparison of the connectivity structure of stroke patients to a group of healthy controls, and show that it outperforms the current practice.

2 State of the Art

2.1 Problem Statement: Comparing Functional Brain Connectivity

We consider S subjects, represented by the correlation matrices between brain-activity time series extracted from n ROIs: $\{\Sigma^s \in \mathbb{R}^{n \times n}, s = 1 \dots S\}$. The challenge is to give a probabilistic description of the population of correlation matrices so as to find the significant differences between subjects or groups. The current practice in functional-connectivity studies is to compare the coefficients of the correlation matrices across subjects (see for instance [3,6]). This procedure can be expressed as a univariate additive linear model on the correlation matrix:

$$\Sigma^s = \Sigma^* + \mathbf{d}\Sigma^s \quad (1)$$

where Σ^* is a covariance matrix representative of the mean effect, or the group, and $\mathbf{d}\Sigma^s$ encode subject-specific contributions.

However, with this description it is difficult to isolate significant contributions to $\mathbf{d}\Sigma^s$. Indeed, for interpretation, some coefficients are zeroed out, eg by thresholding a test statistic, as in [3], which eventually leads to a non positive definite matrix, for which it is impossible to write a multivariate normal likelihood. As a result, the subject-variability description learned on a population cannot give probabilistic tests on new subjects.

2.2 Recent Developments on the Covariance-Matrix Manifold

The mathematical difficulty stems from the fact that the space of SPD matrices, Sym_n^+ , does not form a vector space: $\mathbf{A}, \mathbf{B} \in Sym_n^+ \not\Rightarrow \mathbf{A} - \mathbf{B} \in Sym_n^+$. The Fisher information matrix of the multivariate normal distribution can be

used to construct a metric on a parametrization of covariance matrices [7] and thus define Sym_n^+ as a Riemannian manifold that is well-suited for performing statistics on covariances [8]. Local differences on this manifold can be approximated by vectors of the tangent space: if \mathbf{B} is close enough to \mathbf{A} , the application: $\phi_{\mathbf{A}} : \mathbf{B} \rightarrow \log(\mathbf{A}^{-\frac{1}{2}}\mathbf{B}\mathbf{A}^{-\frac{1}{2}})$ maps locally the bipoint $\mathbf{A}, \mathbf{B} \in Sym_n^+ \times Sym_n^+$ to $\overrightarrow{\mathbf{A}\mathbf{B}} \in Sym_n$, the space of symmetric matrices[1]. A convenient parametrization of $\overrightarrow{\mathbf{W}} \in Sym_n$ is $\text{Vec}(\mathbf{W}) = \{\sqrt{2} w_{i,j}, j < i, w_{i,i}, i = 1 \dots n\}$ that forms an orthonormal basis of the tangent space [8]. Finally, $\|\overrightarrow{\mathbf{A}\mathbf{B}}\|_{\mathbf{A}}^2 = \|\log(\mathbf{A}^{-\frac{1}{2}}\mathbf{B}\mathbf{A}^{-\frac{1}{2}})\|_2^2$ gives the intrinsic norm of $\overrightarrow{\mathbf{A}\mathbf{B}}$ on the Sym_n^+ manifold, according to the metric around point \mathbf{A} : the distance between \mathbf{A} and \mathbf{B} in the manifold.

3 Matrix-Variate Random Effects Model for Covariances

Multi-subject probability distribution for covariance matrices. Using the Riemannian parametrization of Sym_n^+ , we describe the individual correlation matrix population as a distribution of matrices scattered around a covariance matrix representative of the group, Σ^* . As this distribution must be estimated with a small number of observations compared to the feature space, we model it using the probability density function that minimizes the information with a given mean on the manifold, the generalized Gaussian distribution [8]:

$$p(\Sigma) = k(\sigma) \exp\left(-\frac{1}{2\sigma^2} \|\overrightarrow{\Sigma^*\Sigma}\|_{\Sigma^*}^2\right), \tag{2}$$

where σ encodes an isotropic variance on the manifold and k is a normalization factor. Given multiple observations of Σ corresponding to individual correlation matrices, Σ^s , the maximum likelihood estimate of Σ^* is independent of σ and given by the Fréchet mean of the observations [8], minimizing $\sum_s \|\overrightarrow{\Sigma^*\Sigma^s}\|_{\Sigma^*}^2$.

Parametrization in the tangent space. We express the individual covariance matrices as a perturbation of the group covariance matrix Σ^* :

$$\forall s = 1 \dots S, \quad \Sigma^s = \phi_{\Sigma^*}^{-1}(\mathbf{d}\Sigma^s) = \Sigma^{*\frac{1}{2}} \exp(\mathbf{d}\Sigma^s) \Sigma^{*\frac{1}{2}}, \tag{3}$$

thus, using (2),
$$p(\mathbf{d}\Sigma^s) = k'(\sigma) \exp\left(-\frac{1}{2\sigma^2} \|\mathbf{d}\Sigma^s\|_2^2\right). \tag{4}$$

The parameters of $\text{Vec}(\mathbf{d}\Sigma^s)$ follow a normal distribution, with diagonal covariance σ , and the maximum-likelihood estimate of σ is given by $\hat{\sigma}_{MLE}^2 = \frac{1}{S} \sum_s \|\text{Vec}(\mathbf{d}\Sigma^s)\|_2^2$. The model can thus be interpreted as a random-effect model on the parametrization of $\text{Vec}(\mathbf{d}\Sigma^s)$, in the space tangent to the manifold Sym_n^+ .

¹ Note that we do not use the same definition of the mapping as in [8,7], as we are interested in mapping to Sym_n , the tangent space around \mathbf{I}_n , and not the tangent space in \mathbf{A} . It extracts a statistically independent parametrization (Eq. (3) and (4)).

Algorithm 1. Estimation of the group model

- 1: **Input:** individual time series $\mathbf{X}^1 \dots \mathbf{X}^s$.
 - 2: **Output:** estimated group covariance matrix $\widehat{\Sigma}^*$, group variance $\widehat{\sigma}$.
 - 3: **for** $s = 1$ **to** S **do**
 - 4: Compute $\widehat{\Sigma}^s \leftarrow \text{LedoitWolf}(\mathbf{X}^s)$.
 - 5: **end for**
 - 6: Compute $\widehat{\Sigma}^* \leftarrow \text{intrinsic mean}(\widehat{\Sigma}^1 \dots \widehat{\Sigma}^s)$.
 - 7: **for** $s = 1$ **to** S **do**
 - 8: Compute $\mathbf{d}\widehat{\Sigma}^s \leftarrow \widehat{\Sigma}^{*-1/2} \Sigma^s \widehat{\Sigma}^{*-1/2} - \mathbf{I}_n$.
 - 9: **end for**
 - 10: Compute $\widehat{\sigma} \leftarrow \sqrt{\frac{1}{S} \sum_s \|\text{Vec}(\mathbf{d}\widehat{\Sigma}^s)\|_2^2}$.
-

Assuming that the distribution is narrow on the manifold, $\|\mathbf{d}\Sigma^s\|_2 \ll 1$, eq. (3) can be seen as the application of the placement function to move a noise $\mathbf{d}\Sigma^s$ isotropic around \mathbf{I}_n to Σ^* (see [8], section 3.5):

$$\Sigma^s \simeq \Sigma^{*\frac{1}{2}} (\mathbf{I}_n + \mathbf{d}\Sigma^s) \Sigma^{*\frac{1}{2}}. \tag{5}$$

Model estimation from the data. We start from individual time-series of brain activity in selected regions of interest, $\mathbf{X} \in \mathbb{R}^{n \times t}$. We use the Ledoit-Wolf shrinkage covariance estimator [9] for a good bias-variance compromise when estimating correlation matrices from t time points with $n < t < n^2$. From this estimate of individual correlation matrices, we compute the intrinsic mean on Sym_n^+ using algorithm 3 of [10]. Finally, we estimate σ from the residuals of individual correlation matrices in the space tangent in Σ^* (see algorithm 1).

4 Testing Pair-Wise Correlations Statistics

The multivariate probabilistic model for correlations between regions exposed above enables us to define an average correlation matrix of a group, as well as the dispersion of the group in the covariance matrix space. Thus it can be used to test if a subject is different to a control group. However, to aid diagnosis, it is paramount to pin-point why such a subject may be different. In the tangent space, the parameters $d\Sigma_{i,j}^s$ of $\text{Vec}(\mathbf{d}\Sigma^s)$ are mutually independent. We can thus conduct univariate analysis on these parameters to test which significantly differs from the control group. However, the independence of the parameters is true only in the space tangent at the population average Σ^* , of which we only have an estimate $\widehat{\Sigma}^*$. Thus, to account for projection error, we resort to non-parametric sampling of the control population to define a null distribution for the parameters $d\Sigma_{i,j}^s$.

Specifically, we are interested in testing if a difference observed for a subject in one of the $d\Sigma_{i,j}^s$ can be explained by variation of the control population. As the control population is typically small, we generate the null distribution by leave

Algorithm 2. Coefficient-level tests

-
- 1: **Input:** individual time series for controls $\mathbf{X}^1 \dots \mathbf{X}^s$ and a patient \mathbf{X}^k , p-value p , number of bootstraps, m .
 - 2: **Output:** Pair-wise p-values $p_{i,j}$ controlling for the difference in $d\Sigma_{i,j}$ between the patient and the control group.
 - 3: Initialize $P_{i,j}^o \leftarrow$ empty lists, for $i, j \in \{1 \dots n\}, j < i$.
 - 4: **for** 1 **to** m **do**
 - 5: Choose a surrogate patient $\tilde{s} \in 1 \dots S$.
 - 6: Choose a subset \tilde{S} of $\{1 \dots S\} \setminus \tilde{s}$ of S surrogate controls.
 - 7: Compute $\widetilde{\Sigma}^*$ and $\widetilde{\mathbf{d}\Sigma}^s$ for $s \in \tilde{S}$ using algorithm [1](#) on the surrogate controls.
 - 8: Compute $\widetilde{\mathbf{d}\Sigma}^s$ for $s = \tilde{s}$, using $\widetilde{\Sigma}^*$ and eqn [5](#).
 - 9: For all i, j , append to $P_{i,j}^o$ the T test comparing $\widetilde{\mathbf{d}\Sigma}_{i,j}^s$ for $s \in \tilde{S}$ and for $s = \tilde{s}$.
 - 10: **end for**
 - 11: Compute $\widehat{\Sigma}^*$ and $\widehat{\mathbf{d}\Sigma}^s$ for $s \in \tilde{S}$ using algorithm [1](#) on the complete control group.
 - 12: Compute $\widehat{\mathbf{d}\Sigma}^k$, using $\widehat{\Sigma}^*$ and eqn [5](#).
 - 13: For all i, j , compute $t_{i,j}$ the T test comparing $\widehat{\mathbf{d}\Sigma}_{i,j}^s$ for $s \in \tilde{S}$ and for $s = \tilde{s}$.
 - 14: $p_{i,j} = 1 - \text{quantile}(t_{i,j} \text{ in } P_{i,j}^o)$
-

one out: for each control subject, we generate surrogate control populations \tilde{S} by bootstrap from the other control subjects and estimate the corresponding average covariance $\widetilde{\Sigma}^*$. We use $\widetilde{\Sigma}^*$ to project all the individual correlation matrices, including the left out subject, to compute $\widetilde{\mathbf{d}\Sigma}_{i,j}^s$, and we do a one sample T test of the difference between $\widetilde{\mathbf{d}\Sigma}_{i,j}^s$ for the left out subject with regards to the resampled control group \tilde{S} . We record the values of this T test as an estimate of the null distribution $P_{i,j}^o$ of the T test on the corresponding coefficient between the controls and a patient. Finally, we estimate the average covariance $\widehat{\Sigma}^*$ for the complete group of controls and, for each k patient to investigate, we perform a T test of the difference between $\widehat{\mathbf{d}\Sigma}_{i,j}^s$ for the patient and the control group. We use $P_{i,j}^o$ to associate a p-value to each coefficient per subject. We correct for multiple comparisons using Bonferroni correction: for each patient, $\frac{1}{2}n(n-1)$ tests are performed. The procedure is detailed in algorithm [2](#).

5 Algorithm Evaluation on Simulated Data

Algorithm [2](#) relies on approximations of the exact problem for coefficient-level detection of differences. In order to quantify the performance of this detection, we study Receiver Operator Characteristic (ROC) on simulated data: we draw a population of control covariances using eq. [5](#) with the parameters of $\text{Vec}(\mathbf{d}\Sigma)$ normally distributed with deviation σ . For simulated patients, we add differences of amplitude $d\Sigma$ to a few coefficients (~ 20) of this variability noise. For Σ^* and σ , we use the parameters estimated on real data (section [6](#)). We investigate the performance of algorithm [2](#) to recover these differences for a variety of parameters. We observe good recovery for $d\Sigma > \sigma$ (Fig [1](#)), and find that the comparison in the tangent space (eq. [5](#)) outperforms a comparison in $\mathbb{R}^{n \times n}$ (eq. [1](#)).

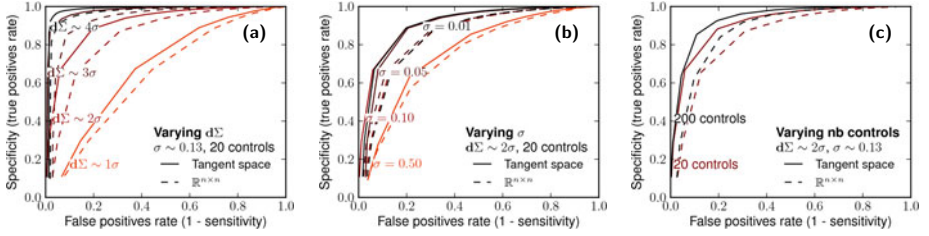


Fig. 1. ROC curves on data simulated according to the variability model given by eqn 5. **(a)** for different values of patient differences $d\Sigma$. **(b)** for different values of control variability σ . **(c)** for different number of controls.

6 Application to Post-stroke Connectivity Reorganization

Standard clinical scores, such as the NIHSS, as well as fMRI studies can be used to assess the consequences of cerebral strokes, but they test specific cognitive functions and have little sensitivity to higher-order cognitive malfunctions. Resting-state functional-connectivity is thus a valuable tool to study post-stroke reorganization. We apply our model to stroke patients.

Resting-state fMRI dataset. After giving informed consent, ischemic-stroke patients, as well as age-matched healthy controls, underwent MRI scanning. Subjects with existing neurology, psychiatry, or vascular pathologies were excluded from the study. 10 patients and 20 controls were scanned during a resting-state task: subjects were given no other task than to stay awake but keep their eyes closed. 2 sessions of 10 minutes of fMRI data were acquired on a Siemens 3T Trio scanner (245 EPI volumes, TR=2.46s, 41 slices interlaced, isotropic 3 mm voxels). After slice-timing, motion correction, and inter-subject normalization using SPM5, 33 ROIs were defined in the main resting-state networks by intersecting a segmentation of the gray matter with correlation maps from seeds selected from the literature. For each subject, the BOLD time series corresponding to the regions were extracted and orthogonalized with respect to confound time series: time courses of the white matter and the cerebro-spinal fluid, and movement regressors estimated by the motion-correction algorithm. Covariance modeling was performed on the resulting 33 time series.

Separating patients from controls with the matrix-variate covariance model. To measure the discriminative power of the matrix-variate model introduced in section 3, we test the likelihood of patient data in a model learned on controls. Specifically, we evaluate by *leave one out* the likelihood of each control in the model learned on the other controls. We compare this value to the average likelihood of patients in the 20 models obtained by *leave one out*. We perform this comparison both using the group model isotropic on the tangent space (eq. 5), and the group model isotropic in $\mathbb{R}^{n \times n}$ (eq. 1). We find that the model in the tangent space separates better patients from controls (Fig 2).

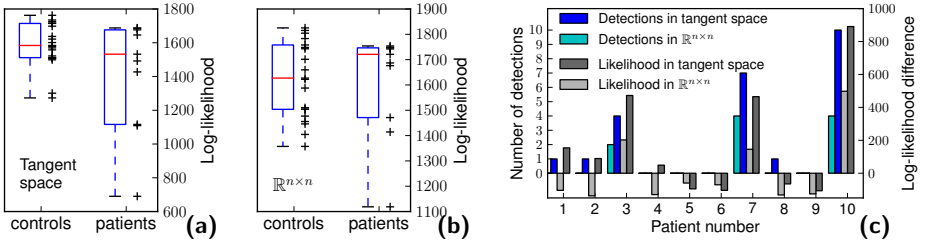


Fig. 2. (a) Likelihood of the controls and the patients in the model parametrized in the tangent space. (b) Likelihood of the controls and the patients in the model parametrized in $\mathbb{R}^{n \times n}$. (c) Number of coefficients detected as significantly different from the control group per patient, for the model parametrized in the tangent space, as well as in $\mathbb{R}^{n \times n}$.

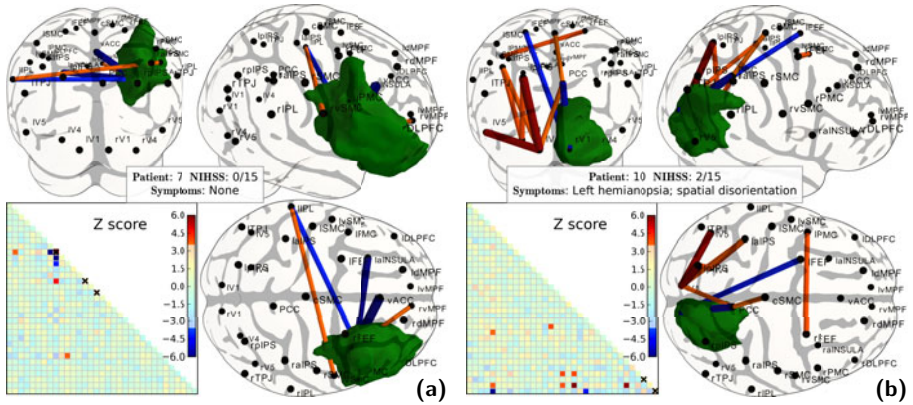


Fig. 3. Significant differences on two patients ($p < 0.05$ uncorrected), represented as connections between regions: increased connectivity appears in red, and decreased in blue. The lesion, manually segmented from anatomical images, is represented in green. ROIs fully covered by a lesion are marked with a black cross on the correlation matrix.

Detected connection differences. We apply algorithm 2 to detect the significant coefficient-level differences for each subject. We compare to a similar univariate procedure applied to the parametrization in $\mathbb{R}^{n \times n}$ given by eq. (11), rather than the tangent space. We find that coefficient-level analysis detects more differences between ROI pairs when applied on the tangent-space parametrization (Fig 2c).

7 Discussion

Interpretation of the tangent space. Projecting on the space tangent to the group mean corresponds to applying a whitening matrix $\Sigma^*^{-\frac{1}{2}}$ learned on the group (eq. 5) that converts the Gaussian process described by the group covariance

to an independent and identically distributed (iid) process. In other words, the *coloring* of the time series common to the group is canceled out to compare subjects on iid coefficients on the correlation matrix.

Probing neurological processes. For certain subjects, both procedures fail to detect a single connection that makes a significant difference. Indeed, the variability of resting-state activity in the control group induces some variability in the projection to the tangent space. For patients with small lesions, this variability is larger than the univariate differences. On the other hand, for patients with important lesions, the functional connectivity analysis reveals profound differences in the correlation structure that reflect functional reorganization. Some express a direct consequence of the lesion, for example when the gray matter in one of two ROIs has been damaged by the lesion, as can be seen on Fig 3a. Others reflect functional reorganization. For instance, patient 10 has a right visual cortex damaged by a focal lesion, but the analysis shows increased connectivity in his left visual cortex (Fig 3b). Functional connectivity analysis thus reveals modifications that go beyond the direct anatomical consequences of the lesion.

8 Conclusion

We have presented a matrix-variate probabilistic model for covariances, and have shown that it can be expressed as a random effect model on a particular parametrization of the covariance matrix. The ability to draw conclusions on the connectivity between pairs of regions is important because it is a natural representation of the problem. We applied this model to the comparison of functional brain connectivity between subjects. We were able to detect significant differences in functional connectivity between a single stroke patient and 20 controls. A controlled detection of network-wide functional-connectivity differences between subjects opens the door to new markers of brain diseases as well as new insights in neuroscience, as functional connectivity can probe phenomena that are challenging to access via stimuli-driven studies.

References

1. Greicius, M.: Resting-state functional connectivity in neuropsychiatric disorders. *Current opinion in neurology* 21, 424 (2008)
2. Vanhaudenhuyse, A., et al.: Default network connectivity reflects level of consciousness in non-communicative brain-damaged patients. *Brain* 133, 161 (2010)
3. Fair, D., et al.: Development of distinct control networks through segregation and integration. *PNAS* 104, 13507 (2007)
4. Uddin, L., Menon, V.: The anterior insula in autism: Under-connected and under-examined. *Neuroscience and Biobehavioral Reviews* 33, 1198–1203 (2009)
5. Sato, J., et al.: An fMRI normative database for connectivity networks using one-class support vector machines. *Human brain mapping* 30, 1068 (2008)
6. Rocca, M., et al.: Altered functional and structural connectivities in patients with MS: a 3-T study. *Neurology* 69, 2136 (2007)

7. Lenglet, C., et al.: Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor MRI processing. *Journal of Mathematical Imaging and Vision* 25, 423–444 (2006)
8. Pennec, X., Fillard, P., Ayache, N.: A Riemannian framework for tensor computing. *International Journal of Computer Vision* 66, 41–66 (2006)
9. Ledoit, O., Wolf, M.: A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivar. Anal.* 88, 365–411 (2004)
10. Fletcher, P., Joshi, S.: Riemannian geometry for the statistical analysis of diffusion tensor data. *Signal Processing* 87, 250–262 (2007)

Understanding the Optics to Aid Microscopy Image Segmentation

Zhaozheng Yin¹, Kang Li², Takeo Kanade¹, and Mei Chen³

¹ Carnegie Mellon University

² Microsoft Corporation

³ Intel Labs Pittsburgh

Abstract. Image segmentation is essential for many automated microscopy image analysis systems. Rather than treating microscopy images as general natural images and rushing into the image processing warehouse for solutions, we propose to study a microscope’s optical properties to model its image formation process first using phase contrast microscopy as an exemplar. It turns out that the phase contrast imaging system can be relatively well explained by a linear imaging model. Using this model, we formulate a quadratic optimization function with sparseness and smoothness regularizations to restore the “authentic” phase contrast images that directly correspond to specimen’s optical path length without phase contrast artifacts such as halo and shade-off. With artifacts removed, high quality segmentation can be achieved by simply thresholding the restored images. The imaging model and restoration method are quantitatively evaluated on two sequences with thousands of cells captured over several days.

1 Introduction

Long-term monitoring of living specimen’s movement and behavior has a wide range of applications in biological discovery. Since transparent specimens such as living cells generally lack sufficient contrast to be observed by common light microscopes, the phase contrast imaging technique was invented to convert the minute light phase variation caused by specimen into changes in light amplitude that can be observed by naked eyes or cameras [15]. Due to the optical principle and some imperfections of the conversion process, phase contrast images contain artifacts such as the halo surrounding the specimen and shade-off (Fig. 3(b), the intensity profile of a large specimen gradually increases from the edges to the center, and even approaches the intensity of the surrounding medium). Over time, biologists have learned how to overcome or even exploit those artifacts for interpreting images. When computer-based microscopy image analysis began to relieve humans from tedious manual labelling [5,8,12,14], it is unsurprising that those artifacts cause the major difficulty in automated image processing. In particular, they hinder the process of segmenting images into cells and background, which is the most critical step in almost all cell analysis and tracking systems.

1.1 Previous Work

In microscopy image segmentation, specimen pixels are segmented from background pixels and then grouped into objects. In the past decades, many image segmentation methods have been invented. For example, thresholding on local intensity value and variation has a long history on cell image segmentation ([13] and references therein). Morphological operations on the gradient or intensity images are widely applied to segment specimen pixels [5,8]. Using the artifact of microscopy images, Laplacian of Gaussian filter is also used to extract object blob and contour [12]. After obtaining binary masks indicating whether each pixel is a specimen or background pixel, connected component labelling or marker-controlled watershed algorithms are often performed to group specimen pixels into specimen objects [5,14].

1.2 Our Proposal

The previous image segmentation methods do not consider the microscopy image formation process and treat them in the same way as general natural images. However, there are apparent differences between natural images and phase contrast microscopy images, such as halo and shade-off. Because of these artifacts, the observed image intensity does not exactly correspond to specimen's optical path length (product of refractive index and thickness). Recently, Li and Kanade [9] proposed an algebraic framework for preconditioning microscopy images captured by differential interference contrast (DIC) microscopes. This inspired us to think about whether understanding the phase contrast optics at an early stage will help segment phase contrast images. In fact, we found that those artifacts are not caused by random processes, instead they can be relatively well modeled by the optical properties of the image formation process in the phase contrast microscope imaging system. In this paper, we derive a linear imaging model for phase contrast microscopy and formulate a quadratic optimization function to restore the "authentic" phase contrast images without halo or shade-off effects. With artifacts removed, high quality segmentation can be achieved by simply thresholding the restored images.

2 Imaging Model of Phase Contrast Microscopy

Phase contrast microscope uses a conjugate pair of condenser annulus and phase plate as shown in Fig. 1(c), where the specialized annulus is placed at the front focal plane of the condenser while the phase plate is at the rear focal plane of the objective lens. Wavefronts illuminate the specimen and divide into two components: one component passes through and around the specimen without deviation (commonly referred to as the S wave); the other component is diffracted, attenuated and retarded because of specimen (D wave). A typical phase retardation caused by living cells in tissue culture is a quarter wave length [10]. The two waves enter the objective lens and combine through interference to produce the particle wave (P wave).

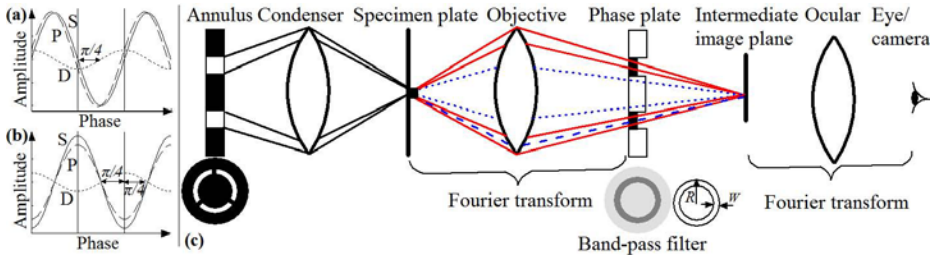


Fig. 1. Wave interaction and phase contrast microscope optics. (a) Without phase contrast, the amplitude difference between P (dash) and S (solid) waves is small; (b) With phase contrast technique, the difference is large enough to be observed. (c) The phase plate separates the surround (solid) and diffracted (dot) lights, speeds up (i.e. advances the phase by $1/4$ wavelength) and attenuates the surround light. A small portion of diffracted light (dash) leaking into the phase ring causes the halo and shade-off artifacts.

Observing the specimen in the microscopy image depends on the intensity difference between specimen and its surrounding background, i.e. the amplitude difference between particle (P) and surround (S) waves. Without any phase contrast technique, the P and S waves have nearly the same wave amplitudes as shown in Fig. 1(a), thus the specimen remains transparent and appears invisible. However, in phase contrast microscope (Fig. 1(c)), the phase plate has an etched ring with reduced thickness to advance the surround wave by a quarter wavelength. The phase ring also has a partially absorbing metallic film to attenuate the surround wave such that the final contrast between P and S waves are easy to be perceived. The diffracted wave spreads over the phase plate. Most of it passes through the phase plate without changes, and it interferes with the surround wave to form the particle wave as shown in Fig. 1(b) where the amplitude difference between P and S waves is now observable and the specimen will appear as dark on a bright background. Unfortunately, some of the diffracted wave will “leak” into the phase ring, which cause the halo and shade-off artifacts.

After illuminating waves pass through the specimen plate, the unaltered surround wave $l_S(\mathbf{x})$ and the diffracted wave $l_D(\mathbf{x})$ (attenuated and retarded) are

$$l_S(\mathbf{x}) = Ae^{i0} \quad (1)$$

$$l_D(\mathbf{x}) = \zeta_c Ae^{-if(\mathbf{x})} \quad (2)$$

where $i^2 = -1$, $\mathbf{x} = \{(\mathbf{x}_j^r, \mathbf{x}_j^c), j = 1, \dots, J\}$ represent 2D locations of J image pixels, A is the illuminating wave’s amplitude, ζ_c and $f(\mathbf{x})$ represent the amplitude attenuation and phase shift caused by the specimen. Our goal is to restore $f(\mathbf{x})$, the “authentic” phase contrast image without artifacts. A thin lens with a large aperture essentially performs a spatial Fourier transform (\mathcal{F}) on the waves from its front focal plane to its rear focal plane [3], thus the waves in front of the phase plate are $L_S(\mathbf{w}) = \mathcal{F}(l_S(\mathbf{x}))$ and $L_D(\mathbf{w}) = \mathcal{F}(l_D(\mathbf{x}))$.

The phase plate functions as a band-pass filter. For the non-diffracted surround wave, the positive phase ring attenuates the wave amplitude and advances its phase by a quarter wave length ($\pi/2$), thus the corresponding transmittance function is

$$T_S(\mathbf{w}) = \zeta_p e^{i\frac{\pi}{2}} = i\zeta_p \tag{3}$$

where ζ_p represents the amplitude attenuation by a phase ring with outer radius R and width W (R and W are provided by microscope manufactures). The diffracted wave spreads over the phase plate with a small portion leaking into the ring. The corresponding transmittance function is

$$T_D(\mathbf{w}) = 1 + (i\zeta_p - 1)[\text{cyl}(\frac{w_r}{R}) - \text{cyl}(\frac{w_r}{R - W})] \tag{4}$$

where $w_r = \sqrt{w_u^2 + w_v^2}$ is the radial frequency and $\text{cyl}(\cdot)$ is a 2D cylinder (or circular) function: $\text{cyl}(t) = 1, \text{ if } 0 \leq t \leq 1; 0, \text{ otherwise. After band-pass filtering, we have the waves immediately after the phase plate as } \tilde{L}_S(\mathbf{w}) = L_S(\mathbf{w})T_S(\mathbf{w}) \text{ and } \tilde{L}_D(\mathbf{w}) = L_D(\mathbf{w})T_D(\mathbf{w}).$

The ocular lens perform another consecutive Fourier transform. Mathematically, the forward and inverse Fourier transforms are identical except for a minus sign. Hence, we have the waves after the ocular lens as $\tilde{l}_S(\mathbf{x}) = l_S(\mathbf{x}) * t_S(\mathbf{x})$ and $\tilde{l}_D(\mathbf{x}) = l_D(\mathbf{x}) * t_D(\mathbf{x})$, where $*$ denotes the convolution operator. $t_S(\cdot)$ and $t_D(\cdot)$ are the inverse Fourier transform of $T_S(\cdot)$ and $T_D(\cdot)$, respectively

$$t_S(\mathbf{x}) = i\zeta_p \delta(\mathbf{x}) \tag{5}$$

$$t_D(\mathbf{x}) = \delta(\mathbf{x}) + (i\zeta_p - 1)\text{airy}(r) \tag{6}$$

where $\delta(\cdot)$ is a Dirac delta function, $r = \sqrt{u^2 + v^2}$ is the radial distance and $\text{airy}(r)$ is an obscured Airy pattern [1]

$$\text{airy}(r) = R \frac{J_1(2\pi Rr)}{r} - (R - W) \frac{J_1(2\pi(R - W)r)}{r} \tag{7}$$

and $J_1(\cdot)$ is the first order Bessel function of the first kind. Fig. 2 shows a sample of obscured Airy pattern where a bright region in the center is surrounded by a series of concentric bright/dark rings.

Substituting l_S, t_S, l_D, t_D into \tilde{l}_S and $\tilde{l}_D(\mathbf{x})$, we get

$$\tilde{l}_S(\mathbf{x}) = i\zeta_p A \tag{8}$$

$$\tilde{l}_D(\mathbf{x}) = \zeta_c A e^{-if(\mathbf{x})} + (i\zeta_p - 1)\zeta_c A e^{-if(\mathbf{x})} * \text{airy}(r) \tag{9}$$

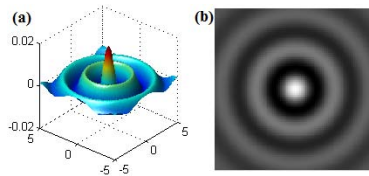


Fig. 2. An obscured Airy pattern. (a) 3D surface view; (b) 2D view.

The first term in Eq.9 is the primary component of the diffracted wave that destructively interferes with the non-diffracted surround wave and generates the contrast for human observation. The second term in Eq.9 comes from the diffracted wave leaking into the phase ring which causes the halo and shade-off artifacts. The intensity of the final observed image is computed as

$$g(\mathbf{x}) = |\tilde{l}_S(\mathbf{x}) - \tilde{l}_D(\mathbf{x})|^2 \quad (10)$$

$$\propto (\delta(r) - \text{airy}(r)) * f(\mathbf{x}) + C \quad (11)$$

where C is a constant. The convolution kernel in Eq.11 represents the point spread function (PSF) of the phase contrast microscope

$$\text{PSF}(u, v) = \delta(u, v) - \text{airy}(\sqrt{u^2 + v^2}) \quad (12)$$

which is a linear operator.

Now we can define the linear imaging model between \mathbf{g} (observed image) and \mathbf{f} (artifact-free phase contrast image to be restored) as

$$\mathbf{g} \approx \mathbf{H}\mathbf{f} + C \quad (13)$$

In practice, we discretize the PSF kernel as a $(2M+1) \times (2M+1)$ matrix (e.g. $M = 5$) and the \mathbf{H} matrix is defined by

$$(\mathbf{H}\mathbf{f})_j = \sum_{u=1}^{2M+1} \sum_{v=1}^{2M+1} \text{PSF}(u, v) \mathbf{f}(\mathbf{x}_j^r + u - M, \mathbf{x}_j^c + v - M) \quad (14)$$

3 Restoring Artifact-Free Phase Contrast Image

Now that the phase-contrast microscopy imaging model is established as in Eq.13, we develop procedures in this section to restore \mathbf{f} from \mathbf{g} . The first step is to remove the background C from \mathbf{g} by flat-field correction [10], after which we re-define \mathbf{g} as the corrected image, i.e. $\mathbf{g} \leftarrow \mathbf{g} - C$. The second and major step is to solve \mathbf{f} from $\mathbf{g} \approx \mathbf{H}\mathbf{f}$. An attempt to solve this by simply inverting \mathbf{H} is known to be highly noise-prone. Instead, we formulate the following constrained quadratic function to restore \mathbf{f}

$$\mathbf{O}(\mathbf{f}) = \|\mathbf{H}\mathbf{f} - \mathbf{g}\|_2^2 + \omega_s \mathbf{f}^T \mathbf{L} \mathbf{f} + \omega_r \|\mathbf{\Lambda} \mathbf{f}\|_p^p \quad (15)$$

where \mathbf{L} and $\mathbf{\Lambda}$ are Laplacian matrix and diagonal matrix defining the local smoothness and sparseness with corresponding weights ω_s and ω_r , and $\|\cdot\|_p$ denotes the l_p -norm. A similarity-based Laplacian matrix is defined as $\mathbf{L} = \mathbf{D} - \mathbf{W}$ where \mathbf{W} is a symmetric matrix whose off-diagonal elements are defined as $\mathbf{W}(i, j) = e^{-(\mathbf{g}_i - \mathbf{g}_j)^2 / \sigma^2}$. Typically, σ^2 is chosen by hand [4] or computed as the mean of all $(\mathbf{g}_i - \mathbf{g}_j)^2$'s. \mathbf{D} is a diagonal degree matrix where $\mathbf{D}(i, i) = \sum_j \mathbf{W}(i, j)$. The image matting Laplacian matrix proposed by Levin *et al.* [7] can also be applied here.

When using l_2 sparseness regularization, $\mathbf{O}(\mathbf{f})$ can be expressed as

$$\mathbf{O}(\mathbf{f}) = \mathbf{f}^T \mathbf{Q} \mathbf{f} - 2\mathbf{b}^T \mathbf{f} + \mathbf{g}^T \mathbf{g} \quad (16)$$

where

$$\mathbf{Q} = \mathbf{H}^T \mathbf{H} + \omega_s \mathbf{L} + \omega_r \mathbf{\Lambda}^T \mathbf{\Lambda} \quad (17)$$

$$\mathbf{b} = \mathbf{H}^T \mathbf{g} \quad (18)$$

A closed-form solution of \mathbf{f} can be achieved by solving a sparse linear system of equations $\mathbf{Q}\mathbf{f} = \mathbf{b}$. Available solvers include conjugate gradient method or Matlab backslash operator (used in [47]). When using l_1 regularization, there is no closed-form solution and only numerical approximation is available. We combine non-negative multiplicative update [11] and re-weighting techniques [2] to solve the l_1 regularized objective function [9].

4 Experiments

Fig. 3 shows several restored samples by solving the l_1 -regularized quadratic problem based on the derived imaging model. We will discuss the effects of different regularization terms and corresponding solvers in Section 4.1. In Fig. 3, it appears that the restored artifact-free images are easier to be segmented because the cells are represented by bright (positive-valued) pixels set on a uniformly black (zero-valued) background. To see if this is the case, we have done quantitative evaluation of segmentation by using restored images in Section 4.2.

4.1 Smoothness and l_p Sparseness Regularizations

When using l_2 sparseness regularization in Eq 15, we can restore a closed-form solution \mathbf{f} (Fig. 4(a)) from \mathbf{g} (Fig 3(d)). This closed-form solution has both positive and negative values corresponding to cell pixels (Fig. 4(b)) and halo/mitotic cell pixels (Fig. 4(c)) respectively. Comparing Fig. 3(i) with Fig. 4(b), we find that l_2 regularization does not enhance sparseness as well as l_1 regularization does. This phenomenon is also discussed in [6]. When there is no sparse constraint, the restored image includes many background pixels (Fig. 4(d)). When there is no smoothness constraint, gaps between object parts such as cell nuclei and membrane may appear (Fig. 4(e)). Without smooth or sparse constraints, the directly solved \mathbf{f} by inverting \mathbf{H} is quite noisy (Fig. 4(f)).

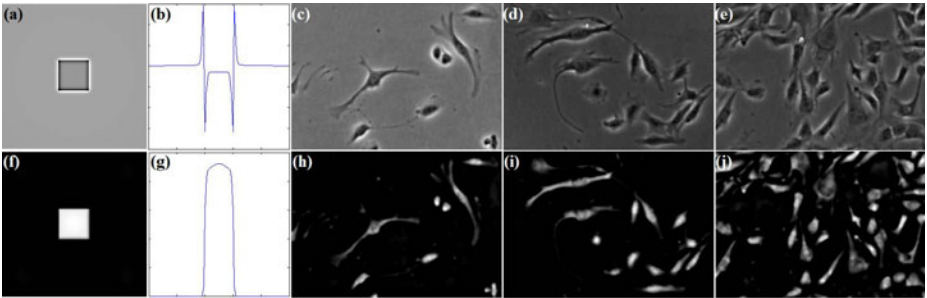


Fig. 3. Restore artifact-free phase contrast images. (a) synthesized phase contrast image; (b) intensity profile of the central row of (a) (or one may think of it as a 1D phase contrast image); (c)-(e) real phase contrast microscopy images with increasing cell densities; (f)-(j) restored artifact-free images corresponding to (a)-(e).

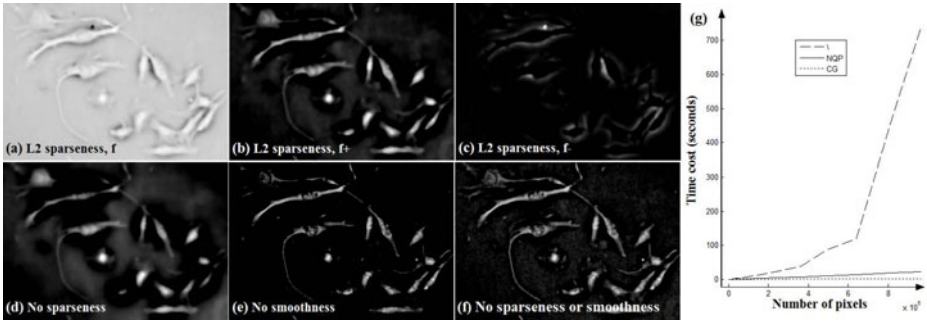


Fig. 4. Sparseness and smoothness. (a) restored image by applying conjugate gradient method (CG) to solve $O(\mathbf{f})$ with l_2 sparseness regularization; (b) and (c) the positive and negative part of (a); (d) restored image without sparseness; (e) restored image without smoothness; (f) restored image without sparseness or smoothness; (g) computational cost comparison.

In terms of computational cost (Fig. 4(g)), the nonnegative quadratic programming solver (NQP) for l_1 -regularized $O(\mathbf{f})$ is slower than conjugate gradient method (CG) for l_2 -regularized $O(\mathbf{f})$. On a workstation with Intel Xeon X5550 CPU and 24G memory, it costs 23 seconds for NQP solver to restore an image with 1000×1000 pixels using Matlab while CG costs 1.9 seconds. The Matlab backslash operator costs much more time and memory. In the following evaluation, we use NQP solver for l_1 -regularized $O(\mathbf{f})$.

4.2 Effects of Restoration on Segmentation

Data. Two sequences were captured at the resolution of 1040×1392 pixels per image. Seq1: C2C12 muscle stem cells proliferated from 30+ to 600+ cells (imaged by ZEISS Axiovert 135TV phase contrast microscope at 5X magnification over 80 hours, 1000 images, Fig. 5(a,c)). Seq2: hundreds of bovine vascular cells migrated to the central image region (imaged by Leica DMI 6000B phase contrast microscope at 10X magnification over 16 hours, 200 images, Fig. 5(f,h)).

Metrics. We denote cell and background pixels as positive (P) and negative (N) respectively. The true positive rate is defined as $\text{TPR} = |\text{TP}|/|P|$ where true positive (TP) stands for those cell pixels correctly labelled by both skilled human and our method. The false positive rate is defined as $\text{FPR} = |\text{FP}|/|N|$ where false positive (FP) are those cell pixels labelled by our method mistakenly. The accuracy is defined as $\text{ACC} = (|\text{TP}| + |N| - |\text{FP}|)/(|P| + |N|)$.

Parameters. We normalized the restored images onto value range $[0, 1]$. Based on a training pair of restored image and its ground truth mask (Seq1 uses the 500th image and Seq2 uses the 100th image), we applied a series of values between zero and one to threshold the restored image and compared with the ground truth to compute TPR and FPR scores, which provided a ROC curve.

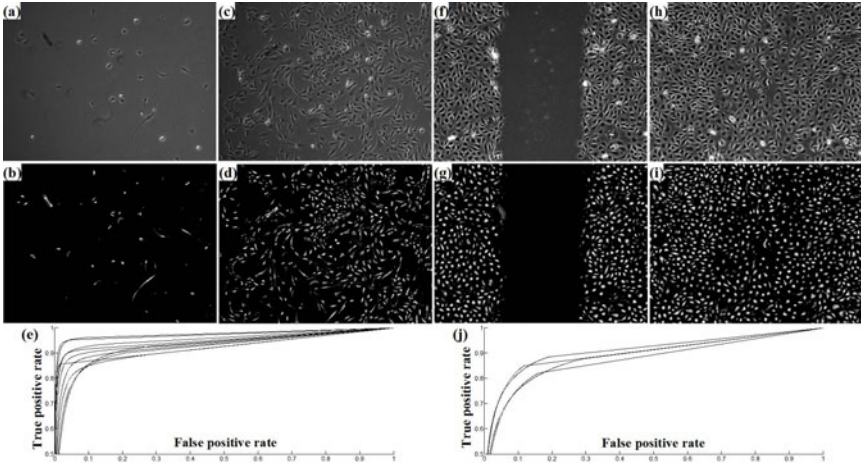


Fig. 5. Quantitative evaluation. (a-b, c-d, f-g, h-i): the pair of observed and restored images; (e) and (j): ROC curves for sequence 1 and 2 respectively.

Different parameter sets in the objective function generate different ROC curves. We searched the optimal parameter set ($\omega_s = 1$ and $\omega_r = .001$ in the evaluation) that has the largest area under the ROC curve (AUC). For the curve with the largest AUC, we searched the threshold that has the highest ACC to segment the restored image into a binary mask (both Seq1 and Seq2 got the best threshold equal to 0.22). We applied the learned parameters to all other images.

Evaluation. We manually labelled every 100^{th} image in Seq1 (2369 annotated cells, 8.6×10^5 cell pixels) and every 50^{th} image in Seq2 (2918 annotated cells, 1.1×10^6 cell pixels). It took human experts about 2 hours to label one thousand cell boundaries in an image. Fig. 5 shows some input and restored images with all ROC curves shown in Fig. 5(e) and (j), where ROC curves deviate gradually from the perfect top-left corner (AUC=1) as cell density increases. The average AUC is 94.2% (Seq1) and 89.2% (Seq2), and the average segmentation accuracy is 97.1% (Seq1) and 90.7% (Seq2).

5 Conclusion

We derived a linear imaging model representing the optical properties of phase contrast microscopes. Using our model, authentic artifact-free phase contrast images are restored by solving a constrained quadratic optimization problem. This work suggests that a better understanding of the optics of microscopes leads to better microscopy image analysis and interpretation. In particular, we have demonstrated that this approach can greatly facilitate the effort on microscopy image segmentation.

References

1. Born, M., Wolf, E.: Principles of Optics, 6th edn. Pergamon Press, Oxford (1980)
2. Candes, E., Wakin, M., Boyd, S.: Enhancing Sparsity by Reweighted ℓ_1 Minimization. *The Journal of Fourier Analysis and Applications* 14(5), 877–905 (2008)
3. Gaskill, J.: *Linear Systems, Fourier Transforms, and Optics*. Wiley, Chichester (1978)
4. Grady, L.: Random Walks for Image Segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 28(11), 1768–1783 (2006)
5. House, D., et al.: Tracking of Cell Populations to Understand their Spatio-Temporal Behavior in Response to Physical stimuli. In: *MMBIA Workshop* (2009)
6. Kim, S., et al.: An Interior-Point Method for Large-Scale ℓ_1 -Regularized Least Squares. *IEEE Journal of Selected Topics in Signal Processing* 1(4) (2007)
7. Levin, A., et al.: A Closed Form Solution to Natural Image Matting. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 30(2), 228–242 (2008)
8. Li, K., et al.: Cell Population Tracking and Lineage Construction with Spatiotemporal Context. *Medical Image Analysis* 12(5), 546–566 (2008)
9. Li, K., Kanade, T.: Nonnegative Mixed-Norm Preconditioning for Microscopy Image Segmentation. In: Prince, J.L., Pham, D.L., Myers, K.J. (eds.) *IPMI 2009*. LNCS, vol. 5636, Springer, Heidelberg (2009)
10. Murphy, D.: *Fundamentals of Light Microscopy and Electronic Imaging*. Wiley, Chichester (2001)
11. Sha, F., et al.: Multiplicative Updates for Nonnegative Quadratic Programming. *Neural Computation* 19(8), 2004–2031 (2007)
12. Smith, K., et al.: General Constraints for Batch Multiple-Target Tracking Applied to Large-Scale Videomicroscopy. In: *Proceedings of CVPR* (2008)
13. Wu, K., Gauthier, D., Levine, M.: Live Cell Image Segmentation. *IEEE Trans. on Biomedical Engineering* 42(1), 1–12 (1995)
14. Yang, F., et al.: Cell Segmentation, Tracking, and Mitosis Detection Using Temporal Context. In: Duncan, J.S., Gerig, G. (eds.) *MICCAI 2005*. LNCS, vol. 3749, pp. 302–309. Springer, Heidelberg (2005)
15. Zernike, F.: How I discovered phase contrast. *Science* 121, 345–349 (1955)

Complete Valvular Heart Apparatus Model from 4D Cardiac CT

Saša Grbić^{1,2}, Razvan Ionasec^{1,2}, Dime Vitanovski², Ingmar Voigt², Yang Wang², Bogdan Georgescu², Nassir Navab¹, and Dorin Comaniciu²

¹ Computer Aided Medical Procedures, Technical University Munich, Germany

² Integrated Data Systems, Siemens Corporate Research, Princeton, USA

Abstract. The cardiac valvular apparatus, composed of the aortic, mitral, pulmonary and tricuspid valve, is an essential part of the anatomical, functional and hemodynamic mechanism of the heart and the cardiovascular system as a whole. Valvular heart diseases often involve multiple dysfunctions and require joint assessment and therapy of the valves. In this paper, we propose a complete and modular patient-specific model of the cardiac valvular apparatus estimated from 4D cardiac CT data. A new constrained Multi-linear Shape Model (cMSM), conditioned by anatomical measurements, is introduced to represent the complex spatio-temporal variation of the heart valves. The cMSM is exploited within a learning-based framework to efficiently estimate the patient-specific valve parameters from cine images. Experiments on 64 4D cardiac CT studies demonstrate the performance and clinical potential of the proposed method. To the best of our knowledge, it is the first time cardiologists and cardiac surgeons can benefit from an automatic quantitative evaluation of the complete valvular apparatus based on non-invasive imaging techniques. In conjunction with existent patient-specific chamber models, the presented valvular model enables personalized computation modeling and realistic simulation of the entire cardiac system.

1 Introduction

The heart valves are an essential anatomical structure regulating the heart chambers hemodynamics and the blood flow between the heart and the human body. Valvular heart disease (VHD) is a frequent symptom, with a prevalence of 2.5% among the global population often obliged to complex management and therapeutical procedures. Moreover, valve operations are the most expensive and most precarious cardiac interventions, with an average cost of \$141,120 and 4.9% in-hospital death rate in the US [1].

Due to the strong anatomical, functional and hemodynamic inter-dependency of the heart valves, VHDs do not affect only one valve, but rather several valves are impaired. Recent studies demonstrate strong influence of pulmonary artery systolic pressure on the tricuspid regurgitation severity [2]. In [3,4] the simultaneous evaluation of aortic and mitral valves is encouraged, given the fibrous aortic-mitral continuity, which anchors the left side valves and facilitates the reciprocal

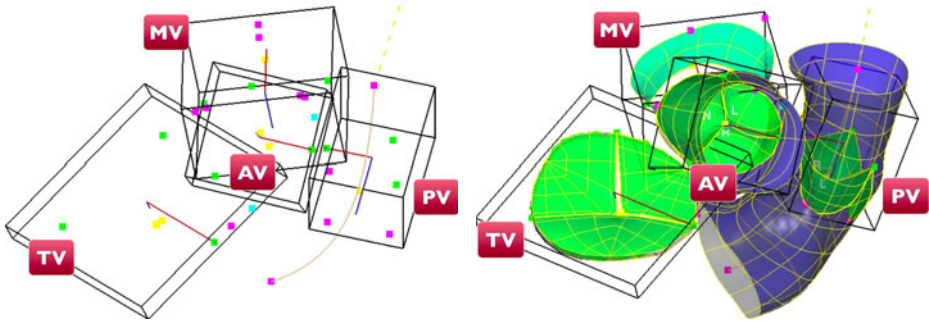


Fig. 1. Complete heart valves model consisting of aortic valve (AV), mitral valve (MV), pulmonary valve (PV) and tricuspid valve (TV). **Left:** Similarity transform illustrated as a bounding box and anatomical landmarks. **Right** Complete mesh surface model.

opening and closing motion during the cardiac cycle. Moreover, in patients with mitral and tricuspid valve regurgitation, joint surgery is recommended to minimize the risk for successive heart failure or reduced functional capacity. Complex procedures mostly performed in patients with congenital heart diseases, such as the Ross operation, in which the aortic valve is replaced with the pulmonary, require comprehensive quantitative and qualitative evaluation of the heart valves. Morphological and functional assessment of the complete heart valve apparatus is crucial for clinical decision making during diagnosis and severity assessment as well as treatment selection and planning.

An increased holistic view of the heart, demanded by clinicians is in perfect accordance with the tremendous scientific effort worldwide, such as the Virtual Physiological Human project [5], geared towards multi-scale physiological modeling and simulation, which will promote personalized, preventive and predictive healthcare. However, the majority of cardiac models to date focus on representation of the left or right ventricle [6], while very few include the left and right atrium [7], but none explicitly handles the valves. Recently introduced models of the aortic-mitral coupling [8,9] do not incorporate the right side valves. A critical component for a patient-specific computation model of the entire heart and realistic cardiovascular simulations, which was not reported yet in the literature, is a personalized and complete representation of the valvular apparatus. In this paper we propose a complete and modular model of the heart valves comprising the anatomy of the aortic, mitral, tricuspid and pulmonary valves as well as their morphological, functional and pathological variations. The patient-specific valvular apparatus is estimated non-invasively from 4D Computed Tomography images, using a discriminative learning-based approach. Global valve location and motion, as well as the non-rigid movement of anatomical valvular landmarks, are computed within the Marginal Space Learning (MSL) [10] and Trajectory Spectrum Learning (TSL) [8] frameworks. A novel anatomical Constrained Multi-linear Shape Model (cMSM) is introduced to capture complex spatio-temporal statistics, and in conjunction with robust boundary detectors,

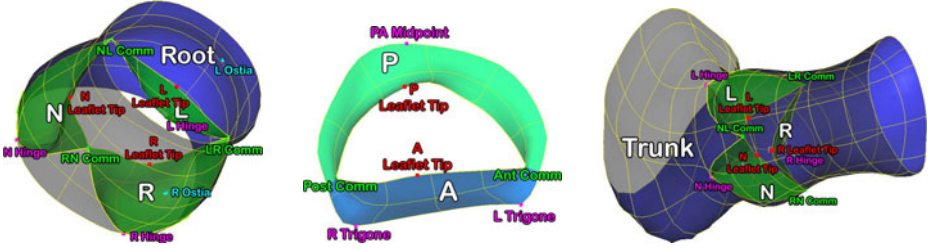


Fig. 2. Anatomical Landmark Model and Complete Valve Model of the aortic valve (left), mitral valve (middle) and pulmonary valve (right)

it is applied to extract the complete valvular anatomy and motion. Our approach enables for integrated quantification and treatment management of heart valves, and together with a patient specific chamber model [10], it will enable comprehensive heart simulations [11].

2 Physiological Model of the Heart Valves

In this section we introduce the complete heart valves model, which includes the aortic, mitral, tricuspid and pulmonary valves, and captures morphological, functional and pathological variations. To reduce anatomical complexity and facilitate effective estimation, the heart valve model is represented on two abstraction layers: Anatomical Landmark Model - which represents the global location and motion of each valve, as well as the motion of the corresponding anatomic landmarks; Complete Valve Model - which parameterizes the full anatomy and dynamics of the valves using dense surface meshes.

Anatomical Landmark Model. The global dynamic variation of each valve is parametrized as a temporal dependent similarity transform (see Fig. 1).

$$\mathbf{B}_t = \{(x_1, x_2, x_3), (q_1, q_2, q_3), (s_1, s_2, s_3)\} \quad t \in 1 \dots T \quad (1)$$

(x_1, x_2, x_3) is the translation, (q_1, q_2, q_3) the quaternion representation of the rotation, (s_1, s_2, s_3) the similarity transform scaling factors and t the temporal position in the cardiac cycle. A set of 33 anatomical landmarks, described in the next paragraph, are used to parameterize the complex and synchronized motion pattern of all valves, which explains the non-linearities of the hemodynamic movements. Thereby, each landmark is described by a T time-step trajectory in a three dimensional space, normalized by the temporal dependent similarity transform \mathbf{B} :

$$\mathbf{L}_n(\mathbf{B}) = \{l_1, l_2, \dots, l_T\} \quad n \in 1 \dots 33 \quad l_i \in \mathbb{R}^3 \quad (2)$$

Complete Valve Model. The final valves model is completed with a set of 13 dense surface meshes. Each mesh is sampled along anatomical grids of vertices defined through the landmarks:

$$V_q(\mathbf{L}, \mathbf{B}) = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K\} \quad q \in 1 \dots 13 \quad \mathbf{v}_i \in \mathbb{R}^3 \quad (3)$$

where \mathbf{v}_i are the vertices, and K is the total number of vertices of mesh q .

Aortic valve: The aortic valve is constrained by 11 landmarks (3 commissures, 3 hinges, 3 leaflet tips and 2 ostias) and four anatomical structures (aortic root, N-, L- and R-leaflet). The aortic root is constrained by the hinge and commissure plane and each leaflet is spanned between two commissures and one hinge, as in [\[8\]](#), (see Fig. [2](#) (left)).

Mitral valve: The mitral valve is composed of 7 landmarks (3 trigones, 2 commissures and 2 leaflet tips). The anterior leaflet is defined by two trigons, one leaflet tip and two commissures and the posterior leaflet by three trigons, one leaflet tip and one commissure, as in [\[8\]](#), (see Fig. [2](#) (middle)).

Pulmonary valve: The pulmonary valve is consisting of 9 landmarks (3 commissures, 3 hinges and 3 leaflet tips) and four anatomical structures (pulmonary root, N-, L- and R-leaflet), as in [\[8\]](#), (see Fig. [2](#) (right)).

Tricuspid valve: The function of the tricuspid valve is to regulate the blood flow from the right atrium to the right ventricle, staying closed during systole and opened during diastole. The model is constrained by four surface geometries (annulus, septal-, anterior- and posterior leaflet) (Fig. [3](#)) and six anatomical landmarks (three commissures and three leaflet tips) which are corresponding to vertices on the meshes. The tricuspid annulus is represented as a surface mesh constrained by the three commissures. The tricuspid leaflets, the A, S and P, are modeled as hyperbolic paraboloids. The vertices on the leaflet border have fixed correspondences to vertices on the tricuspid annulus. Each leaflet is defined by two commissures and one leaflet tip.

3 Patient-Specific Model Estimation

A hierarchical approach is utilized to estimate model parameters, introduced in the previous section, from 4D cardiac CT images. First, robust machine learning techniques are applied to estimate the global valves and anatomic landmarks parameters introduced in Eq. [1](#) and [2](#). Second, we present the novel anatomical constrained Multi-linear Shape Model (cMSM), which effectively captures the complex spatio-temporal variation of all valves. Lastly, the cMSM is applied in a learning-based framework to estimate the complete valve model described in Eq. [3](#).

3.1 Landmark Location and Motion Estimation

The global motion and landmarks parameters are estimated within the Marginal Space Learning framework [\[7\]](#), similar as in [\[8\]](#).

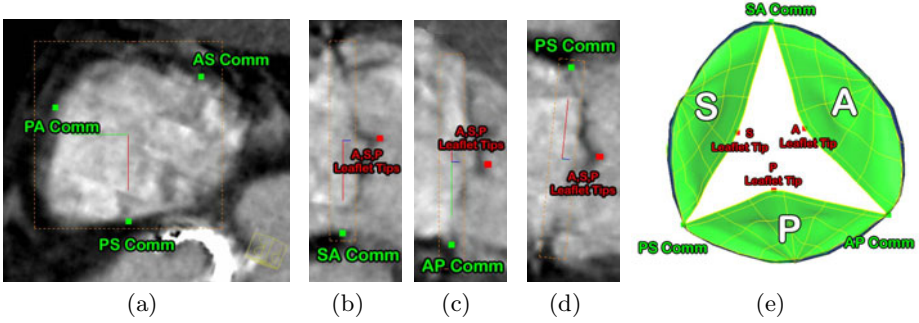


Fig. 3. Tricuspid valve model. (a)-(d) shows anatomical landmarks (commissures and leaflet tips) and (e) is showing the complete surface mesh (annulus and leaflets).

3.2 Constrained Multi-linear Shape Model

Multilinear modeling enables the decomposition of a shape space in a temporal and spatial component in contrast to active shape models (ASM) where both are coupled. In this paper we present a MPCA (Multi-linear Principle Component Analysis) and MICA (Multi-linear Independent Component Analysis) shape model of all valves which is conditioned by anatomical measurements.

Shape Space. In order to construct the shape model all shapes \mathbf{V} are aligned by calculating the mean sequence model and aligning them using general procrustes analysis (GPA). The normalized shapes are represented as third-order tensors $\mathbf{D} \in \mathbb{R}^{(S \times T \times P)}$, where S is the number of patients, T is the frame number inside a multi phase sequence and P represents the number of shape points.

$$\mathbf{D} = \mathcal{Z} \times_1 \mathbf{U}_{\text{patient}} \times_2 \mathbf{U}_{\text{motion}} \times_3 \mathbf{U}_{\text{points}}$$

As mentioned by [12] the motion subspace due its non-Gaussian properties is decomposed using ICA and the patient and points space using PCA. We use the fixed point algorithm to perform the Independent Component Analysis [13].

$$\begin{aligned} \mathcal{D} &= \mathcal{Z} \times_1 \mathbf{U}_{\text{patient}} \times_2 \mathbf{U}_{\text{motion}} \mathbf{W}^{-1} \mathbf{W} \times_3 \mathbf{U}_{\text{points}} \\ &= (\mathcal{Z} \times_2 \mathbf{W}_1) \times_1 \mathbf{U}_{\text{patient}} \times_2 \mathbf{U}_{\text{motion}} \mathbf{W}^{-1} \times_3 \mathbf{U}_{\text{points}} \\ &= \mathcal{S} \times_1 \mathbf{U}_{\text{patient}} \times_2 \mathbf{C}_{\text{motion}} \times_3 \mathbf{U}_{\text{points}} \end{aligned}$$

Constrained Model Estimation. A set of anatomical measurements \mathbf{M} (m_1, m_2, \dots, m_R) extracted from the non-linear valve model is used to condition a surface parameterization \mathbf{V}_q ($\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K$) [14]. Assuming the joint multivariate distribution ($\mathbf{V}_q | \mathbf{M}$) follows a Gaussian distribution a conditioned surface \mathbf{V}_q^M , containing the anatomical measurements \mathbf{M} , can be estimated as follows:

$$\mathbf{V}_q^M = \mu_{\mathbf{V}_q} + \Sigma_{\mathbf{V}_q \mathbf{M}} \Sigma_{\mathbf{M} \mathbf{M}}^{-1} (\mathbf{M} - \mu_{\mathbf{M}})$$

where μ_{V_q} is the mean surface parameterization from all training sets of the valve surface V_q , μ_M the mean of the measurements M in the training set, $\Sigma_{V_q M}$ the covariance matrix between V_q and M . The constrained surface V_q^M is used to reconstruct the dynamic motion surface model of the whole sequence. Therefore we first estimate the patient modes $\mathbf{u}_{\text{patient}}$ and then use them to reconstruct $V_q(L, B)$.

$$\mathbf{u}_{\text{patient}} = V_q^M \mathbf{T}_{(1)}^{-1} \quad \mathcal{T} = \mathcal{S} \times_2 \mathbf{C}_{\text{motion}} \times_3 \mathbf{U}_{\text{points}}$$

where $\mathbf{T}_{(1)}^{-1}$ is the pseudo-inverse of the tensor \mathcal{T} flattened along the first mode and $\mathbf{C}_{\text{motion}}$ the one dimensional motion mode. The complete surface model for the complete sequence, introduced in Eq. 3, can be extracted by a tensor multiplication:

$$V_q(L, B) = \mathcal{S} \times_1 \mathbf{u}_{\text{patient}} \times_2 \mathbf{C}_{\text{motion}} \times_3 \mathbf{U}_{\text{points}}$$

3.3 Complete Valve Model Estimation

The comprehensive surface model $V_q(L, B)$ of each anatomical structure q , is estimated by adopting the constrained Multi-linear shape method using anatomical measurements M defined between the landmarks $L_1 \cdots L_{33}$. E.g. for the aortic root we use three measurements $M = \{m_1, m_2, m_3\}$ (m_1 -inter-commissure distance, m_2 -hinge-commissure plane distance, m_3 -hinge-commissure plane angle). The shape prediction $V_q(L, B)$ is locally refined using a boundary detector, trained using PBT [15] and Steerable features [16].

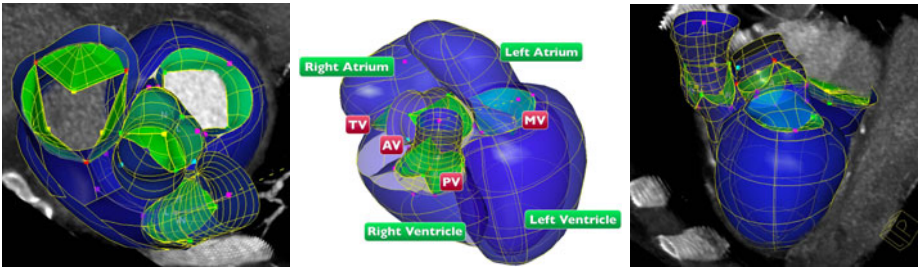


Fig. 4. Examples of estimated personalized model from a multiphase CT sequence

4 Experimental Results

The accuracy of the proposed method was evaluated using 64 multi phase CT data sets, containing 640 volumes. The data sets comprise a variety of cardiovascular diseases and due to different acquisition protocols they have heterogeneous image quality, resolution and sizes. The ground-truth for training and testing

Table 1. System precision for valve model estimation averaged on all valves. **Left:** Evaluation of the Anatomical Landmark Model. **Right:** Comprehensive Surface assessment.

	Mean	STD	Median
Similarity T. [mm]	5.67	2.36	5.29
Landmark Motion [mm]	2.61	1.17	2.41

	Mean	STD	Median
Tracking by Detection [mm]	1.52	0.98	1.47
ML PCA-ICA [mm]	1.39	0.91	1.32
cML PCA-ICA [mm]	1.24	0.91	1.18

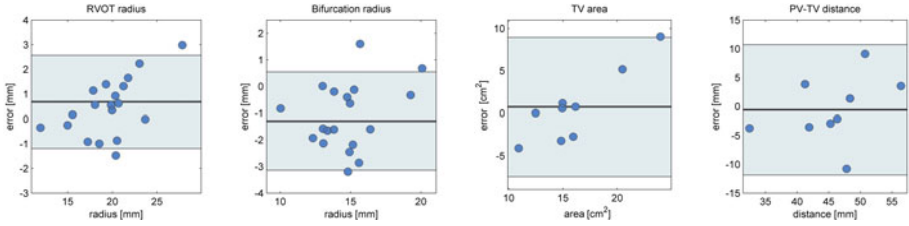


Fig. 5. Bland altman plots for a) right ventricle output tract diameter, b) pulmonary valve bifurcation diameter, c) tricuspid valve area and d) distance between pulmonary and tricuspid valve

was obtained through an incremental annotation process guided by experts, which includes the manual placement of anatomical landmarks and delineation of valves surface over the entire cardiac cycle of each study. Each evaluation is done using three-fold cross validation.

The estimation of the Anatomical Landmark Model is performed in two stages. First the global dynamic motion is evaluated by the Euclidian distance between the corner points of the ground-truth and estimated box in each frame. Second, the non-linear landmark model error is computed as the mean Euclidian distance of all detected landmarks compared to the expert annotation. The complete valve model, containing the dense surface meshes, is measured as the mesh-to-mesh distance. Results shown in table 1 corroborate that our constrained ML PCA-ICA shape estimation approach achieves best performance, compared to a regular ML PCA-ICA method and a standard frame-wise estimation procedure (tracking by detection). Within three minutes a complete personalized dynamic model of all valves is estimated with an average accuracy of 1.24 mm. The full valvular model together with the four chambers of the heart is illustrated in Fig. 4. Important clinical parameters are extracted from the personalized model in the right heart. They include right-ventricle outflow tract (RVOT) radius, Bifurcation radius, tricuspid valve area and a joint measurement of the two valves, the pulmonary and tricuspid valve distance. Quantitative comparison is shown in Fig. 5 by comparing ground truth measurements and the estimated, demonstrating a strong correlation.

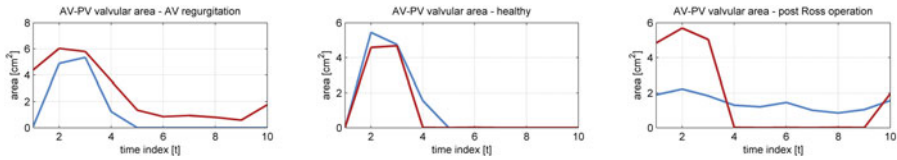


Fig. 6. Measurements of aortic (AV) and pulmonary valve (PV) area obtained from a patient with aortic valve regurgitation (**left**), a healthy patient (**middle**) and a post Ross operation patient (**right**). The red graph is representing the aortic valve and the blue the pulmonary.

Finally we show quantitative comparison between a patient suffering from aortic valve regurgitation, a healthy patient and a post-operative patient who underwent a Ross operation. An important clinical measurement, the valvular area, extracted from the personalized aortic and pulmonary valve model, demonstrated in Fig. 6, confirms a successful outcome since no regurgitation is observed at the aortic valve.

5 Conclusion

In this paper, we propose a novel personalized model for quantitative and qualitative evaluation of the complete heart valve apparatus in 4D CT. It is capable to delineate the full anatomy and dynamics needed to depict a large variation of valve pathologies, especially diseases affecting several valves. Its hierarchical approach using state of the art machine learning algorithms in combination with a constrained Multi-linear shape space enables an patient specific model estimation within three minutes and an accuracy of 1.24 mm. Clinical validation shows strong correspondence between expert- and model estimated anatomical measurements. The proposed method could spark research in many areas such as examining interconnections of multiple valves and hemodynamic simulation of the complete cardiac apparatus.

References

1. Lloyd-Jones, D., et al.: Heart disease and stroke statistics–2009 update: a report from the american heart association statistics committee and stroke statistics subcommittee. *Circulation* 119, 21–181 (2009)
2. Mutlak, D., et al.: Functional tricuspid regurgitation in patients with pulmonary hypertension. In: *CHEST*, pp. 115–121 (2009)
3. Lansac, E., et al.: Dynamic balance of the aortomitral junction. *J. Thorac. Cardiovasc. Surg.* 123, 911–918 (2002)
4. Timek, T., et al.: Aorto-mitral annular dynamics. *Ann. Thorac. Surg.* 76, 1944–1950 (2003)
5. Clapworthy, G., et al.: Editorial. *Philosophical Transactions of the Royal Society A: Mathematical. Physical and Engineering Sciences* 366(1878), 2975–2978 (2008)

6. Fritz, D., et al.: Segmentation of the left and right cardiac ventricle using a combined bi-temporal statistical model. In: SPIE Medical Imaging, pp. 605–614 (2006)
7. Zheng, Y., et al.: Four-chamber heart modeling and automatic segmentation for 3d cardiac ct volumes using marginal space learning and steerable features. *IEEE TMI* 27(11), 1668–1681 (2008)
8. Ionasec, R.I., et al.: Personalized Modeling and Assessment of the Aortic-Mitral Coupling from 4D TEE and CT. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5762, pp. 767–775. Springer, Heidelberg (2009)
9. Veronesi, F., et al.: A study of functional anatomy of aortic-mitral valve coupling using 3D matrix transesophageal echocardiography. *Circ. Cardiovasc. Imaging* 2(1), 24–31 (2009)
10. Zheng, Y., et al.: Fast automatic heart chamber segmentation from 3d ct data using marginal space learning and steerable features. In: ICCV (2007)
11. Mihalef, V., et al.: Patient-specific modeling of left heart anatomy, dynamics and hemodynamics from high resolution 4d ct. In: ISBI (2010)
12. Zhu, Y., et al.: A dynamical shape prior for lv segmentation from rt3d echocardiography. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5761, pp. 206–213. Springer, Heidelberg (2009)
13. Hyvrine, I., Karhunen, J.: Independent Component Analysis. John Wiley, Chichester (2001)
14. Blanc, R., et al.: Conditional variability of statistical shape models based on surrogate variables. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5762, pp. 84–91. Springer, Heidelberg (2009)
15. Tu, Z.: Probabilistic boosting-tree: Learning discriminative methods for classification, recognition, and clustering. In: ICCV, pp. 1589–1596 (2005)
16. Freeman, W.T., Adelson, E.H.: The design and use of steerable filters. *IEEE Trans. Pattern Anal. Machine Intell.* 13(9), 891–906 (1991)

Improving Deformable Surface Meshes through Omni-Directional Displacements and MRFs

D. Kainmueller, H. Lamecker, H. Seim, S. Zachow, and H.-C. Hege

Zuse Institute Berlin, Germany
kainmueller@zib.de

Abstract. Deformable surface models are often represented as triangular meshes in image segmentation applications. For a fast and easily regularized deformation onto the target object boundary, the vertices of the mesh are commonly moved along line segments (typically surface normals). However, in case of high mesh curvature, these lines may intersect with the target boundary at “non-corresponding” positions, or even not at all. Consequently, certain deformations cannot be achieved. We propose an approach that allows each vertex to move not only along a line segment, but within a surrounding sphere. We achieve globally regularized deformations via Markov Random Field optimization. We demonstrate the potential of our approach with experiments on synthetic data, as well as an evaluation on 2x106 coronoid processes of the mandible in Cone-Beam CTs, and 56 coccyxes (tailbones) in low-resolution CTs.

1 Introduction

Deformable surface models are widely used for image segmentation [1]. Among the many different representations of surfaces, polygonal meshes are advantageous in many respects, such as flexibility or topology preservation [2]. The deformation of the model is often driven by minimizing an energy that consists of an image term that measures how well the model is aligned with features in the image, and a regularization term that controls the smoothness of the model. During energy minimization, the vertices of the mesh are iteratively displaced.

A fundamental challenge of this approach is how to keep the mesh valid, i.e. as regular as possible [3]. One way to confront this problem is to allow vertex displacements only in surface normal direction. Additional measures are adaptive step-size control, adaptive remeshing, special regularization or mesh surgery. The drawbacks of these approaches are that they are often difficult to implement, slow down the convergence of the method drastically, or even lead to situations where the deformation gets stuck. This is especially the case in regions of high curvature. Hierarchical approaches may be able to circumvent this problem: First the model is adapted with strong regularization, e.g. by using shape priors such as statistical shape models [4], which is then gradually relaxed. Although this increases robustness, the last bit of accuracy may only be achieved with a deformation model free of a-priori shape constraints (see e.g. [5]). This is the type of deformation model we focus on.

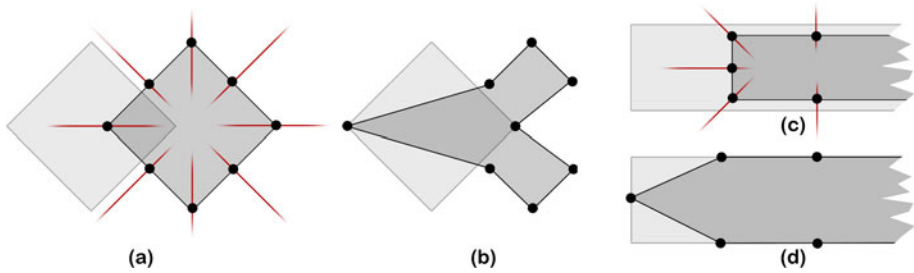


Fig. 1. Normals on a cube (a) and on a tip-like structure (c). Unregularized deformations along normals: No target boundary points (b) or non-corresponding target boundary points (d) found for most vertices.

Normal (or other one-directional) displacements of mesh vertices may either lead to very few features being detected in the image (low *visibility*), or many “wrong” features (in terms of *correspondence*), as depicted in Fig. 1. This induces large distortions of the mesh. Remeshing may restore mesh regularity, while implausible shape deformations are to be remedied by regularization. In summary, certain deformations can hardly be achieved with one-directional displacements, given that directions are not known a-priori. E.g., parallel movements of tip-like structures cannot be achieved along surface normals.

This paper contributes a solution to this problem: We propose a method that allows arbitrary displacement directions at each mesh vertex. This enlarges visibility, while correspondence is likely to be improved, too. However, a larger set of image features may be found per vertex, so global regularization is required to cope with highly inconsistent neighboring displacements. Our method allows displacements to a discrete set of points within a sphere around each vertex. This discrete nature allows us to formulate the segmentation problem as a Markov Random Field (MRF), as shown in Sec. 2. The MRF can be solved efficiently [6], yielding a globally regularized deformation of the mesh. MRFs have been applied to many problems, e.g. image-to-image registration [7], due to their capability of finding good optima. Global regularization has been proposed for one-directional displacements using graph cuts [8]. However, this approach is not applicable in our case as it requires a total order on displacements in terms of “above/below”.

We illustrate on synthetic and real data that omni-directional displacements combined with MRF optimization can handle parallel movements of meshes with high curvature, where previous approaches based on normal displacements fail.

2 Method

We denote the set of vertices $v \in \mathbf{R}^3$ of the deformable surface mesh as V , and the set of pairs of adjacent (i.e. edge-connected) vertices as $N \subset V \times V$. Each vertex can be moved by adding a vector $l \in L$, where $L \subset \mathbf{R}^3$ is a discrete set

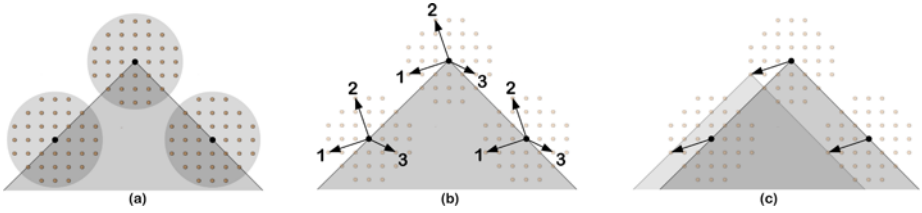


Fig. 2. Three vertices (black dots) on a 2D contour. (a) Omni-directional displacements to yellow/gray dots. (b) Exemplary “same” displacements shown by black arrows with corresponding numbers. (c) Same displacement for all vertices leads to parallel translation.

of possible displacements. We call a position $v + l$ *sample point*, and a mapping $V \rightarrow L, v \mapsto l_v$ that assigns a displacement to each vertex *displacement field*.

2.1 Omni-Directional Displacements

We propose to extend the range of motion for a vertex from a line segment to a sphere centered at its current position. Therefore, we define L as a set of displacements that are uniformly distributed within a sphere (see Fig. 2a). The sphere radius is a parameter of our method. The set of displacements is interpreted in world coordinates for all vertices (see Fig. 2b and 2c). As discussed in Sec. 2.2, this is important for regularization.

2.2 Objective Function

For each displacement $l \in L$ and vertex $v \in V$, a scalar cost $c(v, l) \geq 0$ encodes whether sample point $v + l$ is believed to lie on the object boundary. The stronger the belief, the lower should be the cost. In other words, $c(v, l)$ serves as a penalty for the case that v is displaced by l . We calculate $c(v, l)$ from the image $I : \mathbf{R}^3 \rightarrow \mathbf{R}$. It depends on $I(v + l)$, $\nabla_{n_v} I(v + l)$ (where n_v denotes the surface normal at v), and application-specific parameters (see Sec. 3). In general, our objective function accepts any $c : V \times L \rightarrow \mathbf{R}_0^+$, so c can be defined as appropriate. E.g., if the surface mesh might locally lie perpendicular to the object boundary, derivatives in directions other than n_v may be considered as well.

For each two displacements l_1, l_2 , a scalar distance value $d(l_1, l_2) \geq 0$ serves as a penalty for the case that l_1 and l_2 occur on adjacent vertices. The distance function $d : L \times L \rightarrow \mathbf{R}_0^+$ is supposed to take care of regularization. It has to satisfy $d(l_1, l_2) = 0 \Leftrightarrow l_1 = l_2$, but does not have to be a metric (see Sec. 2.3). It can, e.g., be the Euclidean norm $\|l_2 - l_1\|$ to some power.

We define the objective function of the mesh adaptation problem as follows:

$$\sum_{v \in V} c(v, l_v) + \sum_{(v, w) \in N} d(l_v, l_w) \quad (1)$$

We are looking for the displacement field that minimizes Eq. 1. Note that interpreting displacements in world coordinates yields distance-penalties for scaling the mesh, while parallel translations are not penalized (see Fig. 2c). We consider this beneficial as we expect our initial meshes (as well as its local features) to have approximately correct scale. Alternatively, if scaling should not be penalized, one could interpret displacements in local coordinate systems per vertex.

2.3 Optimal Displacement Field

We encode the objective function in Eq. 1 as an MRF, with vertices being represented by MRF-nodes, mesh edges by MRF-edges, and displacements by the possible states (also called *labels*) of the nodes. Cost $c(v, l)$ defines the unary potential of node v in state l , and distance $d(l_1, l_2)$ defines the binary potential of two adjacent nodes in states l_1, l_2 . The MRF-state with minimal sum of potentials yields the desired displacement field. We optimize the MRF by a method named FastPD [6]. FastPD can deal with non-metric distance functions d as specified in Sec. 2.2. It is guaranteed to find an approximately optimal solution.

3 Results

To evaluate our MRF-based method, we applied it to three types of 3D data: (1) Synthetic binary images, (2) synthetic binary images with various amount of noise, and (3) clinical image data. On synthetic binary images and clinical image data, we also computed results with a globally regularized method [8] (GraphCuts) and a locally regularized method [5] (FreeForm), both employing vertex normals as displacement directions. GraphCuts computes the displacement field with minimum sum of costs while respecting a hard constraint on the difference between the lengths of adjacent displacements. FreeForm takes the minimum cost displacement for each vertex and regularizes locally via a small displacement toward the centroid of the respective adjacent vertices.

For the computation of the costs $c(v, l)$, we used the strategy proposed in [9]: If the intensity $I(v + l)$ lies within a certain window $[i_0, i_1]$, costs are inversely proportional to $\nabla_{n_v} I(v + l)$. Otherwise costs are set to a constant, high value. The thresholds i_0 and i_1 are parameters of the strategy. As distance function d , we used $d(l_1, l_2) = ||l_2 - l_1||^3$ in all experiments. Whenever we employ GraphCuts or FreeForm, we use the same cost function as for the respective MRF experiment, and normal segments with the length of the respective sphere diameter. In contrast to the GraphCuts- and MRF based adaptations, all FreeForm adaptations were performed iteratively, with 30 steps. All GraphCuts experiments were performed with a difference constraint of 2 sample points (i.e. lengths of adjacent displacements can be at most 2 sample points apart).

MRF optimization with FastPD took less than 10 seconds in all our experiments. The computation of the MRF's unary potentials $c(v, l)$ was more time-consuming, taking up to 10 minutes depending on the number of vertices and labels. FastPD required memory up to 8GB for the experiments on clinical data. All experiments were performed on a 3GHz core with 8GB main memory.

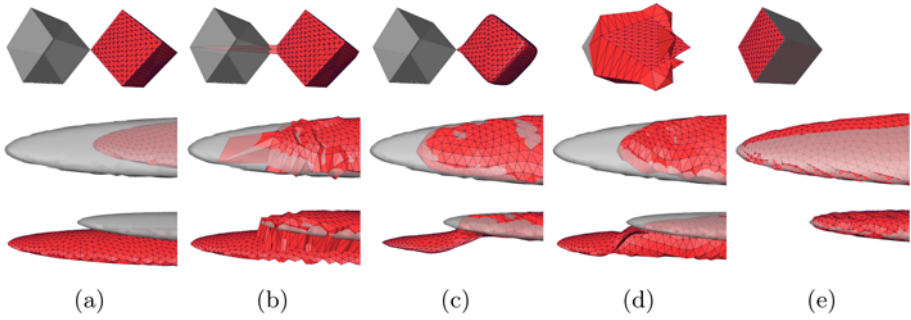


Fig. 3. Deformable mesh (red/dark grey mesh) and target object (transparent gray surface). (a) Initial situation. (b) Displacements along normals without regularization. (c) FreeForm approach. (d) GraphCuts approach. (e) MRF approach.

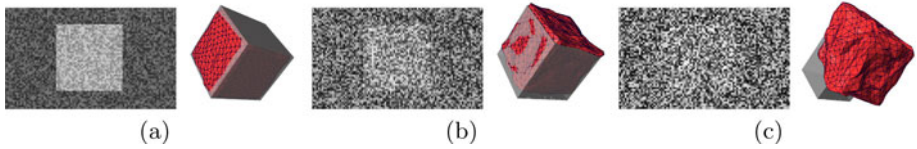


Fig. 4. Performance of MRF approach in the presence of noise. Random noise with range (a) $[-0.5..0.5]$, (b) $[-2.5..2.5]$ and (c) $[-5..5]$. Slices of the image data and respective adaptation result (red/dark grey mesh). Grey surface: ideal target object.

Synthetic Images. We performed experiments on binary images (i.e. intensities $\in \{0, 1\}$) of a cube and a thin ellipsoid. As initial meshes, we used triangulated cube and tip surfaces with ideal shape, but shifted pose (see Fig. 3(a)). The cube mesh had 770, the ellipsoid 1797 vertices. One sphere diameter (or surface normal segment, respectively) was covered by 53 sampling points for the cube, and 63 for the ellipsoid. For all experiments on synthetic data, we chose sphere radii such that the target object boundary was located completely within a band around the initial mesh that has this radius. We set the cost function parameters to $i_0 = 0.1$ and $i_1 = 1.1$. The results of MRF-, FreeForm- and GraphCuts adaptation are shown in Fig. 3(c), 3(d) and 3(e), respectively. Fig. 3(b) shows the results of adding normal displacements without any regularization.

We added various amounts of random noise to the binary cube image and performed MRF based adaptation as before. The cube was detected correctly for noise with ranges $[-0.5..0.5]$ and $[-2.5..2.5]$, and failed for $[-5..5]$. Fig. 4 shows slices of the noisy image data and the respective adaptation results.

Clinical Data. In a quantitative evaluation on 106 mandible Cone-Beam CTs and 50 pelvis CTs we compared MRF, FreeForm and GraphCuts results to gold standard surfaces obtained from manual segmentations. Initial meshes were

automatically generated by Statistical Shape Model adaptation [9,10]. For the mandible surfaces (gold standard as well as adaptation result), we extracted the right coronoid processes as the region of the mesh that lies above 1/2 of the extension of the mandible in transversal direction, between 1/3 and 2/3 of extension in dorsoventral direction, and above 2/3 in longitudinal direction. Extraction of the left coronoid process worked analogously. We identified the tip point as the upmost vertex in longitudinal direction. For the sacrum, we extracted the coccyx as the region of the mesh that lies below 1/3 of the extension of the sacrum in longitudinal direction. We identified the coccyx tip as the vertex with minimum $3 \cdot \text{longitudinal} - 1 \cdot \text{dorsoventral}$ coordinate. As error measures we assessed the tip-to-tip distances (tip2tip), tip-to-surface distances from result tip to gold standard surface (tip2surfRtG) and the other way round (tip2surfGtR), and roots mean square surface distances (rmsRtG and rmsGtR).

All mandible meshes had 8561 vertices, all sacrum meshes 6161. The number of samples along a sphere diameter (or normal segment) was 39 for the mandible and 45 for the sacrum. The sphere diameters were 15 mm (mandible) and 25 mm (sacrum). We set the cost function parameters to $i_0 = 350$ and $i_1 = 800$ (mandible) and $i_0 = 120$ and $i_1 = 520$ (sacrum). For the mandible, we performed a second MRF based adaptation (mrfZ) with a slightly different cost function: We added a small extra cost to sample points with lower longitudinal coordinate, thus slightly preferring the a motion in upward direction.

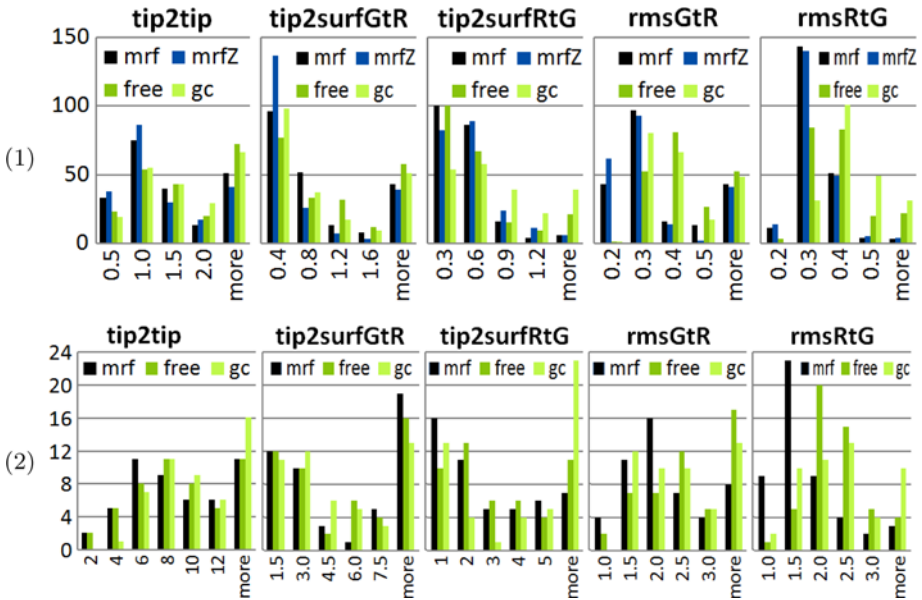


Fig. 5. Row 1: Coronoid process results on 106 individuals (i.e. 212 cases). Row 2: Coccyx results on 50 individuals. X-axis: Error measure in mm (see Sec. 3). Y-axis: Frequency. MRF (mrf, mrfZ), FreeForm (free) and GraphCuts (gc).

Evaluation results are plotted as histograms in Fig. 5. As measurements are not normally distributed, we performed the Wilcoxon signed-rank test to assess significant differences (level 0.05). For the mandible, both MRF results are significantly better than FreeForm and GraphCuts in terms of tip2tip, rmsGtR and rmsRtG, while mrfZ performs significantly better in terms of tip2surfGtR, too. For the sacrum, MRF results are significantly better than GraphCuts in terms of all but the rmsGtR measure. However, compared to FreeForm, the error difference is significant only for the two rms measures.

4 Discussion

Experiments on synthetic binary images show that our MRF-based approach is able to handle parallel translations, in contrast to a globally and a locally regularized approach (GraphCuts and FreeForm) that employ normal displacements. Experiments on noisy synthetic images show that the MRF approach is able to produce well-regularized displacement fields in the presence of noise. However, for a very low signal-to-noise ratio, it may fail to detect the target object. Real-world experiments show that the MRF approach is able to produce very accurate segmentations of tip-like structures. On the mandible tips, the MRF approach clearly outperforms the GraphCuts and FreeForm approach. Here, normal directions often exhibit the “wrong-visibility” problem, see Fig. 6(a-c), which our new method resolves. However, on the sacrum tips, more experiments need to be performed to draw decisive conclusions. At least, the “non-visibility” problem for normals can be resolved in a few exemplary cases, see Fig. 6(d,e).

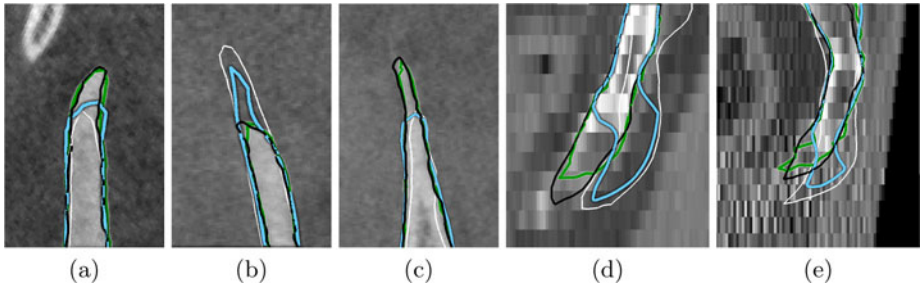


Fig. 6. Exemplary results on clinical data. (a-c) Coronoid processes of the mandible. (d,e) Coccyx tips of the sacrum. Contours: Black: Gold standard. White: Initial mesh. Green/gray: MRF result. Blue/light gray: FreeForm result.

5 Conclusion

We proposed a method that allows omni-directional displacements for all vertices of a surface mesh during deformable model adaptation. We achieve global regularization by encoding the adaptation problem as a Markov Random Field, which

we then optimize with a fast approximate solver. In an evaluation on synthetic as well as clinical data, we showed that this approach can outperform traditional mesh adaptation along line segments (normals) in regions with high curvature (tips) in terms of segmentation accuracy. In this paper, we focused on parallel translations and tip-like structures. A closer investigation of our MRF based method in situations where scaling and rotation is desired, as well as an extended quantitative evaluation that considers all regions on mandible and pelvis shall be performed in future work. Furthermore, computational performance (both in terms of memory and speed) shall be enhanced, possibly involving a hybrid deformation model that employs omni-directional displacements in regions with high curvature and normal displacements elsewhere. Last but not least, a question of interest is if the MRF based method can be extended to simultaneous adaptation of multiple, adjacent meshes, i.e. multi-object segmentation.

Acknowledgements. D. Kainmueller is funded by the DFG Collaborative Research Center SFB760. H. Lamecker is funded by the German Research Center MATHEON. Thanks to M. Zinser (Universitätsklinikum Köln, Germany) and M. Heller (Julius Wolff Institut, Charité Berlin, Germany) for providing image data and manual segmentations.

References

1. McInerney, T., Terzopoulos, D.: Deformable models in medical image analysis: A survey. *Medical Image Analysis* 1, 91–108 (1996)
2. Montagnat, J., Delingette, H., Ayache, N.: A Review of Deformable Surfaces: Topology, Geometry and Deformation. *Image Vis. Comput.* 19, 1023–1040 (2001)
3. Park, J.Y., McInerney, T., Terzopoulos, D., Kim, M.H.: A non-self-intersecting adaptive deformable surface for complex boundary extraction from volumetric images. *Computers & Graphics* 25(3), 421–440 (2001)
4. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active Shape Models - Their Training and Application. *Comput. Vis. Image Underst.* 61(1), 38–59 (1995)
5. Kainmueller, D., Lange, T., Lamecker, H.: Shape Constrained Automatic Segmentation of the Liver based on a Heuristic Intensity Model. In: *3D Segmentation in the Clinic: A Grand Challenge*, pp. 109–116 (2007)
6. Komodakis, N., Tziritas, G., Paragios, N.: Performance vs computational efficiency for optimizing single and dynamic MRFs: Setting the state of the art with primal-dual strategies. *Comput. Vis. Image Underst.* 112(1), 14–29 (2008)
7. Glocker, B., Komodakis, N., Tziritas, G., Navab, N., Paragios, N.: Dense image registration through MRFs and efficient linear programming. *MIA* 12(6), 731–741 (2008); Special issue on information processing in medical imaging 2007
8. Li, K., Wu, X., Chen, D.Z., Sonka, M.: Optimal Surface Segmentation in Volumetric Images - A Graph-Theoretic Approach. *IEEE TPAMI* 28(1), 119–134 (2006)
9. Seim, H., Kainmueller, D., Heller, M., Lamecker, H., Zachow, S., Hege, H.C.: Automatic Segmentation of the Pelvic Bones from CT Data Based on a Statistical Shape Model. In: *Proc. VCBM*, pp. 93–100 (2008)
10. Kainmueller, D., Lamecker, H., Seim, H., Zinser, M., Zachow, S.: Automatic Extraction of Mandibular Nerve and Bone from Cone-Beam CT Data. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009*. LNCS, vol. 5762, pp. 76–83. Springer, Heidelberg (2009)

Multiplicative Jacobian Energy Decomposition Method for Fast Porous Visco-Hyperelastic Soft Tissue Model

Stéphanie Marchesseau¹, Tobias Heimann¹, Simon Chatelin², Rémy Willinger²,
and Hervé Delingette¹

¹ Asclepios Research Project, INRIA Sophia Antipolis, France

² University of Strasbourg, IMFS-CNRS, Strasbourg, France

Abstract. Simulating soft tissues in real time is a significant challenge since a compromise between biomechanical accuracy and computational efficiency must be found. In this paper, we propose a new discretization method, the *Multiplicative Jacobian Energy Decomposition* (MJED) which is an alternative to the classical Galerkin FEM (Finite Element Method) formulation. This method for discretizing non-linear hyperelastic materials on linear tetrahedral meshes leads to faster stiffness matrix assembly for a large variety of isotropic and anisotropic materials. We show that our new approach, implemented within an implicit time integration scheme, can lead to fast and realistic liver deformations including hyperelasticity, porosity and viscosity.

1 Introduction

The simulation of soft tissue deformation has attracted a growing interest in the past 15 years both in the biomechanics and the medical image analysis communities. Modeling in silico the deformation of soft tissues is of high interest in particular for surgical gesture training[1] and therapy planning[2]. In this paper, we focus on the simulation of liver deformation in the context of surgery training. In such case, it is crucial that soft tissue deformation is simulated in real-time, i.e. at a minimum of 25 frames per second for visual feedback. Furthermore, in surgery simulation, there are additional constraints of numerical stability during the occurrence of contact between soft tissue and (virtual) surgical instruments.

To simulate soft tissues efficiently and realistically, some authors have relied on the *Total Lagrangian Explicit Dynamic* (TLED) algorithm[3,4] to simulate deformations with explicit time integration schemes. However, the main limitation of these explicit schemes is that they require very small time steps to keep the computation stable, especially for stiff materials. Indeed, it is necessary to iterate multiple times to propagate applied forces from a node to the whole mesh. Therefore, with such approaches, it is difficult to produce realistic simulations of contact with rigid objects, such as surgical tools. Implicit integration schemes require the evaluation of a global stiffness matrix and the solution of linear system of equations at each time step.

In this paper, we first introduce the *Multiplicative Jacobian Energy Decomposition* (MJED): a general algorithm to implement hyperelastic materials based on total Lagrangian FEM with implicit time integration schemes. Then we propose a realistic biomechanical model of the liver which combines hyperelasticity, viscoelasticity as well as poroelasticity. The viscoelasticity of our liver model is based on Prony series, the parameters of which have been experimentally estimated through a dynamic strain sweep testing. Finally, we take into account the porous medium of the liver parenchyma through a poro-elastic model which computes the fluid pressure and the resulting applied pressure on the solid phase.

2 Multiplicative Jacobian Energy Decomposition

Under large deformation, linear elasticity is no longer valid and the liver behavior is better represented as an hyperelastic material. Since we are using implicit time integration schemes, it is necessary at each time step to compute hyperelastic forces and stiffness matrices with a discretization method. The Finite Element Method is a widely used approach to this end, however the constraint of real-time simulation is not always satisfied. The objective of this section is to introduce a fast discretization method suitable for all hyperelastic materials.

To discretize the liver geometry, we use tetrahedral linear finite elements. T_P is the rest tetrahedron (with vertices P_i) which is transformed under the deformation function $\phi(X)$ into the tetrahedron T_Q (with vertices Q_i). Any hyperelastic material is fully determined by its strain energy function W_h which describes the amount of energy necessary to deform the material. This strain energy function is defined in a way which is invariant to the application of rigid transformations: it involves the invariants of the *Cauchy-deformation tensor* defined as $\mathbf{C} = \nabla\phi^T \nabla\phi$. There are numerous invariants of \mathbf{C} (see [5] for detailed explanation) but the one commonly used are the following: $I_1 = \text{tr}\mathbf{C}$, $I_2 = \frac{1}{2}((\text{tr}\mathbf{C})^2 - \text{tr}\mathbf{C}^2)$, $I_4 = a^T \mathbf{C} a$ (where a is the main fiber direction), and the Jacobian $J = \det \nabla\phi$.

2.1 Decomposition of Strain Energy

We decouple in the strain energy, the invariants of \mathbf{C} from J so as to avoid complex derivative expressions and matrix inversion of \mathbf{C} . Instead of computing the force and stiffness matrix using the classical Galerkin FEM [6], we compute them directly using the Rayleigh-Ritz method by deriving the energy with respect to the nodal position:

$$F_i = - \left(\frac{\partial W_h}{\partial Q_i} \right)^T \quad \text{and} \quad \mathbf{K}_{ij} = \left(\frac{\partial^2 W_h}{\partial Q_j \partial Q_i} \right) \quad (1)$$

It is important to note that the approach developed in this section is completely equivalent to the classical FEM one but leads to more efficient computation. A comparison with the open source software FEBio proved this equivalence. We propose to write the strain energy functions as a sum of terms $W_h^k = f^k(J)g^k(\bar{I})$

or exponentials of W_h^k , where $\tilde{I} = (I_1, I_2, I_4, \dots)$. This decomposition applies to every material models we studied so far (Costa, Veronda Westmann, Boyce Arruda, StVenant Kirchhoff, NeoHookean, Ogden, Mooney Rivlin). This gives the following expression of the tetrahedron strain energy:

$$W_h = V_0 \sum_{k=1}^n f^k(J)g^k(\tilde{I}) + V_0 \exp \left(\sum_{k=n+1}^{n'} f^k(J)g^k(\tilde{I}) \right) \quad (2)$$

Using this exact decomposition of strain energy enables complex material formulation to be computed more efficiently with only a sum of reasonably simple terms. With this decomposition, getting $f^{k'}(J)$ requires a 1D derivation, and getting $\mathbf{S}_h^k = 2 \frac{\partial g^k(\tilde{I})}{\partial \mathbf{C}}$ requires to combine well-known derivatives of the invariants (such as $\frac{\partial I_1}{\partial \mathbf{C}} = \mathbf{Id}$ or $\frac{\partial I_2}{\partial \mathbf{C}} = \mathbf{Id}I_1 - \mathbf{C}$ where \mathbf{Id} is the 3×3 identity matrix). For instance, the nodal force expression becomes:

$$F_{h,i} = -V_0 \sum_{k=1}^n \left(f^{k'}(J)g^k(\tilde{I}) \left(\frac{\partial J}{\partial Q_i} \right)^T + f^k(J) \nabla \phi \mathbf{S}_h^k D_i \right) \quad (3)$$

where D_i are gradients of shape functions, called *shape vectors*.

2.2 Formulation of the Stiffness Matrix

Implicit time integration schemes require the computation of the tangent stiffness matrix at each time step. This naturally involves elasticity tensors computed as the derivative of \mathbf{S}_h^k with respect to \mathbf{C} for each tetrahedron and at each time step. The MJED leads to far simpler expressions of those tensors because \mathbf{S}_h^k is independent of J . Furthermore, in many common material models, the term containing those elasticity tensors can be precomputed. The full expression of the stiffness matrix includes 6 terms. Due to space constraints, we only focus below on the term involving the fourth order elasticity tensor: $\mathbf{R}_h^k = f^k(J) \left(\frac{\partial \mathbf{S}_h^k}{\partial Q_j} D_i \right)^T \nabla \phi^T$

which requires the computation of the tensor $\frac{\partial \mathbf{S}_h^k}{\partial \mathbf{C}} : \mathbf{H}$ where \mathbf{H} is a symmetric matrix. In all cases, this tensor can be written as a sum of two kinds of terms, $\beta_1^k \mathbf{A}_1^k \mathbf{H} \mathbf{A}_1^k$ or $\beta_2^k (\mathbf{H} : \mathbf{A}_2^k) \mathbf{A}_2^k$ where β_u^k are scalars, \mathbf{A}_u^k are symmetric matrices, and $\mathbf{A} : \mathbf{B} = \text{tr}(\mathbf{B}^T \mathbf{A})$ for any two matrices \mathbf{A}, \mathbf{B} . Therefore, the term \mathbf{R}_h^k is a combination of two terms:

$$f^k(J) \nabla \phi \mathbf{L}_1^k(i, j) \nabla \phi^T \quad \text{and} \quad f^k(J) \nabla \phi \mathbf{L}_2^k(i, j) \nabla \phi^T \quad (4)$$

where $\mathbf{L}_1^k(i, j)$ and $\mathbf{L}_2^k(i, j)$ are linear matrices depending on the shape vectors D_i, D_j , the matrices \mathbf{A}_u^k and the scalars β_u^k . This formulation leads to an optimization for the assembly of the stiffness matrix for two reasons. First, no fourth order tensors are required, only scalars and symmetric matrices are involved in the computation. Second, except for the Ogden model, the matrices \mathbf{A}_u^k are constant and therefore matrices $\mathbf{L}_1^k(i, j)$ and $\mathbf{L}_2^k(i, j)$ can be precomputed for each tetrahedron before the simulation.

2.3 Coping with Highly Compressed Elements

In case of high compression, the volumetric terms $f^k(J)$ in the strain energy become dominant. This makes the stiffness matrix singular and thus leads to numerically unstable computations. We propose to *regularize* a second term of the stiffness matrix $\mathbf{G}_h^k = f^{k''}(J) g^k(\tilde{I}) \frac{\partial J}{\partial Q_j} \otimes \frac{\partial J}{\partial Q_i}$ by replacing it with the following expression : $\mathbf{G}_h^k = f^{k''}(J) g^k(\tilde{I}) \left((1-b) \frac{\partial J}{\partial Q_j} \otimes \frac{\partial J}{\partial Q_i} + b \frac{\partial J}{\partial Q_j} \cdot \frac{\partial J}{\partial Q_i} \mathbf{Id} \right)$. The closer b is to 1, the more \mathbf{G}_h^k resembles to a diagonal matrix. In practice, we set $b = (1 - J)$ if $0 \leq J \leq 1$, $b = 0$ if $J \geq 1$ and $b = 1$ if $J \leq 0$. With this technique, it is even possible to handle inverted elements when the strain energy remains finite at $J = 0$.

3 Modeling Visco - Poro - Hyperelasticity

3.1 Visco-Hyperelasticity Based on Prony Series

To accurately model the viscoelasticity of the liver, we propose to rely on Prony series [4]. In this method, time-dependent stresses are added to the hyperelastic stress tensor \mathbf{S}_v . Time-dependence is given by $\alpha(t) = \alpha_\infty + \sum_i \alpha_i \exp(-t/\tau_i)$ with the condition $(\alpha_\infty + \sum_i \alpha_i) = 1$. The visco-hyperelastic SPK tensor \mathbf{S}_v can be written as:

$$\mathbf{S}_v = \mathbf{S}_h - \sum_i \gamma_i \quad \text{where} \quad \gamma_i = \int_0^t \alpha_i \left(1 - \exp\left(-\frac{t' - t}{\tau_i}\right) \right) \frac{\partial \mathbf{S}_h}{\partial t'} dt' \quad (5)$$

After a discretization over time this results in the recursive formula: $\gamma_i^n = a_i \mathbf{S}_h^n + b_i \gamma_i^{n-1}$ where $a_i = \Delta t \alpha_i / (\Delta t + \tau_i)$ and $b_i = \tau_i / (\Delta t + \tau_i)$. Δt is the time step used for discretization and has to be the same as the time step for the solvers of the simulation.

3.2 Poro-Elasticity

Following Kerdok’s model [7], we propose to model the liver as a fluid-filled sponge as it filters the blood through its parenchyma. The proportion of free-fluid (blood, water, ...) in the liver parenchyma in the reference configuration is set to a constant f_w , $1 - f_w$ represents the initial ratio of the solid phase (e.g. hepatic cells...). We introduce the effective volumetric Jacobian $J^* = (f_w + J - 1)/f_w$, and define the volumetric Cauchy stress following Hencky’s elasticity : $\sigma_{\text{Heq}} = K_0 f_w \ln(J^*)$ where K_0 is the bulk modulus of the material. With this model, when J gets close to $1 - f_w$, the solid phase of the liver is completely compressed and the resulting stress is infinite. To avoid instabilities due to this infinite stress, we substitute σ_{Heq} when $J \leq J_0$ by its tangent curve at $J_0 = 1 - f_w + K_0/K_{\text{lim}}$ where K_{lim} is a bulk modulus and represents the slope of the tangent (see Fig 1). The fluid phase of the liver also applies some volumetric stresses due to the transient response of the fluid through the porous liver parenchyma.

A straightforward way of modeling the porous behavior is through the linear Darcy’s law. In this setting, variation of fluid pressure P_{fluid} is governed by the variation of volume change and a diffusive process:

$$\frac{1}{K_{\text{lim}}}\dot{P}_{\text{fluid}} = \kappa \nabla^2 P_{\text{fluid}} - \frac{\dot{J}}{J} \quad (6)$$

where κ is the permeability parameter, kept constant to decrease the computational cost. Finally, the total Cauchy stress response in the volumetric part is defined by summing the solid and the fluid terms: $\sigma_p = \sigma_{\text{heq}} \mathbf{Id} - P_{\text{fluid}} \mathbf{Id}$. The simulated fluid pressure field during a deformation under gravity force is shown in Fig. 1. Highest pressure in the fluid occurs when the liver is compressed either by the gravity pressure (diffusion starts at the top) or by elastic reaction (diffusion starts at the bottom).

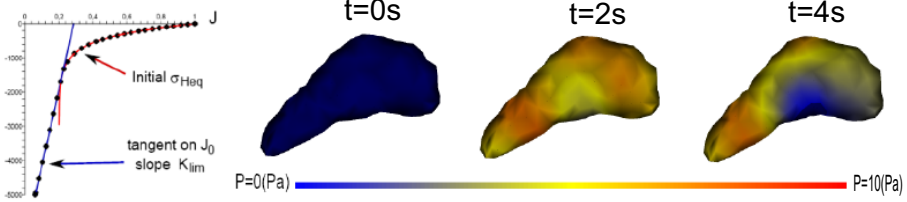


Fig. 1. (Left) Representation of the static Cauchy stress before and after substitution. The dot curve represents the new stress. Here $f_w = 0.8$. (Right) Pressure field of the porous component on a liver under gravity. Using $f_w = 0.5$, $K_0 = 400$ Pa, $K_{\text{lim}} = 2200$ Pa and $\kappa = 20$ m⁴/Ns.

4 Results and Validation

4.1 Computation Time of the Hyperelastic Implementation

Decreasing computation time of the hyperelastic term is essential to reach real-time simulation as this term represents around 60% of the total time needed in one step (see Fig. 2). In order to validate the MJED method, we compared our implementation with the classical FEM method explained in [6], referred to as "Standard FEM", implemented in SOFA¹. We measured the time elapsed for the computation of the nodal forces and the stiffness matrices averaged over 100 iterations. We simulated the deformation of a cube with 20 700 tetrahedra and 4300 nodes. The results are given in Fig. 2. For all implemented models the proposed strategy is more efficient than the standard FEM, up to five times as fast for St Venant Kirchhoff material.

¹ SOFA is an Open Source medical simulation software available at www.sofa-framework.org

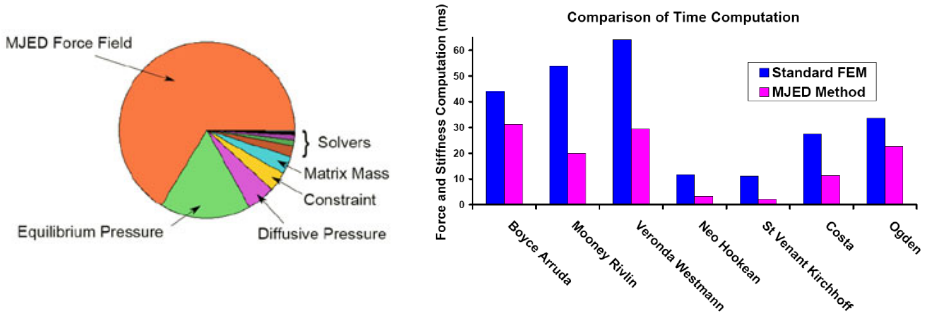


Fig. 2. (Left) Break-down of computational time between the components during one time step. (Right) Comparison of the computation times of nodal forces and stiffness matrices between two different discretization methods averaged over 100 iterations.

4.2 Visco-Elasticity Validation

To calibrate the visco-elastic parameters of our liver model, tests on at least 60 samples from 5 animals were performed on porcine livers. Dynamic viscoelastic behavior of hepatic tissue was investigated using *in vitro* Dynamic Mechanical Analysis (DMA) without perfusion in order to capture only the viscoelasticity. Dynamic Frequency Sweep tests were performed on a dedicated stress-controlled AR2000 rheometer (TA-Instruments, New Castle, DE, USA) in the linear viscoelastic strain range of the samples (see Fig.3). From the results, the Dynamic Modulus G can be obtained as a function of the frequency or function of the time, and the viscoelastic behavior can be modeled after fitting a generalized Maxwell model with two modes of relaxation to those measurements: $G(t) = G_0(g_\infty + g_1 \exp(-t/\tau_1) + g_2 \exp(-t/\tau_2))$ where $G_0 g_\infty = G_\infty$ is called the equilibrium modulus, g_1, g_2, τ_1, τ_2 are parameters such as $g_\infty + g_1 + g_2 = 1$.

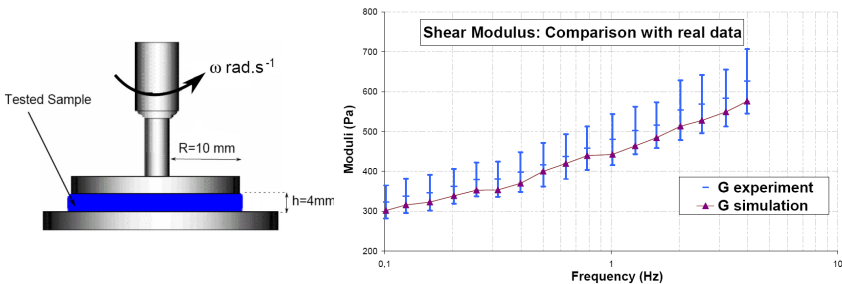


Fig. 3. (Left) Rheometer: Lower plate is fixed whereas upper is sinusoidally rotating. (Right) Comparison of the simulated values with the data obtained by DMA testing. The moduli are given on a X-log scale. The material is St Venant Kirchhoff with $G_0 = 770 \text{ Pa}$, $(\alpha_1, \alpha_2) = (0.235, 0.333)$ and $(\tau_1, \tau_2) = (0.27 \text{ s}, 0.03 \text{ s})$.

From the rheological experiments we derive the shear modulus G_0 required in the hyperelastic term and the Prony series parameters for the viscous term. To check the validity of those parameters, we compared *in silico* simulations with the performed *in vitro* rheological tests. Dynamic frequency sweep tests are simulated using similar geometries and boundary conditions as in the real DMA tests. We estimated the values of the Dynamic Modulus and compared them with experimental data. Figure 3 shows the results: the simulation manages to capture the viscous behavior of the liver for small deformations with a mean relative error of 5%. Given the fitting errors and the standard deviation of the values obtained with the DMA tests, this mean error is reasonably good.

4.3 Complete Liver Model

To analyze the influence of each component in the complete model, several simulations were performed using the same liver mesh (1240 vertices and 5000 tetrahedra) with Boyce Arruda hyperelastic material [7]. The liver mesh was segmented from a CT scan image and meshed with tetrahedra by the GHS3D software [2]. An Euler implicit time integration scheme was used, the linear equation was solved with a conjugated gradient algorithm. As boundary conditions, several nodes of the liver were fixed along the vena cava and suspensive ligament. The liver then deformed under the action of gravity. All computations were performed on a laptop PC with a Intel Core Duo processor at 2.80 GHz (simulations are available in the video clip). Adding viscosity to hyperelasticity increases the amplitude of the oscillations as the material becomes less stiff (see Fig 4). The frame rate is around 7 FPS against 7.5 FPS for hyperelasticity alone. We did not reach the 25 FPS needed for real-time interaction. However, the implicit integration scheme allows larger time step (0.3s for instance) which speeds up the simulation and makes user interactions efficient. High amount of extension and compression are possible which may be somewhat unrealistic, therefore the porous component is necessary to control the amount of viscosity. The maximum amplitude with

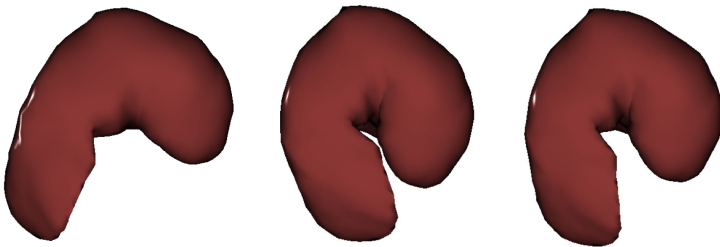


Fig. 4. Comparison of the maximum amplitudes under gravity. (Left) Hyperelastic Liver, (Middle) Visco-Hyperelastic liver, (Right) Visco-Poro-Hyperelastic.

² GHS3D is a mesh generator for tetrahedral elements, developed at INRIA, France.

porosity is in between the hyperelasticity alone and the visco-hyperelasticity. The addition of this component decreases the computational efficiency (6 FPS) since a semi-implicit integration scheme is used for the porous component. Because of the fast variation of the explicit term \dot{J}/J , the time step has to be decreased to 0.15s. On our laptop PC, the simulation is still fluid enough to allow user interactions.

5 Conclusion

In this paper, we have introduced an efficient method to assemble stiffness matrices for complex biomechanical material models which compares favorably with the standard FEM method. We have also proposed a poro-visco-hyperelastic liver model suitable for real-time interaction which is, up to our knowledge, among the most realistic ones. Several model parameters have been identified from rheometric tests performed on porcine livers and a validation study has been successfully performed to reproduce those tests. Despite those advances, further research is needed to achieve realistic liver surgery simulations including the realistic contact with neighboring structures, the influence of breathing and cardiac motion, the simulation of hepatic resection, bleeding and suturing.

Acknowledgment. This work is partially funded by the European PASSPORT project (Patient-Specific Simulation for Pre-Operative Realistic Training of Liver Surgery) FP7- ICT-2007- 223894.

References

1. Delingette, H., Ayache, N.: Soft tissue modeling for surgery simulation. In: Computational Models for the Human Body, pp. 453–550. Elsevier, Amsterdam (2004)
2. Hawkes, D., Barratt, D., Blackall, J., Chan, C., Edwards, P., Rhode, K., Penney, G., McClelland, J., Hill, D.: Tissue deformation and shape models in image-guided interventions: a discussion paper. *Medical Image Analysis* 9(2), 163–175 (2005)
3. Miller, K., Joldes, G., Lance, D., Wittek, A.: Total lagrangian explicit dynamics finite element algorithm for computing soft tissue deformation. *Communications in Numerical Methods in Engineering* 23, 121–134 (2006)
4. Taylor, Z., Comas, O., Cheng, M., Passenger, J., Hawkes, D., Atkinson, D., Ourselin, S.: On modelling of anisotropic viscoelasticity for soft tissue simulation: numerical solution and gpu execution. *Medical Image Analysis* 13, 234–244 (2009)
5. Weiss, J., Maker, B., Govindjee, S.: Finite element implementation of incompressible, transversely isotropic hyperelasticity. *Computer Methods in Applied Mechanics and Engineering* 135, 107–128 (1996)
6. Zienkiewicz, C., Taylor, R.: *The Finite Element Method. Solid Mechanics*, vol. 2. Butterworth-Heinemann, Butterworths (2000)
7. Kerdok, A.E.: Characterizing the Nonlinear Mechanical Response of Liver to Surgical Manipulation. PhD thesis, Harvard University (2006)

Fast Ultrasound Image Simulation Using the Westervelt Equation

Athanasios Karamalis¹, Wolfgang Wein^{2,*}, and Nassir Navab¹

¹ Computer Aided Medical Procedures, Technische Universität München, Germany
`karamali@cs.tum.edu`

² White Lion Technologies AG, Munich, Germany

Abstract. The simulation of ultrasound wave propagation is of high interest in fields as ultrasound system development and therapeutic ultrasound. From a computational point of view the requirements for realistic simulations are immense with processing time reaching even an entire day. In this work we present a framework for fast ultrasound image simulation covering the imaging pipeline from the initial pulse transmission to the final image formation. The propagation of ultrasound waves is modeled with the Westervelt equation, which is solved explicitly with a Finite Difference scheme. Solving this scheme in parallel on the Graphics Processing Unit allows us to simulate realistic ultrasound images in a short time.

Keywords: Ultrasound Simulation, Westervelt Equation, Finite Difference Method, GPU.

1 Introduction

The realistic simulation of medical ultrasound has applications in fields such as ultrasound system development. Here the quality of ultrasound images depends highly on numerous system parameters including for example transducer shape, focusing strategies and active aperture size. To speed up prototyping and lower development costs engineers simulate the effects of different system parameterizations before moving on to the actual system assembly [6]. A more recent application domain is the simulation of High Intensity Focused Ultrasound (HIFU), also referred to as Therapeutic Ultrasound. The nonlinear propagation of ultrasound in tissue produces high-frequency components that are absorbed more rapidly by the tissue. Simulating these effects is crucial for the correct assessment of the ultrasound dose required for therapy and is subject to ongoing research [2]. In terms of education, the physics and instrumentation of ultrasound are complex and require in-depth knowledge to understand the impact of different system parameterizations on the final image [15]. A fast simulation can demonstrate the results of different system parameterizations and provide more insights into the underlying mechanisms of the imaging modality. Last but

* Work conducted while at Siemens Corporate Research, Princeton, NJ, USA.

not least, ultrasound simulation has recently been utilized for multimodal image registration between CT and ultrasound [13]. A similarity measure is evaluated between the real ultrasound images and ones simulated from CT. Improving the simulation could also improve the accuracy and robustness of the registration.

The overall processing time is a decisive factor for the simulation of ultrasound. Simulating a single ultrasound scan line using the Westervelt equation and a Finite Difference scheme was reported to take about 1 hour on a desktop PC [5], with the simulation of a complete image requiring probably more than a day of processing time. Furthermore, alternative simulation approaches like Field II [6], one of the most widely used linear ultrasound simulation packages, can require up to two days of processing time on a modern desktop PC for generating a single 128 scan line ultrasound image [10].

In Pinton et al. [9] a thorough comparison was presented between the Westervelt equation and alternatively proposed methods for modeling ultrasound wave propagation. They presented simulated ultrasound images by solving the Westervelt equation with a Finite-Difference Time-Domain scheme, whereas, details on the Radio-Frequency (RF) processing for the image formation were omitted.

In this work we focus on the realistic simulation of ultrasound wave propagation and the subsequent generation of ultrasound images in acceptable time. For this purpose the ultrasound imaging pipeline was implemented from the initial pulse transmission to the final image formation. The wave propagation is modeled using the Westervelt equation, which is explicitly solved with a Finite Difference scheme. In order to achieve fast simulation times the Finite Difference scheme was implemented on the Graphics Processing Unit (GPU), which has already demonstrated its potential for accelerating parallel computations and efficiently solving these schemes [3][8].

2 Wave Propagation

The propagation of ultrasound waves and their interaction with different media was modeled using the Westervelt Partial Differential Equation (PDE), also referred to as the nonlinear full-wave equation [2][5]. It describes the propagation of waves and additionally models thermal attenuation and nonlinearity. The reader interested in an accuracy analysis of the Westervelt equation is referred to Huijssen et al. [5], which includes comparisons to an analytical solution, the Khokhlov-Zabolotskaya-Kuznetsov equation and water tank measurements. The Westervelt equation is given as follows:

$$\nabla^2 p - \frac{1}{c_0^2} \frac{\partial^2 p}{\partial t^2} + \frac{\delta}{c_0^4} \frac{\partial^3 p}{\partial t^3} + \frac{\beta}{\rho_0 c_0^4} \frac{\partial^2 p^2}{\partial t^2} = 0, \quad (1)$$

where p [Pa] is the acoustic pressure, c_0 [ms^{-1}] is the propagation speed, ρ_0 [kgm^{-3}] is the ambient density, δ [m^2s^{-1}] is the diffusivity of sound, and β is the coefficient of nonlinearity. Thus, the first two terms are identical to the D'Alembertian operator on p , the third term is the loss term due to thermal conduction and the fourth term describes the nonlinearity.

One thing to note is that various simulation approaches model sound waves as rays, taking into account the physics from optics [10,13]. This simplification results in faster processing times, but reduces significantly the realism of the simulation. Ray based simulation approaches are not able to fully model the complex effects modeled by wave based approaches. These effects include interference, scattering, diffraction etc. which are common in medical ultrasound propagation and contribute tremendously to the formation of the final ultrasound image. For more details on ultrasound image characteristics and artifacts see Zagzebski [15].

The Westervelt equation is numerically solved with the Finite Difference method [13,8]. The basic idea behind this method is to evaluate the PDE equation, more specifically calculate the wave amplitude, on sampling points of a computational grid. The grid can have complex shapes, while we use a regular 2D grid with equidistant sampling points. For the calculation of the wave amplitude the partial derivatives of the equation are substituted with their finite difference representations and the equation is solved for the future timestep. A thorough analysis of higher order Finite Difference schemes for solving the acoustic wave equation is presented by Cohen and Joly [1]. Fourth-order accurate in space and second-order accurate in time schemes have demonstrated good results [12], and are therefore used in our work. The finite differences for equation (1) are given as follows:

$$\frac{\partial p}{\partial t} \approx \frac{p_{i,j}^n - p_{i,j}^{n-1}}{\Delta t}, \quad \frac{\partial^2 p}{\partial t^2} \approx \frac{p_{i,j}^{n+1} - 2p_{i,j}^n + p_{i,j}^{n-1}}{\Delta t^2}, \quad (2)$$

$$\frac{\partial^3 p}{\partial t^3} \approx \frac{6p_{i,j}^n - 23p_{i,j}^{n-1} + 34p_{i,j}^{n-2} - 24p_{i,j}^{n-3} + 8p_{i,j}^{n-4} - p_{i,j}^{n-5}}{(2\Delta t)^3}, \quad (3)$$

$$\frac{\partial^2 p}{\partial x^2} \approx \frac{-p_{i+2,j}^n + 16p_{i+1,j}^n - 30p_{i,j}^n + 16p_{i-1,j}^n - p_{i-2,j}^n}{12\Delta x^2}, \quad (4)$$

$$\frac{\partial^2 p}{\partial y^2} \approx \frac{-p_{i,j+2}^n + 16p_{i,j+1}^n - 30p_{i,j}^n + 16p_{i,j-1}^n - p_{i,j-2}^n}{12\Delta y^2}, \quad (5)$$

where i, j are the axial and the lateral indices of the discrete computational grid, n is the timestep, $\Delta x, \Delta y$ are the spatial discretization steps and Δt is the temporal discretization step. Thus, an explicit solution is calculated for each sampling point based on the wave amplitudes at sampling points of the previous timesteps.

To model the interaction of the waves with heterogeneous media, different coefficients are used for the speed of sound, the ambient density, the diffusivity of sound and the nonlinearity. The most important effect in ultrasound imaging is the reflection of the waves, which is caused by the difference in speed of sound between media.

3 Ultrasound Image Simulation

In our framework we implemented the basic procedure to synthesize B-mode ultrasound images. Other than the wave propagation, the image simulation process

involves transmission and reception of ultrasound pulses and processing of the resulting echoes for forming the final image. For more details on ultrasound physics, instrumentation and image formation see Hedrick et al. [4] and Szabo [11].

3.1 Ultrasound Transmission and Reception

For each simulation run sound waves are emitted at selected points on the computational grid. Various transducer geometries can be simulated by selecting the appropriate points on the grid, with a linear transducer modeled in this work. Modifying the wave amplitude at these points introduces a wave disturbance that propagates through the grid with the simulation of consecutive timesteps. The shape of the emitted pulse is of crucial importance. Non-modulated sinusoidal or Gaussian shaped pulses for instance can cause grid disturbances even after the pulse transmission has ceased. We use a 6 cycle sinusoidal pulse modulated by a Gaussian shaped envelope, commonly used by ultrasound systems [4]. The pulse is formulated as follows:

$$s(t) = \left(A \cdot \sin \left(\frac{2\pi}{l} t \right) \right) \left(\alpha e^{-\frac{(x - \mu)^2}{2\sigma^2}} \right), \quad (6)$$

where $s(t)$ is the pulse amplitude at timestep $t \in [0..l]$, l is the pulse length, A is the maximal pulse amplitude, α is the Gaussian amplitude, x is the random variable, μ the mean, and σ is the standard deviation. The echoes are recorded for each timestep at the positions where the grid was perturbed by the pulse transmission. Ultrasound images are formed from multiple scan lines with the total number of scan lines playing an important role for the overall spatial resolution of the final ultrasound image. Each scan line in the ultrasound image corresponds to the echoes received along an ultrasound beam, which brings us to topic of ultrasound beamforming.

Beamforming refers to the constructive/destructive interference of waves emitted by multiple transducer elements. Triggering a group of transducer elements at a time results in high acoustic intensities along the center axis of the group. Figure 1(a) schematically shows the process of focusing using a group of elements and figure 1(b) shows the simulation run, with our framework, for beam focusing at a low depth. Generating a narrow beam is desirable in ultrasound system development as it improves the spatial resolution of the ultrasound scan. Additionally, triggering the transducer with the appropriate time delays allows to position and steer the beam [4].

Multiple scan lines are acquired for the image formation by transmitting a beam, receiving the echoes and moving the active element group until the entire transducer element surface is covered.

3.2 Radio Frequency Processing

The result of the previously described simulation is Radio Frequency (RF) data acquired for each element of an active element group for each scan line. Before forming an image, the raw data needs to be processed. The RF processing

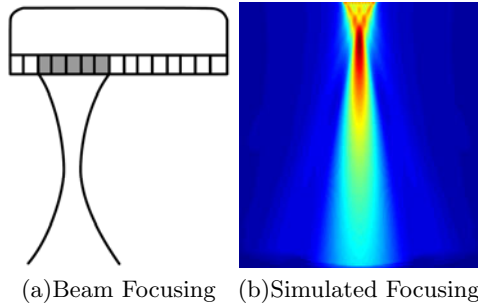


Fig. 1. Image (a) schematically demonstrates ultrasound beam focusing by triggering a group of elements with different time delays. Image (b) shows the maximum wave amplitudes of a simulation run using a focus scheme for low depths in a medium with uniform speed of sound.

pipeline varies slightly between different ultrasound system vendors, but the basic principles are common and are implemented in this framework.

An ultrasound scan line is formed by combining the RF data acquired at each element of an active element group. For this we apply the Delay and Sum beamforming algorithm [12], which can among others be defined as:

$$d(t) = \sum_{i=0}^N W_i \cdot e_i(t + \tau), \quad (7)$$

where $d(t)$ is the final signal response of an ultrasound beam at timestep t , N is the number of active transducer elements, W_i is a weighting function (in our case a Hanning window), $e_i(t)$ is the signal response of each active transducer element i , and τ is the transmission delay expressed in timesteps.

The received signal contains noise that is mainly manifested in the low and high frequency parts of its spectrum [11]. Therefore, the low and high frequency components are removed with a bandpass filter, in our case a Butterworth filter.

Afterwards the signal goes through the process of demodulation, which results in a signal that retains its overall pulse response but contains much less high frequency modulations. Demodulation is performed by finding the envelope of the rectified signal and is implemented by taking the absolute of the Hilbert Transform of the signal.

As a pulse traverses through the medium it is attenuated and reflectors at greater depth appear weaker than reflectors at smaller depth. This is compensated by applying a so-called Time-Gain Compensation (TGC), which amplifies echoes based on their reception time (depth). In our implementation the signal is convolved with a simple linear function $f \in [1..n]$, where n is the amplification factor for the maximal depth.

Last but not least, the resulting signal $h(t)$ is compressed by decreasing its dynamic range (ratio of strongest to weakest signal). This is usually done with a logarithmic scaling (also referred to as log-compression):

$$h_c(t) = \log(h(t) + c), \quad (8)$$

where c is the compression coefficient. At this point the ultrasound image is formed by combining the processed RF lines into a single image.

4 Results

The proposed simulation framework was utilized for generating 2D ultrasound images out of two synthetic datasets, one showing a fetus, figure 2(a), and the other one showing multiple anechoic regions embedded in a highly scattering medium, figure 2(c). The fetus dataset is a modified version of the one presented in Jensen and Munk [7] and the phantom dataset is generated using Rayleigh noise, with similar phantoms being used for testing real ultrasound imaging systems [4]. The mediums had following characteristics: $\delta = 4.5 \cdot 10^{-6} [m^2 s^{-1}]$, $\beta = 6$ and $\rho_0 = 1100 [kg m^{-3}]$, which are common for human tissue [14].

The corresponding simulated ultrasound images are shown in figure 2(b) and 2(d). They clearly demonstrate a realistic speckle pattern, interference effects and beam focusing artifacts. The spatial resolution is high at the center of the focal zone and decreases with increasing distance from the focal zone, an effect also observed in real ultrasound imaging. Furthermore, interference of echoes is strongly evident in the anechoic regions of the phantom dataset. Following parameters were used for simulating the presented ultrasound images: 11 transducer elements formed the active group, $\lambda/2$ elements spacing was used, 192 scan lines were processed, and the discretization steps were set to $\Delta x = \Delta y = 5 \cdot 10^{-3} [m]$ and $\Delta t = 5.5 \cdot 10^{-7} [s]$. For the fetus dataset 6000 timesteps were evaluated and for the phantom dataset 8600, because of the increased depth.

The simulation of the ultrasound wave propagation for generating the raw RF data is performed on the GPU using C++, OpenGL, and the GL shading language (GLSL). Implementing the Finite Difference scheme on the GPU is relatively straight forward with explicit and implicit solvers presented in Harris [3] and Krüger et al. [8] respectively. Switching from GLSL to C-like GPU programming languages like CUDA or OpenCL might improve the performance since they offer more elaborate shared memory features.

After simulating the scan lines, the resulting RF data is processed on the CPU, as there are no computationally expensive tasks involved. For a 2048^2 grid and 192 scan lines the RF simulation on the GPU required 55 minutes for the fetus dataset and 78 minutes for the phantom dataset, with the image formation on the CPU requiring 19 seconds and 24 seconds respectively. The performance was evaluated on a desktop PC with an Intel Core 2 2.66 GHz with 4 GB RAM and a NVIDIA GeForce GTX 280 with 1GB VRAM. In comparison, the framework presented in Pinton et al. [9] required 32 hours processing time on a 56 PC cluster with 118 GB RAM for a 3D simulation. Details on the processing time of the 2D simulation were omitted, but should be in the same range given the provided parameterizations. Furthermore, one simple 2D image was simulated from a synthetic dataset containing only a single anechoic region.

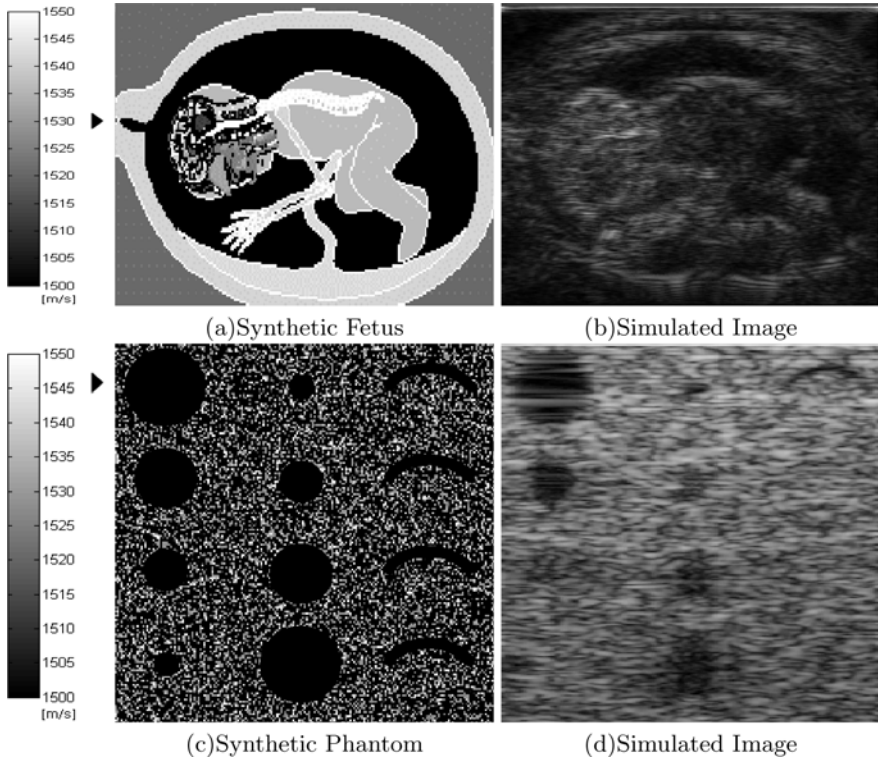


Fig. 2. Image (a) shows the synthetic fetus dataset and image (c) the synthetic phantom dataset. The intensity values correspond to speed of sound values in the range of 1500-1550 [m/s]. The center of the focal zones are marked with a small triangle on the left side. Image (b) and (d) show the simulated ultrasound images.

5 Conclusion

In this work we presented a framework for fast ultrasound image simulation, covering the imaging pipeline from the initial pulse transmission to the final image formation. Our implementation on the GPU simulates realistic ultrasound images in under 80 minutes, avoiding the cumbersome use of PC clusters. The considerably lower simulation time, compared to other implementations, has practical implications for most simulation related application domains like ultrasound system development. Particularly, our approach has strong implications for future intra-operative simulation of HIFU treatment, as the simulation of a single focal zone is computed in less than 30 seconds. Furthermore, simulating on the GPU allows for an interactive visualization of the wave propagation during the simulation at almost no computational cost, which is of interest to both education and system development applications.

In our current implementation the simulation grid is extended to prevent reflections at the grid boundary from interfering with the region of interest in the

simulation grid. Implementing Absorbing Boundary Conditions, similar to [9], could notably improve the overall performance of the simulation as less grid cells would need to be evaluated. This would pave the way for utilizing our framework for 3D ultrasound simulation in acceptable time.

Acknowledgments. We would like to thank Christian Wachinger for valuable discussions. This work was partly funded by the European Project PASSPORT and Siemens Corporate Research, Princeton, NJ, USA.

References

1. Cohen, G., Joly, P.: Construction and Analysis of Fourth-Order Finite Difference Schemes for the Acoustic Wave Equation in Nonhomogeneous Media. *SIAM Journal on Numerical Analysis* 33(4), 1266–1302 (1996)
2. Hallaj, I.M., Cleveland, R.O.: FDTD Simulation of Finite-Amplitude Pressure and Temperature Fields for Biomedical Ultrasound. *The Journal of the Acoustical Society of America* 105, 7–12 (1999)
3. Harris, M.: Fast Fluid Dynamics Simulation on the GPU. *GPU Gems* 1, 637–665 (2004)
4. Hedrick, W.R., Hykes, D.L., Starchman, D.E.: *Ultrasound Physics and Instrumentation*. Mosby (2005)
5. Huijssen, J., Bouakaz, A., Verweij, M.D., de Jong, N.: Simulations of the Nonlinear Acoustic Pressure Field without using the Parabolic Approximation. In: *IEEE Symposium on Ultrasonics*, vol. 2, pp. 1851–1854 (2003)
6. Jensen, J.A.: Field: A Program for Simulating Ultrasound Systems. *Medical and Biological Engineering and Computing* 34, 351–352 (1996)
7. Jensen, J.A., Munk, P.: Computer Phantoms for Simulating Ultrasound B-mode and CFM Images. *Acoustical Imaging* 23(75-80) (1997)
8. Krüger, J., Westermann, R.: Linear Algebra Operators for GPU Implementation of Numerical Algorithms. *GPU Gems* 2 (2005)
9. Pinton, G.F., Dahl, J., Rosenzweig, S., Trahey, G.E.: A Heterogeneous Nonlinear Attenuating Full-Wave Model of Ultrasound. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 56, 474–488 (2009)
10. Shams, R., Hartley, R., Navab, N.: Real-Time Simulation of Medical Ultrasound from CT Images. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part II. LNCS*, vol. 5242, pp. 734–741. Springer, Heidelberg (2008)
11. Szabo, T.L.: *Diagnostic Ultrasound Imaging: Inside Out*. Academic Press, London (2004)
12. Thomenius, K.E.: Evolution of Ultrasound Beamformers. In: *IEEE Ultrasonics Symposium*, pp. 1615–1622 (1996)
13. Wein, W., Brunke, S., Khamene, A., Callstrom, M.R., Navab, N.: Automatic CT-Ultrasound Registration for Diagnostic Imaging and Image-Guided Intervention. *Medical Image Analysis* 12, 577–585 (2008)
14. Yongchen, S., Yanwu, D., Jie, T., Zhensheng, T.: Ultrasonic Propagation Parameters in Human Tissues. In: *IEEE Ultrasonics Symposium*, pp. 905–908 (1986)
15. Zagzebski, J.A.: *Essentials of Ultrasound Physics*. Mosby (1996)

Accounting for Anisotropic Noise in Fine Registration of Time-of-Flight Range Data with High-Resolution Surface Data

L. Maier-Hein¹, M. Schmidt², A.M. Franz¹, T.R. dos Santos¹, A. Seitel¹,
B. Jähne², J.M. Fitzpatrick^{3,*}, and H.P. Meinzer¹

¹ Div. Medical and Biological Informatics, German Cancer Research Center
l.maier-hein@dkfz-heidelberg.de

² Heidelberg Collaboratory for Image Processing, University of Heidelberg, Germany

³ Dept. Electrical Engineering and Computer Science, Vanderbilt University, USA

Abstract. Time-of-Flight (ToF) sensors have become a considerable alternative to conventional surface acquisition techniques such as laser range scanning and stereo vision. Application of ToF cameras for the purpose of intra-operative registration requires matching of the noisy surfaces generated from ToF range data onto pre-interventionally acquired high-resolution surfaces. The contribution of this paper is two-fold: Firstly, we present a novel method for fine rigid registration of noisy ToF data with high-resolution surface meshes taking into account both, the noise characteristics of ToF cameras and the resolution of the target mesh. Secondly, we introduce an evaluation framework for assessing the performance of ToF registration methods based on physically realistic ToF range data generated from a virtual scene. According to experiments within the presented evaluation framework, the proposed method outperforms the standard ICP algorithm with respect to correspondence search and transformation computation, leading to a decrease in the target registration error (TRE) of more than 70%.

1 Introduction

A growing number of applications in the field of computer-assisted medical interventions depend on accurate and fast 3D surface acquisition. To date, however, the estimation of a range map by image analysis or other techniques is still a challenging and time-consuming task. A novel fast and robust alternative for distance measurements are Time-of-Flight (ToF) cameras, which provide range images in addition to gray-scale intensity images with high update rates [2]. Processing ToF data typically involves registration of the acquired distance images with pre-operative volume data.

The Iterative Closest Point (ICP) algorithm [3] is a widely used method for geometric alignment of 3D models. Given two roughly aligned shapes represented by two point sets, it iteratively (1) establishes point correspondences given the

* Thanks to R. Balachandran for the source code of the algorithm described in [1].

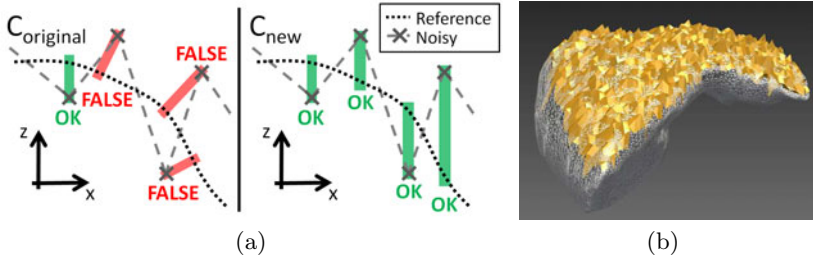


Fig. 1. (a) Schematic illustration of the establishment of point correspondences with the original ICP (left) and the anisotropic ICP (right). Reference mesh (represented by dotted black line), sparse noisy mesh (represented by gray crosses) and correspondences (boxes) are shown for $C_{original}$ (standard closest point operator) and C_{new} (new closest point operator with less weight given to the direction z). (b) Noisy submesh registered to a reference liver mesh via the proposed registration method.

current alignment of the data and (2) computes a rigid transformation accordingly. Although widely used, the algorithm implicitly assumes that the input points are observed with zero-mean, identical and isotropic Gaussian noise. The statistical errors related to the process of generating 3D points from ToF range data, however, lead to highly anisotropic error distributions in the point data. This, in turn, may result in wrong assignments during a correspondence search, as illustrated in Fig. 1(a).

In a related paper, we proposed a new anisotropic variant of the ICP [4] which allows for definition of a covariance matrix for each point in the noisy input set. In this paper, we extend the algorithm, such that a covariance matrix can be defined for each point in *both* point sets to be aligned and show how to apply the derived variant of the ICP for fine registration of noisy ToF images with high-resolution surface data (sec. 2.1). The performance of the proposed method is assessed with a novel evaluation framework that provides physically realistic ToF range data generated from a virtual scene (sec. 2.2).

2 Materials and Methods

2.1 Method for Fine Registration of ToF Range Data

Let $X = \{\mathbf{x}_1, \dots, \mathbf{x}_{N_X}\}$ be a sparse point set acquired with a ToF camera to be registered with a dense reference set $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_{N_Y}\}$ (e.g., the set of vertices from a surface mesh). Given two points \mathbf{x} and \mathbf{y} , whose localization errors are assumed to be independent and can be represented by two zero-mean Gaussian distributions with covariance matrices $\Sigma_{\mathbf{x}}$ and $\Sigma_{\mathbf{y}}$ (rank 3), we define the anisotropically weighted distance between them as:

$$d_{new}(\mathbf{x}, \mathbf{y}) = \|\mathbf{W}_{\mathbf{x}\mathbf{y}}(\mathbf{x} - \mathbf{y})\|_2 \quad (1)$$

where a weighting matrix $W_{\mathbf{x}\mathbf{y}} = (\Sigma_{\mathbf{x}} + \Sigma_{\mathbf{y}})^{-\frac{1}{2}}$ accounts for the variance in the input data similar as in [14,5]. The closest point in a set Y to a given point \mathbf{x} is then defined as:

$$C_{new}(\mathbf{x}, Y) = \arg \min_{\mathbf{y}_i \in Y} \|W_{\mathbf{x}\mathbf{y}_i}(\mathbf{x} - \mathbf{y}_i)\|_2 \tag{2}$$

Based on the provided covariance matrices $\Sigma_{\mathbf{p}}$ for each point $\mathbf{p} \in X \cup Y$, the aim of the anisotropic ICP algorithm is to find a rotation matrix R and a translation vector \mathbf{t} such that the following error metric is minimized:

$$e(R, \mathbf{t}) = \frac{1}{N_X} \sum_{i=1}^{N_X} \|W_i(R\mathbf{y}_{idx(i)} + \mathbf{t} - \mathbf{x}_i)\|_2^2 \tag{3}$$

where $idx(i) = \arg \min_j d_{new}(\mathbf{x}_i, \tilde{\mathbf{y}}_j)$ with $\tilde{\mathbf{y}}_j = R\mathbf{y}_j + \mathbf{t}, j = 1, \dots, N_Y$ and $W_i = (R\Sigma_{\mathbf{y}_{idx(i)}}R' + \Sigma_{\mathbf{x}_i})^{-\frac{1}{2}}$.

Similar as in [4], this is achieved by iteratively (1) establishing point correspondences with the new closest point operator C_{new} given the current alignment of the data and (2) computing a rigid transformation for mapping the current corresponding points using an extension of the recently published algorithm by Balachandran and Fitzpatrick [1]. The latter minimizes an anisotropically weighted FRE, where the weighting matrix for two corresponding points is a function of their covariance matrices as depicted above. It can be shown, that the cost function given in eq. [3] decreases in each iteration of the proposed anisotropic ICP (proof by contradiction).

To apply the anisotropic ICP for registration of two surfaces represented by the point sets X and Y , it is necessary to define a covariance matrix $\Sigma_{\mathbf{p}}$ for each $\mathbf{p} \in X \cup Y$ representing the localization uncertainty in that point. We propose the following three variants for defining $\Sigma_{\mathbf{x}}$ for a point $\mathbf{x} \in X$ on the ToF surface:

ANISOTROPIC: We assume that the localization error for each point occurs primarily along the ray connecting the imaged object point \mathbf{x} with the center of the associated chip pixel and set the remaining two principal axes orthogonal to that ray. Let $V_{\mathbf{x}}$ be the matrix whose columns represent the normalized principal axes. We then set $\Sigma_{\mathbf{x}} = V_{\mathbf{x}}S_{\mathbf{x}}^2V_{\mathbf{x}}'$ with $S_{\mathbf{x}} = \text{diag}(s_{\mathbf{x}1}, s_{\mathbf{x}2}, s_{\mathbf{x}3})$. The standard deviation $s_{\mathbf{x}1}$ along the ray is set to a constant value depending on the distance of the examined object to the camera (default: 3mm), whereas $s_{\mathbf{x}2}$ and $s_{\mathbf{x}3}$ represent the lateral statistical error and are set to the empirically determined value $s_{\mathbf{x}2} = s_{\mathbf{x}3} = 0.1 \text{ mm}$. Note that in this variant, the covariance matrices can be computed prior to image acquisition.

ANISOTROPIC_SIM: A variant of ANISOTROPIC where the variance along the first principal axis is determined with a ToF simulator: For a given illumination power, reflectivity of the imaged object, as well as optics characteristics, the standard deviation can be determined as a function of the measured distance $d(\mathbf{x})$ of \mathbf{x} to the camera: $s_{\mathbf{x}1} = f(d(\mathbf{x}))$. This function can be determined by computing $f(d_i)$ for a set of distances d_i using the simulator presented in sec. [2.2]

and fitting a spline to the output data. The performed simulations can be seen as a training for the method.

ANISOTROPIC_EST: A variant of ANISOTROPIC where the variance along the first principal axis is estimated from the depth values corresponding to \mathbf{x} from the last n samples (default: $n = 10$). Note that the standard deviation needs to be computed from a relatively small amount of data to avoid motion artefacts.

When registering the *mean* image of a set of n ToF images, the standard deviations computed for a single image are divided by \sqrt{n} .

To account for the resolution of the mesh represented by Y , we set the first two principal axes $\mathbf{v}_{\mathbf{y}1}$, $\mathbf{v}_{\mathbf{y}2}$ for a point $\mathbf{y} \in Y$ in that mesh orthogonal to the normal vector \mathbf{n} of the corresponding vertex. The standard deviation along these axes is set to the radius of a circle with area $A_V(\mathbf{y})$, where $A_V(\mathbf{y})$ is the area of the Voronoi region corresponding to \mathbf{y} . The rationale behind this procedure is that the variance along the surface should increase with an increasing size of the triangles associated with this vertex to ensure that corresponding vertices in two meshes to be registered do not need to correspond to exactly the same anatomical location. Note, however, that different spreads along different directions are not accounted for. The standard deviation $s_{\mathbf{y}3}$ along the vertex normal can either be set to a value close to 0, thus representing a noise-free target mesh, or to a value inversely proportional to the intensity gradient in that point, thus reflecting the segmentation error (provided that the mesh was generated from volume data). We then set $\Sigma_{\mathbf{y}} = V_{\mathbf{y}} S_{\mathbf{y}}^2 V_{\mathbf{y}}'$ with $S_{\mathbf{y}} = \text{diag}(s_{\mathbf{y}1}, s_{\mathbf{y}2}, s_{\mathbf{y}3})$ and $V_{\mathbf{y}} = [\mathbf{v}_{\mathbf{y}1} \ \mathbf{v}_{\mathbf{y}2} \ \mathbf{n}]$.

2.2 Evaluation Framework

The evaluation framework consists of the following components (cf. Fig. 2):

ToF Surface Generator: This component generates a noisy ToF mesh from a given high-resolution surface mesh based on a chosen camera type (default: CamCube 2.0; PMDTechnologies GmbH, Siegen, Germany), a pose of the camera relative to the object and (optionally) the object's reflectivity (per vertex). It consists of three modules. The *ideal depth map generator* provides means for placing the input mesh relative to the chosen ToF camera in a virtual scene and generates a corresponding ideal depth map based on the intrinsic camera parameters. In the second step, the *ToF camera simulator* generates a noisy depth map from the ideal depth map, taking into account realistic sensor properties like the quantum efficiency, dark currents, system amplification factors as well as the digitization process. It applies a previously proposed concept for ToF *sensor* simulation [6] and additionally simulates the attenuation of illumination intensity with depth as well as the lateral statistical error caused by non-ideal optics. The latter is achieved by convoluting the input of the virtual camera with a Gaussian blur kernel prior to surface generation. Systematic errors, such as pixel offsets and the periodical depth deviation (*wiggling error*) are not simulated because unlike statistical errors, they can be compensated for by a calibration

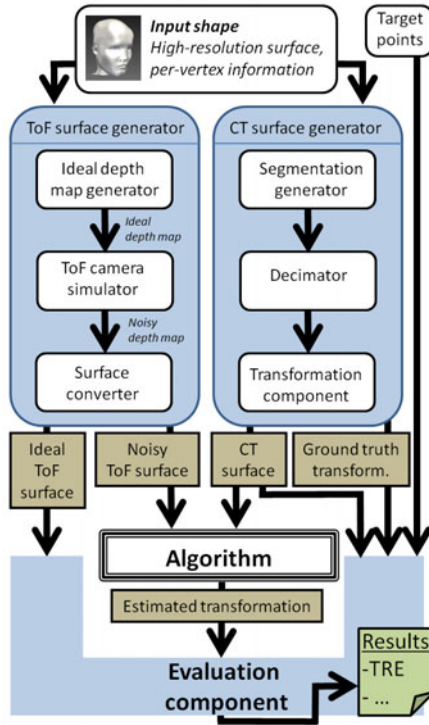


Fig. 2. Evaluation framework developed for this study (cf. sec. 2.2)

procedure based on reference data [2]. Finally, the *surface converter* generates a surface from the noisy depth map based on the intrinsic camera parameters. All points with a corresponding distance value above a certain threshold are excluded, hence, an object segmentation is automatically generated.

CT Surface Generator: This component generates a surface representing the reference input mesh for the algorithm to be evaluated. Segmentation errors can be simulated by adding noise to the vertices based on the intensity gradient in that point, and a decimator can be applied for modifying the resolution of the mesh. Finally, an initial pose relative to the ToF surface in the world coordinate system can be chosen.

Evaluation Component: Based on a set of provided target points, the output of the other components and the transformation computed by the algorithm to be evaluated, this module computes the target registration error (TRE), which we define as the root-mean-square (RMS) distance between the ground truth target positions and the corresponding target positions according to the transformation computed by the algorithm. Furthermore, the translation error and the rotational error are determined. If the algorithm under evaluation yields vertex

correspondences (such as the ICP), the percentage of correct vertex correspondences is additionally determined. Ground truth correspondences are calculated by transforming the *ideal* depth map into a 3D point cloud and determining the closest vertex in the simulated CT mesh for each ToF mesh point.

2.3 Experiments

Artificial Setting: The goal of the first experiment was to demonstrate with a simple example that the proposed algorithm is better suited for coping with anisotropic noise than the original ICP algorithm. For this purpose, an object consisting of two sides arranged perpendicular to each other (imagine a book opened with an angle of 90°) was placed in the world coordinate system such that the direction of view of the virtual ToF camera was almost perpendicular to one of the two sides of the object and almost parallel to the other side (a small portion of which was still visible). The idea behind this was, that - similar as in the example shown in Fig. 1(a) - wrong assignments with the standard ICP should primarily occur in one direction (towards the other side) and could not be compensated for by wrong assignments in the opposite direction. Based on an ideal depth map generated by the *ideal depth map generator*, a noisy depth map was generated by the *ToF camera simulator* with a simulated exposure time of 25 ms and a modulation contrast of the light source of 0.5. All sensor parameters (quantum efficiency, dark currents, etc.) were set to values measured for the CamCube 2.0 ToF system using methods proposed by Erz *et al.* [7]. The illumination power and aperture of the optics were set to values which led to 90% pixel saturation on average. The optics influence was simulated as a Gaussian point spread function with a full width at half maximum (FWHM) of $20 \mu\text{m}$. Next, a noisy partial surface was generated with the *surface converter*, and both, the original ICP and the proposed variant were used to register the noisy surface to the high-resolution surface mesh, starting from the ideal position and assuming a noise-free reference mesh ($s_{y3} = 10^{-20}$). For both algorithms, the percentage of correct correspondences as well as the TRE for a set of points distributed in the target object were determined.

Medical Setting: To evaluate the performance of the proposed ICP in the medical context, three high-resolution surface meshes were generated from medical imaging data: a liver with little depth variation (LIVER1), a liver with more depth variation (LIVER2) and a face mesh (FACE). For the livers, the targets were distributed within the organ itself. In the case of the face, the targets were distributed in the corresponding brain segmentation. For each mesh, an ideal ToF image was created with an approximately ventral view of the camera on the patient. These images were averaged to obtain a mean depth map which was then converted into a noisy partial surface using the same parameters as in the *Artificial setting*. Both, the standard ICP and the proposed variant of the ICP were then used to register the noisy ToF mesh (between 1800 and 3000 points) with the original mesh (approximately 10 000 points), starting with a small known initial misalignment. For each image and each method, the TRE and the percentage of correct point correspondences was computed.

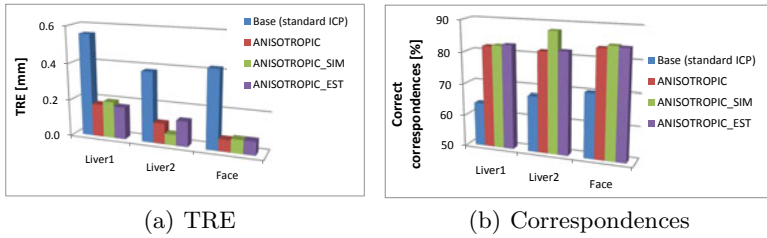


Fig. 3. Summary of the results for the *medical setting* showing (a) the TRE and (b) the percentage of correct matches after convergence for the different registration methods and algorithms

3 Results

Registration of the artificial object with the standard ICP yielded a TRE of 2.9 mm compared to 0.5 mm (ANISOTROPIC), 0.4 mm (ANISOTROPIC_SIM), and 0.4 mm (ANISOTROPIC_EST) with the proposed algorithm. In the former case (standard ICP), the percentage of correct correspondences was 23.1% in the ground truth position and 0.8% after ICP convergence, which could be increased to $(30 \pm 1)\%$ and $(27 \pm 1)\%$ respectively when applying the new registration method (averaged over the three variants). The registration results for the medical imaging data are shown in Fig. 3. Again, the different variants of the proposed algorithm performed similarly, yielding a mean decrease of the TRE of $(73 \pm 2)\%$, and a mean increase of the percentage of correct correspondences of $(24 \pm 2)\%$ on average. A typical registration result is shown in Fig. 1(b).

4 Discussion

We introduced a new method for fine registration of partial ToF surfaces with high-resolution surface meshes based on an anisotropic variant of the ICP algorithm and assessed its performance with a novel evaluation framework. To our knowledge, we are the first to present a framework for ToF registration algorithms based on physically realistic ToF sensor simulations. The simulated sources of errors include various errors related to the process of generating a ToF depth image from a given scene such as the influence of interfering non-modulated background light and dark currents, as well as errors caused in the process of mesh segmentation and mesh decimation. As the framework generates ideal ToF surfaces in addition to the noisy ones, it is also suitable for evaluation of ToF preprocessing algorithms (such as denoising methods).

According to our experiments, the proposed method is better suited for fine registration of ToF range data than the standard ICP, reducing the TRE by over 70% and increasing the percentage of correct matches by more than 20% on medical imaging data. Although the percentage of correct matches was always below 70% for the standard ICP, the absolute value of the TRE was quite low

(< 1 mm), suggesting that wrong assignments may “average out” because they occur equally in all directions. In an artificial setting, however, we showed that depending on the geometry of the imaged object, this compensation is not always possible, thus causing a large TRE.

In this study, the proposed variants for defining the covariance matrices as input for the anisotropic ICP yielded similar results. This can be attributed to the fact that we used the same reflectivity for all vertices and hence, similar standard deviations in each pixel. Furthermore, it should be noted that the training and testing conditions for ANISOTROPIC_SIM were similar. Future studies should account for these aspects.

In conclusion, the proposed registration method is well suited for dealing with anisotropic noise in the context of ToF range data registration, yielding a low TRE and a high percentage of correct vertex correspondences. Future work includes (1) assessment of convergence speed and accuracy with respect to the initial pose, (2) comparison with other variants of the ICP as well as with other methods for point registration in the presence of anisotropic noise, (3) evaluation on *in-vitro* and *in-vivo* data, and (4) validation of the extended ToF simulator.

References

1. Balachandran, R., Fitzpatrick, J.M.: Iterative solution for rigid-body point-based registration with anisotropic weighting. In: SPIE Medical Imaging, vol. 7261, p. 72613D (2009)
2. Kolb, A., et al.: Time-of-Flight sensors in computer graphics. In: Eurographics State of the Art Reports, pp. 119–134 (2009)
3. Besl, P.J., et al.: A method for registration of 3-d shapes. IEEE T. Pattern Anal. 14, 239–256 (1992)
4. Maier-Hein, L., et al.: Iterative closest point algorithm in the presence of anisotropic noise. In: Bildverarbeitung für die Medizin (BVM), pp. 231–235. Springer, Heidelberg (2010)
5. Danilchenko, A., Fitzpatrick, J.M.: General approach to error prediction in point registration. In: SPIE Medical Imaging, vol. 7625, p. 76250 F (2010)
6. Schmidt, M., Jähne, B.: A physical model of Time-of-Flight 3D imaging systems, including suppression of ambient light. In: Kolb, A., Koch, R. (eds.) Dyn3D 2009. LNCS, vol. 5742, pp. 1–15. Springer, Heidelberg (2009)
7. Erz, M., Jähne, B.: Radiometric and spectrometric calibrations, and distance noise measurement of ToF cameras. In: Kolb, A., Koch, R. (eds.) Dyn3D 2009. LNCS, vol. 5742, pp. 28–41. Springer, Heidelberg (2009)

Self-encoded Marker for Optical Prospective Head Motion Correction in MRI*

Christoph Forman^{1,2}, Murat Aksoy¹,
Joachim Hornegger², and Roland Bammer¹

¹ Department of Radiology, Stanford University, Stanford, California, USA

² Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany

Abstract. The tracking and compensation of patient motion during a magnetic resonance imaging (MRI) acquisition is an unsolved problem. For brain MRI, a promising approach recently suggested is to track the patient using an in-bore camera and a checkerboard marker attached to the patient's forehead. However, the possible tracking range of the head pose is limited by the locally attached marker that must be entirely visible inside the camera's narrow field of view (FOV). To overcome this shortcoming, we developed a novel self-encoded marker where each feature on the pattern is augmented with a 2-D barcode. Hence, the marker can be tracked even if it is not completely visible in the camera image. Furthermore, it offers considerable advantages over the checkerboard marker in terms of processing speed, since it makes the correspondence search of feature points and marker-model coordinates, which are required for the pose estimation, redundant. The motion correction with the novel self-encoded marker recovered a rotation of 18° around the principal axis of the cylindrical phantom in-between two scans. After rigid registration of the resulting volumes, we measured a maximal error of 0.39 mm and 0.15° in translation and rotation, respectively. In in-vivo experiments, the motion compensated images in scans with large motion during data acquisition indicate a correlation of 0.982 compared to a corresponding motion-free reference.

1 Introduction

Patient motion during data acquisition remains a challenging problem in MRI. The consequences are often significant image artifacts which lower the diagnostic confidence of the image data. Recent publications have proposed methods to reduce or compensate the impact of motion on the images. Techniques using PROPELLER or spiral sequences correct patient motion with alternative data acquisition strategies [1,2]. Navigator echos are added to MR sequences

* This work was supported in part by the National Institutes of Health (Grant numbers: 1R01EB008706, 5R01EB002711, 1R01EB006526, 1R21EB006860, P41RR09784), Lucas Foundation, Oak Foundation, Bavarian California Technology Center, and GE Healthcare.

to compensate retrospectively and prospectively patient’s motion during the scan [3,4,5]. Another MR based motion correction method was introduced by Ooi et al. [6], using the response of active markers in form of small coils attached to the forehead of the patient. For a data acquisition independent approach, external tracking systems were proposed. In order to transfer the detected motion of these systems to motion, which actually occurred in the scanner image plane, a cross-calibration of both frame of references is required. External optical systems outside the scanner bore were used to track a marker attached to the patient’s head [7]. Drawback of this system is, that it requires a line of sight on the marker inside the scanner. Aksoy et al. [8] introduced a motion correction system with an in-bore camera. In this approach, an MR-compatible camera is mounted on the head coil, tracking the position and orientation of a checkerboard marker attached to the patients forehead. One essential constraint of this method is that once the marker is occluded by another object or is partly outside the camera’s field of view (FOV), no motion correction is possible anymore. The restricted space inside the scanner bore entails camera-marker distances between 5 and 7 cm. Additionally, the shape of different coils may occlude parts of the camera FOV. Thus, the restriction of the marker detection delimits the possible tracking range of the patient’s head position. To overcome this limitation, we developed the self-encoded marker with additional codes for each feature point.

2 Materials and Methods

The motion correction system was implemented on a GE Signa 1.5T whole body system (GE Healthcare, Milwaukee, WI). Fig. 1 shows the setup of the system. The MR-compatible camera is mounted on the 8 channel head coil. Infrared diodes attached to the camera body illuminate the scene inside the scanner bore. An independent tracking computer processes the captured camera images as described in [8,9]. The patient’s pose at the beginning of each scan serves as initial point of reference to describe the motion throughout the scan. For each camera image, the detected motion of the optical system is transformed into motion, which actually occurred in the scanner image plane. This requires an initial cross-calibration of the tracking system with the MR scanner. The pose updates in form of translation and rotation are relative to the patient’s initial position at the first data acquisition. They are sent in real-time via network connection to the MR sequencer. Assuming rigid head motion, these updates are directly used by the sequencer to adjust the gradients and radio frequencies before each data acquisition. That way, the slice position and orientation is determined by scanned anatomy and not by the scanner geometry. Once a large difference in rotation or translation between two subsequent pose updates was detected, the current acquisition data was disregarded and repeated to compensate the latency of the entire scan.

Basis of the self-encoded marker is a checkerboard pattern. Adjacent corners of neighboring quads on this pattern describe the feature points of this marker. For the checkerboard marker all features are required to establish the point correspondences of detected feature points in the camera image and their model

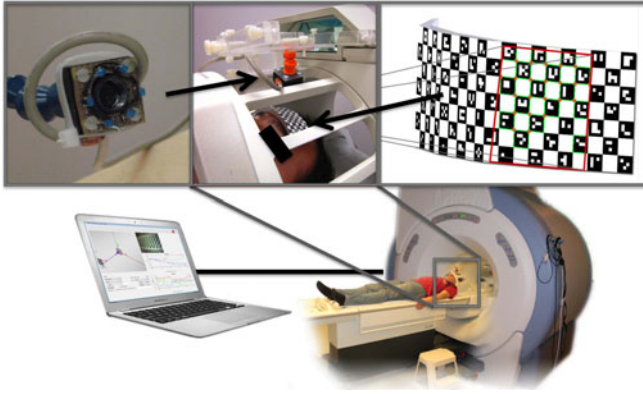


Fig. 1. Setup of the optical motion correction system: The MR compatible camera is mounted on the 8 channel head coil. The signal is processed by an external tracking computer, which sends the pose updates to the MR scanner via network connection.

coordinates in the marker-model geometry. Within the black quads of the self-encoded marker, 2-D barcodes similar to the ARTag marker [10] are embedded. These unique codes identify each feature of the pattern, specify its position on the marker geometry and consequently define the aforementioned point correspondences. In contrast to the ARTag marker we are using a 10 bit encoding instead of 36 bit for the embedded code. The redundancy in the ARTag marker permits a verification of the code. Instead of verifying each quad independently, we compare the recognized code of each quad and its neighboring quads with a map containing all positions of the codes on the marker. That way, every bit of the embedded code can be used for the encoding, which leads to a more robust detection of the marker in the in-bore camera image. Based on the defined point correspondences and the known intrinsic camera parameters, the marker pose is estimated by homography (planar marker) or direct linear transformation (3-D marker). Even if only parts of the self-encoded marker are visible to the camera, its pose can still be determined. Furthermore, different feature points of the self-encoded marker can be used for a robust tracking of the marker position and orientation. That way, we overcome the limitation of the checkerboard marker to the camera's FOV.

For the pose estimation of the marker, first, the captured camera image is converted into a binary image by thresholding. The outline of the black quads is detected by quadrangular contours in this image. Using the boundary of each quad we sample the interior into a 5×5 grid. Then, the embedded code in the inner 3×3 cells is classified by thresholding into a binary code. By means of a-priori knowledge of the marker layout, the recognized codes are verified as mentioned above. That way, erroneous detected quads can be eliminated. Finally, the relative pose of the marker to the camera is estimated using the point correspondences of detected features and marker-model points.

3 Experiments and Results

Phantom and in-vivo experiments were performed using an axial 3D spoiled gradient recalled (SPGR) sequence with $TR = 9.5$ ms, $TE = 4.1$ ms, flip angle = 20° , slice thickness = 1.5 mm, FOV = 24 cm, and a resolution of $192 \times 192 \times 96$.

3.1 Phantom Experiment

We evaluated the accuracy of the optical motion correction system with a cylindrical phantom and two subsequent MRI scans. In-between both scans, the static phantom was manually rotated about its principal axis by 18° . The first scan used as a reference, was compared to the motion corrected second scan. Assuming an ideal motion correction system, we expected an identical image of the phantom in both scans. For qualitative evaluation, Fig. 2 shows both scans and the difference image. We performed this experiment with both markers. The difference images indicate a discrepancy of the structure at the top of the phantom, which is caused by the phantom being not completely filled with water. While the structure of the phantom was rotated, the water remained at the same position. For quantitative evaluation of the residual mismatch, retrospective rigid registration was performed. This registration resulted in a remaining offset of:

$$\begin{aligned} \mathbf{t}_{self-enclosed} &= (-0.36, 0.10, -0.39) \quad [\text{mm}] \\ \mathbf{r}_{self-enclosed} &= (0.11, 0.00, 0.15) \quad [^\circ] \\ \mathbf{t}_{checkerboard} &= (0.89, 0.09, -0.79) \quad [\text{mm}] \\ \mathbf{r}_{checkerboard} &= (-0.35, -0.03, -0.29) \quad [^\circ] \end{aligned}$$

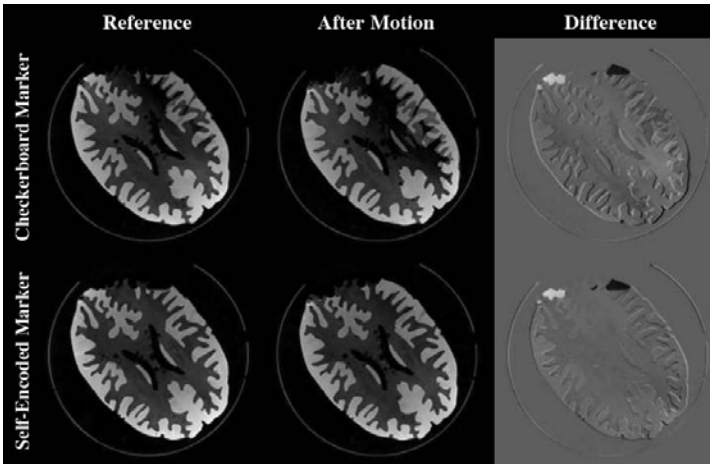


Fig. 2. The reference scan is compared to a scan with correction after 18° rotation. Both difference images are with a contrast enhancement by a scaling factor of 2.

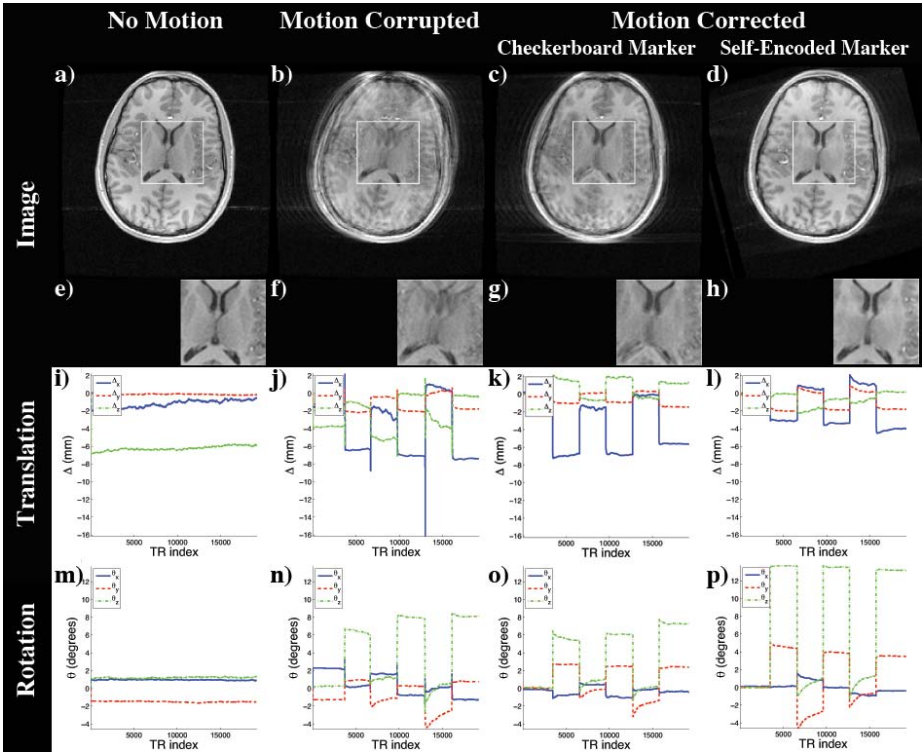


Fig. 3. Images of the scan comparing self-encoded and checkerboard marker: (a) Reference Scan; Scan with motion and (b) no correction, motion correction using the pose updates of the (c) checkerboard and the (d) self-encoded marker; (e-h) Magnification of window in (a-d); Detected translation (i-l) and rotation (m-p)

3.2 In-vivo Experiments

In-vivo experiments were performed on a healthy volunteer to evaluate the position estimates of the self-encoded marker for motion correction. For each scan, the obtained pose estimates relative to the initial head position were recorded in a log file.

In the first experiment, the pose estimates of the reference checkerboard marker and the novel self-encoded marker were compared. Four scans were obtained for this experiment. In order to track the head motion during data acquisition, first the checkerboard marker was attached to the forehead. In the first scan, the volunteer was instructed to maintain a stationary head position to create a motion-free reference image. For the following scans, the subject was asked to perform a similar head rotation every 30 seconds in order to assure a comparable motion pattern. In the second scan, the motion-correction system was turned off and the position estimates of the checkerboard marker were recorded. The obtained pose updates of this marker were used in the third scan to adapt

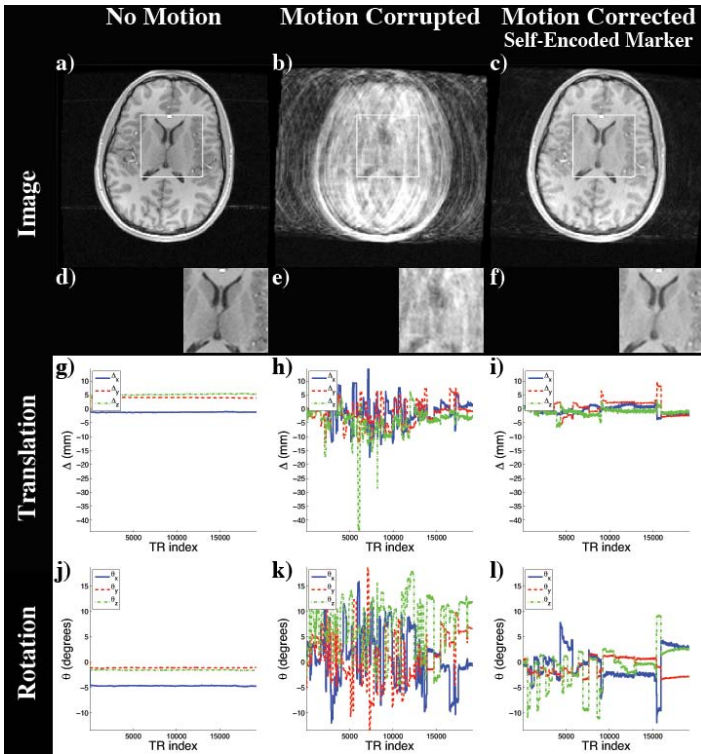


Fig. 4. Images of the scan simulating an uncooperative patient: (a) Reference Scan; Scan with random motion and (b) no correction and (c) prospective motion correction using the self-encoded marker; (d-f) Magnification of window in (a-c); Detected translation (g-i) and rotation (j-p)

the scanner for motion, while in the last scan the self-encoded marker was attached to the forehead to track the volunteers head motion. Fig. 3 shows the resulting images of the performed scans. Without correction, the MRI images exhibited motion artifacts. Using the pose updates of the checkerboard marker, these artifacts were reduced. However, inaccuracies of the marker became apparent in a mismatch of the scanned anatomical structure. In this camera setup the tracking range of the checkerboard marker was restricted to 6° . The self-encoded marker was able to extend it to a head rotation of 13° , which is maximal without touching the coil. The improvement in accuracy of the self-encoded marker compared to the checkerboard marker was measured by Pearson’s correlation coefficient [11]. Whereas the correlation of reference and motion-corrupted image resulted in a coefficient of 0.908, the optical tracking system using the pose updates of the checkerboard marker improved this value to 0.936. Using the self-encoded marker for the tracking of the volunteer showed a correlation of 0.971.

In the second experiment, the volunteer was asked to simulate an uncooperative behavior. First, a reference image was acquired. Then, in the following scans, the volunteer performed a random trembling motion for the entire scan. The pose estimates of the self-encoded marker were used to describe the head position over time during the data acquisitions. For the second scan, the detected translation and rotation were recorded whereas the scanner was also adapting for motion in the third scan. Due to the performed random motion, it was not possible to repeat the experiment with the identical motion pattern. Continuous motion has a strong impact on the resulting images, see Fig. 4. The motion induced artifacts corrupted the entire anatomical structure of the brain. While adapting the scanner geometry based on the detected head pose of the volunteer, the system was able to recover the structure of the brain. The effects of motion resulted in a correlation coefficient of 0.858, while the pose updates of self-encoded marker were able to improve this factor to 0.982.

3.3 Performance

We compared the performance of the self-encoded marker with 76 features to a checkerboard marker with 20 features. The captured image of the in-bore camera had a resolution of 640×480 pixel. For the analysis of the computational time the software run on a Intel Core2Duo CPU (2.26GHz). The entire process of feature detection and point correspondence search took 17.2ms for the self-encoded marker and 28.6ms for the checkerboard marker (acceleration factor 1.7x). Compared to a common checkerboard detection algorithm, this factor was achieved with a optimized detection of the self-encoded marker, since its outcome is not crucial for the generation of the point correspondences.

4 Discussion

A crucial limitation of existing in-bore tracking systems for prospective motion correction in MRI is the narrow FOV of the camera. In this study, we introduced a novel marker design with embedded 2-D barcodes that identify each feature on the pattern. Recognizing these codes in the captured camera image the tracking algorithm is able to estimate the pose of the self-encoded marker in situations where the marker is only partly visible.

We compared the accuracy of the self-encoded and checkerboard marker in a phantom experiment. The rotation of the phantom in-between two scans was compensated by the motion correction system in the second scan. Both resulting MR images were rigidly registered, which showed an improved accuracy for the self-encoded marker with a maximal offset of 0.39mm and 0.15° for translation and rotation, respectively. In the first in-vivo experiment, we compared the pose estimates of both markers for motion correction. This comparison study indicated the restricted range of motion that can be tracked with the checkerboard marker. The tracking range was extended from 6° with the checkerboard marker to 13° using the self-encoded marker. Furthermore, replacing the checkerboard

marker with the self-encoded marker increased the correlation of the resulting motion compensated MR images from 0.936 to 0.971 compared to a motion-free reference. In case of an uncooperative patient the motion compensation based on the pose updates of the self-encoded marker was able to recover the scanned anatomical structure. The correlation of the motion compensated scan resulted in a coefficient of 0.982 compared to a reference scan without motion.

In order to provide an estimate of the patient's head position for every data acquisition step, the total latency of the prospective motion correction system must not exceed the repetition time. Although we were able to accelerate the processing time of the self-encoded marker by a factor of 1.7, there is still room for further improvements.

References

1. Pipe, J.G.: Motion Correction With PROPELLER MRI: Application to Head Motion and Free-Breathing Cardiac Imaging. *Magn. Reson. Med.* 42(5), 963–969 (1999)
2. Sarty, G.E.: Single TrAjjectory Radial (STAR) imaging. *Magn. Reson. Med.* 51(3), 445–451 (2004)
3. Thesen, S., Heid, O., Mueller, E., Schad, L.R.: Prospective Acquisition Correction for Head Motion With Image-Based Tracking for Real-Time fMRI. *Magn. Reson. Med.* 44(3), 457–465 (2000)
4. Aksoy, M., Liu, C., Moseley, M., Bammer, R.: A Self-navigated Spiral In & Out Pulse Sequence Design for Retrospective Motion Correction. In: Proceedings of the 14th Annual Meeting of the ISMRM, Seattle, Washington (2006)
5. White, N., Roddey, C., Shankaranarayanan, A., Han, E., Rettmann, D., Santos, J., Kuperman, K., Dale, A.M.: PROMO: Real-Time Prospective Motion Correction in MRI Using Image-Based Tracking. *Magn. Reson. Med.* 63(1), 91–105 (2010)
6. Ooi, M.B., Krueger, S., Thomas, W.J., Swaminathan, S.V., Brown, T.R.: Prospective real-time correction for arbitrary head motion using active markers. *Magn. Reson. Med.* 62(4), 943–954 (2009)
7. Dold, C., Zaitsev, M., Speck, O., Firlie, E.A., Hennig, J., Sakas, G.: Prospective Head Motion Compensation for MRI by Updating the Gradients and Radio Frequency During Data Acquisition. In: Duncan, J.S., Gerig, G. (eds.) MICCAI 2005. LNCS, vol. 3749, pp. 482–489. Springer, Heidelberg (2005)
8. Aksoy, M., Newbould, R., Straka, M., Holdsworth, S., Skare, S., Santos, J., Bammer, R.: A Real Time Optical Motion Correction System Using a Single Camera and 2D Marker. In: Proceedings of the 16th Annual Meeting of the ISMRM, Toronto, Canada (2008)
9. Zaitsev, M., Dold, C., Sakas, G., Hennig, J., Speck, O.: Magnetic Resonance Imaging of Freely Moving Objects: Prospective Real-time Motion Correction Using an External Optical Motion Tracking System. *NeuroImage* 31(3), 1038–1050 (2006)
10. Fiala, M.: ARTag, a Fiducial Marker System Using Digital Techniques. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp. 590–596 (2005)
11. Edwards, A.L.: An Introduction to Linear Regression and Correlation. W. H. Freeman and Co., San Francisco (1993)

Robust 3D Visual Tracking for Robotic-Assisted Cardiac Interventions

Rogério Richa, Antônio P.L. Bó, and Philippe Poignet

LIRMM - UMR 5506 CNRS - UM2
161 Rue Ada 34392 Montpellier, France
{rogerio.richa, antonio.bo, philippe.poignet}@lirmm.fr

Abstract. In the context of minimally invasive cardiac surgery, active vision-based motion compensation schemes have been proposed for mitigating problems related to physiological motion. However, robust and accurate visual tracking is a difficult task. The purpose of this paper is to present a hybrid tracker that estimates the heart surface deformation using the outputs of multiple visual tracking techniques. In the proposed method, the failure of an individual technique can be circumvented by the success of others, enabling the robust estimation of the heart surface deformation with increased spatial resolution. In addition, for coping with the absence of visual information due to motion blur or occlusions, a temporal heart motion model is incorporated as an additional support for the visual tracking task. The superior performance of the proposed technique compared to existing techniques individually is demonstrated through experiments conducted on recorded images of an *in vivo* minimally invasive CABG using the DaVinci robotic platform.

Keywords: robotic assisted cardiac surgery, robust visual tracking, stereo.

1 Introduction

Recently, great developments have been made in the field of Minimally Invasive Surgery (MIS). While the benefits of this modality of surgery for the patient are numerous, mobility and visibility difficulties hinder the precise execution of the surgical gestures. For tackling some of these difficulties, surgical robots have been developed for assisting surgeons by improving the ergonomics, visualization and dexterity issues related to the minimally invasive procedure. However, current surgical platforms do not offer solutions for restoring the tactile feedback and physiological motion still needs to be manually compensated by the surgeons.

In this context, active vision-based motion compensation schemes [1] have been proposed for mitigating problems related to physiological motion during surgery. Such systems are particularly useful in beating heart interventions such as the off-pump minimally invasive coronary bypass artery grafting (CABG). Furthermore, the estimation of the beating heart motion using the visual feedback from the endoscope is a practical solution since no additional sensors are required in the MIS workspace.

In the literature, several techniques for tracking the heart motion using vision have been proposed. Mainly, visual cues such as salient features [2,3], texture [4] or shading [5] have been explored. However, most techniques display poor performance in the presence of the large heart surface deformations, illumination variations and specular reflections. Furthermore, occlusions by surgical tools are not handled in most cases.

The purpose of this paper is to present a hybrid tracker that estimates the 3D temporal and spatial deformation of a selected region of interest on the heart surface using stereo endoscopic images. The method comprises the estimation of a parametric deformable model for representing the heart surface based on the most reliable outputs from multiple visual tracking techniques. In addition, for coping with the absence of visual information due to motion blur or occlusions (e.g. by specular reflections, surgical instruments), a temporal heart motion model based on a time-varying dual Fourier series is incorporated as an additional support for the visual tracking task. Hence, the proposed technique is able to robustly track large regions of interest on the heart with high spatial resolution while naturally handling eventual occlusions. The superior performance of the proposed technique compared to existing techniques individually is demonstrated through *in vivo* experiments conducted on recorded images of a minimally invasive CABG using the DaVinci robotic platform.

2 Methods

2.1 The Hybrid Visual Tracking Method

In the context of MIS, the assumptions on which most tracking algorithms are based are often violated due the presence of large soft-tissue deformations, illumination variations and specular reflections, making continuous and accurate visual tracking with a single technique a difficult task. In this study, the use of multiple visual cues is proposed for increasing tracking robustness and spatial resolution. Using multiple methods, the failure of an individual technique can be circumvented by others, enabling the estimation of the heart surface deformation with superior robustness and spatial resolution. Furthermore, all tracking techniques used in the scheme run in parallel and can be implemented in a computationally efficient manner.

Our objective is the estimation of the deformation of a region of interest on the heart surface, manually chosen by the surgeon from any of the stereo images. For this purpose, rectified images of a calibrated stereo endoscope were used. Three different tracking methods are employed: two feature-based approaches – the modified Lucas-Kanade tracker proposed in Stoyanov et al. [2] and the SIFT [6] – and the region-based iterative registration technique proposed in Richa et al. [4]. Since the proposed framework is modular, additional methods can be easily incorporated for increased tracking quality.

Initially, for choosing the most reliable estimate of the heart surface deformation given the outputs of multiple methods, a quality evaluation step is executed. For this purpose, the alignment error is chosen as a measure of tracking quality

and the normalized cross correlation (NCC) in a subregion around each feature or control point is computed between both left and right stereoscopic images and the reference image. If the image alignment error drops below a defined threshold ($\tau < 0.90$), the 3D coordinates relative to a given estimate is treated as unreliable. The threshold τ was chosen empirically and is set very high for avoiding the need of fine tuning and false matches.

Next, using the set of reliable estimates of the 3D motion of the region of interest, a dense model of the heart surface deformation is computed. Here, the Thin-Plate Spline (TPS) model proposed in [4] is employed since it offers a good approximation of the heart surface shape in comparison with similar parametric deformable models. However, it can also be replaced by models that handle surface topological changes [7] or consider the biomechanical tissue properties [8]. The TPS is an interpolating function $m : \mathbb{R}^2 \rightarrow \mathbb{R}$ of a 2D point \mathbf{x} , defined by the basis function $B(r) = r^2 \log(r^2)$, a $(n+3)$ parameter vector $(w_1, \dots, w_n, a_1, b_2, c_3)$ and a set of control points $\mathbf{c} = (\tilde{x}, \tilde{y})$:

$$m(\mathbf{x}) = c_1 + a_2x + b_3y + \sum_{i=1}^n w_i B(\|\mathbf{c}_i - \mathbf{x}\|) \quad (1)$$

The TPS model is computed using the reliable tracking estimates as control points \mathbf{c} . By stacking three TPS functions $[m^x(\mathbf{x}) \ m^y(\mathbf{x}) \ m^z(\mathbf{x})]$, the 3D position of any point on the reference image of the region of interest selected by the surgeon can be computed in 3D by back-projection using the parameters of the calibrated stereo cameras.

The 3D deformable model has two functions: it provides a reliable dense model of the heart surface deformation and interpolates the unreliable tracking measurements. The latter consists in one of strong points of the proposed method, enabling tracking to recover from failures (due to large deformations, local minima problems). In our works, we found that no regularization or constraints on the TPS surface deformation were necessary (consequence of the alignment error threshold). This represents an advantage since *ad hoc* assumptions about the heart surface deformations are avoided. Figure 1 summarizes the different steps of the hybrid tracking method.

2.2 Temporal Modeling of the Beating Heart Motion

Due to motion blur, large specular reflections or the motion of the surgical instruments, visual information from the tracked region of interest may not be available for certain periods of time. For circumventing such problems, the quasi-periodic beating heart motion can be modeled and occlusions can be bridged using the motion predicted by the model.

As proposed in [9], the heart motion can be considered as the sum of the respiratory and cardiac motions, which can be represented as a dual non-stationary Fourier series. Given the 3D coordinates $\mathbf{p} = [x \ p \ y \ p \ z \ p]$ of a given point on the heart surface, the motion dynamics p of a Cartesian coordinate at a given instant t can be parameterized as the sum of two Fourier series, such that:

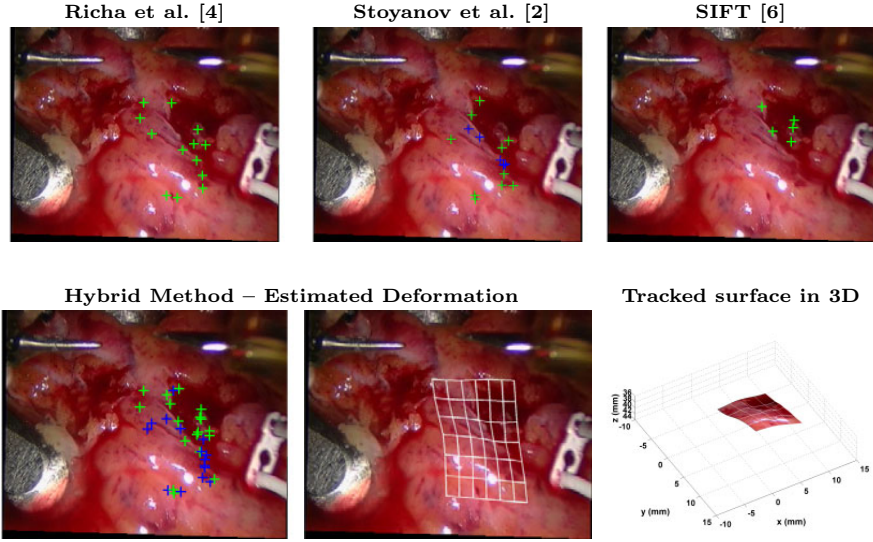


Fig. 1. (Top row) The output of the tracking techniques used in the study from the left camera of the stereo endoscope. Lost features from [2] are marked in blue. Notice that only two reliable SIFT matches were available when large deformations occurred. (Bottom left) The reliable correspondences are marked in green, while interpolated unreliable estimates are marked in blue. (Bottom center) Interpolating TPS surface (Bottom right) Estimated 3D surface.

$$\begin{aligned}
 p(t, \mathbf{f}) = & \sum_{h=1}^{H_r} \left[a_h \sin \left(h \sum_{k=t_0}^t \omega_r(k) \right) + b_h \cos \left(h \sum_{k=t_0}^t \omega_r(k) \right) \right] + \\
 & c_r + \sum_{h=1}^{H_c} \left[d_h \sin \left(h \sum_{k=t_0}^t \omega_c(k) \right) + e_h \cos \left(h \sum_{k=t_0}^t \omega_c(k) \right) \right], \quad (2)
 \end{aligned}$$

where H_r and H_c are the number of harmonics for modeling the respiratory and cardiac components respectively, ω_r and ω_c are the respiratory and cardiac frequencies, $\sum_{k=t_0}^t \omega$ is the sum of all estimated ω starting from t_0 and \mathbf{f} is the corresponding vector containing the Fourier series parameters:

$$\mathbf{f} = [a_1, \dots, a_{H_r}, b_1, \dots, b_{H_r}, c_r, d_1, \dots, d_{H_c}, e_1, \dots, e_{H_c}]^T \quad (3)$$

Consequently, a point of interest (POI) \mathbf{p} can be modeled with $\gamma = 3 \cdot [2 \cdot ({}^x H_r + {}^x H_c) + 1 + 2 \cdot ({}^y H_r + {}^y H_c) + 1 + 2 \cdot ({}^z H_r + {}^z H_c) + 1]$ parameters plus the respiratory and cardiac frequencies, which are shared among all coordinates and points. The number of harmonics H_r and H_c among the xyz directions may vary due to differences in their motion complexity. For recursively estimating the Fourier series parameters, the Extended Kalman Filter (EKF) is employed. The EKF state vector \mathbf{y} for estimating the trajectory of φ POIs $\mathbf{p} = [{}^x p \ {}^y p \ {}^z p]$ is composed of $(\varphi \cdot \gamma + 2)$ parameters, such that:

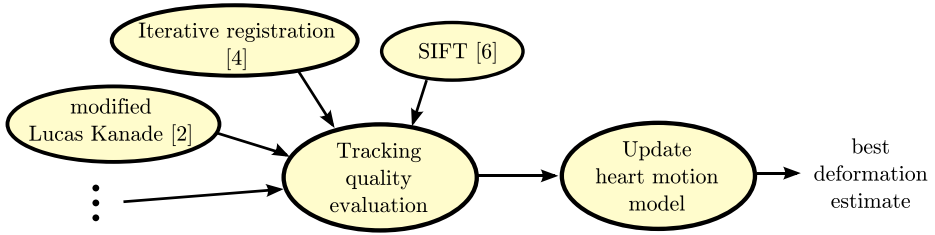


Fig. 2. A schematic overview of the hybrid visual tracking algorithm

$$\mathbf{y} = [{}^1\mathbf{f}_x, {}^1\mathbf{f}_y, {}^1\mathbf{f}_z, {}^2\mathbf{f}_x, {}^2\mathbf{f}_y, {}^2\mathbf{f}_z, \dots, {}^\gamma\mathbf{f}_x, {}^\gamma\mathbf{f}_y, {}^\gamma\mathbf{f}_z, \omega_r, \omega_c]^\top; \quad (4)$$

where $[{}^i\mathbf{f}_x, {}^i\mathbf{f}_y, {}^i\mathbf{f}_z]$ are the parameter vectors of the i -th estimated POI \mathbf{p}_i . In the correction step of the filter, the most reliable tracking outputs are used for updating the temporal heart motion model. For more details on the predictive EKF design, see [9].

In this framework, the predicted heart motion from a moment preceding a tracking loss can be used for bridging the disturbance. As described in section 2, tracking failures are detected by the tracking quality evaluation step which indicates if the number of tracked features is insufficient for estimating the current shape of the heart. The computation of future position estimates at a given instant t can be done in a straightforward manner using equation (2), considering a stationary system within the prediction horizon. Furthermore, spatially close points can be clustered for reducing the computational effort when estimating a large number of POI.

3 Experiments and Results

The diagram in Figure 2 summarizes all steps involved in the estimation of the heart surface motion. For evaluating the performance of the proposed hybrid tracking concept, two sets of experiments on recorded images of a minimally invasive CABG using the DaVinci robotic platform have been conducted. The first set aims to compare the performance of the hybrid tracker with the individual techniques ([4],[6]) while the second focuses on the prediction quality of the predictive EKF described in section 2.2.

3.1 Comparative Study

For the method proposed in Richa et al. [4], 14 control points distributed on textured parts of the region of interest are used. The maximal number of iterations for the minimization loop are fixed to 20. Although the method is robust to illumination changes and specular reflections, the image blur caused by large inter-frame motion can induce large errors in the estimation of the heart surface deformation (see figure 3).

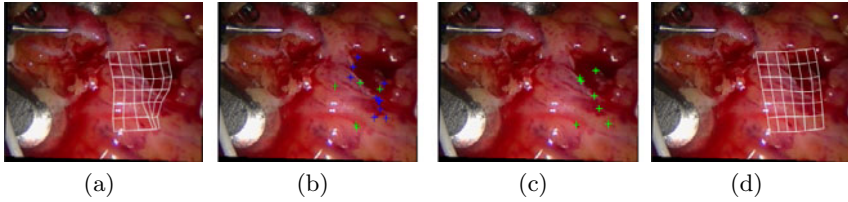


Fig. 3. (a) Tracking errors induced by motion blur and specular reflections in [4]. (b-c) Using the motion of stable features obtained with techniques [2] and [6] marked in green, respectively, a more reliable estimate of the heart surface deformation shown in (d) can be computed.

In Stoyanov et al. [2], tracking loss caused by large specular reflections and tissue deformation is the main performance issue. Figure 4 (top) indicates the significant drop in the number of tracked features while tracking (from the 30 Shi-Tomasi features initially detected within the reference region of interest, only 5 are tracked after 6 seconds of tracking).

Although a very high number of SIFT features are detected during the initialization (384 features within the region of interest), very few matches are available when to large tissue deformations occur (see Figure 4). Furthermore, ‘tracking-by-detection’ methods such as the SIFT are unsuitable for performing motion compensation since continuous tracking is not possible.

The proposed hybrid tracking concept offers the possibility of overcoming the failure of an individual technique using the output of others. In the example in figure 1, the hybrid tracker uses 14 control points [4], 15 Lucas-Kanade features [2] and 384 SIFT features [6]. Using the interpolating TPS surface, problems due to lost features or convergence errors can be circumvented. In addition, ‘tracking-by-detection’ methods such as the SIFT can be incorporated for increased tracking quality. Figure 4 shows the number of active features during tracking, demonstrating the superior performance of the proposed technique compared to existing techniques individually.

3.2 Improvements Using the Temporal Heart Motion Model

For evaluating the performance of the temporal heart motion model, the prediction errors for 0.2, 1 and 3 second prediction horizons are evaluated using the recorded tracked coordinates of the POI on the heart highlighted in Figure 5. The error is calculated as the Euclidean distance $\|\mathbf{d} - \mathbf{p}\|$ between the predicted \mathbf{d} and true \mathbf{p} positions of the POI for all xyz coordinates. The obtained root mean square and peak prediction errors at every motion sample were (0.73mm/1mm), (0.86mm/1.55mm) and (1.00mm/2.03mm) for the 0.2, 1 and 3 second prediction horizons, respectively. The low prediction errors attest the capability of the prediction scheme to overcome both short occlusions by specular reflections or motion blur and long occlusions by surgical instruments (it is important to remark that although the predicted motion is accurate enough to restart tracking,

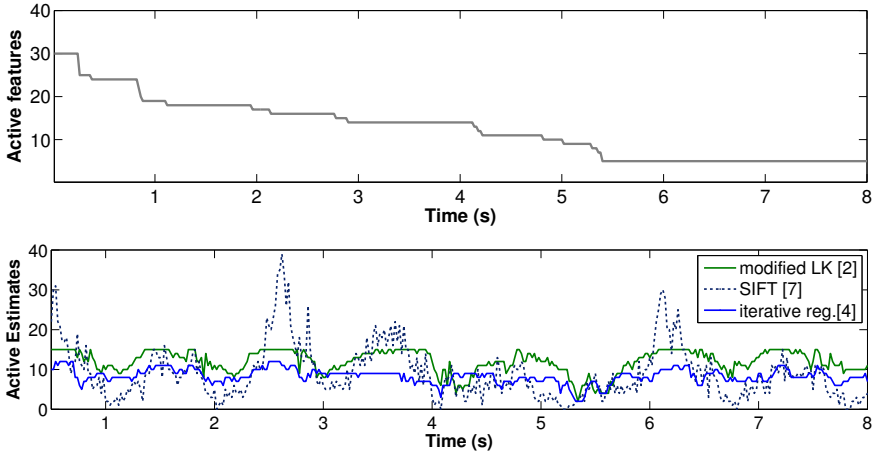


Fig. 4. (Top) The number of active tracked features using the method proposed in [2]. (Bottom) Active motion estimates during tracking using the hybrid method.

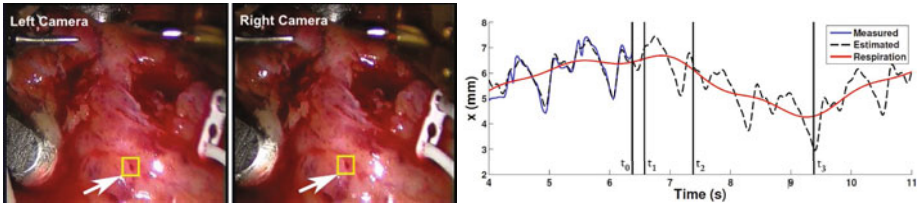


Fig. 5. The estimated dual Fourier series model of the x coordinate of the tracked feature on the heart surface highlighted on the left at $t_0 = 6.4s$. The 0.2s, 1s and 3s prediction horizons (t_1, t_2, t_3 respectively) used for evaluating the quality of the predicted heart motion are defined in the plots.

it cannot be used for motion compensation since the prediction errors can exceed the minimal precision requirements of cardiac surgery [9].

3.3 Computational Requirements

The computational burden introduced by the tracking quality evaluation and the EKF update is negligible compared to the computational time required by a single tracking technique. In addition, all techniques used in this study [2,4,6] have successfully been implemented in Graphics Processor Units (GPU) and tracking speeds over 100 Hz have been reported. Therefore, it is expected that the even though the proposed method incorporates an additional computational burden, tracking at high speeds is possible. This is a great advantage since the future deployment on a prototype surgical platform is envisaged.

4 Conclusion

In this paper, we presented a hybrid tracker for estimating the 3D motion of the heart surface using stereo endoscopic images. The method uses multiple visual trackers working in parallel for tracking with increased robustness and spatial resolution. For coping with tracking failures and occlusions, the temporal heart motion dynamics have also been incorporated as an additional support for the visual tracking task. Experiments on recorded *in vivo* images of a minimally invasive CABG using the DaVinci robotic platform attest the superior performance of the proposed tracker in comparison with existing techniques individually. The future implementation of the tracking method in a control scheme is envisaged.

References

1. Nakamura, Y.K., Kawakami, H.: Heartbeat synchronization for robotic cardiac surgery. In: Proceedings of IEEE International Conference on Robotics and Automation (ICRA 2001), Seoul, Korea, vol. 2, pp. 2014–2019 (May 2001)
2. Stoyanov, D., Mylonas, G.P., Deligianni, F., Darzi, A., Yang, G.Z.: Soft-tissue motion tracking and structure estimation for robotic assisted mis procedures. In: Duncan, J.S., Gerig, G. (eds.) MICCAI 2005. LNCS, vol. 3750, pp. 139–146. Springer, Heidelberg (2005)
3. Mountney, P., Yang, G.Z.: Soft tissue tracking for minimally invasive surgery: Learning local deformation online. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part II. LNCS, vol. 5242, pp. 364–372. Springer, Heidelberg (2008)
4. Richa, R., Poignet, P., Liu, C.: Three-dimensional motion tracking for beating heart surgery using a thin-plate spline deformable model. The International Journal of Robotics Research (IJRR) – Special Issue on Robot Vision 29(2-3), 218–230 (2010)
5. Visentini-Scarzanella, M., Mylonas, G.P., Stoyanov, D., Yang, G.Z.: *i*-BRUSH: A Gaze-Contigent Virtual Paintbrush for Dense 3D Reconstruction in Robotic Assisted Surgery. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5761, pp. 353–360. Springer, Heidelberg (2009)
6. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60(2), 91–110 (2004)
7. Stoyanov, D., Yang, G.Z.: Soft-tissue deformation tracking for robotic assisted minimally invasive surgery. In: Proceedings of IEEE International Conference of the Engineering in Medicine and Biology Society (EMBS 2009), Boston, USA, pp. 254–257 (2009)
8. Bogatyrenko, E., Hanebeck, U.D., Szabó, G.: Heart surface motion estimation framework for robotic surgery employing meshless methods. In: Proceedings of IEEE Conference on Intelligent Robots and Systems (IROS 2009), St. Louis, USA, pp. 67–74 (2009)
9. Richa, R., Bó, A.P.L., Poignet, P.: Beating heart motion prediction for robust visual tracking. In: Proceedings of IEEE Conference on Robotics and Automation (ICRA 2010), Anchorage, USA, pp. 4579–4584 (2010)

Real-Time Stereo Reconstruction in Robotically Assisted Minimally Invasive Surgery

Danail Stoyanov, Marco Visentini Scarzanella, Philip Pratt,
and Guang-Zhong Yang

Institute of Biomedical Engineering,
Imperial College London, London SW7 2AZ, UK
{danail.stoyanov,marco.visentini-scarzanella02,
p.pratt,g.z.yang}@imperial.ac.uk
<http://vip.doc.ic.ac.uk>

Abstract. The recovery of 3D tissue structure and morphology during robotic assisted surgery is an important step towards accurate deployment of surgical guidance and control techniques in minimally invasive therapies. In this article, we present a novel stereo reconstruction algorithm that propagates disparity information around a set of candidate feature matches. This has the advantage of avoiding problems with specular highlights, occlusions from instruments and view dependent illumination bias. Furthermore, the algorithm can be used with any feature matching strategy allowing the propagation of depth in very disparate views. Validation is provided for a phantom model with known geometry and this data is available online in order to establish a structured validation scheme in the field. The practical value of the proposed method is further demonstrated by reconstructions on various *in vivo* images of robotic assisted procedures, which are also available to the community.

1 Introduction

In robotically assisted Minimally Invasive Surgery (MIS), recovering the underlying 3D structure of the operating field *in vivo* is important for registering pre-operative data to the surgical field-of-view for providing dynamic active constraints and motion compensation [1]. Tomographic intra-operative imaging modalities can potentially provide anatomically co-registered information about the 3D shape and morphology of the soft tissues but their deployment in operating theatres is a significant challenge [2]. Currently, the most practical method of recovering the 3D structure of the operating site *in situ* is through optical techniques using a stereo laparoscope. This information can subsequently be used to align multimodal information within a global reference 3D coordinate system and enhance robotic instrument control. However, the recovery of 3D geometry from stereo in real-time during robotic procedures is difficult due to tissue deformation, partial occlusion due to instrument movement, and specular inter-reflections.

The recovery of 3D information from stereo images is one of the classic problems in computer vision. Given a calibrated stereo rig, the task is to identify the unique correspondence of image primitives across the stereo image pair. Recent review articles

provide a good summary of progress in the field [3, 4] and establish a benchmarking framework with ground truth data [3]. However, these methods, as well as the data used, are not always well suited to MIS settings where the scene is complicated by the large disparity discontinuities and occlusions arising from surgical instruments. The presence of view-dependent reflectance characteristics and a lack of fronto-parallel surfaces with unique colors and texture further complicate the issue. Thus far, for robotically assisted MIS, a number of stereoscopic techniques for recovering 3D shape and morphology have been proposed [5-9]. Many methods assume a geometric surface model of the tissue and use this constraint to track tissue 3D structure and morphology with a particular focus on cardiac procedures [7, 8, 10]. Such techniques have demonstrated the feasibility and potential of stereo vision in MIS but they are constrained by smooth surface parameterizations without adequate handling of instrument occlusions. Furthermore, there has been no extensive validation and comparative assessment of the existing methods.

In this study, we propose a technique for building a semi-dense reconstruction of the operating field in MIS that can operate in real-time. The method initially starts with a sparse 3D reconstruction based on feature matching across the stereo pair and subsequently propagates structure into neighboring image regions. The method is validated by experiments on phantom data with known ground truth. Qualitative validation is also provided for *in vivo* robotic MIS images. All the data used in this study is available for access online (<http://ubimon.doc.ic.ac.uk/dvs/m857.html>) to facilitate the community to establish a structured validation framework for stereo vision in robotic surgery.

2 Method

2.1 Feature Matching

The first step of the proposed method is to recover a sparse set of matches across the stereo-laparoscopic image pair using a feature based technique. The most commonly used image features are based on points where the image intensity gradient is high in both the vertical and horizontal directions [11]. Such salient points can be reliably matched across the stereo pair to recover a sparse set of 3D points in the surgical site [12]. The method works effectively with stereo-laparoscopic images because the vergence of the cameras creates a zero disparity region that assists convergence. Furthermore, the technique can be adapted to compensate for linear illumination changes and to incorporate the expected disparity as a starting solution and prior information about the 3D surface. Efficient implementations of this method have been reported to operate at very high frame-rates on GPU accelerated hardware [13].

It is important to note that for the proposed stereo propagation method, any feature matching strategy can be deployed. As the technique can be used to match images temporally and recover structure with a moving imaging device more complex and robust feature detectors and descriptors can be used to accommodate wider variations in perspective distortion and tissue deformation [14].

2.2 Structure Propagation

With a sparse set of 3D points established in the surgical field-of-view, it is possible to propagate 3D information to cover a semi-dense portion of operating field domain. This is necessary since the sparse 3D structure is usually not sufficient to describe the tissue geometry in detail. During the first stage of the proposed propagation algorithm, all features correspondences are used as seed matches. They are sorted subject to the correlation score between their respective templates and stored using a priority queue structure where popping and element retrieves and removes the queue element with highest priority. The algorithm then proceeds to propagate structure around the matches with highest correlation scores on a best-first basis by popping the priority queue as proposed by Lhuillier *et al* [15]. New matches are simply added to the queue as the algorithm iterates until no more matches can be retrieved.

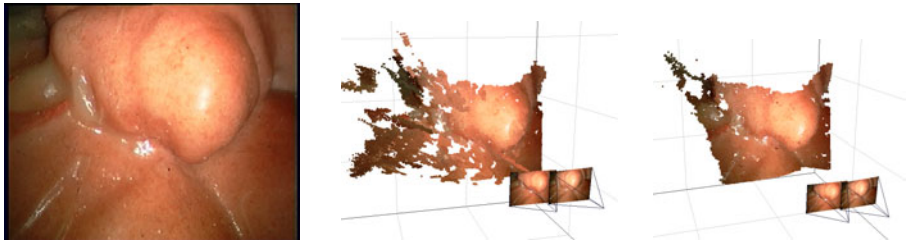


Fig. 1. Example image from a stereo-laparoscope and the corresponding stereo reconstructions using the proposed method with a sum of squared difference metric and with the ZNCC, which clearly performs more effectively

The spatial neighborhood of a seed match $(\mathbf{x}, \mathbf{x}')$, where \mathbf{x} and \mathbf{x}' denote image positions in the left and right images respectively, is defined as $N(\mathbf{x}, \mathbf{x}')$ and it is used to enforce a 2D disparity gradient limit as a smoothness constraint. Rather than the 1D regions typically enforced by epipolar geometry rectification. For each image, the spatial neighborhoods around a respective seed match are defined by $N(\mathbf{x}) = \{\mathbf{u}, \mathbf{u} - \mathbf{x} \in [-N, N]^2\}$ and $N(\mathbf{x}') = \{\mathbf{u}', \mathbf{u}' - \mathbf{x}' \in [-N, N]^2\}$ where $(\mathbf{u}, \mathbf{u}')$ denotes a candidate pair of pixels. The scheme means that the points considered for propagation are within a spatial window of $(2N + 1) \times (2N + 1)$ pixels centered at the respective seed locations. The full match propagation neighborhood is then denoted as:

$$N(\mathbf{x}, \mathbf{x}') = \left\{ (\mathbf{u}, \mathbf{u}'), \mathbf{u} \in N(\mathbf{x}), \mathbf{u}' \in N(\mathbf{x}'), \left\| (\mathbf{u} - \mathbf{u}') - (\mathbf{x} - \mathbf{x}') \right\| \leq \gamma \right\} \quad (1)$$

This defines all the possible candidate matches around the seed correspondence $(\mathbf{x}, \mathbf{x}')$ and the strategy can be easily adapted to the 1D search space of rectified images [15]. The term γ is used to control the smoothness of the disparity map and we determine this adaptively depending on the color similarity and proximity between the seed and candidate pixels. Such a scheme was shown to be effective by Yoon *et al*

[16] for locally adapting the weight of support windows in stereo aggregation. To avoid the heavy computational load of the CIE Lab color space the method was adapted to use the RGB space [17]. In this study, we only propagate information into regions of similar color weighed by the proximity of the spatial neighborhood. By defining $\gamma = \beta^{-1} \Sigma |I_i(\mathbf{x}) - I_i(\mathbf{u})| + \lambda^{-1} \|\mathbf{x} - \mathbf{u}\|$ and stopping the propagation if the value of γ we enforce a consistency and crude segmentation to the propagation process. The values of β and λ control the weight given to color similarity and proximity respectively and were kept constant this study. The dissimilarity measure used during propagation is the zero mean normalized cross correlation (ZNCC), which is less prone to illumination bias in homogeneous regions while it is also more indicative in regions with discriminative texture. This observation is illustrated in Fig 1 where the proposed algorithm was applied using the traditional sum of squared differences metric and the ZNCC.

2.3 Parallelization of Propagation

To improve the computational performance of the method, it is possible to exploit modern GPGPU technology to calculate multiple correlation windows and propagate structure over multiple pixels. The simplest parallelization strategy is to execute the correlation searches during propagation in parallel. This can be implemented to exploit the large number of concurrent threads that run on modern graphics hardware. In order to maximize the throughput on the graphics hardware, we implemented in this study the propagation of each pixel as a kernel in the NVIDIA language CUDA, including left-right consistency checking. Integral images were used to keep the running time invariant to the correlation window size as for example shown in the stereo algorithm by Veskler [18].

3 Experiments and Results

The proposed method was implemented using C++ and the NVIDIA® CUDA language for GPGPU performance enhancement. The CPU implementation of our propagation technique was able to operate at ~10fps for images of resolution 360 x 288 on a Hewlett-Packard P® xw4600 desktop workstation with an Intel® Pentium® 2.5 GHz Dual-Core Processor and 4Gb of RAM. On the same workstation our CUDA implementation using an NVIDIA Quadro® FX 5800 card was able to operate at 15fps. We believe this can be significantly improved given optimization of our CUDA code to maximize the use of the GPU cores.

For validation a phantom model of the heart (Chamberlain Group, MA, USA) was embedded with high contrast CT markers and scanned using a Siemens SOMATOM CT scanner. Registration between the camera and CT coordinate systems was performed using the contrast fiducials in the CT coordinate frame and their observed 3D reconstructions from stereo images captured using the daVinci® surgical system. The absolute orientation algorithm by Horn [19] followed by non-linear refinement was used to compute the aligning transformation.

3.1 Phantom Experiment

To validate the stereo reconstruction accuracy of the proposed method we recorded two datasets using the previously described experimental setup. Ground truth information was obtained using the CT data to generate disparity maps for each laparoscopic image as shown in Fig 2. While this method for generating the ground truth has embedded registration error associated with it, it is representative of real clinical scenario.

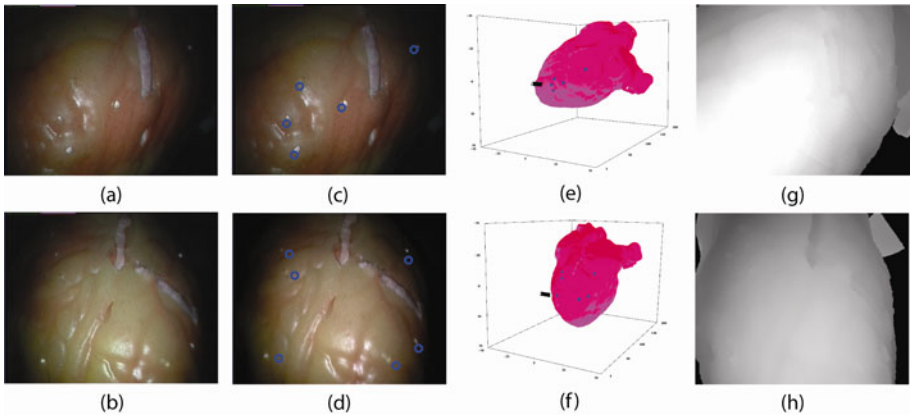


Fig. 2. (a-b) Images of the phantom heart model used in this study taken with the daVinci surgical system; (c-d) fiducial points located in the image space; (e-f) 3D reconstruction of the heart model from CT data registered in the calibrated stereo-laparoscope coordinate system; (g-h) ground truth disparity maps for the images generated by projecting the CT model into the stereo laparoscope images

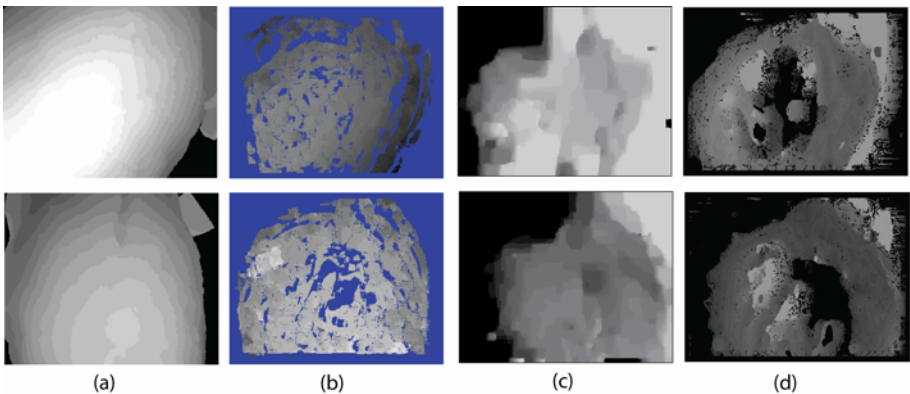


Fig. 3. (a) Ground truth disparity images, the same as shown in Fig 2 but discretized to integer disparity levels; (b) disparity map generated with the proposed technique; (c) disparity map generated with a global belief propagation (BP) algorithm [20]; (d) disparity map generated with a real-time technique [22]

Table 1. Summary of the disparity reconstruction error for the phantom model datasets used in this study. Different stereo techniques are compared with the approach proposed in this work and the mean disparity error and standard deviations are reported.

Method	Heart 1 Disparity Error	Heart 2 Disparity Error
Proposed	0.89 [± 1.13]	1.22 [± 1.71]
BP [20]	9.95 [± 5.22]	9.59 [± 2.77]
RT [22]	12.58 [± 4.57]	9.32 [± 2.80]
CUDA [21]	10.12 [± 4.21]	9.92 [± 3.43]

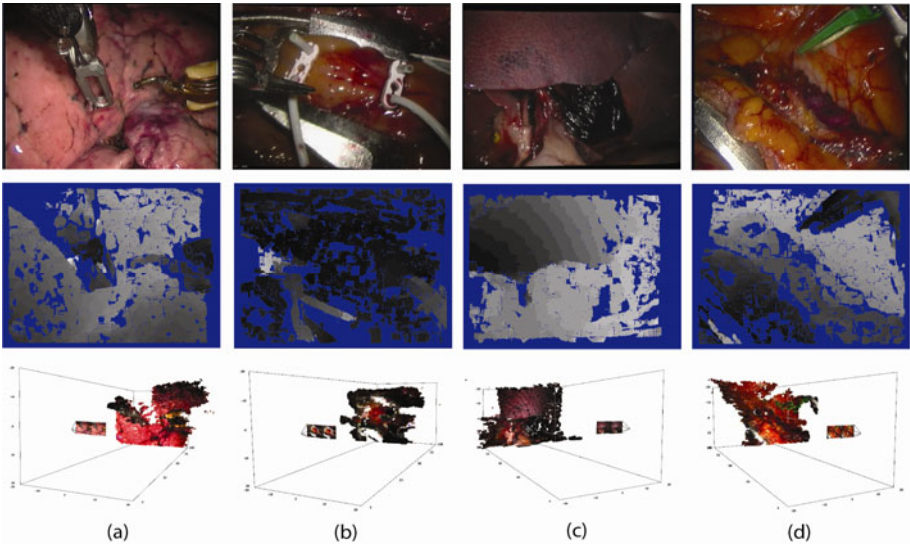


Fig. 4. (top row) Examples of *in vivo* robotic assisted MIS images taken with the daVinci® surgical system; (middle row) the corresponding disparity maps for each image computed with the method proposed in this study, where light colors indicate further away from the camera; (bottom row) 3D renditions of the corresponding reconstruction of the surgical field of view in the camera coordinate space.

The proposed technique was used to derive the 3D structure and disparity map of the scene and this was measured against the known ground truth information. The results for this experiment are shown in Fig 3 where the performance of our technique is compared to several dense computational stereo algorithms. The selection of comparison algorithms was based on the availability of the source code for the techniques and their suitability for efficient real-time implementation. It is clear that the disparity map generated by our approach yields a more consistent result than the other methods. A quantitative summary of the results is provided in Table 1 where it is evident that the proposed method outperforms the other approaches measured. It is important to note that the proposed technique does not recover as dense a result as the other

approaches. Thereby perhaps by discarding matches of low reliability their results could be improved to closer match the performance of our technique.

3.2 Qualitative *In Vivo* Data Experiments

To qualitatively evaluate the performance of the proposed method on *in vivo* images, we have used several datasets taken from different surgical procedures using the daVinci® surgical system. The results for the disparity map and the corresponding 3D renditions of reconstructions are shown in Fig 4. It is evident that the proposed technique effectively captures the 3D geometry of the surgical site. It copes well with the presence of large disparity discontinuities due to the surgical instruments and with large specular reflections. However, there are errors, particularly at occlusion boundaries, which need to be addresses further. It is important to note that we do not explicitly cater for specular reflections or model occlusion, for example by using detection, and we do not employ any surgical instrument tracking. By incorporating such strategies in our method we believe that results can be improved significantly.

4 Discussion

In this article, we have presented a real-time stereo reconstruction framework for robotic assisted MIS. The proposed technique relies on propagating a sparse set stereo correspondences into a semi-dense 3D structure by using a best-first principle growing scheme. By incorporating constraints into the propagation framework to consider uniqueness, consistency and disparity smoothness the algorithm produces robust, semi-dense 3D reconstructions of the operating field. We have validated the effectiveness of the approach using phantom data with known ground truth. This data is available online (<http://ubimon.doc.ic.ac.uk/dvs/m857.html>) together with an executable of the approach to help the development and benchmarking of future work in the field. In our future work, we hope to extend the validation database to more complex phantom models with known ground truth and to include *in vivo* images with manually labeled disparity information.

References

- [1] Stoyanov, D., Lerotic, M., Mylonas, G., Chun, A.J., Yang, G.-Z.: Intra-operative Visualizations: Perceptual Fidelity and Human Factors. *J. Display Technol.* 4, 491–501 (2008)
- [2] Taylor, R.H., Stoianovici, D.: Medical Robotics in Computer-Integrated Surgery. *IEEE Trans. Robot. Autom.* 19, 765–781 (2003)
- [3] Scharstein, D., Szeliski, R.: A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *Int. J. Comput. Vision* 47, 7–42 (2002)
- [4] Brown, M.Z., Burschka, D., Hager, G.D.: Advances in Computational Stereo. *IEEE Trans. Pattern Anal. Mach. Intell.* 25, 993–1008 (2003)
- [5] Devernay, F., Mourgues, F., Coste-Maniere, E.: Towards endoscopic augmented reality for robotically assisted minimally invasive cardiac surgery. In: *Medical Imaging and Augmented Reality* (2001)

- [6] Mourgues, F., Devernay, F., Malandain, G., Coste-Manière, È.: 3D reconstruction of the operating field for image overlay in 3D-endoscopic surgery. In: International Symposium on Augmented Reality (2001)
- [7] Lau, W.W., Ramey, N.A., Corso, J., Thakor, N.V., Hager, G.D.: Stereo-Based Endoscopic Tracking of Cardiac Surface Deformation. In: International Conference on Medical Image Computing and Computer Assisted Intervention, pp. 494–501 (2004)
- [8] Richa, R., Poignet, P., Liu, C.: Efficient 3D Tracking for Motion Compensation in Beating Heart Surgery. In: International Conference on Medical Image Computing and Computer Assisted Intervention, vol. II, pp. 684–691 (2008)
- [9] Hager, G., Vagvolgyi, B., Yuh, D.: Stereoscopic Video Overlay with Deformable Registration. In: Medicine Meets Virtual Reality (2007)
- [10] Stoyanov, D., Darzi, A., Yang, G.-Z.: A Practical Approach Towards Accurate Dense 3D Depth Recovery for Robotic Laparoscopic Surgery. *Comput. Aided Surg.* 10, 199–208 (2005)
- [11] Shi, J., Tomasi, C.: Good features to track. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 593–600 (1994)
- [12] Stoyanov, D., Mylonas, G.P., Deligianni, F., Darzi, A., Yang, G.-Z.: Soft-tissue Motion Tracking and Structure Estimation for Robotic Assisted MIS Procedures. In: International Conference on Medical Image Computing and Computer Assisted Intervention, pp. 139–146 (2005)
- [13] Sinha, S.N., Frahm, J.-M., Pollefeys, M., Genc, Y.: Feature Tracking and Matching in Video Using Programmable Graphics Hardware. *Mach. Vision and Appl.* (2007)
- [14] Mountney, P., Lo, B.P.L., Thienjarus, S., Stoyanov, D., Yang, G.-Z.: A Probabilistic Framework for Tracking Deformable Soft Tissue in Minimally Invasive Surgery. In: International Conference on Medical Image Computing and Computer Assisted Intervention, pp. 34–41 (2007)
- [15] Lhuillier, M., Quan, L.: Robust dense matching using local and global geometric constraints. In: International Conference on Pattern Recognition, vol. 1, pp. 968–972 (2000)
- [16] Yoon, K.-J., Kweon, I.S.: Adaptive Support-Weight Approach for Correspondence Search. *IEEE Trans. Pattern Anal. Mach. Intell.* 28, 650–656 (2006)
- [17] Yang, Q., Wang, L., Yang, R., Stewenius, H., Nister, D.: Stereo Matching with Color-weighted Correlation, Hierarchical Belief Propagation and Occlusion Handling. *IEEE Trans. Pattern Anal. Mach. Intell.* 31, 492–504 (2009)
- [18] Veksler, O.: Fast Variable Window for Stereo Correspondence using Integral Images. In: International Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 556–564 (2003)
- [19] Horn, B.K.P.: Closed-form solution of absolute orientation using unit quaternions. *J. Optic. Soc. America* 4, 629–642 (1987)
- [20] Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient Belief Propagation for Early Vision. *Int. J. Comput. Vision* 70 (2006)
- [21] Fung, J., Mann, S., Aimonè, C.: OpenVIDIA: Parallel GPU Computer Vision. In: ACM Multimedia, pp. 849–852 (2005)
- [22] <http://people.csail.mit.edu/demirdji/download/index.html>

Prostate Brachytherapy Seed Reconstruction Using C-Arm Rotation Measurement and Motion Compensation*

Ehsan Dehghan¹, Junghoon Lee², Mehdi Moradi³, Xu Wen³,
Gabor Fichtinger¹, and Septimiu E. Salcudean³

¹ Queen's University, Canada

² Johns Hopkins University, USA

³ University of British Columbia, Canada

Abstract. During prostate brachytherapy, C-arm fluoroscopy images are used for a qualitative assessment of the procedure. Three dimensional reconstruction of the implanted seeds can be used for intraoperative dosimetry and quantitative assessment. Accurate C-arm pose estimation is necessary for 3D localization of the seeds. We propose to measure the C-arm rotation angles and computationally compensate the inevitable C-arm translational motion to estimate the pose. We compensate the oscillation, sagging and wheel motion of the C-arm using a three-level optimization algorithm, without which the reconstruction can fail. We validated our approach on simulated and 10 data sets from 5 patients and gained on average 99.1% success rate, 0.33 mm projection error and computation time of less than one minute per patient, which are clinically excellent results.

1 Introduction

Low dose rate prostate brachytherapy is an effective treatment for localized prostate cancer which entails permanent placement of radio-active capsules or “seeds” inside the prostate. The seeds are delivered to preoperatively determined positions using needles under real-time visual guidance from ultrasound. The quality of the treatment depends on the accurate placement of the seeds to deliver sufficient radiation to the cancer while sparing the healthy tissue. However, seed misplacements and consequent complications are common. C-arm fluoroscopy images are taken during the procedure to visually assess the implant (see Fig. 1(a)). Three dimensional reconstruction of the seeds has several benefits such as intraoperative dosimetry modifications and quantitative assessment, and has been proposed previously [1,2,3].

Seed reconstruction entails solving a seed matching problem – assigning the shadows of a seed in different images to each other- which has been solved using simulated annealing [1], Hungarian algorithm [2] and linear programming [3].

* This work was supported by an Ontario Ministry of Research and Innovation post-doctoral fellowship and NIH R21CA120232-01. The authors are grateful to Dr. W. J. Morris from British Columbia Cancer Agency for providing the clinical data.

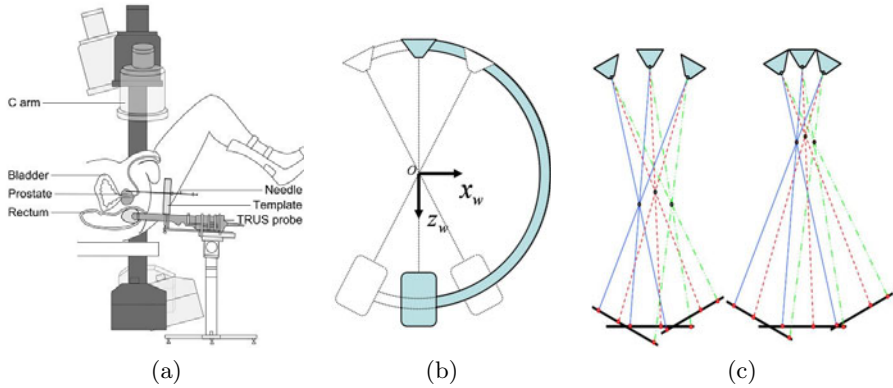


Fig. 1. (a) Brachytherapy procedure, (b) rotation of the C-arm around the primary axis (PA) and the world coordinate system, and (c) the scaling problem. In (c), the left C-arms have the same intrinsic parameters, rotation angles, X-ray images, reconstruction cost and matching as the right ones; however, the seed cloud is scaled.

Seed matching and reconstruction are performed using known C-arm pose, estimated using radio-opaque beads or fiducials [4,5,6], or obtained from optical and electromagnetic trackers [7]. The fiducials and beads may interfere with the anatomy in the images, require segmentation and can limit the clinical working volume. Auxiliary trackers are expensive and further complicate the surgery as they need calibration steps, line of sight and space in the operating room. It has been suggested that implanted seeds can be used for pose estimation [1,8]. However, to be computationally feasible, they need an initial estimation of the pose, conveniently by using external trackers and fiducials.

If C-arm images are taken by rotation of the C-arm around a fixed axis (see Fig. 1(b)), known rotation angles will yield the relative pose. C-arm devices are available with or can be easily equipped with a joint encoder. A simple digital protractor can also be used to measure the C-arm angles. However, oscillation and sagging of the C-arm, especially in the elevational direction are significant due to the C-arm weight. The movements in the two other perpendicular directions caused by wheel motion are much smaller but are often significant. If unattended, C-arm translational movements cause inaccuracies in the pose estimation which may lead to reconstruction failure.

We demonstrate that in the case of a C-arm with small angle span around a single axis, sole measurement of rotation angles combined with a computational algorithm to compensate for C-arm oscillation, sagging and wheel motion suffices for a successful and accurate reconstruction. However, the reconstruction is prone to failure without such a motion compensation. By measuring the angles using joint encoders or protractors, we obviate the need for full pose tracking using fiducials or external trackers. Considering the simplicity of the implementation, high speed and accuracy of reconstruction, this approach is especially suitable for clinical translation.

2 Methods

In order to solve the matching problem, we rely on a linear programming approach introduced by Lee *et al.* [3] which, for sake of completeness, we outline in Section 2.1. The motion compensation method is discussed in detail in Section 2.2.

2.1 Seed Localization Using Linear Programing

Assume that three projection images of N implanted seeds are available and $N_i, i \in \{1, 2, 3\}$ seed shadows are segmented in each image. The matching problem can then be written as:

$$\begin{aligned} & \min_{x_{ijk}} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \sum_{k=1}^{N_3} c_{ijk} x_{ijk}, \\ & \text{s.t.} \begin{cases} \sum_{j=1}^{N_2} \sum_{k=1}^{N_3} x_{ijk} \geq 1, & \forall i \\ \sum_{i=1}^{N_1} \sum_{k=1}^{N_3} x_{ijk} \geq 1, & \forall j \\ \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} x_{ijk} \geq 1, & \forall k \\ \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \sum_{k=1}^{N_3} x_{ijk} = N, \\ x_{ijk} \in \{0, 1\}, \end{cases} \end{aligned} \tag{1}$$

where c_{ijk} is the cost of matching seed shadows p_i^1, p_j^2 and p_k^3 from the first, second and third images, respectively, and x_{ijk} is a binary variable showing the correctness of such an assignment. The cost c_{ijk} is the symbolic distance between lines L_i^1, L_j^2 and L_k^3 which connect the seed shadows p_i^1, p_j^2 and p_k^3 to their corresponding X-ray source locations in 3D space. Equation (1) can be written as a constrained linear programing problem by concatenating the x_{ijk} and c_{ijk} into vectors. The unknowns that correspond to a cost higher than a specified threshold are removed and the linear programing problem with reduced dimensions is solved to find the correct seed matching solution [3].

2.2 Motion Compensation Algorithm

It has been shown that higher accuracy in C-arm pose estimation will result in more accurate matching and vice versa [11,8]. Therefore, the matching and the motion compensation problems can be solved iteratively in a loop in which reconstructed seed positions can be used to improve on C-arm pose estimation.

Assume a world coordinate system $Ox_w y_w z_w$ centered at the center of rotation of the C-arm, with unit vectors x_w and z_w as shown in Fig. 1(b) and y_w perpendicular to the plane of rotation and aligned with the craniocaudal axis of the patient. The coordinates of the X-ray source corresponding to image i are q_i in $Ox_w y_w z_w$. Measured rotation angles and known intrinsic parameters, such as focal length and source to center distance, yield good initial estimates for q_i and the corresponding pose.

The goal of our motion compensation algorithm is to find the optimal position adjustments (offsets) δ_n for the source positions to solve the following problem:

$$\min_{x_{ijk}, \delta_n} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \sum_{k=1}^{N_3} c_{ijk}(\delta_n) x_{ijk}, \quad n \in \{1, 2, 3\}, \tag{2}$$

subject to the constraints from Eq. (1).

In order to compensate the motion of the C-arm, one may try to find the positions of the source corresponding to the second and third images (henceforth called second and third source positions) relative to the position of the source corresponding to the first image (henceforth, the first source position). It is known that this problem can be solved in 3D up to a scale [18], which means that the reconstructed volume can arbitrary shrink or expand (see Fig 1(c)). In [8], the distance between two points on a radio-opaque fiducial was used to find the scaling factor. However, we do not use an external fiducial here.

Through observation, it can be seen that C-arm motion along x_w is less significant compared to the motion along y_w and especially z_w , along which we expect the largest motion. In order to avoid the scaling problem, we will constrain the C-arm motion to the Oy_wz_w plane.

First-Level Optimization: In the first level of optimization, we find matching seeds in the images to provide an initial offset estimate. Since the fluoroscopy images are generally taken by rotation of the C-arm around the y_w axis only, the seeds located at the top or at the bottom of one image appear at the top or at the bottom of other images. We use this observation and select $n = 5$ seed shadows from the top and n seeds from the bottom of the images and solve the matching problem for them. Since these n seed shadows do not necessarily correspond to n seeds in 3D, the matching problem is solved as in Eq. (1), while the fourth constraint is relaxed. We fix the first source in 3D space and optimize the 2D offsets (to avoid scaling) for the other two sources to minimize the overall reconstruction cost for $p = 4$ seeds (2 from the top and 2 from the bottom of the image) which have the least reconstruction cost. The matching and the 2D offset calculation are iteratively solved until there is no change in the matching.

Second-Level Optimization: At first, seed matching is solved with the given 2D offset parameters from the first level and the 3D positions of the seeds are calculated. Then, the positions of the three sources are adjusted in 3D space to minimize the reconstruction error for the given 3D seeds. The matching and source position adjustments are iteratively performed until either the reconstruction cost or the change in the reconstruction cost between two iterations is less than an empirically assigned threshold (<0.1 mm for the former and $<0.1\%$ for the latter). The 3D position of source i at iteration $(k + 1)$ is calculated as:

$$q_i^{k+1} = \left(\sum_{j=1}^N (I - v_{ij}v'_{ij}) \right)^{-1} \sum_{j=1}^N (I - v_{ij}v'_{ij})s_j^k, \quad (3)$$

where, s_j^k is the position of the j^{th} seed at iteration k , v_{ij} is the unit vector from the source q_i^k to the shadow of seed j on image i , and I is a 3×3 identity matrix.

Since in each iteration, the position of the seeds in 3D space is fixed, the shrinkage or expansion is very small. However, as the algorithm iterates the scaling factor can become significant. It should be noted that the matching problem has a solution independent of the scaling factor (see Fig 1(c)). Therefore,

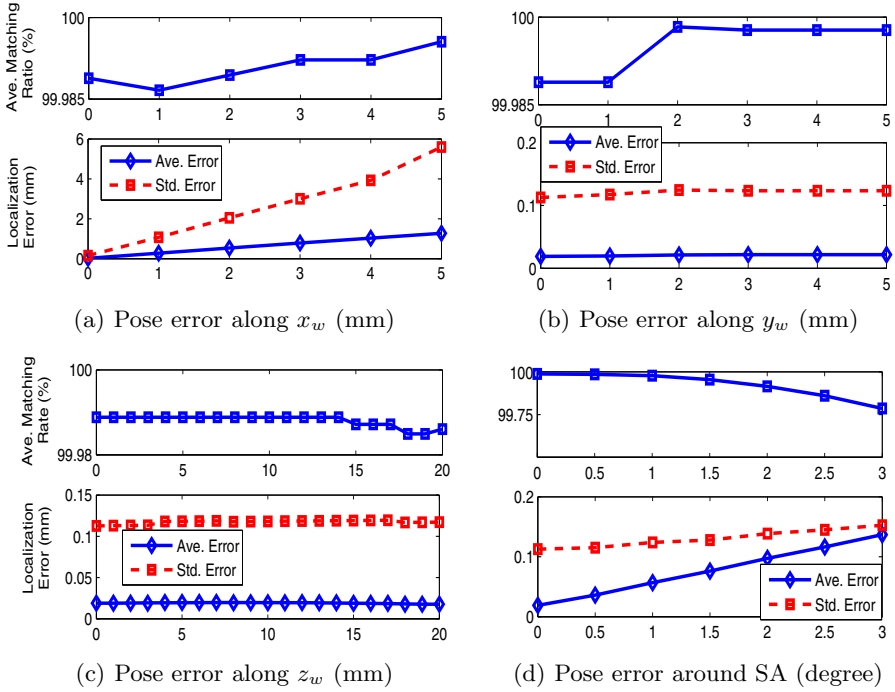


Fig. 2. Simulation results, showing the matching ratio and localization error for variable pose errors

we exploit the advantages of a 3D motion compensation in this level to increase the likelihood of finding the correct matching.

Third-Level Optimization: At this point in the workflow, the correct matching is available, but seed positions may be scaled as a result of 3D motion compensation. In the third level, once more, we assume that pose error along x_w is negligible. The source positions are initialized at their ideal positions, using the rotation angle readings and C-arm’s intrinsic parameters. Then, the 2D offset parameters are calculated, assuming that the first source is fixed. Note that we are only interested in the relative positions of the seeds, since the seed cloud will be registered to the prostate anatomy.

3 Results and Discussion

First, the motion compensation algorithm was tested on simulated data. We synthesized 3D positions of seeds, based on realistic dosimetry plans of four patients with 100, 108, 110 and 130 seeds. The seed shadows were simulated by rotating the C-arm around the primary axis (PA) at 0° , $\pm 5^\circ$ and $\pm 10^\circ$ while keeping

the secondary axis (SA) fixed at 180° . The seed locations were reconstructed for every combination of three images out of the available five for each patient, while translational and rotational errors were added to the third C-arm pose in each set. The added pose error was 0-5 mm along x_w and y_w , 0-20 mm along z_w and 0-3 degrees around SA. There were on average 1.6 hidden seeds per image with a maximum of 14. The reconstructed seeds were compared against the ground truth in term of localization errors, defined as the distance between the true and reconstructed seed positions after a rigid rotation and translation of the reconstructed seed cloud. The average matching ratio and localization error are shown in Fig. 2. As it can be seen the reconstruction algorithm shows consistently successful performance over a wide range of C-arm translational position errors, while the performance without motion compensation decreases with measurement errors [3]. It can be seen in Figs. 2(b) and 2(c) that the localization error and the matching ratio are almost constant for the errors added along z_w and y_w , since we compensated the motion along these two directions. The matching ratio for pose error along x_w is almost constant as shown in Fig. 2(a), since the second-level optimization finds the correct matching using a 3D offset optimization. However, since a 2D optimization is used in the third level, the localization error increases with the pose error along x_w . As expected, the localization and matching errors increase with the measurement error around SA. However, the matching ratio remains well above clinically acceptable threshold, which is customarily around 98%.

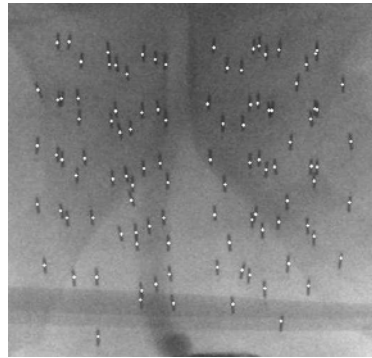
The motion compensation algorithm was also tested on patient data. Five patients were chosen with implanted ^{125}I seeds. Five images were taken for each patient at angles approximately 0° , $\pm 5^\circ$ and $\pm 10^\circ$ around y_w . Accurate rotation angles were measured using a digital protractor attached to the C-arm, which does not interfere with the image or the working space. The images were taken using a GE OEC[®] 9800 with motorized joints. This particular C-arm is a solid machine with a fixed axis of rotation. However, it has a heavy intensifier which makes sagging problem more severe. The rotation around the SA was set to be constant; however, a variation of 1° was observed according to the C-arm joint angle readings and was taken into account during seed reconstruction. We reconstructed the seeds using manually segmented images taken at $(0^\circ, \pm 5^\circ)$ and $(0^\circ, \pm 10^\circ)$ for each patient (10 data sets in total). There were an average of 1.45 and maximum of 6 hidden seeds in the images. We assumed that the C-arm intrinsic parameters do not change due to small span of rotation. After reconstruction the seeds were projected on the images. Projection error, measured as the distance of the centroid of the seed and the projected location of the seed is reported in Table 1. Figure 3 is a sample of projected seeds on an image, showing very small errors. The images were meticulously inspected to detect any mismatching of seeds. The matching success rate and the reconstruction time are reported in Table 1. In the second data set of the third patient, 8 seeds showed ambiguous assignments and were conservatively considered as mismatches. Therefore, the reported matching ratio is smaller or equal to the real matching ratio.

Table 1. Clinical results. “F” stands for reconstruction failure

Patient Number	Number of seeds	App. Angles (degree)	Match rate	Ave. Proj. Err. (mm)	Std. Proj. Err. (mm)	Comp. time(s)	Match rate w/o MC
1	105	(0, ± 5)	100%	0.32	0.20	45	<95%
		(0, ± 10)	100%	0.37	0.25	53	F
2	102	(0, ± 5)	100%	0.22	0.12	32	<97%
		(0, ± 10)	100%	0.31	0.18	39	F
3	100	(0, ± 5)	100%	0.31	0.19	43	F
		(0, ± 10)	92%	0.32	0.19	57	F
4	115	(0, ± 5)	99.1%	0.51	0.28	60	<90%
		(0, ± 10)	100%	0.33	0.22	28	F
5	113	(0, ± 5)	100%	0.23	0.18	48	97.3%
		(0, ± 10)	100%	0.36	0.23	56	F

The reconstruction ratio without motion compensation is also reported in Table 1. Reconstructions with a success rate smaller than 80% are considered as failure. As it can be seen, in 6 out of 10 data sets the reconstruction without motion compensation failed, and in the remaining 4 cases the matching ratio was considerably lower than with motion compensation (the projection error was significantly larger and is not reported here). Comparison of the reconstruction results with and without motion compensation proves the necessity of motion compensation when only the C-arm joint angles are measured.

The algorithm was implemented using MATLAB on a PC with an Intel 2.33 GHz Core2 Quad CPU and 3.25 GB of RAM. The first-level optimization is fast as we select in total 10 seeds from each image. The second level is the most time consuming part of the algorithm since the seed matching problem should be solved for several iterations. The seed matching problem can be solved in approximately 10s or less per iteration, depending on the number of seeds. The second-level optimization never took more than 10 iterations both in the simulations and patient study. The initial 2D offsets from the first level decrease the number of necessary iterations in the second level, that in turn decreases the overall computational time. The third level has a closed form solution and can be solved very fast. The motion compensation and seed reconstruction took less than one minute for each clinical case. Detailed clinical study with additional patient data is in progress.

**Fig. 3.** Projected seeds overlaid on a C-arm fluoroscopy image

4 Conclusion and Future Work

In conclusion, we introduced a motion compensation algorithm that combined with C-arm joint angle measurement can be used to estimate the C-arm pose for brachytherapy seed reconstruction. For joint angle measurement an off-the-shelf digital protractor was used. This removed the need for full pose tracking with fiducials or external trackers. The clinical study showed the feasibility of the method to be used in the operating room. With an average matching ratio above 99.1%, average projection error of less than 0.33 mm and a computation time less than one minute, the algorithm is suitable for clinical application.

The optimization on the third degree of freedom was ignored to avoid the scaling effect. Investigation on using the length of a ^{125}I seed to find the scaling factor and compensate the motion in 3D is part of the future work. The seeds were segmented manually. The effects of segmentation errors and application of automatic segmentation methods (e.g. [9]) will be inspected in the future.

References

1. Tubic, D., Zaccarin, A., Beaulieu, L., Pouliot, J.: Automated seed detection and three-dimensional reconstruction. II. Reconstruction of permanent prostate implants using simulated annealing. *Medical Physics* 28(11), 2272–2279 (2001)
2. Jain, A.K., Zhou, Y., Mustafa, T., Burdette, E.C., Chirikjian, G.S., Fichtinger, G.: Matching and reconstruction of brachytherapy seeds using the Hungarian algorithm (MARSHAL). *Medical Physics* 32, 3475–3492 (2005)
3. Lee, J., Labat, C., Jain, A.K., Song, D.Y., Burdette, E.C., Fichtinger, G., Prince, J.L.: Optimal matching for prostate brachytherapy seed localization with dimension reduction. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009*. LNCS, vol. 5761, pp. 59–66. Springer, Heidelberg (2009)
4. Navab, N., Bani-Hashemi, A., Mitschke, M., Holdsworth, D.W., Fahrig, R., Fox, A.J., Graumann, R.: Dynamic geometrical calibration for 3-D cerebral angiography. In: *SPIE Medical Imaging*, vol. 2708, pp. 361–370 (1996)
5. Brack, C., Gotte, H., Gosse, F., Moctezuma, J., Roth, M., Schweikard, A.: Towards accurate X-ray camera calibration in computer-assisted robotic surgery. In: *Proc. Int. symp. Computer Assisted Radiology*, pp. 721–728 (1996)
6. Jain, A.K., Mustafa, T., Zhou, Y., Burdette, C., Chirikjian, G.S., Fichtinger, G.: FTRAC – A robust fluoroscope tracking fiducial. *Medical Physics* 32(10), 3185–3198 (2005)
7. Peters, T., Cleary, K. (eds.): *Image-Guided Interventions: Technology and Applications*. Springer, Heidelberg (2008)
8. Jain, A., Fichtinger, G.: C-arm tracking and reconstruction without an external tracker. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) *MICCAI 2006*. LNCS, vol. 4190, pp. 494–502. Springer, Heidelberg (2006)
9. Lam, S.T., Marks II, R.J., Cho, P.S.: Prostate brachytherapy seed segmentation using spoke transform. In: *Proc. SPIE.*, vol. 4322, pp. 1490–1500 (2001)

Reconstructing Geometrically Consistent Tree Structures from Noisy Images

Engin Türetken^{1,*}, Christian Blum^{2,**}, Germán González¹, and Pascal Fua¹

¹ Computer Vision Lab., Ecole Polytechnique Fédérale de Lausanne, Switzerland

² ALBCOM, Universitat Politècnica de Catalunya, Barcelona, Spain

Abstract. We present a novel approach to fully automated reconstruction of tree structures in noisy 2D images. Unlike in earlier approaches, we explicitly handle crossovers and bifurcation points, and impose geometric constraints while optimizing a global cost function. We use manually annotated retinal scans to evaluate our method and demonstrate that it brings about a very substantial improvement.

1 Introduction

Tree-like structures, such as vascular networks, dendritic trees, or bronchial networks, are pervasive in biological imagery. With the advent of modern acquisition techniques that produce endless streams of 2D and 3D imagery, there has been renewed interest in automated delineation as a means of exploiting this data. Of particular interest are topologically accurate delineations, which are critical for diagnosis and analysis purposes. However, despite many years of sustained effort, automated delineation techniques remain fragile and error-prone.

In earlier work [1], we showed that robustness could be improved by exploiting the global tree topology early in the algorithm. However, this method suffers from the fact that the tree-growing algorithm it uses makes all its decisions based on local image evidence without regard to tree shape. As a result, it still makes topological mistakes, such as those highlighted by circles in Fig. 1.

In this paper, we address this issue by incorporating into the tree reconstruction algorithm shape priors that enforce geometric consistency between pairs of graph nodes and edges. This involves solving a *k-cardinality arborescence problem*, which is known to be NP-hard [2]. Nevertheless, we will show that good approximate solutions can be obtained by extending the metaheuristic approach of [3]. We also explicitly model crossovers and bifurcations, which are particularly troublesome in 2D projections of 3D volumes and can result in spurious branches, gaps, and other topological errors if not handled properly.

* This work was supported in part by the Swiss National Science Foundation and in part by the MicroNano ERC project.

** Christian Blum acknowledges support from the *Ramón y Cajal* program of the Spanish Government.

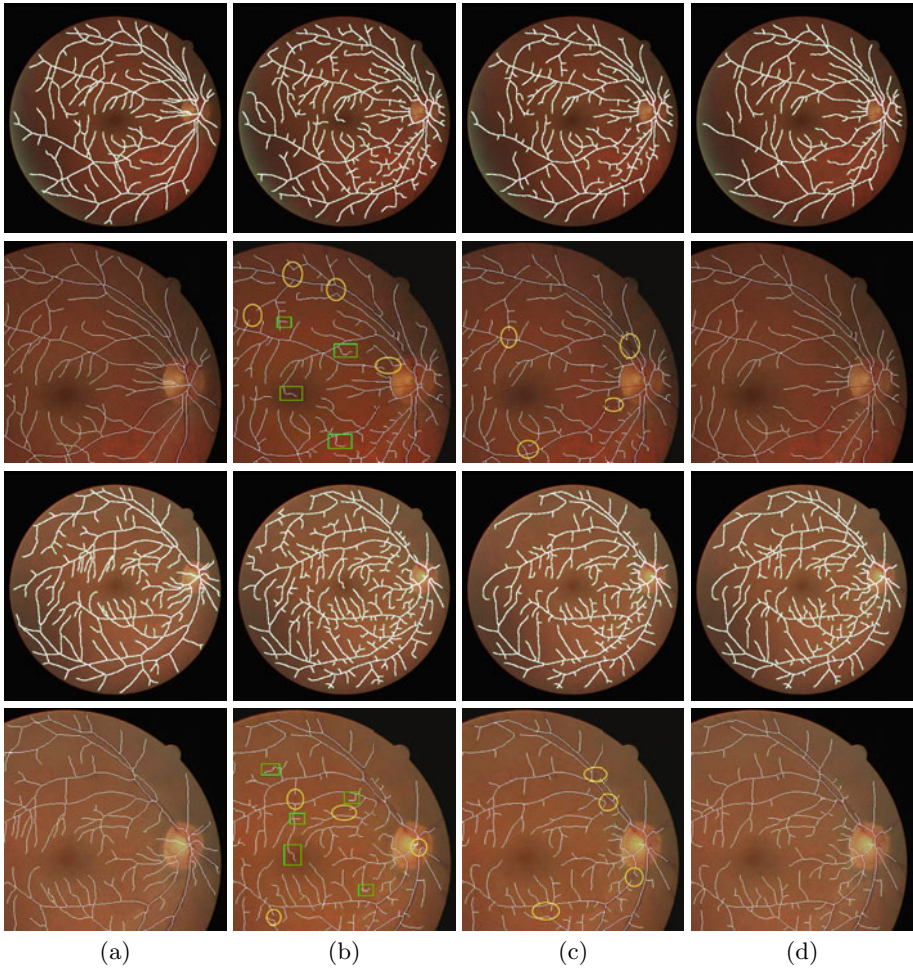


Fig. 1. The first and third rows depict vascular tree reconstructions from two different retinal scans. In the second and fourth rows, we show enlarged versions of the same images. (a) Original images with manually outlined blood vessels overlaid in white. (b) Trees reconstructed using a minimum spanning tree (MST) approach. (c) Trees reconstructed using our earlier method [1]. (d) Trees reconstructed using the proposed technique. Note the false positives and the false negatives, highlighted respectively by the green rectangles and the yellow circles in columns (b) and (c), have disappeared from column (d).

Introducing geometry constraints and handling crossovers results in a robust fully automated delineation technique that we demonstrate on vascular trees of retinal fundus images. As shown in Fig. 1, it yields a significant improvement, which is also supported by quantitative evaluation.

2 Related Work

Most automated delineation techniques start by computing a *tubularity* image in which pixels likely to belong to filaments have high scores. This image can then be thresholded and its skeleton computed [4]. This tends to produce disconnected components and artifacts on noisy data, which often require considerable post-processing and analysis for a correct tree to be produced.

Alternatively, the tubularity scores can be used to find seed points and recursively trace high-tubularity paths [5,6]. Although computationally efficient, this technique lacks robustness since cumulative tracing errors can result in large topological ones. More global methods avoid this problem by using more of the image evidence and optimizing a global objective function [7,8]. However, while such methods produce smooth tree components, they do not guarantee their spatial connectedness. Furthermore, they are computationally intensive, which limits their applicability to large datasets.

By contrast, methods that sample local maxima of the tubularity image and then connect these samples into a spanning tree [1,9], guarantee connectivity. However, they do not take into account global tree geometry, such as smoothness along the edges or branching factors, which can play an important role in improving topological accuracy, avoiding over-fitting, and speeding up convergence. They also fail to explicitly account for bifurcations and crossovers, which can easily lead to mistakes. While post-processing pruning [9] can sometimes eliminate some spurious branches, it does not allow for recovery from other topological mistakes. This is the problem we address in this paper by introducing more global geometry constraints early in the algorithm to prevent such mistakes.

3 Method

Our method consists of the following steps:

1. We compute a tubularity value at each pixel [10], which encodes how likely it is to be on the centerline of an elongated linear structure.
2. To avoid having to compute a tree that spans individual image pixels, which would be prohibitively expensive, we select high-probability pixels such as those of Fig. 2(b) that are as evenly spaced as possible and treat them as the graph nodes to be linked.
3. We compute the most probable paths between pairs of nearby nodes, such as those of Fig. 2(c), and treat them as the edges of our directed graph. We assign them probabilistic costs that are lowest when all pixels along them are likely to lie in the middle of a filament.
4. We compute the lowest-cost arborescence among those that span K edges for a wide range of $K < N$, where N is the number of nodes. This is known as the K -Cardinality Arborescence Problem. Even though it is NP-Hard, approximate solutions can nevertheless be computed efficiently and fast.
5. Among these arborescences, we select the one that optimizes a global objective function.

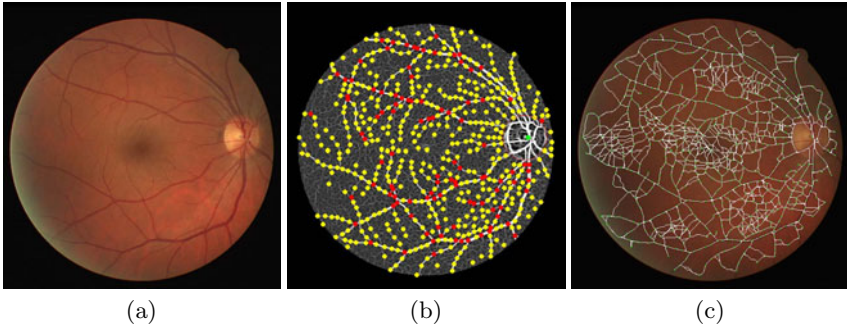


Fig. 2. Graph construction. (a) Original image. (b) Sampled points. The green point and the red circle represent the root node and optic disk region respectively. The red points are potential crossovers and the yellow ones are ridge points. (c) Graph built from the most probable paths.

The above workflow is the same as the one we introduced in our earlier work [1] with two key differences: First, in [1], we did not take into account the geometric properties of the trees when constructing them. Here, we incorporate into our objective function geometric terms that favor trees whose branches are smooth and along which the width remains consistent. Second, we explicitly model bifurcations and crossovers.

To solve the associated minimum arborescence problem, we had to substantially modify the tree building procedure [3] we used in our previous work. This produced very significant performance improvements, as shown in Fig. 1 and further discussed in Section 4.

3.1 Sampling

We first compute the response of the Rotational Features introduced in [10] at different scales and orientations at each image location x_i , and retain the maximum value as our tubularity score $f(x_i)$. We then map this score to a posterior probability that x_i belongs to a filament centerline, $P(x_i|f(x_i))$, by fitting a sigmoid function. We keep the orientation ϕ_i and the width estimates w_i of each point.

We then use a two-step approach to sampling local maxima of this probability image. First, we threshold and skeletonize it, and use the combined cross-point number method [11] to detect potential crossovers and bifurcations, which we will refer to as *landmarks* such as those depicted by the red dots in Fig. 2(b). For each landmark, we create two colocated nodes to prevent problems at tree reconstruction time, such as those depicted by Fig. 3(a). The sampling of bifurcations such as the one in Fig. 3(b) is useful to avoid overcounting pixels when scoring the trees that we reconstruct. In the second step, we sort the remaining

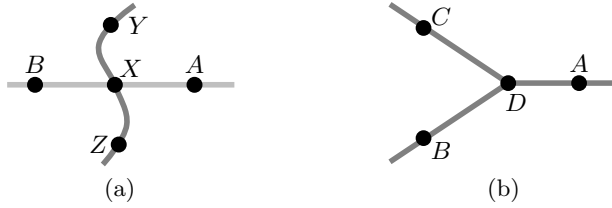


Fig. 3. Two cases that can lead to reconstruction errors. (a) A crossover with a single sample point that can be used to build either the horizontal \overline{AXB} branch or the vertical \overline{YXZ} one, but not both. (b) If there is no sample point at the bifurcation D , incorporating both the \overline{AB} and \overline{AC} paths into the graph will result in counting twice the pixels in \overline{AD} . This is avoided by introducing a sample point at D .

pixels according to their probability of belonging to a filament, select the most probable one, eliminate all those within a certain radius, and iterate. This produces the regularly spaced samples shown as yellow dots in Fig. 2(b).

In the specific case of retinal scans, we know *a priori* that the root of the tree we want to build is located in the optic disk, which is depicted by the red outline in Fig. 2(b). We use a variant of [12] to detect the optic disk and remove all samples that reside within it except the one nearest to its center, which will be assigned as the tree root.

3.2 Building the Graph

The procedure described above returns a set V of nodes. We construct a directed graph $G = (V, E)$ such as the one of Fig. 2(c) by linking all pairs of samples v_m and $v_n \in V$ that are within a certain distance of each other—except colocated ones—by a Dijkstra path $e_{mn} \in E$ that minimizes the integral of the negative log of the posterior probabilities $P(x_i|f(x_i))$ introduced in Section 3.1.

In [1], we showed that such paths are maximal probability paths between the vertices and that their total costs c_{mn}^{nl} can be treated as sum of negative log likelihoods along them. We also showed that if we assume that all geometric arrangements of edges are equally probable, a near-optimal tree can be obtained by minimizing

$$F^{nl}(T(k)) = \sum_{e_{mn} \in T(k)} c_{mn}^{nl} \quad , \quad (1)$$

where $T(k)$ denotes a tree of cardinality k . However, in reality, not all trees are equally plausible. Those whose branches are smooth, conform to the underlying image orientation, and whose widths vary slowly and consistently, are much more likely to be correct than others.

To account for this, we exploit four geometric terms to capture the underlying relations between parts of the tree structure:

1. **Edge Direction Similarity** (Φ_e). We model the angular difference between pairs of adjacent directed edges by a von Mises distribution (circular normal distribution).
2. **Width Consistency** (Φ_w). We model the width differences of pair of vertices connected through directed edges by an asymmetric Gaussian distribution, since for most datasets a decrease in width is more probable than an increase along a directed path from the root vertex.
3. **Orientation Consistency** (Φ_o). For pairs of connected vertices, we measure the angular deviations of their orientations from the direction of the line between them. The deviations are modeled using a von Mises distribution.
4. **Tortuosity** (Φ_t). We compute tortuosity values of the paths corresponding to edges and represent them by a Gaussian distribution. For the vascular reconstructions obtained in this paper, the tortuosity measure that we used is the ratio of the path length to the linear distance between the endpoints.

We estimate the parameters of these distributions using maximum likelihood estimation (MLE). Given the individual terms, the combined pairwise potential can be written as

$$\begin{aligned} \Phi(e_{mn}|e_{rm}, \mathbf{w}_{rmn}, \phi_{rmn}) &= \Phi_e(e_{mn} | e_{rm}) + \Phi_w(e_{mn} | w_m, w_n) + \\ &\Phi_o(e_{mn} | \phi_m, \phi_n) + \Phi_t(e_{mn} | l(v_m, v_n)) \end{aligned} \quad (2)$$

where $\mathbf{w}_{rmn} = (w_r, w_m, w_n)$ and $\phi_{rmn} = (\phi_r, \phi_m, \phi_n)$ denote triplets of width and orientation estimates, and $l(v_m, v_n)$ the estimated path between vertices v_m and v_n . The criterion we optimize then becomes

$$F^{\text{nl}}(T(k)) = \sum_{e_{mn} \in T(k)} c_{mn}^{\text{nl}} + \sum_{\substack{e_{rm} \in T(k) \\ e_{mn} \in T(k)}} \Phi(e_{mn} | e_{rm}, \mathbf{w}_{rmn}, \phi_{rmn}) \quad (3)$$

When k is given, finding the tree that minimizes $F^{\text{nl}}(k)$ yields very good results. In practice, however, k is unknown and the criterion of Eq. 3 suffers from one severe drawback: since the cost of additional edges is always positive, it systematically favors low values of k . To overcome this difficulty, when scoring the final tree, we count not only the cost of including some edges in the tree but also of discarding those that do not belong to it. This turns out to be equivalent to replacing the c_{mn}^{nl} , which are the sums of negative log likelihoods along the paths, by c_{mn}^{lr} computed by summing log likelihood ratios $-\log(P(x_i|f(x_i))/(1 - P(x_i|f(x_i))))$ along the paths. We then take the global score to be

$$F^{\text{lr}}(T(k)) = \sum_{e_{mn} \in T(k)} c_{mn}^{\text{lr}} + \sum_{\substack{e_{rm} \in T(k) \\ e_{mn} \in T(k)}} \Phi(e_{mn} | e_{rm}, \mathbf{w}_{rmn}, \phi_{rmn}) \quad (4)$$

In effect, replacing the log likelihoods by the log likelihood ratios amounts to penalizing graph edges with high probabilities that are left out of the final tree and overcomes the bias for low values of k [1].

In summary, for each cardinality $k=2 \dots N$, we build the tree that minimizes Eq. 3, and assign to it the score of Eq. 4. The final result is the tree that minimizes such score among all cardinalities,

$$\hat{\mathbf{T}}^* = \underset{T(k) \in \{T(2), \dots, T(N)\}}{\operatorname{argmin}} F(T(k)). \quad (5)$$

3.3 Estimating the Optimal Tree

Both to force trees to grow from the estimated root and to take advantage of the pairwise terms to guide the reconstruction process, we extended the ant colony optimization based algorithm presented in 3. Due to space limitations, here we only sketch the modifications we made.

First, to operate on directed graphs, we modify the neighborhood structure to only contain those edges that point away from the leaves of the current arborescence. Second, we fix the root node and initiate the search from it. At each growing step, we compute the effective cost of an edge in this neighborhood by taking into account the pairwise potentials introduced in Section 3.2. An edge is then stochastically selected based on this cost and the number of times it has been previously selected in the search. Finally, we set a limit on the bifurcation factor of a node in order avoid false connections on crossovers.

The above procedure is run several times until the obtained minimum costs for all cardinalities stabilize. Finally, we pick among all these arborescences of different cardinalities the one that minimizes the criterion of Eq. 4. The algorithm is very efficient since it requires only a few minutes to converge to topologically sound reconstructions, such as the ones shown in Fig. 1.

4 Results

The retinal scans of Fig. 1 belong to the DRIVE database [13]. We manually outlined the vascular trees and treat them as ground truth. These ground truths are used to compare the quality of the reconstructions of the presented method as shown in Fig. 1(d), against a spanning tree as given in Fig. 1(b), and the work of [1], Fig. 1(c).

To compare the methods quantitatively, we use the metric introduced by the DIADEM challenge [14], which is specifically designed to compare topology of a reconstructed tree against ground truth. The metric returns a number between 0 and 1 that measures the topological distance between trees by matching their branching and end points, and then analyzing the connecting paths. Mistakes close to the tree root are penalized more heavily than the ones closer to the leaves since they produce more severe topological changes.

Fig. 4(a) shows that the method proposed here yields a very substantial improvement over the other methods with respect to the DIADEM metric, which is consistent with the qualitative results presented Fig. 1(d). The optimal tree cardinalities, illustrated by the diamonds in Fig. 4(b), are automatically obtained by minimizing the score of Eq. 4.

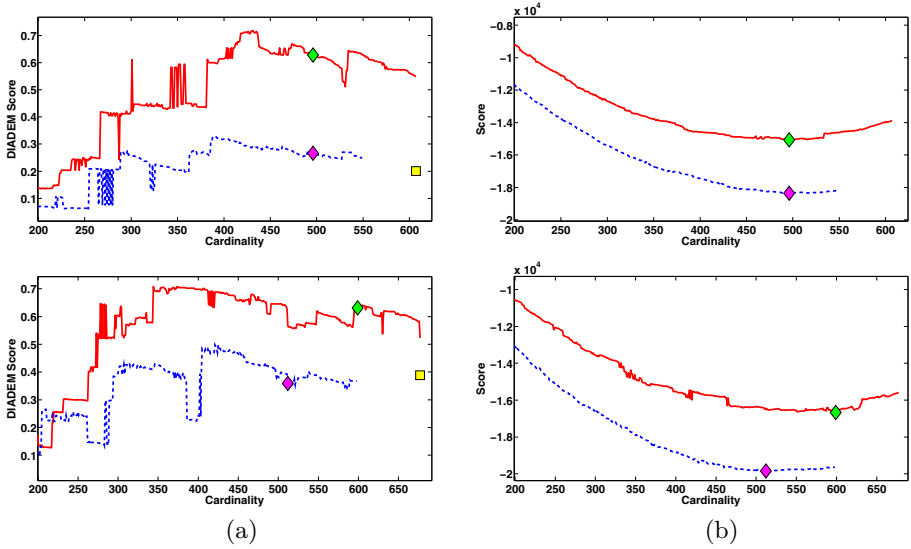


Fig. 4. Quantitative evaluation of the reconstructions in the two retinal scans of Fig. 1 (a) DIADeM scores as a function of the tree cardinality for our earlier approach [1] in blue dotted line, and for the one presented here in red solid line, which is substantially better. The yellow squares represent the scores of the standard MST. (b) Corresponding scores computed using the criterion of Eq. 4. In both cases, the diamonds denote the selected cardinality, taken to be the one that minimizes this criterion.

5 Conclusion

We have presented an algorithm for automatic tree reconstruction that enforces geometric constraints such as smoothness and width consistency along the branches, while explicitly handling crossovers and bifurcations. This yields a substantial qualitative and quantitative improvement in the reconstructions of retinal vascular trees at an acceptable computational cost.

In future work, we will extend our approach to other domains such as 3D dendrite delineation and generalize the type of constraints we can impose.

References

1. Gonzalez, G., Türetken, E., Fleuret, F., Fua, P.: Delineating Trees in Noisy 2D Images and 3D Image-Stacks. In: CVPR, San Francisco, CA (June 2010)
2. Chimani, M., Kandyba, M., Ljubić, I., Mutzel, P.: Obtaining optimal k-cardinality trees fast. *J. Exp. Algorithmics* 14, 2.5–2.23 (2009)
3. Blum, C., Blesa, M.J.: New metaheuristic approaches for the edge-weighted k-cardinality tree problem. *Computers & Operations Research* 32(6), 1355–1377 (2005)

4. Leandro, J., Soares, J., Cesar, R., Jelinek, H.: Blood vessels segmentation in non-mydratic images using wavelets and statistical classifiers. In: Brazilian Symposium on Computer Graphics and Image Processing, p. 262 (2003)
5. Yedidya, T., Hartley, R.: Tracking of blood vessels in retinal images using kalman filter. In: DICTA, Washington, DC, USA, pp. 52–58. IEEE Computer Society, Los Alamitos (2008)
6. Can, A., Shen, H., Turner, J., Tanenbaum, H., Roysam, B.: Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms. *TITB* 3(2), 125–138 (1999)
7. Fan, D.: Bayesian inference of vascular structure from retinal images. PhD thesis, Dept. of Computer Science, U. of Warwick, Coventry, UK (2006)
8. Sun, K., Sang, N., Zhang, T.: Marked point process for vascular tree extraction on angiogram. In: Yuille, A.L., Zhu, S.-C., Cremers, D., Wang, Y. (eds.) *EMMCVPR 2007*. LNCS, vol. 4679, pp. 467–478. Springer, Heidelberg (2007)
9. Gonzalez, G., Fleuret, F., Fua, P.: Automated Delineation of Dendritic Networks in Noisy Image Stacks. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) *ECCV 2008, Part IV*. LNCS, vol. 5305, pp. 214–227. Springer, Heidelberg (2008)
10. Gonzalez, G., Fleuret, F., Fua, P.: Learning rotational features for filament detection. In: *CVPR*, Miami, FL, pp. 1582–1589 (June 2009)
11. Aibinu, A., Iqbal, M., Shafie, A., Salami, M., Nilsson, M.: Vascular intersection detection in retina fundus images using a new hybrid approach. *Computers in Biology and Medicine* 40(1), 81–89 (2010)
12. Huang, K., Yan, M.: Robust Optic Disk Detection in Retinal Images using Vessel Structure and Radon Transform. In: *SPIE*, vol. 6144 (2006)
13. Staal, J., Abramoff, M., Niemeijer, M., Viergever, M., van Ginneken, B.: Ridge based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging* 23, 501–509 (2004)
14. HHMI: Diadem challenge (2010), <http://diademchallenge.org/>

Dual-Model Automatic Detection of Nerve-Fibres in Corneal Confocal Microscopy Images

M.A. Dabbah¹, J. Graham¹, I. Petropoulos², M. Tavakoli², and R.A. Malik^{2,*}

¹ Imaging Sciences and Biomedical Engineering (ISBE),
The University of Manchester, Oxford Rd,
Manchester, M13 9PT, UK

{m.a.dabbah,jim.graham}@manchester.ac.uk

² Cardiovascular Research Group, The University of Manchester,
46 Grafton St., Manchester, M13 9NT, UK

{ioannis.petropoulos,mitra.tavakoli,rayaz.a.malik}@manchester.ac.uk

Abstract. Corneal Confocal Microscopy (CCM) imaging is a non-invasive surrogate of detecting, quantifying and monitoring diabetic peripheral neuropathy. This paper presents an automated method for detecting nerve-fibres from CCM images using a dual-model detection algorithm and compares the performance to well-established texture and feature detection methods. The algorithm comprises two separate models, one for the background and another for the foreground (nerve-fibres), which work interactively. Our evaluation shows significant improvement ($p \approx 0$) in both error rate and signal-to-noise ratio of this model over the competitor methods. The automatic method is also evaluated in comparison with manual ground truth analysis in assessing diabetic neuropathy on the basis of nerve-fibre length, and shows a strong correlation ($r = 0.92$). Both analyses significantly separate diabetic patients from control subjects ($p \approx 0$).

1 Introduction

Diabetic Peripheral Neuropathy (DPN) is one of the most common long-term complications of diabetes. The accurate detection and quantification of DPN are important for defining at-risk patients, anticipating deterioration, and assessing new therapies. Current methods of detecting and quantifying DPN, such as neurophysiology, lack sensitivity, require expert assessment and focus only on large nerve-fibres whereas the earliest signs of neuropathy are likely to be found among small nerve-fibres. On the other hand, small nerve-fibre damage is currently assessed using skin/nerve biopsy, which is highly invasive and is not suitable for repeated investigations.

However, recent research [15,10,8] using Corneal Confocal Microscopy (CCM) suggests that this non-invasive, and hence reiterative, test might be an ideal surrogate endpoint for human diabetic neuropathy. These studies demonstrate that measurements made by CCM accurately quantify corneal nerve fibre morphology. The measurements reflect the severity of DPN and relate to the extent of

* This work is supported by a JDRF scholar grant 17-2008-1031.

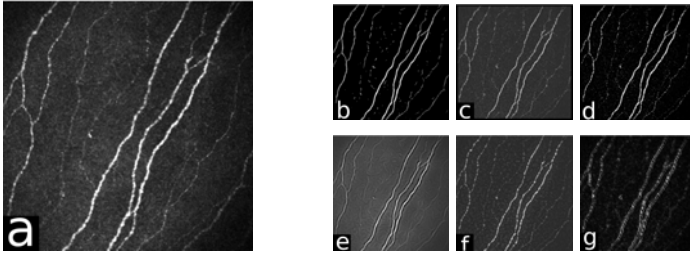


Fig. 1. An illustration of the methods' responses. (a) the CCM image, (b) Dual-model, (c) Linop, (d) Hessian, (e) 2D Gabor, (f) Monogenic and (g) DTCWT.

intra-epidermal nerve-fibre loss seen in skin biopsy. However, the major limitation preventing extension of this technique to wider clinical practice is that analysis of the images using interactive image analysis is highly labour-intensive and requires considerable expertise to quantify nerve-fibre pathology. To be clinically useful as a diagnostic tool, it is essential that the measurements be extracted automatically.

The first critical stage in analysis of CCM images (an example is shown in Figure 1(a)) is the detection of nerve-fibres. This is challenging as the nerve-fibres often show poor contrast in the relatively noisy images. The literature on this topic is not extensive, although the problem has a superficial similarity to other, more widely investigated, applications, such as detection of blood-vessels in retinal images. Ruggeri *et al.* [17] describe a heuristic method that was adapted from retinal analysis. In [2] we conducted a preliminary comparison of methods for contrast enhancement of nerve-fibres, comparing a Gabor wavelet with a well-established line detector.

This paper presents a dual-model algorithm for automatic detection and measurement of nerve-fibres in CCM images. Using a 2D Gabor wavelet and a Gaussian envelope, the dual-model of foreground (nerve-fibres) and background is constructed and applied to the original CCM image. The detection relies on estimating the correct local and dominant orientation of the nerve-fibres. Identifying low-contrast fibrous structures is a commonly encountered problem in several areas of investigation. Examples include mammography, retinopathy, angiography and detection of asbestos fibres. A number of methods have been developed and successfully applied in these applications. We evaluate our dual-model in comparison with some of these methods and with appropriate well-established feature detectors. While our analysis focuses on CCM images, our results suggest that this may be an appropriate contrast enhancement method in other application domains. In addition to the evaluation of the nerve-fibre detection responses, we have also evaluated the clinical utility of the method by a comparison with manual analysis.

2 Linear-Structure and Feature Detection

A method of linear structure detection (Line Operator - Linop), originally developed for detection of asbestos fibres [4] has also been shown to be effective

in detecting ducts in mammograms [18]. Linop exploits the linear nature of the structures to enhance their contrast by computing the average intensity of pixels lying on a line passing through the reference pixel for multiple orientations and scales. The largest values are chosen to be corresponding to the line, the strength of which is determined by the difference with the average intensity of the similarly oriented square neighbourhood.

In a preliminary study [2], we use the 2D Gabor filter [9] to detect nerve-fibres in CCM images. The filter is a band-pass filter that consists of a sinusoidal plane wave with a certain orientation and frequency, modulated by a Gaussian envelope. This spatial domain enhancement is based on the convolution of the image with the even-symmetric Gabor filter that is tuned to the local nerve-fibre orientation.

Frangi *et al.* [6] used a multiscale decomposition of the Hessian matrix to detect and measure blood vessels in Digital Subtraction Angiography images. They derived a discriminant function based on the eigenvalues and eigenvectors that has maximum response for tube-like structures. The external energy is used to attract the curve towards points which have a high likelihood of lying on a central vessel axis.

The Dual-Tree Complex Wavelet Transform (DTCWT) [11] is an extension of the Discrete Wavelet Transform (DWT), which provides a sparse representation and characterisation of structures and texture of the image at multiresolutions. The DTCWT utilises two DWT decompositions (trees) with specifically selected filters that gives it the properties of approximate shift-invariance and good directionality. The key feature of the DTCWT operation lies in the differences between the filters in the two trees.

The Monogenic signal [5] (a variant of a 2D analytic signal) is an extension of the analytic signal using quaternionic algebra in an attempt to generalise the method so it is capable of analysing intrinsically 2D signals e.g. structures within images. The Monogenic signal is based on the Riesz transform, which is a 2D generalization of the Hilbert transform used in the conventional analytic signal. The Monogenic signal is defined as the combination of the original signal and the Riesz-transformed one in the algebra of quaternions.

3 Dual-Model Nerve-Fibre Detection

In order to quantify the CCM images the nerve-fibres have to be detected. These captured images of nerve-fibre structures could suffer from several types of corruption due to some acquisition conditions, and nerve-fibres may appear faint due to small size or being only partly in the focus plane. Therefore, a nerve-fibre contrast enhancement algorithm is needed to exploit the linear structure of the nerve-fibres and distinguish them from the background noise. All of the methods described in the previous section are capable of providing this enhancement. In the next section we describe our approach.

3.1 Nerve-Fibre Contrast Enhancement

For this purpose the foreground model $\mathcal{M}_{\mathcal{F}}$ is an even-symmetric and real-valued Gabor [9,3] wavelet and the background model $\mathcal{M}_{\mathcal{B}}$ is a two-dimensional Gaussian envelope.

$$\mathcal{M}_{\mathcal{F}}^{(x_\theta, y_\theta)} = \left(\cos \left(\frac{2\pi}{\lambda} x_\theta + \phi \right) \right) e^{\left\{ -\frac{1}{2} \left(\frac{x_\theta^2}{\sigma_x^2} + \frac{\gamma^2 y_\theta^2}{\sigma_y^2} \right) \right\}} \quad (1)$$

$$\mathcal{M}_{\mathcal{B}}^{(x_\theta, y_\theta)} = \alpha e^{\left\{ -\frac{1}{2} \left(\frac{x_\theta^2}{\sigma_x^2} + \frac{\gamma^2 y_\theta^2}{\sigma_y^2} \right) \right\}} \quad (2)$$

$$x_\theta = x \cos \theta + y \sin \theta \quad (3)$$

$$y_\theta = -x \sin \theta + y \cos \theta \quad (4)$$

The x and y axes of the dual-model coordinate frame x_θ and y_θ are defined by a rotation of θ , which is the dominant orientation of the nerve-fibres in a particular region within the image (see Section 3.2). This dual-model is used to generate the positive response $\mathcal{R}_{\mathcal{P}} = \mathcal{M}_{\mathcal{F}} + \mathcal{M}_{\mathcal{B}}$ and the negative response $\mathcal{R}_{\mathcal{N}} = \mathcal{M}_{\mathcal{F}} - \mathcal{M}_{\mathcal{B}}$ that are applied to the original CCM image and can be represented as in Equations (5) and (6) respectively.

$$\mathcal{R}_{\mathcal{P}}^{(x_\theta, y_\theta)} = \left[\cos \left(\frac{2\pi}{\lambda} x_\theta + \phi \right) + \alpha \right] e^{\left\{ -\frac{1}{2} \left(\frac{x_\theta^2}{\sigma_x^2} + \frac{\gamma^2 y_\theta^2}{\sigma_y^2} \right) \right\}} \quad (5)$$

$$\mathcal{R}_{\mathcal{N}}^{(x_\theta, y_\theta)} = \left[\cos \left(\frac{2\pi}{\lambda} x_\theta + \phi \right) - \alpha \right] e^{\left\{ -\frac{1}{2} \left(\frac{x_\theta^2}{\sigma_x^2} + \frac{\gamma^2 y_\theta^2}{\sigma_y^2} \right) \right\}} \quad (6)$$

The equations of $\mathcal{R}_{\mathcal{P}}$ and $\mathcal{R}_{\mathcal{N}}$ assume that the Gaussian envelope of both responses are identical i.e. they have the same variances $\sigma_{(x,y)}^2$ and the same aspect ratio γ . The magnitude of the Gaussian envelope α defines the threshold in which a nerve-fibre can be distinguished from the background image. The value of α can be set empirically to control sensitivity and accuracy of detection. The wavelength λ defines the frequency band of the information to be detected in the CCM image. Its value might be computed for a subregion within the image that has significant variability of nerve-fibre width. However for simplicity, λ is chosen to be a global estimate of the entire image based on empirical results.

This in turn enhances the nerve-fibres that are oriented in the dominant direction, and decreases anything that is oriented differently by increasing the contrast between the foreground and the noisy background, whilst effectively reducing noise around the nerve-fibre structure as shown in Figure 1(b). This pixel-wise operation adjusts the models to suit the local neighbourhood characteristics of the reference pixel at $f^{(i,j)}$ by modifying the parameters of the foreground and background models. The dot products of the models and the reference pixel's neighbourhood (Equations 7 and 8) are then combined to generate the final enhanced value of this particular reference pixel $g^{(i,j)}$ (Equation 9).

$$\Gamma_p^{(i,j)} = \left\langle f_\omega^{(i,j)}, \mathcal{R}_{\mathcal{P}} \right\rangle \quad (7)$$

$$\Gamma_n^{(i,j)} = \left\langle f_\omega^{(i,j)}, \mathcal{R}_{\mathcal{N}} \right\rangle \quad (8)$$

$$g^{(i,j)} = \frac{\Gamma_p^{(i,j)}}{1 + e^{(-2k\Gamma_n^{(i,j)})}} \quad (9)$$

The neighbourhood area of the reference pixel is defined by the width ω . The sharpness of the transition of the enhanced image value at a particular pixel $g^{(i,j)}$ is controlled by k . A larger k amounts to a sharper transition when $\Gamma_n = 0$.

3.2 Nerve-Fibre Orientation Estimation

In CCM images, the nerve-fibres flow in locally consistent orientations everywhere. In addition, there is a global orientation that dominates the general flow. This orientation field describes the coarse structure of nerve-fibres. Using the least mean square (LMS) algorithm [7], the local orientation of the block centred at certain pixel is computed as in [16].

Since the orientations vary at a slow rate, a low-pass Gaussian filter is applied globally in order to further reduce errors at near-nerve-fibre and non-nerve-fibre regions. The LMS produces a stable smooth orientation field in the region of the nerve-fibres; however when applied on the background of the image, i.e. between fibres, the estimate is dominated by noise due to the lack of structure and uniform direction.

4 Experimental Results and Analysis

The evaluation has been conducted on a database of 525 CCM images captured using the HRT-III¹ microscope from 69 subjects (20 controls and 49 diabetic patients). The resolution is $1.0417\mu\text{m}$ and the field of view is $400 \times 400\mu\text{m}^2$ of the cornea. For each individual, several fields of view are selected manually from near the centre of the cornea that show recognisable nerve-fibres. Using the Neuropathy Disability Score (NDS) [1], 48 patients were categorised into four groups according to severity of neuropathy (asymptomatic: $0 \geq \text{NDS} \leq 2$ ($n = 26$), mild: $3 \geq \text{NDS} \leq 5$ ($n = 9$), moderate: $6 \geq \text{NDS} \leq 8$ ($n = 10$) and severe: $9 \geq \text{NDS} \leq 10$ ($n = 3$)).

The performance of all methods is obtained by validating the extracted nerve-fibres in comparison with an expert manual delineation using *CCMetric*:². Only the raw response of each method is taken into account without any further post-processing operations or shade correction methods as shown in Figure (II). Binary images are obtained by a simple uniform thresholding operation that is followed by a thinning operation to achieve a one-pixel-wide skeleton image.

4.1 Comparison of Nerve-Fibre Detection Methods

Three measures have been used in order to quantify the evaluation: the false-positive (FPR), the true-positive (TPR) and the equal-error rate (EER), which is the average of optimal FPR and false-negative rate at minimal difference between both. The measurements are taken by comparing the generated skeleton at different threshold intervals of the methods' responses with the manually delineated "ground-truth". A tolerance of $\pm 3.141\mu\text{m}$ (3 pixels) was allowed in

¹ Heidelberg Engineering Inc., modified to acquire corneal confocal images.

² CCMetrics is a purpose built interactive graphical interface which helps experts to manually delineate nerve-fibres in CCM images.

determining coincidence between the ground-truth and the detected nerve-fibres. The Peak Signal to Noise Ratio (PSNR) is also used to evaluate the performance of all methods. The PSNR is computed with respect to the mean squared error of the detected nerve-fibres from the manual delineation. The practical implementations of the Hessian, the DTCWT and the Monogenic signal were obtained from public domain sources [12,14,13], while the rest are implemented by our research group.

The EER and PSNR values for all the methods are presented in the box-plots in Figure 2 and Table 1. Each data point in Figure 2 corresponds to the evaluation on one of the 525 CCM images in the database. The dual-model shows lower EER and higher PSNR than all other methods (Table 1). These improvements are statistically significant ($p \approx 0$ using three different non-parametric tests). The table also shows that the standard deviations of both EER and PSNR are low for the dual-model, which indicates a more stable and robust behaviour. The closest competitor is Linop. The methods designed for linear structures perform rather better on this test than the more generic DTCWT and Monogenic signal methods.

The superior performance of the dual-model is borne out by the ROC curves of Figure 2, in which the dual model shows improved detection at all operation points.

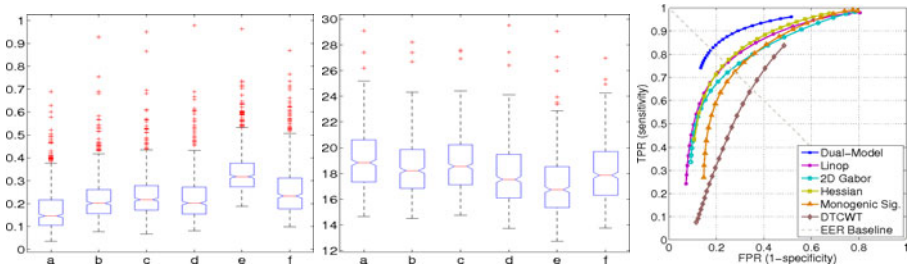


Fig. 2. From left to right, the box-plots of the EER and the PSNR are shown for all methods. The ROC curves are presented at the far right. The box-plots indicate the upper and the lower quartiles as well as the median (the bar) of the EER and PSNR values respectively; whiskers show the extent of the rest of the data while crosses indicate outliers for (a) dual-model, (b) Linop, (c) 2D Gabor, (d) Hessian, (e) DTCWT and (f) Monogenic.

Table 1. A comparison of mean EER and PSNR and their standard deviations

	Dual-Model	Linop [4]	2D Gabor [29]	Hessian [6]	DTCWT [11]	Monog. [5]
EER[%]	17.79 ± 10.58	22.65 ± 10.76	24.15 ± 10.74	23.14 ± 11.53	34.17 ± 10.43	26.50 ± 12.58
PSNR _[dB]	19.08 ± 2.16	18.51 ± 2.09	18.80 ± 2.11	17.93 ± 2.27	17.00 ± 2.23	18.11 ± 2.20

4.2 Assessment of Clinical Utility Results

In previous studies, using manual measurement of nerve-fibres, several features have been used to quantify the CCM images, including nerve-fibre length (NFL):

the total length of nerve-fibres measured in an image, nerve-fibre density: the total number of nerve-fibres per unit area and branch density: the number of fibre branches per unit area. Of these nerve-fibre length proved to be the most discriminating, and we use this measure here to compare automated with manual measurement of the nerve-fibre images.

The box-plots in Figure 3 show a strong similarity between the manual and the automated analysis. However the scale of the NFL has slightly changed from (3.68–33.91) for the manual analysis to (5.67–26.53) for the automated analysis. ANOVA analysis results in a p -value for discrimination among these groups which is slightly higher for the automated than the manual analysis, though both are significant ($p \approx 0$). The automated NFL measurements show a very strong correlation ($r = 0.92$) with the manual NFL values, which indicates that the automated system is successfully identifying the correct nerve-fibres. The coefficient of variation $c_v = \frac{\sigma}{\mu}$ of the manual analysis is 0.34, reducing for the automated analysis to 0.29, which indicates more reliability and robustness of the results.

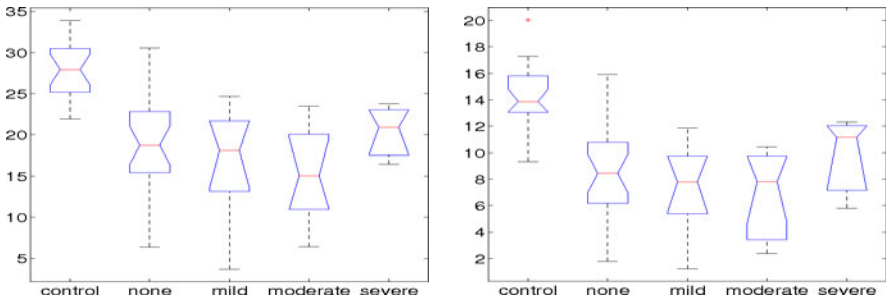


Fig. 3. Box-plots showing the NFL scores for each of the NDS groups calculated manually (left) and automatically (right)

5 Conclusion

The analysis of CCM images requires the identification of fibre-like structures with low contrast in noisy images. This is a requirement shared by a number of imaging applications in biology, medicine and other fields. A number of methods have been applied in these applications, and we have compared some of these, and more generic methods with a dual-model detection algorithm devised for this study. The comparison used a large set of images with manual ground truth. In terms of both error-rates (pixel misclassification) and signal-to-noise ratio, the dual model achieved highest performance. It seems reasonable to propose that this filter is likely to prove equally useful in applications of a similar nature. The question of the clinical utility of the method was also addressed in this paper. The evaluation has shown that the automatic analysis is consistent with the manual ground truth with a correlation of ($r = 0.92$). Similarity in grouping control and patient subjects between manual and automated analysis was also achieved with ($p \approx 0$). Therefore, it is sound to conclude that the automated analysis, which can be much quicker, is a potentially more reliable and practical alternative to

manual analysis due to its consistency and immunity to the inter/intra-observer variabilities.

References

1. Abbott, C.A., Carrington, A.L., Ashe, H., Bath, S., Every, L.C., Griffiths, J., Hann, A.W., Hussein, A., Jackson, N., Johnson, K.E., Ryder, C.H., Torkington, R., Ross, E.R.E.V., Whalley, A.M., Widdows, P., Williamson, S., Boulton, A.J.M.: The north-west diabetes foot care study: incidence of, and risk factors for, new diabetic foot ulceration in a community-based patient cohort. *Diabetic Medicine* 19(5), 377–384 (2002)
2. Dabbah, M.A., Graham, J., Tavakoli, M., Petropoulos, Y., Malik, R.A.: Nerve fibre extraction in confocal corneal microscopy images for human diabetic neuropathy detection using gabor filters. In: *Medical Image Understanding and Analysis (MIUA)*, pp. 254–258 (July 2009)
3. Daugman, J.G.: Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research* 20(10), 847–856 (1980)
4. Dixon, R.N., Taylor, C.J.: Automated asbestos fibre counting. In: *Machine Aided Image Analysis*, pp. 178–185. Institute of Physics, London (1979)
5. Felsberg, M., Sommer, G.: The monogenic signal. *IEEE Transactions on Signal Processing* 49(12), 3136–3144 (2001)
6. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. In: Wells, W.M., Colchester, A.C.F., Delp, S.L. (eds.) *MICCAI 1998*. LNCS, vol. 1496, pp. 130–137. Springer, Heidelberg (1998)
7. Hong, L., Wan, Y., Jain, A.: Fingerprint image enhancement: algorithm and performance evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(8), 777–789 (1998)
8. Hossain, P., Sachdev, A., Malik, R.A.: Early detection of diabetic peripheral neuropathy with corneal confocal microscopy. *The Lancet* 366(9494), 1340–1343 (2005)
9. Jain, A.K., Farrokhnia, F.: Unsupervised texture segmentation using gabor filters. *Pattern Recognition* 24(12), 1167–1186 (1991)
10. Kallinikos, P., Berbanu, M., O'Donnell, C., Boulton, A., Efron, N., Malik, R.: Corneal nerve tortuosity in diabetic patients with neuropathy. *Investigative Ophthalmology & Visual Science* 45(2), 418–422 (2004)
11. Kingsbury, N.: Complex wavelets for shift invariant analysis and filtering of signals. *Applied and Computational Harmonic Analysis* 10(3), 234–253 (2001)
12. Kingsbury, N.: Dual-Tree Complex Wavelet Transform Pack (June 2002), <http://www-sigproc.eng.cam.ac.uk/~ngk/>
13. Kovsesi, P.: An implementation of Felsberg's monogenic filters (August 2005), <http://www.csse.uwa.edu.au/~pk/research/matlabfns/>
14. Kroon, D.J., Schrijver, M.: Hessian based Frangi Vesselness filter (October 2009), <http://www.mathworks.co.uk/>
15. Malik, R.A., Kallinikos, P., Abbott, C.A., van Schie, C.H.M., Morgan, P., Efron, N., Boulton, A.J.M.: Corneal confocal microscopy: a non-invasive surrogate of nerve fibre damage and repair in diabetic patients. *Diabetologia* 46(5), 683–688 (2003)
16. Rao, A.R.: *A taxonomy for texture description and identification*. Springer, New York (1990)
17. Ruggeri, A., Scarpa, F., Grisan, E.: Analysis of corneal images for the recognition of nerve structures. In: *IEEE Conference of the Engineering in Medicine and Biology Society (EMBS)*, pp. 4739–4742 (September 2006)
18. Zwiggelaar, R., Astley, S., Boggis, C., Taylor, C.: Linear structures in mammographic images: Detection and classification. *IEEE Transactions on Medical Imaging* 23(9), 1077–1086 (2004)

Characterizing the Regional Structural Difference of the Brain between Tau Transgenic (rTg4510) and Wild-Type Mice Using MRI

Zhiyong Xie, Dewen Yang, Diane Stephenson, Daniel Morton,
Carol Hicks, Tracy Brown, and Thomas Bocan

Pfizer Inc, Groton, CT, United States

{Zhiyong.Xie, Dewen.Yang, Diane.Stephenson, Daniel.Morton,
Carol.Hicks, Tracy.Brown, Thomas.Bocan}@pfizer.com

Abstract. rTg4510 transgenic mouse model demonstrates features resembling Alzheimer's disease including neurofibrillary degeneration and progressive neuronal loss. We investigated the volumetric differences of brain structures between transgenic and wild-type mice using MR images of fourteen 5.5 month old female mice. Tensor-based morphometry and atlas-based segmentation were applied to MRI images. Severe atrophy of hippocampus and neocortex as well as ventricular dilatation were observed in the transgenic mice. These findings were confirmed by histopathologic evaluation of the same mice. The results suggest that MRI should be useful for evaluating disease-modifying therapies for Alzheimer's disease in the rTg4510 model and comparing treatment responses in mice and humans.

Keywords: Alzheimer, rTg4510, MRI, Brain, Atlas-based segmentation, Tensor-based morphometry.

1 Introduction

Alzheimer's disease is characterized by deposition of neurofibrillary tangles that consist of abnormally hyperphosphorylated tau [1]. Neuronal loss in neocortex and hippocampus are closely associated with the process of neurofibrillary degeneration [2]. Different transgenic mouse models of Alzheimer's neuropathology have been created to advance the understanding of the disease and development of treatments [3]. The rTg4510 mouse model is characterized by conditional overexpression of hyperphosphorylated human P301L mutant tau and profound neurofibrillary pathology, neurodegeneration and behavioral impairment [4]. Spatial memory deficits, progressive increase in neurofibrillary tangles and rapidly progressing neuronal loss have been reported to occur by 5.5 months of age [4].

MRI has been widely investigated as a biomarker to diagnose and estimate the disease progression in Alzheimer's disease [5][6][7]. In this study, we evaluated the sensitivity of in-vivo MRI in characterizing the morphological change of rTg4510 mice at 5.5 months. MR brain images of seven double transgenic rTg4510 female mice and 7 age-matched wild-type female mice were acquired. Tensor-based

morphometry (TBM) was performed on brain images to inspect regional morphological changes. To quantify the changes in different structures, we segmented the hippocampus, neocortex, ventricle, and cerebellum from MR images of each mouse and compared the size of these structures between the two groups. The results indicated that severe atrophy in the neocortex and hippocampus and dilation of lateral ventricles occurred in transgenic rTg4510 mice while the size of cerebellum remained unchanged. These results were confirmed by histopathologic examination performed in the same mice.

2 Material and Methods

2.1 Animals and MR Imaging

All animal handling procedures were carried out in compliance with the NIH Guide for the Care and Use of Laboratory Animals under a protocol approved by the Pfizer Global Research and Development Animal Care and Use Committee.

Seven double transgenic rTg4510 female mice and seven age-matched wild-type (WT) female mice (5 months 5 days to 5 months 12 days of age, 22-31 g) were evaluated. The mice were housed in a ventilated, temperature-controlled room with a 12-hour light/dark cycle.

Brain MRI was performed on a horizontal bore 4.7T magnet (Bruker Biospec 47/40, Bruker-Biospin, Inc). After an initial 5-10 min 2.5% isoflurane anesthesia induction period, all mice were anesthetized and maintained with 1.6-2.0% isoflurane in oxygen delivered via a nose cone. The mice were positioned in prone position with heads fixed to the plastic nose cone with the aid of a tooth bar and ear pins. Body temperature was recorded and maintained at 35.5-37.5C using a water heated animal bed and a water heated blanket (on the top). Respiratory rate was monitored continuously using a small animal monitoring and gating system (SA Instrument, Inc. Stony Brook, NY 11790, USA). RF excitation for imaging was delivered through a 72mm volume coil, and an actively decoupled mouse brain quadrature surface coil placed on the head was used as the receiver. T2-weighted 3D RARE images were acquired with the following parameters: field of view = $16 \times 16 \times 19.2$ mm³, matrix dimensions = $128 \times 128 \times 64$, spatial resolution = $125 \mu\text{m} \times 125 \mu\text{m} \times 300 \mu\text{m}$, TR = 2600 ms TE = 23 ms, RARE factor = 16. The total imaging time was 44m 22s 400ms.

After MR imaging, mice were euthanized using carbon dioxide gas. The brains were collected, fixed by immersion in 10% neutral buffered formalin, and embedded in the coronal plane using a mouse brain mold. Each brain produced about 6 slabs 2mm thick. Tissues were processed routinely, embedded in paraffin blocks, sectioned at a thickness of 5 μm , and stained with hematoxylin and eosin (H&E). All H&E sections of brain were examined qualitatively by light microscopy.

2.2 Image Analysis

The method proposed by Cohen et. al. [8] was used to correct RF inhomogeneity of MR images. Whole brain was segmented from the MR image using an open source software ITK-Snap [9] followed by a manual correction. More specifically, a 3D

geodesic active contour method [10] implemented in ITK-Snap was used for the semi-automatic brain segmentation.

With the intensity-corrected brain images, we performed tensor-based morphometry [11] to characterize the structural difference between two groups. Besides the TBM study, we also segmented four structures (hippocampus, neocortex, lateral ventricle, and cerebellum) from the MR images and measured the structural size to quantify the difference.

2.2.1 Tensor-Based Morphometry

Tensor-based morphometry (TBM) has been widely used to characterize the brain atrophy in Alzheimer's disease in clinical studies [12]. We employed the same method to evaluate the brain morphological differences between transgenic and wild-type mice. We first created a template brain image by registering the brain image of each subject to a pre-selected image, and computing the average of all aligned images. Then we registered all individual brain images to the template, generating a deformation field for each subject. The registration was initialized with a rigid body registration and followed by a nonlinear registration. Rigid body registration was computed by optimizing the mutual information using one-plus-one method [13] implemented in ITK (www.itk.org). Our nonlinear registration method [14], [15] was a spline-based extension to Thirion's Demons technique [16]. It used optical flow to determine the correspondence of voxels which exhibit sufficiently large intensity gradients. Based on the estimated sparse correspondences, a B-spline function of the correspondences over the whole brain volume was determined using weighted scattered data approximation. This two-step algorithm was applied over multiple resolution levels in conventional coarse to fine fashion: both the resolution of the images and the number of spline control parameters were simultaneously adjusted. Specifically, starting with B-spline functions that have a small number of parameters, the algorithm was iterated to match the coarse features of the images. The result was used to initialize the registration at the next resolution level, where the number of spline parameters was increased to allow alignment of the finer features that were apparent in the higher resolution images. This strategy provided a way to incrementally refine the registration and improve the robustness of the method. To correct for variability in intensity, the histograms of the images were matched. The registration method was validated based on the labeled T2 weighted MR images created by RCIBI [16] (<http://www.bnl.gov/medical/RCIBI/mouse/>), and was capable of yielding overlap ratios of greater than 90% for big structures (Neocortex, Cerebellum, Thalamus, etc) and 70% for small structures such as ventricle.

With the deformation field that aligned the brain image of each subject to the template, a 3D Jacobian map can be generated by computing the determinant of local Jacobian matrix at each voxel of template image. The Jacobian determinant was a local measurement of volume difference of each subject relative to the template image. A greater than one value of Jacobian determinant represented voxel expansion during the registration, while a value less than one represented voxel contraction. Since all images were registered to the same template, the Jacobian map can be used to voxel-wise compare the volumetric difference between two groups. To better characterize the local morphological difference, the Jacobian map only included the

nonlinear deformation. Global transformation such as scaling of the whole brain was not taken into account.

To increase the normality of the data distribution, the Jacobian map for each subject was log-transformed and smoothed with a 0.2mm Gaussian kernel. Two tailed t-tests were performed at each voxel of smoothed images to find the regions with significant volumetric difference between the two groups.

2.2.2 Quantification of Size Differences of Different Structures

To verify the results discovered from the TBM analysis and quantify the volumetric differences between two groups, we measured the volume of anatomical structures in the brain images acquired from each mouse. In both clinical and preclinical studies, hippocampus, neocortex, and lateral ventricle have been used as biomarkers to predict Alzheimer's disease progression [17], [18], [19], [20], [21], [22]. Our TBM study also showed the morphological differences in these regions. Cerebellum was formerly thought to be relatively unaffected in the AD brain. Recent studies revealed that the cerebellum also underwent degenerative changes in Alzheimer's disease [23], [24]. In this study, we segmented hippocampus, neocortex, lateral ventricle, and cerebellum from MRI images of each mouse and compared the size of these structures between groups. The segmentation was done by applying an atlas-based segmentation followed by a manual correction. Based on the T2-weighted MR template and 20 labeled structures from RCIBI [25], we first registered the template to the brain image of each animal, and mapped the label of the template to each subject with computed transformation. Computer-generated segmentation was then reviewed by a trained expert and correction was made if necessary. Fig 1 shows an example of such segmentation.

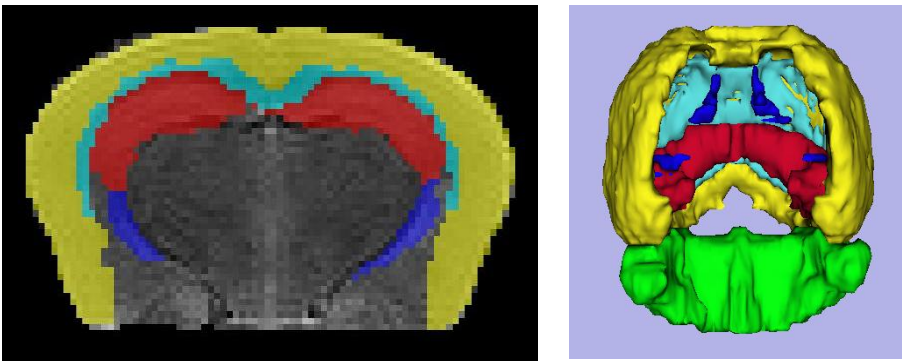


Fig. 1. An example of image segmentation. Left picture shows one slice of a coronal brain image with overlap of colored labels of different structures (yellow = cortex; light blue = corpus callosum & external capsule; red = hippocampus; dark blue = ventricles; green = cerebellum). 3D visualization of four structures is shown on the right picture.

3 Results

The result of TBM analysis showed significant atrophy of neocortex and hippocampus and dilatation of lateral ventricles in the rTg4510 transgenic mice. Figure 2

demonstrates one slice of a t-score map generated from the voxel-wise two-tailed t-tests of Jacobian maps between two groups. The blue color represents regions of volume contraction in transgenic mice and red color indicates the regions of volume expansion. Color bars give the scale of the t-scores. $t=3.05$ corresponds to $p=0.01$ for an uncorrected two-tailed t-test.

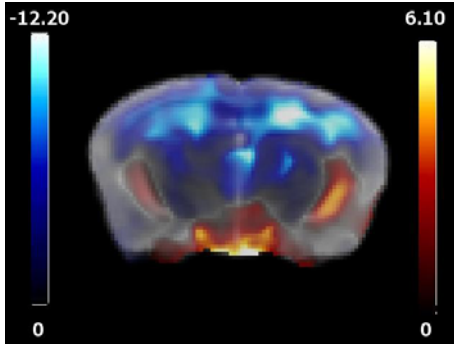


Fig. 2. t-score map of voxel-wise volumetric comparisons between rTg4510 and wild-type mice in a coronal plane at the level of the dorsal hippocampus. The blue color represents regions of structural contraction in transgenic mice and red color indicates the regions of structural expansion.

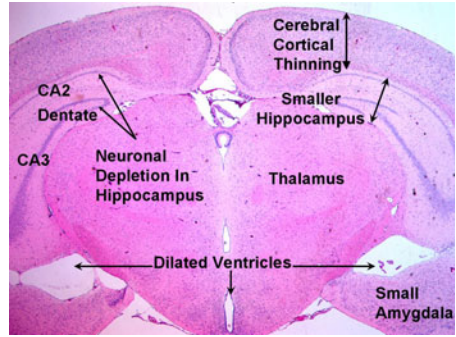


Fig. 3. A representative histopathology image of the brain of an rTg4510 mouse. Thinner neocortex, smaller hippocampus and smaller amygdala were observed. Ventricular dilation indicated additional loss of brain parenchyma.

Based on the segmentation of hippocampus, neocortex, lateral ventricle, and cerebellum, we computed the volume of different structures for each animal. Figure 4 illustrates the mean volume difference between two groups for hippocampus, neocortex, and lateral ventricle. The hippocampus and neocortex of transgenic mice were significantly smaller than those of wild-type mice (p value for two-tailed t-tests were less than 0.0001). Lateral ventricle size of transgenic mice was significantly larger than wild-type mice (p value for two-tailed t-test was 0.023). No significant difference was observed in the volume of the cerebellum between the two groups.

When we inspected the individual data, we found that all transgenic mice had a smaller hippocampus and neocortex than wild-type mice. Animal to animal variability in rTg4510 mice was observed in the ventricular volume measurement. 4 out of 7 rTg4510 mice had enlarged lateral ventricles (hydrocephalus) while the other three were similar in size to the wild-type mice. Figure 5 shows one example of a transgenic mouse with enlarged ventricles. The left picture is a slice of a T2-weighted image of the transgenic mouse. The ventricular dilatation (hydrocephalus, shown as bright structures in this picture) is evident in the rTg4510 brain which was not observed in the corresponding slice of a wild-type mouse (right picture, Figure 4). Hydrocephalus likely reflects expansion of the fluid-filled ventricles in response to loss of brain mass.

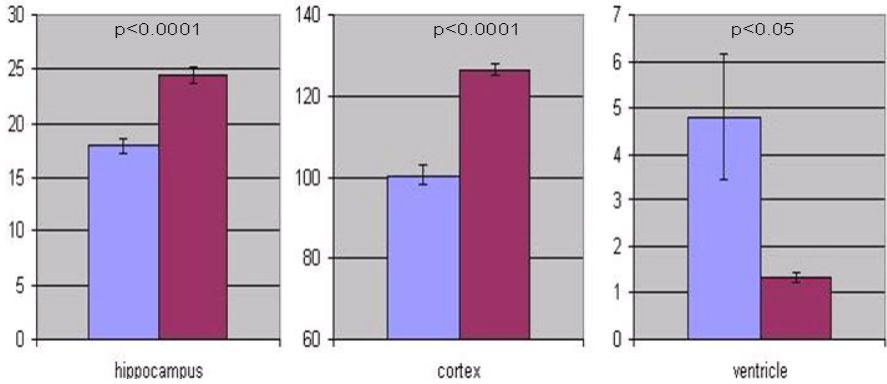


Fig. 4. Volume comparison of three structures between transgenic and wild-type mice. Blue bar represented the average volume of rTg4510 mice and the red bar indicated the average volume of wild-type mice (unit: mm³). Statistical values between the groups are illustrated on the top of each picture.

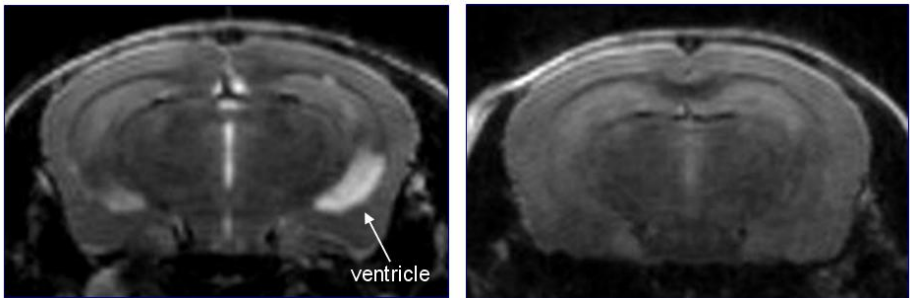


Fig. 5. MR images of a transgenic mouse (left) and a wild-type mouse (right). Ventricular dilatation (high intensity areas in the left panel) can be clearly observed in the image of transgenic mouse.

From the histopathology study, reduced cellularity in the pyramidal cell layers of the CA1 and CA2 regions and granular cell layer of the dentate gyrus was the most consistent difference between rTg4510 mice and control mice (illustrated in figure 3). Decreased cellularity of pyramidal cell layers and decreased thickness of the hippocampus were observed in 6 of 7 rTg4510 mice, with one mouse appearing to have hippocampal thickness comparable to controls while still exhibiting neuronal loss. The thicknesses of neocortex and amygdala in the sections with hippocampus were qualitatively reduced in 3 of the rTg4510 mice. Four Tg4510 mice had minimal to mild dilatation of the ventricles. No significant cellular changes in the cerebellum were observed. These histopathology results correlated with the MRI findings in the same animals and with published reports demonstrating neuropathologic changes in the same mice [22].

4 Discussion

The present study shows that in-vivo MRI is capable of detecting cortical and hippocampal atrophy and ventricular dilatation of rTg4510 at 5.5 months of age. This time point corresponds with the age which reportedly exhibits neurodegeneration in both published reports [4] as well as in our own histopathology examination. These results suggest that noninvasive imaging can be used as a potentially translatable biomarker for progressive neuronal degeneration in this model.

In MRI analysis, TBM provides a qualitative way to characterize the location of atrophy and dilatation as a group, while the segmentation of different structures gives a quantitative way to measure the atrophy and dilatation of anatomical structures for an individual mouse. The 2 methods provide complementary information to monitor the disease progression in transgenic mice.

Future studies will include longitudinal MRI measurements in parallel with behavioral evaluation to investigate the power of MRI in detecting the disease progression at different stages. Since the rTg4510 mutant model was designed with a promoter that can be repressed with tetracycline treatment, it is possible to turn off transgene expression of tau and ask whether MRI can detect reversibility of the phenotype.

References

1. Iqbal, K., Grundke-Iqbal, I.: Alzheimer neurofibrillary degeneration: Significance, etiopathogenesis, therapeutics and prevention. *J. Cell Mol. Med.* 12, 38–55 (2008)
2. Bondareff, W., Mountjoy, C.Q., Roth, M., Hauser, D.L.: Neurofibrillary degeneration and neuronal loss in alzheimer's disease. *Neurobiology of Aging* 10(6), 709–715 (1989)
3. Duff, K., Suleman, F.: Transgenic mouse models of Alzheimer's disease: How useful have they been for therapeutic development? *Briefings in Functional Genomics and Proteomics* 3(1), 47–59 (2004)
4. SantaCruz, K., Lewis, J., Spires, T., et al.: Tau suppression in a neurodegenerative mouse model improves memory function. *Science* 309, 476–481 (2005)
5. Du, A.T., Schuff, N., Chao, L.L., Kornak, J., Jagust, W.J., Kramer, J.H., Reed, B.R., Miller, B.L., Norman, D., Chui, H.C., Weiner, M.W.: Age effects on atrophy rates of entorhinal cortex and hippocampus. *Neurobiol. Aging* 27, 733–740 (2006)
6. Jack Jr., C.R., Slomkowski, M., Gracon, S., Hoover, T.M., Felmlee, J.P., Stewart, K., et al.: MRI as a biomarker of disease progression in a therapeutic trial of milameline for AD. *Neurology* 60, 253–260 (2003)
7. McEvoy, L.K., Fennema-Notestine, C., Roddey, J.C., Hagler Jr., D.J., Holland, D., Karow, D.S., Pung, C.J., Brewer, J.B., Dale, A.M.: Alzheimer disease: quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment. *Radiology* 251, 195–205 (2009)
8. Cohen, M.S., DuBois, R.M., Zeineh, M.M.: Rapid and effective correction of RF inhomogeneity for high field magnetic resonance imaging. *Human Brain Mapping* 10(4), 204–211 (2000)
9. Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C., Gerig, G.: User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* 31(3), 1116–1128 (2006)

10. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. *International Journal of Computer Vision* 22, 61–79 (1997)
11. Ashburner, J., Friston, K.J.: Morphometry. In: Frackowiak (ed.) *Human Brain Function*, pp. 707–725. Academic Press, London (2004)
12. Hua, X., Leow, A.D., Parikshak, N., Lee, S., Chiang, M., Toga, A.W., Jack Jr., C.R., Weiner, M.W., Thompson, P.M.: The Alzheimer’s Disease Neuroimaging Initiative: Tensor-based morphometry as a neuroimaging biomarker for Alzheimer’s disease: An MRI study of 676 AD, MCI, and normal subjects. *NeuroImage* 43(3), 458–469 (2008)
13. Styner, M., Gerig, G., Brechbuehler, C., Szekely, G.: Parametric estimate of intensity inhomogeneities applied to MRI. *IEEE Transactions on Medical Imaging* 19(3), 153–165 (2000)
14. Xie, Z., Farin, G.E.: Image registration using hierarchical b-splines. *IEEE Transactions on Visualization & Computer Graphics* 10(1), 85–94 (2004)
15. Xie, Z., Ng, L., Gee, J.C.: Two algorithms for non-rigid image registration and their evaluation. In: *Proc SPIE Medical Imaging, Image Processing* (2003)
16. Thirion, J.: Image matching as a diffusion process: an analogy with Maxwell’s demons. *Med. Image Anal.* 2(3), 243–260 (1998)
17. de Leon, M.J., George, A.E., Stylopoulos, L.A., Smith, G., Miller, D.C.: Early marker for Alzheimer’s disease: the atrophic hippocampus. *Lancet* 2, 672–673 (1989)
18. Mouton, P.R., Martin, L.J., Calhoun, M.E., Dal Forno, G., Price, D.L.: Cognitive decline strongly correlates with cortical atrophy in Alzheimer’s dementia. *Neurobiology of Aging* 19(5), 371–377 (1998)
19. Nestor, S.M., Rupsingh, R., Borrie, M., Smith, M., Accomazzi, V., Wells, J.L., Fogarty, J., Bartha, R.: Ventricular enlargement as a possible measure of Alzheimer’s disease progression validated using the Alzheimer’s disease neuroimaging initiative database. *Brain* 131, 2443–2454 (2008)
20. Grundman, M., Sencakova, D., Jack Jr., C.R., Petersen, R.C., Kim, H.T., Schultz, A., et al.: Brain MRI hippocampal volume and prediction of clinical status in a mild cognitive impairment trial. *Journal of Molecular Neuroscience* 19, 23–27 (2002)
21. Dickerson, B.C., Bakkour, A., Salat, D.H., Feczko, E., Pacheco, J., Greve, D.N., Grodstein, F., Wright, C.I., Blacker, D., Rosas, H.D., Sperling, R.A., Atri, A., Growdon, J.H., Hyman, B.T., Morris, J.C., Fischl, B., Buckner, R.L.: The cortical signature of Alzheimer’s disease: regionally specific cortical thinning relates to symptom severity in very mild to mild AD dementia and is detectable in asymptomatic amyloid-positive individuals. *Cereb. Cortex* 19, 497–510 (2009)
22. Spires, T.L., Orne, J.D., SantaCruz, K., Pitstick, R., Carlson, G.A., Ashe, K.H., Hyman, B.T.: Region-specific dissociation of neuronal loss and neurofibrillary pathology in a mouse model of tauopathy. *Am. J. Pathol.* 168, 1598–1607 (2006)
23. Lerner, A.J.: The Cerebellum in Alzheimer’s Disease. *Dement Geriatr Cogn. Disord.* 8, 203–209 (1997)
24. Thomann, P.A., Santos, V.D., Essig, M., Schroder, J.: P1-299: The cerebellum in mild cognitive impairment and Alzheimer’s disease: A structural MRI study. *Alzheimer’s and Dementia* 4(4) (suppl. 1); Alzheimer’s Association International Conference on Alzheimer’s Disease, Alzheimer’s Association International Conference on Alzheimer’s Disease, p.T307 (2008)
25. Ma, Y., Smith, D., Hof, P.R., Foerster, B., Hamilton, S., Blackband, S.J., Yu, M., Benveniste, H.: In vivo 3D digital atlas database of the adult C57BL/6J mouse brain by magnetic resonance microscopy. *Front. Neuroanat.* 2, 1 (2008)

Spatially Regularized SVM for the Detection of Brain Areas Associated with Stroke Outcome

Rémi Cuingnet^{1,2}, Charlotte Rosso^{1,3}, Stéphane Lehericy^{1,4},
Didier Dormont^{1,4}, Habib Benali², Yves Samson^{1,3}, and Olivier Colliot¹

¹ Université Pierre et Marie Curie-Paris 6, CNRS UMR 7225, Inserm UMR_S 975, Centre de Recherche de l'Institut Cerveau-Moelle (CRICM), Paris, France

² Inserm, UMR_S 678, LIF, Paris, France

³ AP-HP, Urgences Cérébro-Vasculaires, Pitié-Salpêtrière, Paris, France

⁴ Neuroradiology department & CENIR, Pitié-Salpêtrière, Paris, France

Abstract. This paper introduces a new method to detect group differences in brain images based on spatially regularized support vector machines (SVM). First, we propose to spatially regularize the SVM using a graph encoding the voxels' proximity. Two examples of regularization graphs are provided. Significant differences between two populations are detected using statistical tests on the margins of the SVM. We first tested our method on synthetic examples. We then applied it to 72 stroke patients to detect brain areas associated with motor outcome at 90 days, based on diffusion-weighted images acquired at the acute stage (one day delay). The proposed method showed that poor motor outcome is associated to changes in the corticospinal bundle and white matter tracts originating from the premotor cortex. Standard mass univariate analyses failed to detect any difference.

1 Introduction

Diffusion-weighted imaging (DWI) is of considerable interest to the clinical evaluation of acute stroke patients [1]. The location of the lesions has been suggested to represent a better predictor than their global volume [2]. At the subacute or chronic phases, previous studies have shown that damages to the corticospinal tract (CST) [3] and lesions to the primary sensorimotor cortex [2,4] correlated with poor motor outcome. At the acute stage, regional changes in the apparent diffusion coefficients (ADC) were suggested as early quantitative indices of regional irreversible ischemic damage [5]. However, at the acute stage, the spatial pattern of ADC changes associated with motor outcome remains unclear.

Group analyses of differences between populations in brain imaging have widely relied on univariate voxel-wise analyses, such as voxel-based morphometry (VBM) for structural MRI [6] or their equivalent for diffusion imaging (VB-DWI). In such analyses, brain images are first spatially registered to a common stereotaxic space, and then mass univariate statistical tests are performed in each voxel to detect significant group differences. However, the sensitivity of these approaches is limited when the differences are spatially complex and involve a

combination of different voxels or brain structures [7]. Recently, there has been a growing interest in support vector machines (SVM) methods [8,9] to overcome the limits of these univariate analyses. These approaches allow capturing complex multivariate relationships in the data and have been successfully applied to the individual classification of a variety of neurological conditions [10,11,12,13]. Moreover, the output of the SVM can also be analyzed to localize spatial patterns of discrimination, for example by drawing the coefficients of the optimal margin hyperplane (OMH) – which, in the case of a linear SVM, live in the same space as the MRI data [12,13,14]. However, one of the problems with analyzing directly the OMH coefficients is that the corresponding maps are noisy and lack spatial coherence. Moreover only a few of these approaches perform a statistical analysis of the OMH coefficients [14].

In this paper, we propose a new method to detect group differences in brain images based on spatially regularized SVM. In particular, we show how spatial consistency can be directly enforced into the SVM by using Laplacian regularization. We then propose a statistical analysis based on the spatially regularized SVM to detect brain regions which are significantly different between two groups of subjects. The proposed framework is tested on 2D synthetic test images and then applied to the detection of differences between stroke patients with good and poor outcome based on DWI acquired at the acute stage.

2 Spatially Regularized SVM Using the Graph Laplacian

In this section, after some background on SVM, we propose to spatially regularize the SVM using a graph encoding the voxels' proximity. We then give two examples of regularization graphs.

2.1 Background

In this contribution, we consider the case of brain images which are spatially normalized to a common stereotaxic space. These images can be any characteristics extracted from the MRI, such as gray matter concentration maps (in VBM) or ADC maps (in diffusion MRI). Let $(\mathbf{x}_s)_{s \in [1, N]} \in (\mathbb{R}^d)^N$ be the images of N subjects and $(y_s)_{s \in [1, N]} \in \{\pm 1\}^N$ their group labels (e.g. diagnosis). SVMs search for the hyperplane for which the margin between groups is maximal, the OMH. The standard linear SVM solves the following optimization problem [8]:

$$\min_{\mathbf{w} \in \mathbb{R}^d, b \in \mathbb{R}} \underbrace{\sum_{s=1}^N \phi(y_s [\langle \mathbf{w}, \mathbf{x}_s \rangle + b])}_{\text{Empirical Loss}} + \underbrace{\lambda \|\mathbf{w}\|^2}_{\text{Classical Tikhonov Regularization}} \quad (1)$$

where ϕ is the *hinge loss function* ($\phi : u \mapsto (1 - u)^+$), b the bias and $\lambda \in \mathbb{R}^+$.

With a linear SVM, the *feature space* is the same as the *input space*. Thus, when the input features are the voxels of the image, each component of \mathbf{w} also corresponds to a voxel. One can therefore represent the values of \mathbf{w} in the image

space, and use this map to localize differences. However, the map \mathbf{w} can be noisy and scattered (as for example in [12]). This is due to the fact that the regularization term of the standard linear SVM *is not a spatial regularization*. Voxel-based comparisons are subject to registration errors and interindividual variability. Gaussian smoothing is therefore often used as a preprocessing step. However, some image information is lost with the smoothing which, for example, mixes white matter with gray matter voxels. Tissue probability maps could be used to overcome this limitation. More generally, if voxels are connected, meaning for example spatially, anatomically or functionally close, we would like the SVM to consider them as similar. In the next section, we propose to encode this proximity in a graph and to use its Laplacian to spatially regularize the SVM.

2.2 Regularization Based on Diffusion on Graph

Graphs are a natural framework to take spatial information into consideration. Voxels of a brain image can be considered as nodes of a graph which models the voxels' proximity. This graph can be the voxel connectivity (6, 18 or 26) or a more sophisticated graph.

By changing the regularization term of the standard linear SVM equation (II), one can force \mathbf{w} to be smooth with respect to the graph, hence spatially smooth. That is to say, if two voxels are close, the classifier will consider them as similar. The minimization problem becomes:

$$\min_{\mathbf{w} \in \mathbb{R}^d, b \in \mathbb{R}} \underbrace{\sum_{s=1}^N \phi(y_s [\langle \mathbf{w}, \mathbf{x}_s \rangle + b])}_{\text{Empirical Loss}} + \lambda \underbrace{\| \exp \left\{ \frac{1}{2} \beta L \right\} \mathbf{w} \|^2}_{\text{Spatial Regularization}} \quad (2)$$

where L is the graph Laplacian [15] and $\beta \in \mathbb{R}^+$ and $\exp\{\cdot\}$ is the matrix exponential [16]. This new minimization problem (2) is equivalent to an SVM optimization problem [9,17]. The new kernel K is given by: $K(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{x}_1^t e^{-\beta L} \mathbf{x}_2$. Note that $K(\mathbf{x}_1, \mathbf{x}_2) = \langle e^{-\frac{\beta}{2} L} \mathbf{x}_1, e^{-\frac{\beta}{2} L} \mathbf{x}_2 \rangle$.

Our approach differs from the diffusion kernels introduced by Kondor et al. [18]. In our case, the nodes of the graph are the features, here the voxels, whereas in [18], the nodes were the objects to classify. Laplacian regularization was also used in satellite imaging [19] but, again, the nodes were the objects to classify. Our approach can also be considered as a spectral regularization on the graph [20]. To our knowledge, such spectral regularization has not been applied to brain images but only to the classification of microarray data [21].

2.3 Examples of Regularization Graphs

One has now to define the graph depending on type of spatial proximity one wants to enforce. The simplest option is to use the image connectivity (6, 18

¹ For any square matrix M , $\exp\{M\} = \sum_{k=0}^{\infty} \frac{1}{k!} M^k$

or 26). In this case, the regularized SVM would be equivalent to smoothing the data with a Gaussian kernel with standard deviation $\sigma = \sqrt{\beta}$ [20]. But this would mix gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF). Instead, we propose a graph which takes into consideration both the spatial localization and the tissue types. Based on tissue probability maps, in each voxel v , we have the set of probabilities p_v that this voxel belongs to GM, WM or CSF. We considered the following graph. Two voxels are connected if and only if they are neighbors in the image (6-connectivity). The weight $a_{u,v}$ of the edge between two connected voxels u and v is $a_{u,v} = e^{-d_{\chi^2}(p_u, p_v)^2 / (2\sigma^2)}$, where d_{χ^2} is the χ^2 -distance between two distributions. We chose beforehand σ equal to the standard deviation of $d_{\chi^2}(p_u, p_v)$.

3 Statistical Analysis of the Margins

In this section, we propose a statistical analysis to detect brain regions which are significantly different between two groups of subjects, based on the results of the spatially regularized SVM.

The classification function obtained with a linear SVM is the sign of the inner product of the features with \mathbf{w} , a vector orthogonal to the OMH [8,9]. Therefore, if the absolute value of the i^{th} component of the vector \mathbf{w} , $|w_i|$, is small compared to the other components ($|w_j|_{j \neq i}$), the i^{th} feature will have a small influence on the classification. Conversely, if $|w_i|$ is relatively large, the i^{th} feature will play an important role in the classifier. However, one cannot compare directly the weights from two different comparisons. More precisely, let A, A', B, B' be four groups of subjects and let $\mathbf{w}^{(A)}$ and $\mathbf{w}^{(B)}$ be the optimal weights (SVM outputs) for the comparison A versus A' and the comparison B versus B' respectively. If the separation between A and A' is larger than the separation between B and B' then $\|\mathbf{w}^{(A)}\|$ will be smaller than $\|\mathbf{w}^{(B)}\|$. Hence, one cannot directly compare the components $|w^{(A)}_i|$ and $|w^{(B)}_i|$ for significance tests.

SVMs search for the hyperplane for which the margin [2] between groups is maximal (Fig. 1). The margin m is large when there is a large separation between two groups. By combining m and $\frac{|w_i|}{\|\mathbf{w}\|}$, one can simultaneously quantify the separation between groups and the relative influence of the different features. Therefore, we propose to analyze the statistic of $\frac{m|w_i|}{\|\mathbf{w}\|}$. We performed permutation tests on $\frac{m|w_i|}{\|\mathbf{w}\|}$ under the null hypothesis \mathcal{H}_0 of no relationship between the class labels and the global structure of the MR scan. By randomly permuting the subjects labels 20,000 times and training the SVM with this permutation of labels, we estimated for each voxel i the probability distribution of $\frac{m|w_i|}{\|\mathbf{w}\|}$ under \mathcal{H}_0 . Based on these distributions, it is possible to test \mathcal{H}_0 at the voxel level. The false discovery rate (FDR) was used to correct for multiple comparisons [22]. To the best of our knowledge, other statistical analyses of the OMH did not take the margin into account (e.g. [14]).

² For the standard linear SVM [1], the margin m is given by: $m = 2 \|\mathbf{w}\|^{-1}$. As for the spatially regularized version [2], $m = 2 \|\exp(\frac{1}{2}\beta L) \mathbf{w}\|^{-1}$ [9].

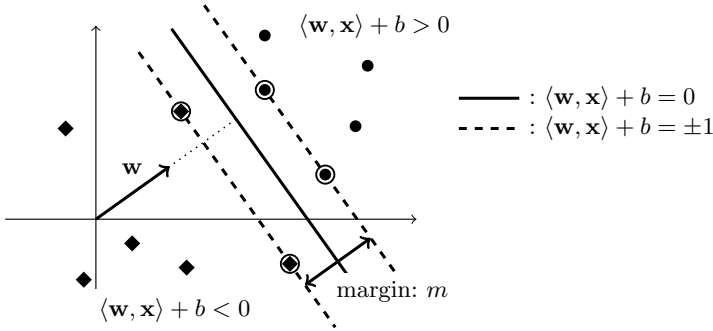


Fig. 1. Illustration of an optimal margin hyperplane with a linear SVM. The support vectors are circled.

4 Experiments and Results

We first tested our method on 2D synthetic test images. We then applied it on real data to the detection of brain areas associated with stroke outcome based on diffusion-weighted MRI acquired at the acute stage.

4.1 Synthetic Images

We first evaluated the ability of the method to detect artificial differences between two groups of 20 2D synthetic images (116×92 with 1.5 mm isotropic voxels) which were constructed as follows. We considered a slice of a WM template. For each of the 40 images, the voxels of the WM were assigned a random number between zero and one, the intensity of the other voxels being null. In each image of the first group, we constructed a hyperintensity h_{green} in the green region of Fig. 2 and h_{red} in the red region such as $(h_{\text{green}} + h_{\text{red}}) \sim \mathcal{N}(2, 0.2)$. Gaussian white noise ($\mathcal{N}(0, 1)$) was added to all images.

We tested six methods: three univariate methods and three SVM methods. We performed three univariate analyses on the voxel intensities: on the raw images, on the images smoothed with a Gaussian kernel and on the images preprocessed by $e^{-\frac{\beta}{2}L}$ (where L was the Laplacian of the graph used in the spatially regularized SVM). We tested three SVM methods: the standard linear SVM on the raw images, the standard linear SVM on the smoothed images and the spatially regularized SVM on the raw images.

All tests were corrected for multiple comparisons with a 5% FDR. The C parameter of the SVM was fixed to one ($\lambda = \frac{1}{2NC}$ [9]). The full width at half maximum (FWHM) of the Gaussian smoothing kernel was three voxels (the red and green regions' widths). β controls the size of the spatial regularization and was chosen to be equivalent to the FWHM of the Gaussian smoothing.

The results are shown on Fig. 2. The univariate analyses did not detect any difference for any images (raw, smoothed and preprocessed with $e^{-\frac{\beta}{2}L}$). The SVM detected only a fraction of the green and red regions on raw images. On smoothed

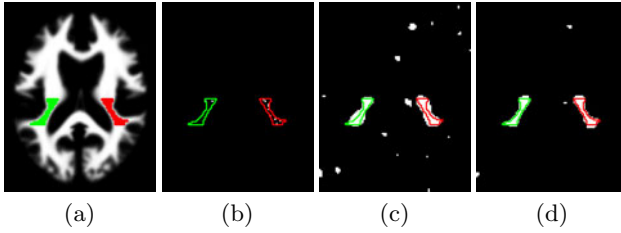


Fig. 2. Synthetic example: (a) WM template and regions to detect; detection with a linear SVM (b) on raw images; (c) on smoothed images; (d) detection with a spatially regularized SVM on raw images. All other analyses detected no difference.

images both regions were detected. The spatially regularized SVM also detected both regions and decreased the number of scattered clusters.

4.2 Brain Areas Associated with Stroke Outcome

Subjects and MRI Acquisition. We included 72 consecutive acute stroke patients (mean age: 60 ± 14 years^[24-81]). Exclusion criteria were symptomatic hemorrhagic transformation on follow-up MRI and death during the follow-up period (90 days). The modified Rankin Scale (mRS) was used to assess outcome at 90 days. Good outcome was defined as independency (mRS 0 to 2; 39 subjects) and poor outcome as severe disability (mRS 3 to 5; 33 subjects).

The median delay between stroke onset and MRI acquisition was 1.2 day³. MR imaging was performed on a 1.5 Tesla MR General Electric Signa. Axial isotropic DWI spin echo EPI included 24 slices of 5 mm thickness, with an interslice gap of 0.5 mm, a 280x210 mm FOV, a 96x64 matrix, TE = 98.9 ms, and TR = 2825 ms. A baseline T2 acquisition and a diffusion-weighted acquisition using a diffusion gradient of 1000 s.mm^{-2} were both acquired within 40 seconds. ADC maps were generated using dedicated commercially-available software (Functool 2, General Electric). ADC maps were then normalized to the Montreal Neurological Institute (MNI) space using SPM5 software⁴. To avoid any effect from lateralization, the template was symmetrized. To put all the lesions on the same side, ADC maps with the infarct lesion in the left hemisphere were flipped to the right for the analysis.

Statistical Analysis and Results. Univariate analyses were done with both a permutation test and a parametric Student's T-test on smoothed images (8-mm FWHM Gaussian filter). We also performed analyses with the spatially regularized SVM on the raw images. As in the previous section, the β parameter was chosen to correspond to the FWHM of the univariate analyses. All tests were corrected with a 5% FDR.

³ All imaging and clinical data were generated during routine clinical workup of the patients in our stroke center. The study was approved by the La Pitié-Salpêtrière Hospital Ethics Committee.

⁴ Statistical Parametric Mapping, Institute of Neurology, London, UK

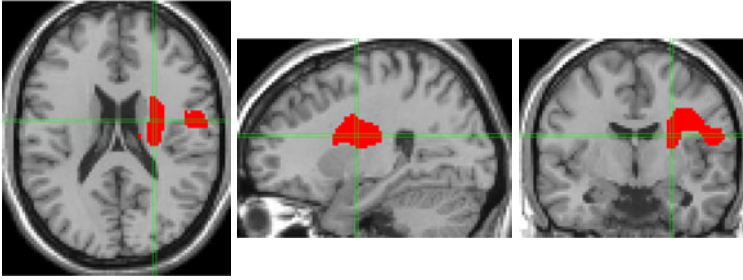


Fig. 3. Group differences of the ADC maps between poor and good motor outcome with a spatially regularized SVM ($z=20$ mm, $x=28$ mm and $y=-8$ mm in the MNI-space)

No regions were detected by the univariate analyses. The spatially regularized SVM detected significant changes predominantly localized in the WM (Fig. 3). The larger part of the detected cluster included the corticospinal tract at the level of the internal capsule. The smaller part was more superficial and was originating from the lower part of the motor and premotor cortex.

5 Conclusion

We proposed a new method based on spatially regularized SVM to detect group differences in brain images. Spatial consistency was directly enforced into the SVM by using the graph Laplacian. This provides a flexible approach to model different types of proximity between voxels. We then proposed to detect differences between groups using a statistical analysis which takes the margin of the SVM into account.

The proposed approach was applied to the detection of brain areas associated with stroke outcome based on DWI acquired at the acute stages. It allowed detecting changes localized in a large WM tract including the corticospinal tract and the lower part of the motor and premotor bundles. Univariate analyses failed to detect any differences. These results, obtained at the acute stage, are in line with previous studies carried out at the subacute or chronic phases [3,24]. Our results suggest that spatially regularized SVM might be useful to analyze MR images acquired in clinical routine as soon as 24 hours post stroke onset.

The proposed approach is not specific to diffusion MRI or stroke patients, and can be applied to other pathologies and other types of data (e.g. anatomical MRI). It has the potential to overcome the limits of traditional mass univariate voxel-wise analyses by detecting complex spatial patterns of alterations.

References

1. Chalela, J., et al.: Magnetic resonance imaging and computed tomography in emergency assessment of patients with suspected acute stroke: a prospective comparison. *The Lancet* 369(9558), 293–298 (2007)
2. Crafton, K., et al.: Improved understanding of cortical injury by incorporating measures of functional anatomy. *Brain* 126(7), 1650–1659 (2003)

3. Domi, T., et al.: Corticospinal tract pre-wallerian degeneration: A novel outcome predictor for pediatric stroke on acute MRI. *Stroke* 40(3), 780–787 (2009)
4. Lo, R., et al.: Identification of critical areas for motor function recovery in chronic stroke subjects using voxel-based lesion symptom mapping. *NeuroImage* 49(1), 9–18 (2010)
5. Rosso, C., et al.: Prediction of Infarct Growth Based on Apparent Diffusion Coefficients: Penumbra Assessment without Intravenous Contrast Material. *Radiology* 250(1), 184–192 (2009)
6. Ashburner, J., Friston, K.J.: Voxel-based morphometry—the methods. *NeuroImage* 11(6), 805–821 (2000)
7. Davatzikos, C.: Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. *NeuroImage* 23(1), 17–20 (2004)
8. Vapnik, V.N.: *The Nature of Statistical Learning Theory*. Springer, Heidelberg (1995)
9. Schölkopf, B., Smola, A.J.: *Learning with Kernels*. MIT Press, Cambridge (2001)
10. Lao, Z., et al.: Morphological classification of brains via high-dimensional shape transformations and machine learning methods. *NeuroImage* 21(1), 46–57 (2004)
11. Fan, Y., et al.: COMPARE: classification of morphological patterns using adaptive regional elements. *IEEE TMI* 26(1), 93–105 (2007)
12. Klöppel, S., et al.: Automatic classification of MR scans in Alzheimer’s disease. *Brain* 131(3), 681–689 (2008)
13. Vemuri, P., et al.: Alzheimer’s disease diagnosis in individual subjects using structural MR images: validation studies. *NeuroImage* 39(3), 1186–1197 (2008)
14. Mourão-Miranda, J., et al.: Classifying brain states and determining the discriminating activation patterns: Support vector machine on functional MRI data. *NeuroImage* 28(4), 980–995 (2005); Special Section: Social Cognitive Neuroscience
15. Chung, F.R.K.: *Spectral Graph Theory*, vol. 92. AMS (1992)
16. Golub, G.H., Van Loan, C.F.: *Matrix computations*. J. Hopkins Univ. Press, Baltimore (1996)
17. Smola, A.J., Schölkopf, B.: On a kernel-based method for pattern recognition, regression, approximation, and operator inversion. *Algorithmica* 22, 211–231 (1998)
18. Kondor, R.I., Lafferty, J.D.: Diffusion kernels on graphs and other discrete input spaces. In: *Proc. International Conference on Machine Learning*, pp. 315–322 (2002)
19. Gómez-Chova, L., et al.: Semi-supervised image classification with Laplacian support vector machines. *IEEE Geo. Rem. Sens. Let.* 5(3), 336–340 (2008)
20. Smola, A., Kondor, R.: Kernels and regularization on graphs. In: Schölkopf, B., Warmuth, M.K. (eds.) *COLT/Kernel 2003*. LNCS (LNAI), vol. 2777, p. 144. Springer, Heidelberg (2003)
21. Rapaport, F., et al.: Classification of microarray data using gene networks. *BMC bioinformatics* 8(1), 35 (2007)
22. Benjamini, Y., Hochberg, Y.: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. B. Met.*, 289–300 (1995)

Sparse Unbiased Analysis of Anatomical Variance in Longitudinal Imaging

Brian Avants, Philip A. Cook, Corey McMillan, Murray Grossman,
Nicholas J. Tustison, Yuanjie Zheng, and James C. Gee

Depts. of Radiology and Neurology
University of Pennsylvania
Philadelphia, PA 19104-6389
avants@grasp.cis.upenn.edu

Abstract. We present a new algorithm for reliable, unbiased, multivariate longitudinal analysis of cortical and white matter atrophy rates with penalized statistical methods. The pipeline uses a step-wise approach to transform and personalize template information first to a single-subject template (SST) and then to the individual's time series data. The first stream of information flows from group template to the SST; the second flows from the SST to the individual time-points and provides unbiased, prior-based segmentation and measurement of cortical thickness. MRI-bias correction, consistent longitudinal segmentation, cortical parcellation and cortical thickness estimation are all based on strong use of the subject-specific priors built from initial diffeomorphic mapping between the SST and optimal group template. We evaluate our approach with both test-retest data and with application to a driving biological problem. We use test-retest data to show that this approach produces (a) zero change when the retest data contains the same image content as the test data and (b) produces normally distributed, low variance estimates of thickness change centered at zero when test-retest data is collected near in time to test data. We also show that our approach—when combined with sparse canonical correlation analysis—reveals plausible, significant, annualized decline in cortical thickness and white matter volume when contrasting frontotemporal dementia and normal aging.

1 Introduction

Longitudinal MRI studies are often underpowered due to the challenges of data collection, data consistency and data analysis. One of the most difficult detection problems in longitudinal neuroimaging is tracking cortical thickness reduction in dementia via clinical resolution T1 MRI. These images typically have resolution near 1 mm^3 whereas the cortex itself is only one to five millimeters thick in this age range and thinner in diseased regions. Power is also reduced by the fact that these studies are testing for “differences of differences,” which increases noise in the measurement and reduces power. Additionally, there are few standard pipelines that exist for processing longitudinal thickness data. Finally, the combined challenges of segmentation, registration and change estimation make longitudinal pipelines difficult to validate. For these reasons, the majority of studies have focused on whole brain or regional, lobe-based methods for

estimating longitudinal atrophy rates such as the Brain Boundary Shift Integral (BBSI) [9][13][12][8][6]. These methods gain stability by integrating over large regions, but lose specificity. Other studies have increased power by using very large cohorts. Shaw, et al. used 1133 images to study differences in cortical thickness change between ADHD and controls [10]. However, studies of rare dementia do not have the luxury of large cohorts and thus few groups have attempted voxel or surface-wise estimates of longitudinal cortical thickness group differences.

Power and accuracy can be increased in longitudinal studies by combining unbiased estimation with feature selection. Yushkevich et al. showed that distributing interpolation equally across all of a subject's images is critical for eliminating the bias associated with longitudinal hippocampal atrophy [14]. This leads to a basic principal in longitudinal studies: all images should undergo an equivalent number of interpolations. While this principal increases accuracy, precision remains relatively low due to the noise associated with these studies. To alleviate this concern, multivariate feature selection may be used to not only reduce the multiple comparisons problem in a principled way [11] but also select a subset of data that contains strong signal that agrees across two independent views of the data. Sparse canonical correlation analysis allows one to select a controllably sparse, mutually informative subset of voxels from paired, complementary sets of measurements taken on the same subject. This subset of voxels may be selected without reference to diagnosis. Thus, after selected, they may be passed to a test for group differences without additional statistical penalty or circularity [7].

Here, we develop a new multivariate, morphometric framework—entitled Sparse Unbiased Analysis of Anatomical Variance (SUA AV)—for detecting voxel-wise group differences in cortical thickness and white matter atrophy rates. SUA AV builds on existing, open-source image registration, MRI-bias correction, segmentation, cortical thickness and sparse estimation tools. We apply and evaluate SUA AV in the challenging problem of using serial T1-weighted MRI to contrast rates of atrophy in cortical thickness and white matter volume between frontotemporal dementia (FTD) and a small set of control data. Tremendous interest exists in sensitive, unbiased approaches for quantifying longitudinal thickness change as such methods are essential tools in treatment trials. SUA AV is designed specifically for these types of tracking studies and achieves power via variable selection and accuracy via application of the latest knowledge in unbiased implementation of image measurement tools.

The most important novel components of SUA AV include: (1) unbiased segmentation based on propagation of tissue priors from a group template, to a single-subject template and, finally, to each time point image; (2) unbiased estimation of cortical thickness change from in vivo resolution T1 MRI; (3) use of multivariate feature selection to increase sensitivity in longitudinal studies of correlated white matter and cortical thickness change. The framework, though advanced, has relatively few requirements in that it only requires a prior-labeled template and serial data where two complementary views/images of each subject's time change are available. In this study, we derive these measures from a single modality, clinical T1-weighted imaging, by pairing an estimate of annualized cortical thickness change with annualized white matter volume change. We evaluate SUA AV with test-retest data and show that results are biologically plausible and sensitive in real clinical data.

2 Methods

2.1 Subjects and Imaging

The input data to our analysis consists of serial T1-weighted MRI collected longitudinally on a Siemens 3T scanner. Each study began with a rapid sagittal T1-weighted scan to determine patient position. A T1-weighted structural acquisition was then acquired with TR (repetition time) = 1620ms, TE (echo time) = 3ms, slice thickness: 1 mm, in-plane resolution: .9766mm x .9766mm and field of view (FOV) $256 \times 256 \times 192$. All data were checked for adequate quality across time points before proceeding to the analysis. An experienced neurologist, who specializes in discrimination of these diseases, grouped the subjects into likely syndromes (—) based on clinical phenotype and cognitive testing. Our analysis included 7 elderly control subjects and 26 FTD subjects matched for sex, age and education. All subjects' follow-up images were acquired approximately one year after the initial visit and there was no significant difference in the set of delay times between the control and FTD longitudinal image sets.

2.2 Sparse Unbiased Analysis of Anatomical Variance

We now describe the components of the SUAAB framework for gaining unbiased estimation of longitudinal neurodegeneration in white matter and cortical thickness and shown in figure 1 and figure 2.

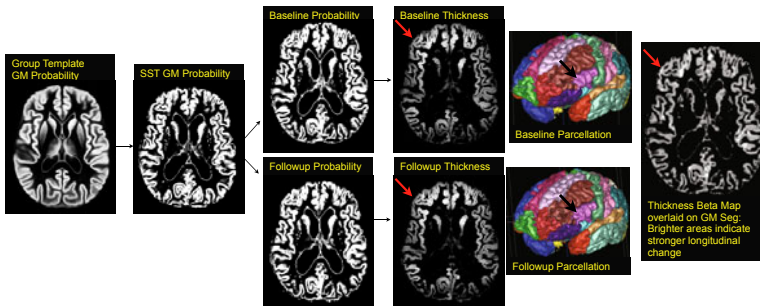


Fig. 1. The SUAAB algorithm's pipeline requires a template that contains cortical priors and cortical labels that are used in a parcellation scheme. These priors are mapped first to a single-subject template (SST) and used to initialize segmentation of the SST. The resulting probability maps are then deformed from SST space to individual space to initialize a prior-constrained segmentation of (in this case) a baseline and follow-up image collected at a one year interval. Thickness maps and cortical parcellation are then computed for each time point image. While parcellation is not explicitly required by SUAAB, we use parcellation, here, to evaluate the stability of the pipeline in test-retest data and to verify the findings provided by the voxel-wise analysis. Red arrows point to a region associated—in the group analysis—with significant atrophy and black arrows point to the corresponding region in the parcellation.

1. **SST Initialization and Optimization:** We construct an optimal, unbiased within-subject template from each subject's serial data by rigidly aligning the first time point image to the group template. Subsequent time points are aligned to the first time point. The result is then averaged to produce an unbiased initialization. Given the initialization, we use the method described in [11], and available in the ANTS toolkit [4], to compute the optimal shape and appearance SST with respect to all time points. This approach guarantees that the mapping for each subject to the group template or to the SST involves the same number of interpolations (one) or compositions (affine + diffeomorphism to go to the SST and affine + diffeomorphism + affine + diffeomorphism to go to the group template).
2. **Mapping to Group Template and Tissue Priors:** Given the SST from the first step, we use standard correlation-based ANTS diffeomorphic registration to find the invertible mapping between the SST and template. This mapping then allows us to deform tissue-priors for cerebrospinal fluid, gray matter and white matter to the SST-space. The white matter tissue probability map is then passed to the N4 algorithm [11] which uses this map to perform a weighted B-Spline based bias correction. The bias corrected image, along with tissue priors, are then passed to a prior-based, parametric Expectation-Maximization, Markov-Random-Field segmentation algorithm that fits the models to the SST. Here, the prior is used only for initialization of the EM-MRF algorithm.
3. **Time Point Segmentation:** The segmentation results in the SST space are deformed from the SST to each individual time point, based on the maps that were computed during the SST optimization. The same segmentation procedure from step 2 is then performed for each time point, but, here, the prior terms are maintained during optimization of the EM-MRF to prevent a large change from the initialization. Given the cortical segmentation, we also parcellate the cortex using the method described in [5], which deforms a template parcellation to the subject and fits the initial parcellation to the cortical data by using the fast marching method.

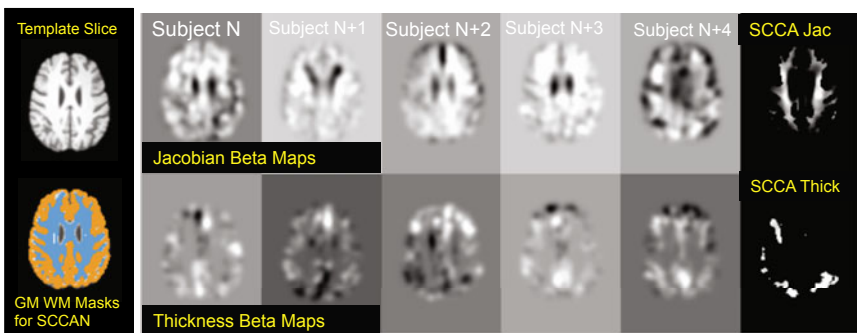


Fig. 2. The SCCAN component of SAAV is illustrated. The template and the masks for each “view” (Jacobian and thickness change) are at left. The top row shows a selection of individual Jacobian change images in the template space, after smoothing. The same subject’s thickness change images, after smoothing, are in the bottom row. A slice from the SCCAN weight maps for each view is shown at right.

The original template parcellation is based on the NIREP data (www.nirep.org). This approach was shown to perform as well as Freesurfer cortical parcellation within the volumetric domain [5].

4. **Thickness Estimation:** Given the segmentation results for each time point from step 3, we measure the cortical thickness at each time point with Diffeomorphic Registration Based Cortical Thickness (DiReCT) [2].
5. **Annualized Atrophy:** We map the resulting cortical thickness images back to the single-subject template and estimate longitudinal change in cortical mm/year—within the SST gray matter mask—by using the beta parameter in a linear regression fit, that is, change in thickness $\approx x_0 + \beta_T * t$ where x_0 is the intercept and t is the time parameter. Solving this regression problem at every voxel provides an image $\beta_T(\mathbf{x})$. When only two time points are present, this image can be computed by subtracting the baseline image from the followup image directly and dividing by the time separation. In the same way, but within the SST white matter mask, we estimate white matter volume change from the Jacobian of the serial deformation in SST space, providing image $\beta_J(\mathbf{x})$. We median filter both $\beta_{J,T}(\mathbf{x})$ images to remove few-voxel registration/interpolation inaccuracies. We then map both of the time-change beta maps (thickness and Jacobian) for each subject to the group template space at two millimeters.
6. **Feature Selection/Dimensionality Reduction:** Once the above steps are completed for all subjects, we locate reliable regions within the derived maps of longitudinal change by applying sparse canonical correlation analysis for neuroanatomy (SCCAN), as in [1]. SCCAN solves a constrained optimization problem that finds projection vectors, here represented as functions of \mathbf{x} in the template image domain, that maximize the relationship between paired modalities/measurements as in figure 2. In this study, we seek *strictly positive* projection vectors $\omega_T(\mathbf{x}) \geq 0, \omega_J(\mathbf{x}) \geq 0$ with $\|\omega_T\|_1 \leq C_T, \|\omega_J\|_1 \leq C_F$ where C is a constant related to the desired sparseness imposed by the L1 penalty, $\|\cdot\|_1$. The positivity constraint guarantees that both projection vectors, $\omega_{T,J}$, will have positive entries and may thus be viewed as a weighted average of the data. The L1 penalty guarantees that only a subset of voxels will enter the model. Under this formulation, SCCAN will maximize:

$$\operatorname{argmax}(\omega_T, \omega_J) : \operatorname{Corr} \left(\underbrace{\mathbf{T}\omega_T}_{\text{Thickness Projections}}, \underbrace{\mathbf{J}\omega_J}_{\text{Jacobian Projections}} \right) - \lambda_T \|\omega_T\|_1 - \lambda_J \|\omega_J\|_1, \quad (1)$$

where \mathbf{T} is a matrix with columns containing voxels from the β_T images and n -subjects number of rows, where \mathbf{J} is a matrix with columns containing voxels from the β_J images and n -subjects number of rows, Corr computes Pearson correlation and the λ are inversely related to the sparseness costs, C . We apply SCCAN to determine regions that have significant covariation between the white matter Jacobian change and thickness change, with a sparseness value that selects a small, informative subset of both white matter and cortex (approximately 20% of each). While both measures may be noisy, we expect the noisy components to be less well correlated than those that show robust, related change.

Note that steps 1—6 do not use diagnosis in any way. The last variable selection step (6) only keeps those voxels that provide maximal information between the two

measurements (here, thickness change and Jacobian change). Thus, if the covariation identified by SCCAN relates to unique disease patterns, detection power will be increased in the supervised, groupwise analysis, by removing uninformative background noise. The retained voxels in both modalities are then assessed by ANCOVA where age and sex effects are factored out to determine the significance of the disease alone.

3 Results

Evaluation 1: Sanity Check. We pass pseudo-longitudinal data to the SUAABV algorithm where three identical images are passed in and treated as if they were collected over three years. If there is any inconsistency in the implementation of the SUAABV pre-processing steps, illustrated in figure 1, SUAABV will yield a non-zero change in thickness and/or Jacobian. As expected, we gained zero differences in segmentation and thickness.

Evaluation 2: Test-Retest. We pass short interval longitudinal data to the SUAABV algorithm where four subjects underwent repeat imaging within one month of the first image. In this case, we use the cortical parcellations gained in our pipeline to provide 32 sample points for each image. We average the cortical thickness within each region over each time point. This approach mirrors that taken in [3] where Fischl and Dale computed a standard deviation of 0.25mm in thickness measurements collected in short-interval repeat sessions. SUAABV produced a 0.1mm standard deviation across these images, indicating a high level of repeatability, given that the image resolution is approximately 1mm³. We also used Monte Carlo simulation to determine if our test-retest distribution of thickness change differs from a randomly sampled Gaussian distribution centered at zero with the same standard deviation (0.104 mm) and same number of samples (4*32=128). We performed over 10,000 simulations where, for each simulation, we performed a t-test of the difference between our distribution and the random Gaussian sampling. 95.2 % of the results showed no difference, thus there is less than a five percent chance that the distribution is not Gaussian and centered at zero. In addition, we performed a Pearson chi-square normality test ($p = 0.2734$) which also showed our test-retest distribution is close to normal.

Evaluation 3: Application to FTD-Elderly Longitudinal Change. Given the results of Evaluation 1 and 2, we test whether SUAABV is able to detect neurobiological effects of dementia expected to occur within the frontal and/or temporal lobes, as in FTD. Thus, the full SUAABV pipeline (with SCCAN) is used to process 26 FTD subjects and 7 elderly subjects. The SCCAN variable selection is used to restrict the ANCOVA test that uses a model of the form: Thickness Change $\approx x_0 + \text{age} + \text{gender} + \text{diagnosis}$. The same model is used for the Jacobian. The p-values associated with the diagnosis predictor are corrected by the false discovery rate and shown, rendered on the template cortex, in figure 3. Middle frontal gyrus (MFG) and inferior frontal gyrus (IFG) on the left survives correction, along with anterior frontal lobe white matter. To eliminate the possibility that misregistration led to these results, we also interrogated the average thickness change across the MFG and IFG regions gained from the parcellation. These regional measures—which do not suffer from interpolation or registration error—also showed significant reduction over time in FTD relative to controls, thus affirming our voxel-wise findings.

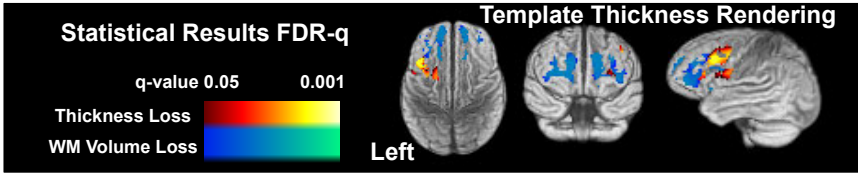


Fig. 3. The SUAAB method provides a restricted set of regions over which to perform statistical testing. Results are FDR corrected where we accept q-value (corrected p-value) < 0.05 as significant with cluster size $> 100 \text{ mm}^3$. Left frontal cortex atrophies at a greater rate (approximately 6-10 % per year) in FTD than controls. The template labeling identifies the cortical regions as middle frontal and inferior frontal gyrus. The white matter loss occurs bilaterally in the anterior frontal lobe. These findings are consistent with what is known about frontotemporal dementia.

4 Discussion

We presented an unbiased method for analyzing group differences in annualized cortical thickness and white matter volume change in dementia. Both unbiased image registration and segmentation algorithms were used. We showed that SUAAB produces zero change when we passed false serial data that contains identical images across time and produces zero-centered Gaussian distributed estimates of change across the brain when SUAAB is passed test-retest data from scanning sessions from the same subject imaged within a month. Finally, we showed that SUAAB produced significant, biologically plausible estimates of cortical and white matter atrophy within left, frontal cortex in FTD that differed from controls.

A critical choice when applying SUAAB is the sparseness penalty term. This term controls the number of voxels that enters into the downstream ANCOVA. The benefit is a vast reduction in the problem of multiple comparisons but this comes at the cost of a loss of detection power in regions that are rejected in this step. Specifically, in our study, regions of cortex that do not relate to white matter change (in the within-subject sense) will not enter into the testing procedure and vice versa. Additionally, the sparseness term should be selected based on a power analysis and *a priori* anatomical knowledge.

SUAAB provides a powerful new way to convert standard, serially collected MRI datasets into a pair of independent but covarying measurements that may be used to identify repeatable and reliable spatial changes in cortical thickness. We also sought to show that SUAAB provides a neuronatomically valid and meaningful way to increase power for longitudinal atrophy studies. While we have taken a high-level approach to evaluation of the full SUAAB pipeline, additional effort will be required to understand, in detail, the criticality of each step in SUAAB. Additionally, a deep understanding of parameter selection and testing on other datasets will be required to verify broad applicability of the algorithm.

References

1. Avants, B.B., Yushkevich, P., Pluta, J., Minkoff, D., Korczykowski, M., Detre, J., Gee, J.C.: The optimal template effect in hippocampus studies of diseased populations. *Neuroimage* (2010) (in press)

2. Das, S.R., Avants, B.B., Grossman, M., Gee, J.C.: Registration based cortical thickness measurement. *Neuroimage* 45(3), 867–879 (April 2009),
<http://dx.doi.org/10.1016/j.neuroimage.2008.12.016>
3. Fischl, B., Dale, A.M.: Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc. Natl. Acad. Sci. USA* 97(20), 11050–11055 (2000),
<http://dx.doi.org/10.1073/pnas.200033797>
4. Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.C., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., Song, J.H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R.P., Mann, J.J., Parsey, R.V.: Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* 46(3), 786–802 (2009),
<http://dx.doi.org/10.1016/j.neuroimage.2008.12.037>
5. Klein, A., Ghosh, S.S., Avants, B., Yeo, B.T.T., Fischl, B., Ardekani, B., Gee, J.C., Mann, J.J., Parsey, R.V.: Evaluation of volume-based and surface-based brain image registration methods. *Neuroimage* (February 2010),
<http://dx.doi.org/10.1016/j.neuroimage.2010.01.091>
6. Knight, W.D., Kim, L.G., Douiri, A., Frost, C., Rossor, M.N., Fox, N.C.: Acceleration of cortical thinning in familial alzheimer's disease. *Neurobiol. Aging* (December 2009),
<http://dx.doi.org/10.1016/j.neurobiolaging.2009.11.013>
7. Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F., Baker, C.I.: Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12(5), 535–540 (2009),
<http://dx.doi.org/10.1038/nn.2303>
8. Krueger, C.E., Dean, D.L., Rosen, H.J., Halabi, C., Weiner, M., Miller, B.L., Kramer, J.H.: Longitudinal rates of lobar atrophy in frontotemporal dementia, semantic dementia, and Alzheimer's disease. *Alzheimer Dis. Assoc. Disord.* (June 2009),
<http://dx.doi.org/10.1097/WAD.0b013e3181a6f101>
9. Scahill, R.I., Schott, J., Stevens, J.: Mapping the evolution of regional atrophy in Alzheimer's disease: unbiased analysis of fluid-registered serial MRI. *Proc. Natl. Acad. Sci. USA* 99, 4135–4137 (2002)
10. Shaw, P., Lalonde, F., Lepage, C., Rabin, C., Eckstrand, K., Sharp, W., Greenstein, D., Evans, A., Giedd, J.N., Rapoport, J.: Development of cortical asymmetry in typically developing children and its disruption in attention-deficit/hyperactivity disorder. *Arch. Gen. Psychiatry* 66(8), 888–896 (2009),
<http://dx.doi.org/10.1001/archgenpsychiatry.2009.103>
11. Tustison, N.J., Gee, J.C.: N4ITK: Nick's N3 ITK implementation for MRI bias field correction. *Insight Journal* (2009)
12. Whitwell, J.L., Jack, C.R., Pankratz, V.S., Parisi, J.E., Knopman, D.S., Boeve, B.F., Petersen, R.C., Dickson, D.W., Josephs, K.A.: Rates of brain atrophy over time in autopsy-proven frontotemporal dementia and Alzheimer disease. *Neuroimage* 39(3), 1034–1040 (2008),
<http://dx.doi.org/10.1016/j.neuroimage.2007.10.001>
13. Whitwell, J.L., Jack, C.R., Parisi, J.E., Knopman, D.S., Boeve, B.F., Petersen, R.C., Ferman, T.J., Dickson, D.W., Josephs, K.A.: Rates of cerebral atrophy differ in different degenerative pathologies. *Brain* 130(Pt. 4), 1148–1158 (2007),
<http://dx.doi.org/10.1093/brain/awm021>
14. Yushkevich, P.A., Avants, B.B., Das, S.R., Pluta, J., Altinay, M., Craige, C., Initiative, A.D.N.: Bias in estimation of hippocampal atrophy using deformation-based morphometry arises from asymmetric global normalization: an illustration in adni 3 t mri data. *Neuroimage* 50(2), 434–445 (2010)

Intra-Patient Supine-Prone Colon Registration in CT Colonography Using Shape Spectrum

Zhaoqiang Lai¹, Jiayi Hu¹, Chang Liu¹, Vahid Taimouri¹, Darshan Pai¹,
Jiong Zhu², Jianrong Xu², and Jing Hua¹

¹ Computer Science, Wayne State University, Detroit, MI 48202, USA

² Radiology, Renji Hospital, Shanghai JiaoTong University, 200127, China
jinghua@wayne.edu

Abstract. CT colonography (CTC) is a minimally invasive screening technique for colorectal polyps and colon cancer. Since electronic colon cleansing (ECC) cannot completely remove the presence of pseudo-polyps, most CTC protocols acquire both prone and supine images to improve the visualization of the lumen wall and to reduce false positives. Comparisons between the prone and supine images can be facilitated by computerized registration between the scans. In this paper, we develop a fully automatic method for registering colon surfaces extracted from prone and supine images. The algorithm uses shape spectrum to extract the shape characteristics which are employed as the surface signature to find the correspondent regions between the prone and supine lumen surfaces. Our experimental results demonstrate an accuracy of 12.6 ± 4.20 mm over 20 datasets. It also shows excellent potential in reducing the false positive when it is used to determine polyps through correspondences between prone and supine images.

Keywords: Colon Registration, Shape Spectrum, CT Colonography, Virtual Colonoscopy.

1 Introduction

Computed tomography colonography (CTC) has received increasing attention as a minimally invasive method for the examination of the colon [1]. CTC has shown promising results in the detection of clinically significant polyps. Using the advanced image technique, doctors can look for polyps throughout the entire colon via fly-through in a virtual colon model which is constructed from patient's abdominal images. In the last decade, many computer aided detection and diagnosis systems have been proposed and actively studied to improve the performance and reliability. In order to better differentiate polyps from pseudo-polyps, and to better view the lumen surface in the presence of fluid, it is common practice to obtain two CT scans of the patient, one with the patient in the prone position and one in the supine. Based on this setup, the registration between the supine and prone colons is required.

Although many methods have been developed for shape registration [2,3], relatively fewer methods have been presented for registration of the supine-prone colons. Li et al. [4] have developed methods to map candidate polyps

between the supine and prone colons using their colon wall positions relative to the colon centerlines. This kind of method only takes into account the local extrema located on the inferior/superior axis which is not reliable due to colon shifts obliquely when the patient changes position. In such a case, some local extrema may not be considered. Nain et al. [5] presented a similar approach for aligning data along the centerlines used dynamic programming. Suh and Wyatt [6] used a piece-wise centerline matching algorithm and an interpolation and extrapolation method of deformation field for deformable registration. However, the deformable model requires the good initial alignment of the two colons and strict constraints. Otherwise, it might create artifacts which, in the worst case, result in pseudo-polyps. Besides the methods based on the centerlines, Näppi et al. [7] developed a directional region growing method for reducing false positive (FP) based on correspondence between the supine and prone datasets.

In this work, we developed a novel method for registration between the supine and prone datasets and reduced the FP based on their correspondence. Without matching the centerlines, we employed shape spectrum to extract the shape characteristics as the signature to find the correspondent regions between the prone and supine lumen surface. The method is simple yet efficient and accurate. Our contributions in this paper can be summarized as follows:

- A novel surface registration method based on shape spectrum is invented and applied to colon surface registration.
- We have applied the algorithm to the real CTC datasets and the experiments demonstrate the excellent accuracy of our registration results.
- We integrate the registration component into the virtual colonoscopy (VC) software for FP reduction, which shows that it is an excellent tool aiding polyp diagnosis in CTC.

2 Methods

Firstly, we discuss how to calculate the shape spectrum on the colon lumen surface to extract the shape characteristics which can be used as the signature for registration purpose. Next we detail how to apply the shape spectrum to the colon registration problem.

2.1 Laplacian Shape Spectrum

For a manifold \mathcal{M} , let Δ denote its Laplace-Beltrami differential operator [8]. Consider the Laplacian eigenvalue equation

$$\Delta\phi = -\lambda\phi, \quad (1)$$

where λ is a real scalar which is called an eigenvalue of Δ and the ϕ is called an eigenvector corresponding to λ . The spectrum is defined to be the eigenvalues arranged increasingly as

$$\lambda_0 = 0 < \lambda_1 < \lambda_2 < \dots < \lambda_i < \dots < +\infty. \quad (2)$$

The Laplace-Beltrami operator is Hermitian, so the eigenvectors corresponding to its different eigenvalues are orthogonal: $\langle \phi_i, \phi_j \rangle = \int_{\mathcal{M}} \phi_i \phi_j = 0$, where $i \neq j$.

Eq. (3) can be solved by the finite element method [9]. For colon surfaces, which are discrete triangle meshes, a discrete Laplace-Beltrami operator K can be applied on it:

$$K(p_i) = \frac{1}{2A_i} \sum_{p_j \in N_1(p_i)} (\cot\alpha_{ij} + \cot\beta_{ij})(p_i - p_j), \tag{3}$$

where $N_1(p_i)$ includes all the vertices, p_j , belonging to one ring neighborhood of a vertex, p_i . α_{ij} and β_{ij} are the two angles opposite to the edge in the two triangles sharing the edge i, j , and A_i is the Voronoi region area of p_i .

For the whole vertices of a triangle mesh, a Laplace-Beltrami matrix can be constructed as:

$$L_{ij} = \begin{cases} -\frac{\cot\alpha_{ij} + \cot\beta_{ij}}{2A_i} & \text{if } i, j \text{ are adjacent,} \\ \sum_{k \in N_1(p_i)} \frac{\cot\alpha_{ik} + \cot\beta_{ik}}{2A_i} & \text{if } i=j, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

where α_{ij} , β_{ij} , and A_i are the same as in Eq. (3) for certain i and j . Then, Eq. (4) turns into the following eigenvalue problem:

$$L\vec{v} = \lambda\vec{v}, \tag{5}$$

where \vec{v} is an n -dimensional vector. Each entry of \vec{v} represents the function value at one of n vertices on the mesh. The equation above can be represented as a generalized eigenvalue problem which is much easier to solve numerically by constructing a sparse matrix M and a diagonal matrix S such that

$$M_{ij} = \begin{cases} -\frac{\cot\alpha_{ij} + \cot\beta_{ij}}{2} & \text{if } i, j \text{ are adjacent,} \\ \sum_{k \in N_1(p_i)} \frac{\cot\alpha_{ik} + \cot\beta_{ik}}{2} & \text{if } i=j, \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

and $S_{ii} = A_i$. Thus, the matrix L is decomposed as $L = S^{-1}M$ and the generalized eigenvalue problem is presented as:

$$M\vec{v} = \lambda S\vec{v}. \tag{7}$$

As defined above, M is symmetric and S is symmetric positive-definite.

The second eigenvector of the Laplacian is called the Fiedler vector and has interesting properties, making it a good permutation vector for numerical computations [10]. Fig. 1(a) and (b) show that it naturally follows the shape of the colon mesh. In other words, the Fiedler vector defines a (1-dimensional) embedding of the surface mesh. We are trying to use the embedding for the registration of the colon surfaces. In next section we will detail how to do the registration.

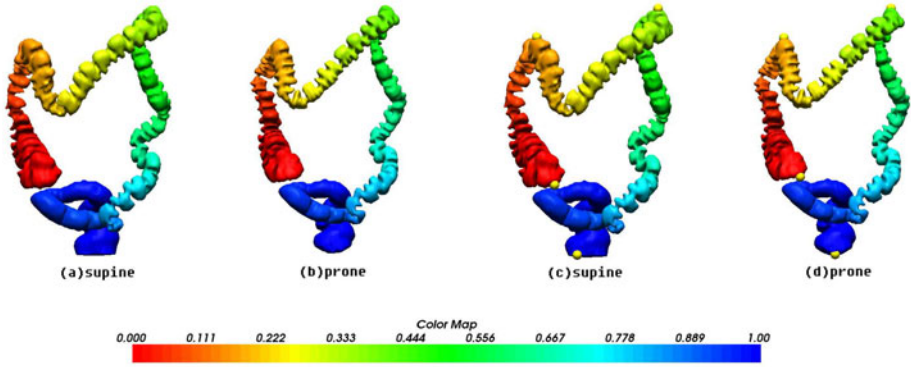


Fig. 1. Illustration of Fiedler vector embedding and landmarks detection. The Fiedler vector gives a natural ordering of the vertex of the colon meshes in (a) and (b). The vector value has been normalized and color map. (c) and (d) show the reliable automatic detected landmarks which are rendering in yellow balls.

2.2 Registration Using Shape Spectrum

First of all, after we obtained the colon surface mesh, we calculated the shape spectrum of the colon mesh. By constructing matrix M and S in Section 2.1, we solved the Eq. 7 to get the eigenvalues and eigenvectors.

To perform the registration efficiently, we first detect reliable anatomical landmarks. Based on the knowledge of colon anatomy, we expect that the mobility of the colon is at its smallest in these landmark regions. Four landmarks are established here: cecum, hepatic flexure, splenic flexure and anus. By using the normalized Fiedler vector value (F_v), the cecum and anus are detected by determination of the smallest and largest F_v . This is done for both supine and prone datasets. For hepatic flexure and splenic flexure, we first process the supine dataset, and then deal with the prone one. For supine dataset, the hepatic flexure is found by detecting the local maximum z-coordinate whose F_v is near 0 and the splenic flexure is found by detecting the local maximum z-coordinate whose F_v is near 1. For prone dataset, we set the neighborhood using the F_v of hepatic flexure in supine as $[F_v - \varepsilon, F_v + \varepsilon]$. At this interval in prone dataset, we detect the local maximum z-coordinate as the hepatic flexure in prone. The splenic flexure in prone is detected with the same strategy. Fig. 1 (c) and (d) show the landmarks both in supine and prone datasets. After this procedure, we have landmarks in sequences: $L_1^S, L_2^S, L_3^S, L_4^S$ represent the cecum, hepatic flexure, splenic flexure and anus, respectively, in supine dataset and $L_1^P, L_2^P, L_3^P, L_4^P$ in prone dataset.

With the landmarks, we register the colon surface using piecewise registration. For each segment in supine between L_i^S and L_{i+1}^S , we will map it to the segment in prone between L_i^P and L_{i+1}^P , $i = 1, 2, 3$. Let

$$F_v^S(L_i^S), F_v^S(L_{i+1}^S) \quad \text{and} \quad F_v^P(L_i^P), F_v^P(L_{i+1}^P) \quad (8)$$

represent the Fiedler vector value at the supine and prone location of L_i^S , L_{i+1}^S and L_i^P , L_{i+1}^P , respectively. Then, for an examined location L_e^S in supine between L_i^S and L_{i+1}^S , the corresponding location L_e^P in prone should have this relation of the F_v :

$$\frac{F_v^S(L_e^S) - F_v^S(L_i^S)}{F_v^S(L_{i+1}^S) - F_v^S(L_i^S)} = \frac{F_v^P(L_e^P) - F_v^P(L_i^P)}{F_v^P(L_{i+1}^P) - F_v^P(L_i^P)}. \quad (9)$$

It is easy to deduce that

$$F_v^P(L_e^P) = F_v^P(L_i^P) + \frac{(F_v^S(L_e^S) - F_v^S(L_i^S)) * (F_v^P(L_{i+1}^P) - F_v^P(L_i^P))}{F_v^S(L_{i+1}^S) - F_v^S(L_i^S)}. \quad (10)$$

Then, the corresponding location in prone is

$$L_e^P = (F_v^P)^{-1}(F_v^P(L_e^P)), \quad (11)$$

where $(F_v^P)^{-1}$ is a mapping to find the locations according to the F_v^P .

3 Experiments and Results

3.1 Data Acquisition and Pre-processing

Each patient was limited to the low-fibre diet beginning 1 day before the scheduled morning CTC. Colonic catharsis was achieved with mannitol on the evening before the examination. Note that, this preparation procedure is close to the prepress VC procedure and is pleasantly accepted by recruited patients. Before the examination, the colon was distended with 1500mL of water-soluble iodinated contrast medium using a manual insufflators with a small rectal catheter. Examinations were performed in supine positions on a 128-MDCT scanner (SO-MATOM Definition AS, SIEMENS, Germany). CT technique consisted of 5.00-mm collimation, 1.375 : 1 pitch, 1-mm reconstruction interval, 120 kVp, and 50-100 mAs. Twenty cases are used for the experiment and all cases were of diagnostic quality, contain mainly fluid, but could contain feces as well. 10 patients (50%) had 20 confirmed polyps: 5 polyps were 5-10 mm, 12 polyps were 3-4 mm and 3 polyp was under 3 mm. 10 patients (50%) were normal.

In the pre-processing step, the colonic lumens are extracted from the CTC datasets by use of a level-set segmentation method [11]. Actually, any other colon segmentation technique could potentially be used for the pre-processing step. Colon lumen surface is extracted after the segmentation.

3.2 Experimental Results

Algorithms described in Sec. 3.1 and Sec. 2.2 are performed to process all the datasets. In order to evaluate the performance of the registration we proposed, a doctor experienced in the interpretation of CTC cases evaluated the 2D and 3D visualizations of the colon by use of our software tool plugged into the VC software platform by Yulon Medical Imaging Company (Fig. 2). Our software

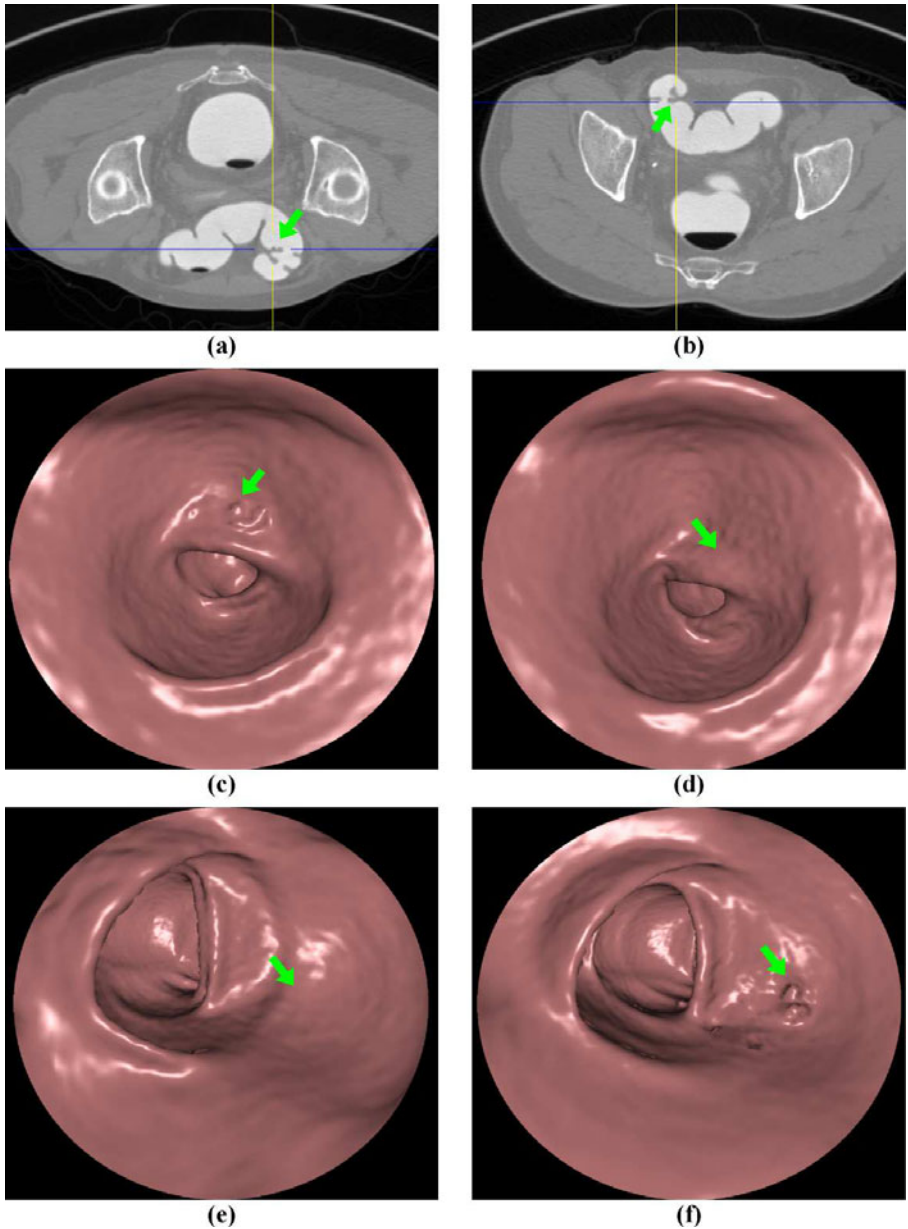


Fig. 2. Illustration of false positive reduction. In (a), in supine position it has a false positive which the green arrow pointing at. In (b), in prone position it confirms that it is a pseudo-polyp but not a true polyp. (c) and (d) is the virtual colonoscopy of (a) and (b). (e) and (f) show another example which has false positives in prone position and can be confirmed as pseudo-polyps in supine position.

Table 1. Results of comparison between shape spectrum method and deformable model

Method	Accuracy(mm)
Shape Spectrum Method	12.6 ± 4.20
Deformable Model	15.4 ± 6.30

tool and algorithm are implemented with C++. The tool displays the 3D VC of two intra-patient colon scans: supine position on the left and prone position on the right. By clicking on either 3D colon lumen, it finds the correspondence region in the other colon lumen, as well as the views of 2D CTC updated simultaneously to show the corresponding location in both colon scans. A doctor can check any location he feels interested in and our system automatically updates the corresponding regions in an opposite position. The application of the method resulted in a registration accuracy of 12.6 ± 4.20 mm over 20 datasets. We compared our method with the deformable model method [6] based on our datasets. The comparison result is shown in Table. 1. Our method outperforms the deformable model approach, and more importantly, our method does not create any artifacts resulting in pseudo-polyps but the deformable model would.

We integrate our method to the VC software for FP reduction. First of all, the polyp candidates are selected by asking doctors to go through all the datasets to find any polyp-like protrusions. 78 polyp candidates from all the CTC datasets are presented, among which 20 polyps are true positives and 58 polyps are false positives. Since it is impossible for any registration algorithm to find exact point-to-point correspondence, in practices, it is common to set a interval, F_v , where the correspondent locations fall in. Specifically, the correspondent interval locations $Interval^P(L_e^S)$ in prone which are the correspondent locations to L_e^S in supine would be

$$Interval^P(L_e^S) = [(F_v^P)^{-1}(F_v^P(L_e^P) - \varepsilon), (F_v^P)^{-1}(F_v^P(L_e^P) + \varepsilon)]. \quad (12)$$

Then we check the polyps candidates in supine position and the correspondent interval locations in prone; and check the polyps candidates in prone position and the correspondent interval location in supine as well. Fig. 2 shows examples of FP reduction. By checking the correspondence, 48 FPs are spotted out, which is 83% reduction compared with independent processing of the datasets, with no true-positive is eliminated.

4 Conclusion

We have developed a novel automatic method for registration of the supine and prone CTC datasets. The application of the method resulted in an accuracy of 12.6 ± 4.20 mm over 20 datasets with the number of false positive reduced by 83% compared with independent processing of the datasets. The experimental results indicate that the method is useful in improving the specificity of the polyps in CTC. In the feature work, we will combine with advanced classification methods for further false positive reduction.

Acknowledgements

This work is supported in part by the National Science Foundation research grants awarded to Dr. J. Hua, including NSF IIS-0915933, NSF IIS-0937586 and NSF IIS-0713315. We would also like to thank Yulonn Medical Imaging Co., Ltd. for providing us the data and the Virtual Colonoscopy software platform.

References

1. Vining, D.J.: Virtual colonoscopy. *Gastrointestinal Endosc. Clin. N. Am.* 7(2), 285–291 (1997)
2. Zou, G., Hua, J., Muzik, O.: Non-rigid Surface Registration Using Spherical Thin-plate Splines. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part I. LNCS*, vol. 4791, pp. 367–374. Springer, Heidelberg (2007)
3. Duan, Y., Hua, J., Qin, H.: Interactive Shape Modeling Using Lagrangian Surface Flow. *The Visual Compute* 21(5), 279–288 (2005)
4. Li, P., Napel, S., Acar, B., Paik, D.S., Jeffrey, R.B., Beaulieu, C.F.: Registration of central paths and colonic polyps between supine and prone scans in computed tomography colonography: Pilot study. *Med. Phys.* 31(10), 2912–2923 (2004)
5. Nain, D., Haker, S., Grimson Jr., W.E.L., Wells, W.M., Ji, H., Kikinis, R., Westin, C.-F.: Intra-patient Prone to Supine Colon Registration for Synchronized Colonoscopy. In: Dohi, T., Kikinis, R. (eds.) *MICCAI 2002. LNCS*, vol. 2489, pp. 573–580. Springer, Heidelberg (2002)
6. Suh, W.J., Wyatt, C.L.: Deformable registration of supine and prone colons using centerline analysis. In: *4th IEEE International Symposium on Biomedical Imaging*, pp. 708–711 (2007)
7. Näppi, J., Okamura, A., Frimmel, H., Dachman, A., Yoshida, H.: Region-based Supine-prone Correspondence for the Reduction of False-positive CAD Polyp Candidates in CT Colonography. *Academic Radiology* 12(6), 695–707 (2005)
8. Rustamov, R.M.: Laplace-Beltrami eigenfunctions for deformation invariant shape representation. In: *Symposium on Geometry Processing*, pp. 225–233 (2007)
9. Reuter, M., Wolter, F.E., Peinecke, N.: Laplace-Beltrami spectra as Shape-DNA of surfaces and solids. *Comput. Aided Des.* 38(4), 342–366 (2006)
10. Lévy, B.: Laplace-Beltrami eigenfunctions: towards an algorithm that understands geometry. In: *IEEE International Conference on Shape Modeling and Applications* (2006)
11. Lankton, S., Tannenbaum, A.: Localizing region-based active contours. *IEEE Transactions on Image Processing* 17(11), 2029–2039 (2008)

A Parameterization of Deformation Fields for Diffeomorphic Image Registration and Its Application to Myocardial Delineation

Hua-mei Chen¹, Aashish Goela^{2,3}, Gregory J. Garvin^{2,4}, and Shuo Li^{1,2,5}

Department of

¹ Medical Biophysics

² Medical Imaging

³ University Hospital

UWO, London, Ontario

⁴ St. Joseph's HealthCare, London, Ontario

⁵ GE HealthCare, London, Ontario

Abstract. This study investigates a new parameterization of deformation fields for image registration. Instead of standard displacements, this parameterization describes a deformation field with its *transformation Jacobian* and *curl of end velocity field*. It has two important features which make it appealing to image registration: 1) it *relaxes* the need of an *explicit* regularization term and the corresponding *ad hoc* weight in the cost functional; 2) explicit constraints on transformation Jacobian such as *topology preserving* and *incompressibility* constraints are straightforward to impose in a *unified* framework. In addition, this parameterization naturally describes a deformation field in terms of *radial* and *rotational* components, making it especially suited for processing cardiac data. We formulate diffeomorphic image registration as a constrained optimization problem which we solve with a *step-then-correct* strategy. The effectiveness of the algorithm is demonstrated with several examples and a comprehensive evaluation of myocardial delineation over 120 short-axis cardiac cine MRIs acquired from 20 subjects. It shows competitive performance in comparison to two recent segmentation based approaches.

Keywords: diffeomorphic image registration, topology preserving, incompressibility constraint, cardiac MRI, myocardial delineation.

1 Introduction

Image registration, which consists of establishing point correspondence between two images, is commonly understood as an *ill-posed* problem when stated as the optimization of a similarity metric alone [1]. Without regularization, this may generate multiple physically non-plausible solutions (i.e., rendering tissue folding/tearing in human anatomy studies). Fig. 1 depicts a typical example, where standard gradient descent optimization over displacements was used to minimize the sum of squared difference between the images S (study) and T (template). Fig. 1(c) shows the transformation that generates S from T . Figs. 1(d) and (e) illustrate the mesh folding phenomenon

from the result. In this study, we investigate a new parameterization of deformation fields which can avoid such undesirable phenomenon (cf. the results in Figs. 1(f) and 1(g)), while affording other important features that make it appealing to image registration. Instead of standard displacements, this parameterization describes a deformation field with its *transformation Jacobian determinant* and *curl of end velocity field*. This relaxes the need of an explicit regularization to produce a physically plausible result because smoothness is implicitly embedded in the solution. Thus, empirical tradeoff between the similarity term and the regularization term, which may cause bias [18], is not necessary. Furthermore, ensuring the deformation to be diffeomorphic, which has spawned a significant amount of recent work [13-16], can be accomplished by directly requiring the transformation Jacobian to be positive. Also, other desirable constraints can be enforced within the *same* framework using explicit restriction on the transformation Jacobian; for instance, the incompressibility constraint requires the Jacobian to be equal to one. In addition, our parameterization naturally describes a deformation field in terms of *radial* and *rotational* components, making it especially suited for processing cardiac data [12]. Fig. 2 illustrates this; it depicts four transformations generated from different pairs of transformation Jacobian m and curl g around the grid center. Obviously, the transformation Jacobian m and curl g are directly related to the radial and rotational components of the deformation field.

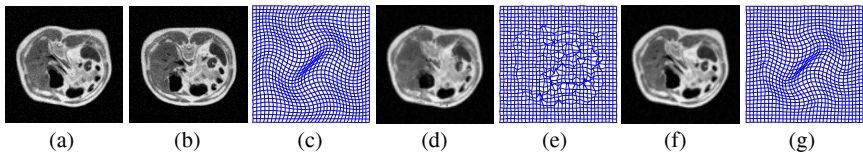


Fig. 1. An illustration of the ill-posed registration problem using conventional deformation parameterization. (a) Study, (b) template, (c) ground truth, (d) deformed template obtained from standard gradient descent optimization over displacements, (e) the deformation field corresponding to (d), (f) deformed template obtained from the proposed parameterization, (g) the deformation field corresponding to (f).

Our parameterization is based on a moving mesh method, originally designed to generate a grid suitable for solving partial differential equations numerically [2-3]. It naturally leads to a formulation of diffeomorphic image registration as a constrained optimization problem which we solve with a *step-then-correct* strategy. Such strategy has been adopted in the Demons algorithm, where unconstrained optimization is followed by Gaussian filtering to impose a smoothness constraint.

The effectiveness of the method is demonstrated with examples of diffeomorphic image registration, registration with an incompressibility constraint, and a comprehensive evaluation of myocardial delineation over 120 short-axis cardiac cine MRIs acquired from 20 subjects. Myocardial delineation is acknowledged as a challenging problem due to large deformations, and has been commonly tackled by segmentation based techniques [6-9]. Our results show favorable performance in comparison to two recent segmentation based approaches. In addition to myocardial delineation, our algorithm yields dense cardiac motion field which is very useful in the diagnosis of cardiovascular diseases [11].

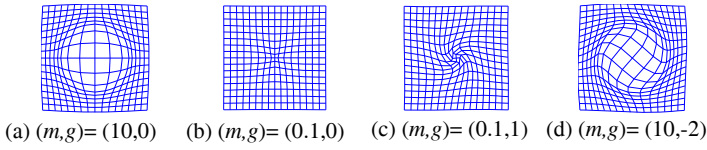


Fig. 2. Geometrical interpretation of the parameterization with (m, g) . It is clear that the transformation Jacobian m and curl g are directly related to the radial and rotational components of the deformation field.

2 Moving Mesh Grid Generation

Our parameterization method is based on the next grid generation theorem [2,3].

Problem 1 For a given continuous monitor function $m(\xi) : \Omega \rightarrow \mathbb{R}$ constrained by

$$\int_{\Omega} m = |\Omega|, \tag{1}$$

find a transformation $\varphi : \Omega \rightarrow \Omega, \partial\Omega \rightarrow \partial\Omega$, such that the transformation Jacobian is equal to the given monitor function m . That is,

$$J_{\varphi}(\xi) = m(\xi). \tag{2}$$

The following theorem provides a solution to Problem 1 [2,3].

Theorem 1 A transformation φ constructed by the following steps satisfies eq. (2).

Step 1: Find a vector field $\eta(\xi)$ which satisfies:

$$\text{div } \eta(\xi) = m(\xi) - 1 \tag{3}$$

with the Neumann boundary condition.

Step 2: Construct an artificial-time dependent velocity vector field via $\eta(\xi)$ by

$$v_i(\xi) = \frac{\eta(\xi)}{t + (1-t)m(\xi)}, t \in [0,1] \tag{4}$$

Step 3: The transformation φ in Problem 1 is the solution of the following ODE evaluated at $t = 1$,

$$\frac{d\psi(\xi, t)}{dt} = v_i(\psi(\xi, t)), t \in [0,1], \tag{5}$$

with $\psi(\xi, t = 0) = \xi$. That is, $\varphi(\xi) = \psi(\xi, t = 1)$.

Proposed parameterization. The solution to Problem 1 is not unique. Generating a unique solution entails constraining the curl of the vector field $\eta(\xi)$, which is the end velocity field $v_i(\xi)$ according to (4), and solving the div-curl system under the Dirichlet boundary condition [4]. We replace eq. (3) by the following div-curl system,

$$\begin{cases} \operatorname{div} \eta(\xi) = m(\xi) - 1 \\ \operatorname{curl} \eta(\xi) = g(\xi) \end{cases} \quad (3')$$

with null boundary condition $\eta(\xi) = 0 \forall \xi \in \partial\Omega$ (i.e., stationary boundary), where $g(\xi)$ is a continuous function over Ω . Hence, a transformation generated by the steps in Theorem 1 can be parameterized with its transformation Jacobian, denoted by $m(\xi)$ ($= J_\varphi(\xi)$), and the curl of the corresponding end velocity field, denoted by $g(\xi)$ ($= \nabla \times \eta(\xi)$). We note that the solution in (3') is always one degree smoother than m and g .

3 Constrained Diffeomorphic Image Registration

With the above parameterization, we formulate diffeomorphic image registration as the following constrained optimization problem.

Problem 2. Given two images S and T , defined over $\Omega \subset \mathbb{R}^2$, find the function pair $(m(\xi), g(\xi))$, $\xi \in \Omega$, that optimizes a similarity measure $E_{Sim}(S, T, \varphi_{m,g})$ between S and T , subject to the constraints:

$$\begin{cases} \int_{\Omega} m = |\Omega| \\ th_{high} > m(\xi) > th_{low}, \quad \xi \in \Omega' \subset \Omega \end{cases} \quad \begin{matrix} (6a) \\ (6b) \end{matrix}$$

where $th_{low} > 0$ ensuring that $\varphi_{m,g}$ is a diffeomorphism.

Note that when $th_{low} \approx th_{high}$ (e.g., $th_{low} = 0.99$ and $th_{high} = 1.01$), inequality (6b) effectively imposes the *incompressibility* constraint in a *sub-region* Ω' of the image domain Ω , which is not always possible with existing methods [17]. The parameterization of φ with m and g is indicated by its subscripts. In the following, we formulate our algorithm for any E_{Sim} , but in the experiments, we use SSD. Next, we present a *step-then-correct* optimization strategy to solve Problem 2.

Algorithm 1. Given an image pair S (study) and T (template) and th_{low} and th_{high} , consider the following steps.

Step 1. Compute unconstrained gradients $\nabla_m E_{Sim}(S, T, \varphi_{m^i, g^i})$ and $\nabla_g E_{Sim}(S, T, \varphi_{m^i, g^i})$.

Step 2. a. Terminate if step size $\delta < \delta_{th}$ or the maximum iteration is reached.

b. Update (m, g) by

$$m^{i+1} = m^i + \delta \cdot \frac{\nabla_m E_{Sim}}{\max |\nabla_m E_{Sim}|} \quad \text{and} \quad g^{i+1} = g^i + \delta \cdot \frac{\nabla_g E_{Sim}}{\max |\nabla_g E_{Sim}|}.$$

Step 3. a. For each pixel location $\xi \in \Omega' \subset \Omega$, impose constraint (6b) by

$$m^{i+1}(\xi) \leftarrow \max(m^{i+1}(\xi), th_{low}) \quad \text{and} \quad m^{i+1}(\xi) \leftarrow \min(m^{i+1}(\xi), th_{high}).$$

b. For each pixel location $\xi \in \Omega$, impose constraint (6a) by

$$m^{i+1}(\xi) \leftarrow m^{i+1}(\xi) \cdot \frac{|\Omega|}{\sum_{\xi \in \Omega} m^{i+1}(\xi)}.$$

Step 4. Use Theorem 1 with (3') to compute $\varphi_{m^{i+1}, g^{i+1}}$ and update E_{Sim} . If it improves, $i \leftarrow i+1$, go to Step 1; otherwise, decrease δ and go to Step 2.

4 Numerical Methods

2D Div-curl solver. We converted div-curl system (3') into two Poisson equations [5]

$$\begin{cases} \Delta \eta^x = m_x - g_y \equiv F^1 & (7a) \\ \Delta \eta^y = m_y + g_x \equiv F^2 & (7b) \end{cases}$$

and used a FFT based Poisson solver to find η efficiently under the null boundary condition. The superscripts denote different components and the subscripts partial derivatives. Euler method with 20 time steps was used in (5) to compute the transformation φ from η via eqs. (4) and (5).

Conversion of unconstrained gradients. Discretizing the image domain and applying the chain rule repeatedly result in

$$\begin{aligned} \frac{\partial E_{Sim}}{\partial m(I_j)} &= \sum_{l \in N(k)} \frac{\partial E_{Sim}}{\partial \varphi^x(I_l)} \frac{\partial \varphi^x(I_l)}{\partial \eta^x(I_l)} \sum_{k \in N(j)} \frac{\partial \eta^x(I_l)}{\partial F^1(I_k)} \frac{\partial F^1(I_k)}{\partial m(I_j)} \\ &+ \sum_{l \in N(k)} \frac{\partial E_{Sim}}{\partial \varphi^y(I_l)} \frac{\partial \varphi^y(I_l)}{\partial \eta^y(I_l)} \sum_{k \in N(j)} \frac{\partial \eta^y(I_l)}{\partial F^2(I_k)} \frac{\partial F^2(I_k)}{\partial m(I_j)}, \end{aligned} \tag{8}$$

and a similar expression for $\frac{\partial E_{Sim}}{\partial g(I_j)}$. In (8), I_j is the j^{th} grid point and the notation $k \in N(j)$ indicates that I_k belongs to some neighborhood of I_j . Each partial derivative term and the size of each neighborhood depend on the actual numerical implementation as explained below.

The terms $\frac{\partial E_{Sim}}{\partial \varphi^x}$ and $\frac{\partial E_{Sim}}{\partial \varphi^y}$ constitute the conventional gradient of a cost function parameterized by displacements, which has been studied extensively. Next, $\frac{\partial \varphi^x}{\partial \eta^x}$ and $\frac{\partial \varphi^y}{\partial \eta^y}$ are approximated by the following Euler 1-step integration

$$\varphi(\xi, 1) = \varphi(\xi, 0) + v_1(\varphi(\xi, 0)) = \xi + \eta(\xi), \tag{9}$$

which leads to $\frac{\partial \varphi^x}{\partial \eta^x} = \frac{\partial \varphi^y}{\partial \eta^y} = 1$. The terms $\frac{\partial \eta^x}{\partial F^1}$, $\frac{\partial \eta^y}{\partial F^2}$ and $N(k)$ are determined by eqs.(7). Since η^x and η^y are the solutions of the Poisson equations $\Delta \eta^x = F^1$ and $\Delta \eta^y = F^2$, the neighborhood of I_k (influence zone of varying $F^1(I_k)$ and $F^2(I_k)$ on η^x and η^y respectively) is the whole image domain and the values $\frac{\partial \eta^x(I_l)}{\partial F^1(I_k)}$ and $\frac{\partial \eta^y(I_l)}{\partial F^2(I_k)}$ are the

values of the response of Δ^{-1} operator at I_l due to an impulse at I_k . Next, $\frac{\partial F^1}{\partial m}$, $\frac{\partial F^2}{\partial m}$ and $N(j)$ are again determined by eqs.(7). Using central finite difference to define the derivatives of m , $N(j)$ can be determined as the 3×3 neighborhood of I_j and the values of $\frac{\partial F^1(I_k)}{\partial m(I_j)}$ and $\frac{\partial F^2(I_k)}{\partial m(I_j)}$ are determined by difference operators D_x and D_y given by

$$D_x = \frac{1}{2} \cdot [1 \quad 0 \quad -1] \text{ and } D_y = D_x^T \tag{10}$$

From the above discussion, we conclude that

$$\begin{cases} \nabla_m E_{Sim} = \Delta^{-1}(\nabla_x E_{Sim}) \otimes D_x + \Delta^{-1}(\nabla_y E_{Sim}) \otimes D_y \end{cases} \tag{11a}$$

$$\begin{cases} \nabla_g E_{Sim} = \Delta^{-1}(\nabla_x E_{Sim}) \otimes -D_y + \Delta^{-1}(\nabla_y E_{Sim}) \otimes D_x \end{cases} \tag{11b}$$

with \otimes denoting the discrete convolution operator and Δ^{-1} the inverse Laplacian operator (Poisson solver).

5 Experimental Results

Diffeomorphic Registration. We first summarize in Table 1 the complete results of the experiment given in Fig. 1. When no noise was added, no mesh folding was observed when the diffeomorphic constraint (DC) was not imposed (e.g., set $(th_{low}, th_{high}) = (-20, 20)$). Results shown in Fig. 1 correspond to this case. However, with Gaussian noise (std = 0.05) added, mesh folding occurred as indicated by the negative Jacobian and the large maximum error in Table I. Setting $(th_{low}, th_{high}) = (0.51, 1.34)$, which were the ground truth values from the deformation shown in Fig. 1(c), not only effectively prevented the mesh from folding, but also reduced registration errors. This suggests that when a proper *prior* knowledge of Jacobian is available, more *reliable and accurate* registration result can be attained. Notice that the min and max transformation Jacobians obtained in this case (i.e., 0.49 and 1.37 respectively) were close to the pre-set values 0.51 and 1.34. No significant differences were observed by setting different (th_{low}, th_{high}) values (e.g., 0.2 and 2).

Table 1. Numerical result of the diffeomorphic registration experiment

	ssd	max error (in pixel)	mean error (in pixel)	max Jacobian	min Jacobian
<i>before registration</i>	3789	7.80	4.75	1.34 (ground truth)	0.51 (ground truth)
<i>noiseless, w/o DC</i>	29	3.17	0.52	1.98	0.06
<i>noisy, w/o DC</i>	250	11.47	0.75	2.01	-1.16
<i>noisy, with DC</i>	246	3.45	0.70	1.37	0.49

Simulated CE-MRI Experiment. The second experiment demonstrates the effectiveness of the proposed method in imposing incompressibility constraint using two

digital phantoms shown in Figs. 3(a) and (b). They simulated pre- and post contrast enhanced (CE) breast MR images. Two sets of thresholds $(th_{low}, th_{high}) = (0.1, 10)$ and $(0.99, 1.01)$ were used in Algorithm 1 to produce the results in Figs. 3(d),3(e),and Figs. 3(f) and 3(g) respectively. Figs. 3(d) and 3(f) are the absolute difference images after registration, and Figs. 3(e) and (g) show the magnified mesh near the simulated cancerous tissues. As a reference, the absolute difference image before registration is provided in Fig. 3(c). The uniformity of the cell sizes in 3(g) indicates that the incompressibility constraint was properly enforced with $(th_{low}, th_{high}) = (0.99, 1.01)$.

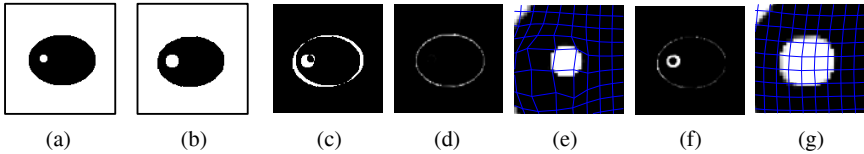


Fig. 3. A simulated contrast enhanced breast MRI experiment to demonstrate the effectiveness of the proposed algorithm in imposing incompressibility constraint

Table 2. Numerical results in terms of Dice Metric (DM) and efficiency

Performance measure	DM (Endo)	DM (Epi)	Average time/frame (sec)
<i>Proposed method</i>	0.93±0.04	0.96±0.01	0.55
<i>Method in [level set]</i>	0.88±0.09	0.94±0.04	6.5
<i>Method in [graph cut]</i>	0.91±0.04	N/A	0.08

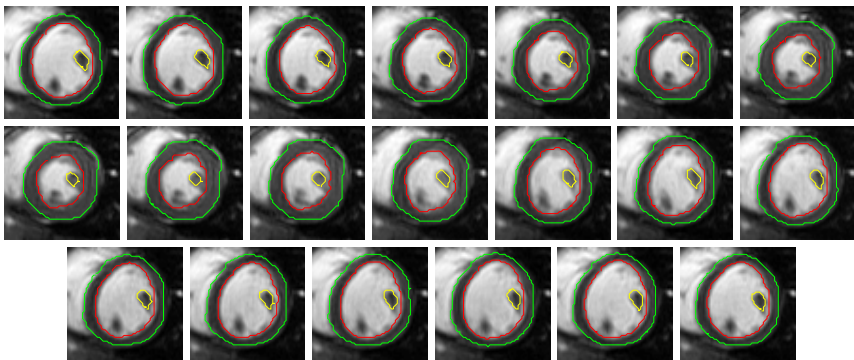


Fig. 4. A complete cardiac cine MRI (left to right, top to bottom) illustrating the tracking of epicardial border (green), endocardial border, and the profile of a papillary muscle (yellow)

Automated Myocardial Borders Delineation. Accurate myocardial borders delineation in cardiac cine MRIs is important in the diagnosis of cardiovascular diseases [6-12]. It yields many significant measures like ejection fraction and radial displacement. In this comprehensive experiment, we used 120 short-axis cardiac cine MRIs clinically

obtained from 20 subjects, including apical, mid-cavity, and basal segments, acquired by 1.5T MRI scanners with fast imaging employing steady state acquisition (FIESTA) mode. For each cine MRI consisting of 20 frames, the first frame was used as the study image S and each of rest frames the template T successively. Each registration result *propagates* to initialize the next registration. As a fair comparison to the segmentation based techniques described in [6-7], our registration based approach did not incorporate a temporal constraint, which can be added by using temporal B-spline basis functions [10], or by recursive Bayesian filtering [11]. The deformation fields obtained were then applied to propagate the manually obtained endo- and epicardium borders from the first frame of each cine MRI. Values of th_{low} and th_{high} were set to be 0.25 and 4 respectively. The maximum number of iteration was fixed at 25 and δ_{th} was set to be 0.01 in all cases. Propagated borders were compared to the ones manually delineated independently by an expert. We also ran the codes described in [6-7] to contrast the accuracies in Table 2, in term of *Dice Metric* (DM), defined as $\frac{2V_{am}}{V_a + V_m}$, where V_a , V_m and V_{am} are the automatically, manually delineated volumes, and the overlap of them respectively. Table 2 shows that our registration based approach yielded the highest *DM* values (higher means closer to manual ones) among the three methods. In addition, our method is very efficient (less than 0.6 sec/frame or 10.5 sec/sequence) which is acceptable for clinical practice. Our platform was HP xw6600 Workstation with 2.83 GHz Intel Xeon CPU with 3.25 GB RAM under Windows XP. All registration codes were written in MATLAB. Although the computational time is higher than the graph cuts based method described in [7], our method produces *dense deformation fields* that can be used to track any features present in the image sequences and provides a means for strain analysis [10], which is not possible using segmentation based approaches. We illustrate this advantage in Fig. 4 by showing a representative cine MRI with not only the automatic delineated epi- and endocardial borders superimposed, but also with the profiles of a papillary muscle (a papillary muscle profile was manually delineated in the first frame). Extension to volume data registration is straightforward, since the underlying moving mesh based grid generation method was originally demonstrated in 3D [2-3].

References

1. Modersitzki, J.: Numerical Methods for Image Registration. Oxford University Press, Oxford (2004)
2. Liao, G., Anderson, D.: A New Approach to Grid Generation. *Applicable Analysis* 44(3), 285–298 (1992)
3. Liu, J.: New Developments of the Deformation Method. PhD dissertation. Department of Mathematics, The University of Texas at Arlington (2006)
4. Zhou, X.: On Uniqueness Theorem of a Vector Function. *Progress in Electromagnetics Research* 65, 93–102 (2006)
5. Jiang, B.-N.: The Least-Squares Finite Element Method: Theory and Applications in Computational Fluid Dynamics and Electromagnetics. Springer, Heidelberg (1998)
6. Ben Ayed, I., Li, S., Ross, I.: Embedding Overlap Priors in Variational Left Ventricular Tracking. *IEEE Transactions on Medical Imaging* 28(12), 1902–1913 (2009)

7. Ben Ayed, I., Punithakumar, K., Li, S., Islam, A.: Left Ventricle Segmentation via Graph Cut Distribution Matching. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5762, pp. 901–909. Springer, Heidelberg (2009)
8. Lynch, M., Ghita, O., Whelan, P.F.: Segmentation of the Left Ventricle of the Heart in 3-D+t MRI Data Using an Optimized Nonrigid Temporal Model. *IEEE Trans. on Medical Imaging* 27(2), 195–203 (2008)
9. Jolly, M.-P.: Automatic Segmentation of the Left Ventricle in Cardiac MR and CT Images. *International Journal of Computer Vision* 70(2), 151–163 (2006)
10. Ledesma-Carbayo, M., et al.: Spatio-Temporal Nonrigid Registration for Ultrasound Cardiac Motion Estimation. *IEEE Transactions on Medical Imaging* 24(9), 1113–1126 (2005)
11. Punithakumar, K., et al.: Heart Motion Abnormality Detection via an Information Measure and Bayesian Filtering. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5762, pp. 373–380. Springer, Heidelberg (2009)
12. Noble, N., et al.: Myocardial Delineation via Registration in a Polar Coordinate System. In: Dohi, T., Kikinis, R. (eds.) MICCAI 2002. LNCS, vol. 2488, pp. 651–658. Springer, Heidelberg (2002)
13. Beg, M., Miller, M., Trounev, A., Younes, L.: Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphism. *International Journal of Computer Vision* 61(2), 139–157 (2005)
14. Rueckert, D., et al.: Diffeomorphic Registration Using B-Splines. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4191, pp. 702–709. Springer, Heidelberg (2006)
15. Ashburner, J.: A Fast Diffeomorphic Image Registration Algorithm. *NeuroImage* 38, 95–113 (2007)
16. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Non-Parametric Diffeomorphic Image Registration with the Demons Algorithm. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part II. LNCS, vol. 4792, pp. 319–326. Springer, Heidelberg (2007)
17. Saddi, K., Ched'hotel, C., Cheriet, F.: Large deformation registration of contrast-enhanced images with volume-preserving constraint. In: *Proc. of SPIE Medical Imaging*, vol. 6512(1) (2007)
18. Rohlfing, T.: Transformation model and constraints cause bias in statistics on deformation fields. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 207–214. Springer, Heidelberg (2006)

Multiple Cortical Surface Correspondence Using Pairwise Shape Similarity

Pahal Dalal¹, Feng Shi², Dinggang Shen², and Song Wang¹

¹ Department of Computer Science and Engineering,
University of South Carolina, Columbia, SC 29208, USA
dalalpk@email.sc.edu, songwang@cec.sc.edu

² Department of Radiology and BRIC,
University of North Carolina, Chapel Hill, NC 27599, USA
fengshi@med.unc.edu, dgshen@med.unc.edu

Abstract. Accurately corresponding a population of human cortical surfaces provides important shape information for the diagnosis of many brain diseases. This problem is very challenging due to the highly convoluted nature of cortical surfaces. Pairwise methods using a fixed template may not handle well the case when a target cortical surface is substantially different from the template. In this paper, we develop a new method to organize the population of cortical surfaces into pairs with high shape similarity and only correspond such similar pairs to achieve a higher accuracy. In particular, we use the geometric information to identify co-located gyri and sulci for defining a new measure of shape similarity. We conduct experiments on 40 instances of the cortical surface, resulting in an improved performance over several existing shape-correspondence methods.

1 Introduction

The cerebral cortex plays a key role in human brain functions such as memory, attention, perceptual awareness, thought, language, and consciousness. Different cortical folding or cortical thickness have been correlated with various brain diseases, including schizophrenia, Alzheimer’s disease, depression, and multiple sclerosis. For identifying the disease-effected cortical regions using neuroimages, it is important to bring all individual cortical surfaces into the common space for comparison (between normal and abnormal groups). To achieve this, one way is to identify the correspondences between different cortical surfaces. Then, the corresponding cortical thickness (or folding features) can be compared at different parts of the brain cortex, thus facilitating the detection of the statistically significant regions that are related to the specific diseases.

Multiple 3D shape correspondence has been studied by many researchers [1,2,4,9]. Some of them correspond each shape instance in the population to a fixed template [1,2] in a pairwise way, which may produce large errors when a target shape instance is substantially different from the template. Other methods consider the entire population simultaneously [4,9,11,12]. Such

groupwise methods usually require complex optimization schemes with high computing complexity and may not be guaranteed to produce the desirable optimal solutions. Specific *brain mapping* models and algorithms have also been developed [5,6,14] to identify corresponded landmarks on cortical surfaces. Most of them are based on geometric and anatomic features and their performances are highly dependent on certain pre-processing steps, such as the accurate extraction of the gyri and sulci.

Recently, Munsell et al. [10] suggested a *shape organization* approach for improving shape correspondence. The basic idea is to organize the shape population into a rooted tree, where each node represents a shape instance and a parent-child pair represents two similar shape instances. Shape correspondence starts from the root and propagates to its children, ending at the leaf nodes. Particularly, in [10], a minimum spanning tree (MST) is constructed, where the tree-edge weight describes the pairwise shape dissimilarity. However, the correspondence error accumulates during the propagation in this approach. While such an accumulation error is not obvious when corresponding simple 2D shape instances [10], it becomes serious in corresponding complex 3D shape instances, such as cortical surfaces, as revealed by the experiments in Section 4. Similar approach of shape organization was also used for the interactive navigation of an image database [8].

In this paper, we develop a new method of shape organization for multiple cortical surface correspondence. As in [10], it organizes the whole shape population into a tree. However, we not only require each parent-child pair to describe similar shape instances, but also control the height of the tree to reduce the accumulation error in the propagation. In this paper, we use the *Freesurfer* software (<http://surfer.nmr.mgh.harvard.edu>) to co-register a pair of cortical surfaces [7], with which we define the pairwise shape similarity and build the correspondence between a parent-child pair for the propagation. In the experiments, we compare the performance of the proposed method to the pairwise method using a fixed template, the groupwise method developed in [11,12], and the MST-based shape-organization method developed in [10].

2 Problem Description

Each cortical surface consists of two hemispheres, which are usually corresponded independently. We represent each hemisphere, or subject, S by (i) *point cloud* extracted from MR images representing the pial surface, (ii) *triangle mesh* constructed from the point cloud, as shown in Fig. 1(a), and (iii) *spherical representation* \tilde{S} , as shown in Fig. 1(c), which is the spherical mapping of the triangle mesh. These can be obtained from a given MR image using the method of [3] available in the *Freesurfer* software. The goal of correspondence is to identify N corresponded landmarks across a population of subjects S_1, S_2, \dots, S_M .

As mentioned above, we use shape organization to facilitate shape correspondence in this paper. In general, it consists of the following steps. First, the given M subjects are organized into a tree, e.g., the ones shown in Fig. 2, where each

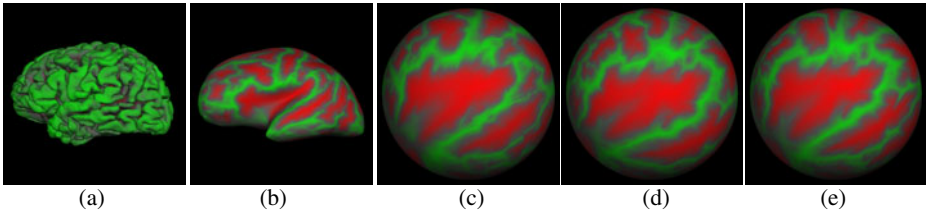


Fig. 1. An illustration of the left hemisphere of cortical surfaces and their co-registration. (a) The pial surface of a subject. (b) The inflated version of the pial surface (a). (c) The spherical mapping, \hat{S} , of (a). (d) Spherical mapping of a second subject. (e) Deformed version of (d) after co-registration with (c).

node represent a subject. Second, we take the subject represented by the root as the template and sample a set of N landmarks on the template. Third, for each child of the root, we treat the subject represented by this node as the target and match the target to the template by identifying N corresponded landmarks on the target. Fourth, each target (with corresponded landmarks) in the previous step is then treated as the new template to match its own child subjects in the tree by identifying the N corresponded landmarks. This pairwise matching process is repeated until propagated to the subjects represented by the leaf nodes, which leads to the final corresponded landmarks across all the subjects.

Clearly, the structure of the constructed tree is important to the performance of the resulting shape correspondence. If we select one subject as the root and set all the other subjects as the children of the root, as shown in Fig. 2(b), the above method is reduced to the widely used pairwise shape correspondence method with a fixed template. The method may lead to large errors when there exist subjects that are substantially different from the template and cannot be accurately corresponded to the template using a pairwise matching method. In [10], an MST is constructed such that each parent-child pair represents very similar subjects. This way, we only need to match very similar subjects, which is relatively simple and can be achieved in high accuracy. However, the constructed MST may have a larger height, as shown in Fig. 2(a), and the correspondence error may accumulate in the propagation and finally affect the consistency of landmarks as revealed in the later experiments.

3 Proposed Method

In this section, we describe a new strategy for constructing the shape-organization tree and apply it for cortical surface correspondence. We focus on addressing the following two problems: (i) measuring the shape dissimilarity between a pair of subjects and (ii) constructing the tree to organize a set of subjects.

3.1 Shape Dissimilarity Using Registration

We use the Freesurfer software to obtain the continuous co-registration between a pair of subjects. Specifically, this software implements the method described in [7]. In registering the subjects S_i and S_j , this method takes as input their spherical representations \tilde{S}_i and \tilde{S}_j and delivers an output spherical representation \tilde{S}'_j , which is a deformed version of \tilde{S}_j on the sphere and is co-registered with \tilde{S}_i . For example, the spherical representation shown in Fig. 1(e) is the deformed version of the subject shown in Fig. 1(d) after co-registration with the subject shown in Fig. 1(c). With this co-registration, we can uniformly sample K co-registered points on both spherical representations \tilde{S}_i and \tilde{S}'_j and define the *dissimilarity* between these two subjects by

$$\delta_{ij} = \sum_{k=1}^K \|c_i(k) - c_j(k)\| \quad (1)$$

where $c_i(k) = 0$ if the curvature value at the k -th sampled point on pial surface i is negative and $c_i(k) = 1$ otherwise. This dissimilarity measure reflects whether the co-registered points from these two subjects show consistent presence of gyri and sulci. Considering the numerical sensitivity involved in computing curvature values, we do not use the exact curvature values in defining this dissimilarity measure.

3.2 Shape Organization Using a Low-Height Tree

In contrast to the MST used in [10], we combine two preferences in constructing the shape-organization tree: (i) the parent-child pair represents similar subjects and (ii) the constructed tree has a low height, which reduces the propagation length and therefore, reduces the accumulation error. We first construct a fully connected graph G , with each node representing a subject and the edge weight between two nodes being the dissimilarity of the subjects represented by these two nodes. We then prune the edges in G sequentially in the descending order of their weights, until one more pruning will make the graph disconnected. From the pruned graph G' , we construct the tree T using the following steps:

1. Find the node with the largest number of incident edges in G' and add it to the initially empty T as the root.
2. Let V_T and \bar{V}_T be the sets of nodes that are in T and not in T respectively.
3. If $\bar{V}_T \neq \emptyset$, for each node in \bar{V}_T with edge links to nodes in V_T in G' , we add this node to T as a child to its linked node in V_T with the smallest edge weight. Note that, in this step, we check every node in \bar{V}_T for possible additions to T without updating V_T and \bar{V}_T . This way, the height of T will only increase by one after all possible node additions in this step.
4. Update V_T and \bar{V}_T and go back to Step 3 until $\bar{V}_T = \emptyset$.

An example of a tree constructed using this method is shown in Fig. 2(c), which will be discussed in more detail in the later experiments.

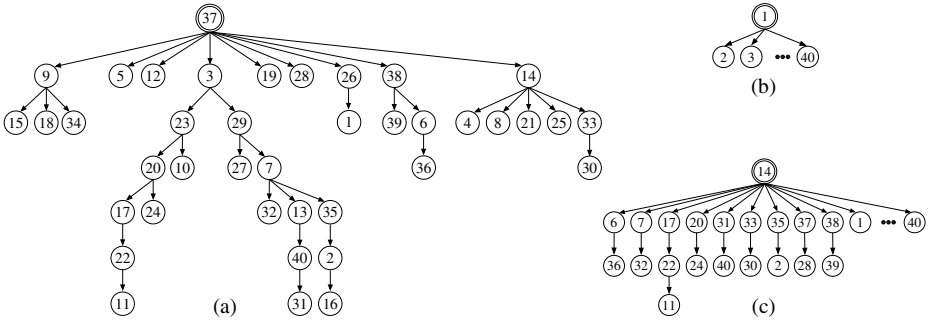


Fig. 2. The shape-organization trees constructed from LONI LPBA40 dataset using (a) the MST-based method in [10], (b) the pairwise method with a fixed template, and (c) the proposed method. In (b) and (c), the ellipsis represents all the unlisted subjects, which are directly linked to the root as its children.

As for the pairwise subject correspondence in the propagation, we use the Freesurfer software to continuously co-register the template and the target subjects. This registration maps the N landmarks on the template to N points on the target and we simply take them as the corresponded landmarks.

4 Experiments

For experiments, we use the LONI Probabilistic Brain Atlas (LPBA40) dataset, which consists of 40 3D MR images of the brain [13]. On each image, 56 anatomic structures have been identified by labeling all the relevant voxels. We extract the cortical surface from these MR images using Freesurfer for all 40 cases and then take the left hemisphere of the extracted cortical surface as the test subjects. For the pairwise shape dissimilarity [11], we uniformly sample the angular-spherical coordinates, together with two poles, to construct $K = 4,952$ co-registered points on each subject.

To evaluate the performance, we use anatomic-structure labels provided with this dataset: we check whether the corresponded landmarks across the population show consistent labels. First, for each identified landmark on each subject, we find its label by searching for the closest labeled voxel in the original image. Second, for each set of corresponded landmarks across the population, e.g., the first landmarks on all 40 subjects, we perform a majority voting to find the label shared by the largest number of subjects. We use this majority label as the true label for these 40 landmarks. This way, for the j -th landmark L_i^j on the subject i , we have its label $r(L_i^j)$ and true label $\hat{r}(L_i^j)$. If $r(L_i^j) \neq \hat{r}(L_i^j)$, we find on image i the closest voxel to L_i^j such that this voxel has a label $\hat{r}(L_i^j)$. We calculate the Euclidean distance between L_i^j and this closest voxel as the error e_i^j for landmark L_i^j . If $r(L_i^j) = \hat{r}(L_i^j)$, we simply set $e_i^j = 0$. We then define the average error on the subject i as $\Delta_i = \frac{1}{N} \sum_{j=1}^N e_i^j$, the average error for the set

of j -th landmarks on all subjects as $\Delta^j = \frac{1}{M} \sum_{i=1}^M e_i^j$, and use these errors to evaluate the correspondence performance.

Figure 2(c) is the low-height tree constructed in our experiment using the proposed method, where subject 14 is chosen as the root and we uniformly sample the angular-spherical coordinates, together with two poles, to construct $N = 4,952$ landmarks on this subject to start the propagation. We also define the total average error as $\Delta = \frac{1}{M} \sum_{i=1}^M \Delta_i = \frac{1}{N} \sum_{j=1}^N \Delta^j$. Figure 3 shows the average errors Δ_i for all 40 subjects and Δ^j for all 4,952 landmarks, with the horizontal lines being the total average error $\Delta = 0.479$.

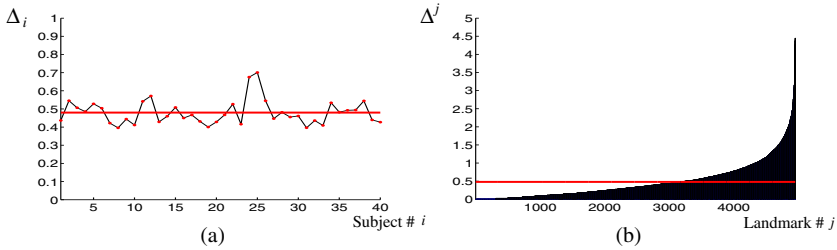


Fig. 3. Average of the error e_i^j in terms of (a) each subject and (b) each set of corresponded landmarks, for the correspondence obtained by the proposed method

To further evaluate the proposed method, we conduct a quantitative comparison with three other methods: (M1) a pairwise method with a fixed template, (M2) the groupwise method developed in [11][12], and (M3) the MST-based shape-organization method developed in [10]. For M1, we choose a subject as the fixed template and then correspond all the other 39 subjects to this template, as illustrated in Fig. 2(b). By selecting different templates, the resulting total average error Δ ranges from 0.509 to 0.684. For the proposed method, the total average $\Delta = 0.479$ (the horizontal lines in Fig. 3) is better than the result from M1 using the best template (subject 29). Note that, in practice, we do not have ground-truth labels and we may not be able to find this best template.

For M2, we downloaded the source code from <http://www.ia.unc.edu/dev/tutorials/InstallLib/index.htm> and also identify $N = 4,952$ landmarks on each subject, by including the *sulcal depth* as an attribute in the similarity metric defined in this code. As shown in Table 1, its resulting total average error $\Delta = 0.791$ is also higher than the proposed method. M2 is a groupwise method and by reading all subjects at once, it takes 11GB of memory in our experiments, while the proposed method takes no more than 2GB of memory.

For M3, we use the same dissimilarity measure in Section 3.1 to construct an MST as shown in Fig. 2(a) and also use uniformly sampled $N = 4,952$ landmarks on the root subject (subject 37) to start the propagation. From Table 1, its resulting total average error $\Delta = 0.532$ is also higher than the proposed method. This is mainly caused by the accumulation error during the propagation. Figure 4

Table 1. Average errors produced by the proposed method and three comparison methods. The variances and standard deviations are calculated over $\Delta_i, i = 1, 2, \dots, M$ and $\Delta^j, j = 1, 2, \dots, N$ respectively. The numbers for M1 come from the best template (subject 29) that leads to smallest total average error Δ .

	In terms of Subject				In terms of Landmark			
	M1(best)	M2	M3	Proposed	M1(best)	M2	M3	Proposed
Total Average (Δ)	0.509	0.791	0.532	0.479	0.509	0.791	0.532	0.479
Variance	0.006	0.073	0.010	0.004	0.333	0.717	0.352	0.266
Standard Deviation	0.082	0.270	0.100	0.067	0.577	0.846	0.594	0.515

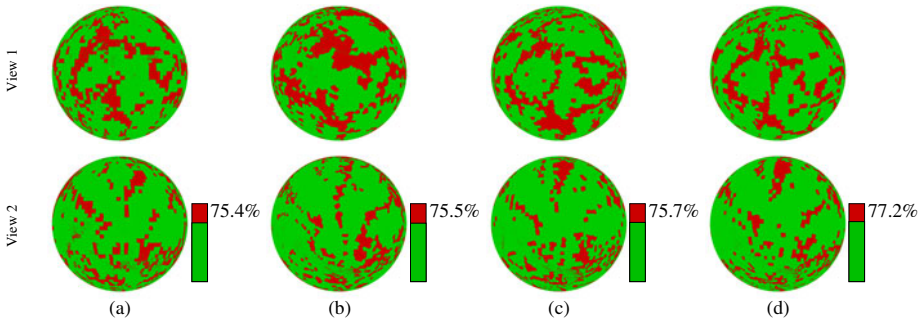


Fig. 4. A visualization of the error e_i^j on the spherical mapping of one subject (subject 18): red area indicates $e_i^j > 0$ and green area indicates $e_i^j = 0$. The top and bottom rows show the views that are the same as and diametrically opposite to the one used in Fig. 3(c), respectively. (a-d) the results from M1, M2, M3, and the proposed method, with green area accounting for 75.4%, 75.5%, 75.7% and 77.2% of the sphere, respectively.

visualizes the error e_i^j for all 4,952 landmarks on subject 18, by using each of the three comparison methods and the proposed method: red area indicates $e_i^j > 0$ and green area indicates $e_i^j = 0$. This is shown on the spherical representation of the subject for clarity. These experiments show that the corresponded landmarks identified by the proposed method show better consistency in terms of the true labels than the other three comparison methods.

5 Conclusion

We have developed a new method for corresponding highly convoluted 3D cortical surfaces. A new shape similarity measure between a pair of cortical surfaces was developed by using Freesurfer co-registration results. The cortical surfaces are then organized into a low-height tree where parent-child pairs represent similar subjects. Multiple shape correspondence was obtained by propagating the pairwise correspondence from the root to the leaves in the constructed tree.

Experiments on 40 LONI LPBA40 images showed that the proposed method produces a better performance than three other existing shape-correspondence methods. The shape dissimilarity measure based on the Freesurfer registration is computationally expensive. In the future, we plan to develop more efficient algorithms to measure such pairwise dissimilarity for shape organization.

Acknowledgement. This work was funded, in part, by AFOSR FA9550-07-1-0250, NSF IIS0951754, NIH EB006733, EB008374, EB009634, and MH088520. Experiments were conducted on a shared memory computer funded by NSF CNS0708391. We would like to thank Ipek Oguz for help with the code of M2.

References

1. Chui, H., Rangarajan, A.: A new algorithm for non-rigid point matching. In: CVPR, pp. 44–51 (2000)
2. Dalal, P., Munsell, B., Wang, S., Tang, J., Oliver, K., Ninomiya, H., Zhou, X., Fujita, H.: A fast 3D correspondence method for statistical shape modeling. In: CVPR (2007)
3. Dale, A., Fischl, B., Sereno, M.: Cortical surface-based analysis I: Segmentation and surface reconstruction. *NeuroImage* 9(2), 179–194 (1999)
4. Davies, R., Twining, C., Cootes, T., Waterton, J., Taylor, C.: A minimum description length approach to statistical shape modeling. *IEEE TMI* 21(5), 525–537 (2002)
5. Essen, D.C.V.: Surface-based approaches to spatial localization and registration in primate cerebral cortex. *NeuroImage* 23, S97–S107 (2004)
6. Fugeras, O., et al.: Variational, geometric, and statistical methods for modeling brain anatomy and function. *NeuroImage* 23, S46–S55 (2004)
7. Fischl, B., Sereno, M., Dale, A.: Cortical surface-based analysis II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage* 9(2), 195–207 (1999)
8. Joshi, S., Horn, J.V., Toga, A.: Interactive exploration of neuroanatomical meta-spaces. *Frontiers in Neuroinformatics* 3 (2009)
9. Kotcheff, A., Taylor, C.: Automatic construction of eigen-shape models by genetic algorithm. In: Duncan, J.S., Gindi, G. (eds.) IPMI 1997. LNCS, vol. 1230, pp. 1–14. Springer, Heidelberg (1997)
10. Munsell, B., Temlyakov, A., Wang, S.: Fast multiple shape correspondence by pre-organizing shape instances. In: CVPR, pp. 840–847 (2009)
11. Oguz, I., Cates, J., Fletcher, T., Whitaker, R., Cool, D., Aylward, S., Styner, M.: Cortical correspondence using entropy-based particle systems and local features. In: ISBI, pp. 1637–1640 (2008)
12. Oguz, I., Niethammer, M., Cates, J., Whitaker, R., Fletcher, T., Vachet, C., Styner, M.: Cortical correspondence with probabilistic fiber connectivity. In: Prince, J.L., Pham, D.L., Myers, K.J. (eds.) IPMI 2009. LNCS, vol. 5636, pp. 651–663. Springer, Heidelberg (2009)
13. Shattuck, D., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K., Poldrack, R., Bilder, R., Toga, A.: Construction of a 3D probabilistic atlas of human cortical structures. *NeuroImage* 39(3), 1064–1080 (2007)
14. Tosun, D., Rettmann, M., Han, X., Tao, X., Xu, C., Resnick, S., Pham, D., Prince, J.: Cortical surface segmentation and mapping. *NeuroImage* 23, S108–S118 (2004)

Cortical Sulcal Atlas Construction Using a Diffeomorphic Mapping Approach

Shantanu H. Joshi¹, Ryan P. Cabeen¹, Bo Sun², Anand A. Joshi¹, Boris Gutman¹, Alen Zamanyan¹, Shruthi Chakrapani¹, Ivo Dinov¹, Roger P. Woods¹, and Arthur W. Toga¹

¹ Laboratory of Neuro Imaging, University of California, Los Angeles CA 90095, USA

² Research Center for Sectional and Imaging Anatomy, Shandong University School of Medicine, 44 Wenhua Xi Road, Jinan 250012, China
sjoshi@loni.ucla.edu

Abstract. We present a geometric approach for constructing shape atlases of sulcal curves on the human cortex. Sulci and gyri are represented as continuous open curves in \mathbb{R}^3 , and their shapes are studied as elements of an infinite-dimensional sphere. This shape manifold has some nice properties – it is equipped with a Riemannian \mathbb{L}^2 metric on the tangent space and facilitates computational analyses and correspondences between sulcal shapes. Sulcal mapping is achieved by computing geodesics in the quotient space of shapes modulo rigid rotations and reparameterizations. The resulting sulcal shape atlas is shown to preserve important local geometry inherently present in the sample population. This is demonstrated in our experimental results for deep brain sulci, where we integrate the elastic shape model into surface registration framework for a population of 69 healthy young adult subjects.

1 Introduction

A surface-based morphometric analysis of the cortex has been shown to have wide reaching applicability for the purpose of mental disease detection, progression, as well as prediction and understanding of normal and abnormal developmental behaviors. Cortical morphometry has three major ingredients: i) surface representation, ii) registration and alignment for construction of atlases, and iii) statistical analysis of deformations or warps explaining the variability of surface features in a given population. Surface registration aims at determining point-to-point correspondences between a pair of surfaces by aligning several homologous features on the two cortical surfaces. These correspondences can be achieved either automatically by using both local and global features as in the case of Dale et al. [2], or in a semi-automated manner, using expertly delineated sulcal and gyral landmarks as in the case of Thompson et al. [10,4]. The underlying idea in both approaches is modeling (either explicitly or indirectly) the sulcal and gyral patterns exclusively based on local geometric features. These features are usually 3D continuous space curves corresponding to the deepest regions of the valleys for sulci, and topmost regions of the ridges for the gyri. The main advantage of using explicit landmarks is the incorporation of expert anatomical knowledge that improves the consistency in matching

of homologous features. This in turn potentially improves statistical power in the neighborhood of the landmarks. Additionally, increasing the number of consistent landmarks also improves the alignment accuracy, thereby allowing more control in the registration process. Previously, landmark curves have mostly been used as boundary conditions for various cortical alignment approaches. Various researchers have modeled the sulci and gyri using different representations. Tao et al. [9] represent sulci using landmark points on curves, and build a statistical model using a Procrustes alignment of sulcal shapes. Vaillant et al. [11] represent cortical sulci by medial surfaces of cortical folds. Although, the advantage of this model is that it represents entire cortical folds, a limitation of this method is the use of unit speed parameterizations of active contours for constructing Procrustes shape averages for sulci. Furthermore, for the two approaches, the shapes are represented by finite features or landmarks, and thus are limited in the characterization of rich geometric detail that manifests in the cortical folds giving the sulci their shapes. Recently, there have been several interesting approaches using continuous representations for sulci [8][13]. For e.g. Auzias et al. [1] model whole sulci using distributions of point sets and use a LDDMM framework for registering not only the surfaces but full MRI volumes. Durrleman et al. [3] use currents for modeling curves and surfaces.

In our work, we represent sulci and gyri by parameterized three-dimensional curves. However, unlike previous approaches, we construct a shape space of such sulcal and gyral curves and build a statistical model of sulci and gyri intrinsically on the shape space. Our approach models the whole curve without the use of landmarks or discrete parametric representations and deals with functional mappings of curve instances on the shape manifold. The main contributions of this paper are as follows: i) an inverse-consistent diffeomorphic framework for matching sulcal shapes, ii) An intrinsic sulcal shape atlas based on the Riemannian metric on the shape manifold, and iii) integration of sulcal curve diffeomorphisms in driving cortical surface registrations. To our knowledge, the proposed framework of direct diffeomorphic three-dimensional sulcal curve mappings have not been used in cortical registration before. This paper is organized as follows. Section 2 outlines the shape modeling scheme for sulci and gyri. It also outlines the procedure for computing statistical shape averages for sulci and gyri for a given population. Section 3 incorporates the sulcal shape model in cortical surface registration, followed by results and conclusion.

2 Diffeomorphic Shape Analysis of Sulci and Gyri

In this section, we describe the modeling scheme used to represent sulcal and gyral shape features. We represent the cortical valleys (sulci), and the ridges (gyri) by open curves. However unlike previous approaches which have used landmarks for representing the sulcal and gyral features, we will use continuous functions of curves for representing shapes.

2.1 Shape Representation

Let β be a 3D, arbitrarily parameterized [5], open curve, such that $\beta : [0, 2\pi] \rightarrow \mathbb{R}^3$. We represent the shape of the curve β by the function $q : [0, 2\pi] \rightarrow \mathbb{R}^3$ as follows,

$$q(s) = \frac{\dot{\beta}(s)}{\sqrt{\|\dot{\beta}(s)\|}} \in \mathbb{R}^3. \tag{1}$$

Here, $s \in [0, 2\pi]$, $\|\cdot\| \equiv \sqrt{(\cdot, \cdot)_{\mathbb{R}^3}}$, and $(\cdot, \cdot)_{\mathbb{R}^3}$ is the standard Euclidean inner product in \mathbb{R}^3 . The quantity $\|q(s)\|$ is the square-root of the instantaneous speed, and the ratio $\frac{q(s)}{\|q(s)\|}$ is the instantaneous direction along the curve. The original curve β can be recovered upto a translation, using $\beta(s) = \int_0^s \|q(t)\| q(t) dt$. In order to make the representation scale invariant, we will normalize the function q as $\mathbf{q} = \frac{q}{\sqrt{(q,q)_{\mathbb{R}^3}}}$. We now

denote $\mathbb{S}^q \equiv \{q|q(s) : [0, 2\pi] \rightarrow \mathbb{R}^3 | \int_0^{2\pi} (q(s), q(s))_{\mathbb{R}^3} ds = 1\}$ as the space of all unit-length, elastic curves. On account of scale invariance, the space \mathbb{S}^q becomes an infinite-dimensional unit-sphere and represents all open elastic curves invariant to translation and uniform scaling. The tangent space of \mathbb{S}^q is easy to define and is given as $T_q(\mathbb{S}^q) = \{w = (w_1, w_2, \dots, w_n) | w(s) : I \rightarrow \mathbb{R}^3 \forall s \in [0, 2\pi] | \int_0^{2\pi} (w(s), q(s))_{\mathbb{R}^3} ds = 0\}$. Here each w_i represents a tangent vector in the tangent space of \mathbb{S}^q . Due to the spherical nature of the shape space, any vector on the shape space can be transformed to a tangent vector by simply subtracting its normal component. We define a metric on the tangent space as follows. Given a curve $q \in \mathbb{S}^q$, and the first order perturbations of q given by $u, v \in T_q(\mathbb{S}^q)$, respectively, the inner product between the tangent vectors u, v to \mathbb{S}^q at q is defined as, $\langle u, v \rangle = \int_0^{2\pi} (u(s), v(s))_{\mathbb{R}^3} ds$. Now given two shapes q_1 and q_2 , the translation and scale invariant shape distance between them is simply found by measuring the length of the geodesic connecting them on the sphere. We know that geodesics on a sphere are great circles and can be specified analytically. Thus given a tangent vector $f \in T_{q_1}(\mathbb{S}^q)$, the geodesic on \mathbb{S}^q between the two points $q_1, q_2 \in \mathbb{S}^q$ along f , for a time t is given by $\chi_t(q_1; f) = \cos(t \cos^{-1}\langle q_1, q_2 \rangle) q_1 + \sin(t \cos^{-1}\langle q_1, q_2 \rangle) f$ where t is a subscript for time. Then the geodesic distance between the two shapes q_1 and q_2 is given by $d(q_1, q_2) = \int_0^1 \sqrt{\langle \dot{\chi}_t, \dot{\chi}_t \rangle} dt$. The quantify $\dot{\chi}_t$ is also referred to as the velocity vector field on the geodesic path χ_t . So far, we have constructed geodesics between a pair of curves directly on the sphere (\mathbb{S}^q). In doing so, we implicitly assumed that the curves were rotationally aligned, as well as the parameterization of the curves was fixed. However the shape of a curve remains unchanged under rotations as well as different parameterizations of the curve. Thus in order to register shapes accurately, the matching should be invariant to rotations as well as reparameterizations. This matching is achieved by constructing the space of elastic shapes, and measuring the ‘‘elastic’’ distance between curves under certain well-defined shape-preserving transformations as explained in the next section.

2.2 Geodesics between Elastic Shapes

In order to match curves elastically, in addition to translation and scaling, we consider the following reparameterizations and group actions on the curve that preserve its shape. A rigid rotation of a curve is considered as a group action by a 3×3 rotation matrix $O_3 \in SO(3)$ on q , and is defined as $O_3 \cdot q(s) = O_3 q(s), \forall s \in [0, 2\pi]$. A curve traveled at arbitrary speeds is said to be reparameterized by a non-linear differentiable map γ (with a differentiable inverse) also referred to as a diffeomorphism.

We define $\mathcal{D} = \{\gamma : \mathbb{S}^1 \rightarrow \mathbb{S}^1\}$ as the space of all orientation-preserving diffeomorphisms. Then the resulting variable speed parameterizations of the curve can be thought of as diffeomorphic group actions of $\gamma \in \mathcal{D}$ on the curve q . This group action is derived as follows. Let q be the representation of a curve β . Let $\alpha = \beta(\gamma)$ be a reparameterization of β by γ . Then the respective velocity vectors can be written as $\dot{\alpha} = \dot{\gamma}\dot{\beta}(\gamma) = \dot{\gamma}q(\gamma)\|q(\gamma)\| = \|\sqrt{\dot{\gamma}}q(\gamma)\|\sqrt{\dot{\gamma}}q(\gamma)$. The reparameterization of q by γ is defined as a right action of the group \mathcal{D} on the set \mathcal{C} and written as $q \cdot \gamma = \sqrt{\dot{\gamma}}(q \circ \gamma)$. Thus we are interested in constructing the shape space as a quotient space of \mathbb{S}^q , modulo shape preserving transformations such as rigid rotations and reparameterizations.

Altogether, the set of curves affected by the group actions above, partition the space \mathbb{S}^q into equivalence classes. We now define the elastic shape space as the quotient space $\mathcal{S} = \mathbb{S}^q / (SO(3) \times \mathcal{D})$. The problem of finding a geodesic between two shapes in \mathcal{S} is same as finding the shortest path between the equivalent classes of the given pair of shapes. Since the actions of the re-parametrization groups on \mathcal{C} constitute actions by isometries, this problem also amounts to minimizing the length of the geodesic path, such that

$$d_e(q_1, q_2) = \min_{O_3 \in SO(3), \gamma \in \mathcal{D}} d(q_1, (O_3 q_2) \cdot \gamma), \tag{2}$$

where d is given by the geodesic distance. In order to optimize Eq. 2 we recognize that for a fixed rotation O_3 , the distance d_e can be obtained by finding the optimal reparameterization $\hat{\gamma}$ between q_1 and q_2 , whereas for a fixed γ , the distance d_e is calculated by finding the optimal rotation \hat{O}_3 . Thus in order to minimize the distance in Eq. 2 we alternate between optimizing over O_3 and γ repeatedly until the process converges. At each step, the optimal rotation \hat{O}_3 is given by $\hat{O}_3 = USV^T$, where, all $U, S, V \in \mathbb{R}^{3 \times 3}$, and given by the singular decomposition of \hat{O}_3 . Furthermore this decomposition is approximated using the \mathbb{L}^2 function given by $USV^T = \int_0^{2\pi} q_1(s)q_2(s)^T ds$. Also, at each iteration, we compute a geodesic path between the starting shape q_1 and the target shape $O_3 q_2 \cdot \gamma$. Upon convergence of this procedure, we also obtain the tangent vector $\dot{\chi}$ along the geodesic path connecting the two shapes.

2.3 Construction of a Statistical Sulcal Atlas

In order to construct a sulcal atlas of a large population of curves in the shape space, we need the notion of a shape average based on the sulcal and gyral curves. Owing to the nonlinearity of the shape space, the computation of an average shape is not straightforward. There are two well known approaches of computing statistical averages in such spaces. The extrinsic shape average is computed by projecting the elements of the shape space in the ambient linear space, where an Euclidean average is computed, and subsequently projected back to the shape space. On the other hand, the intrinsic average, also known as the Karcher mean ([6]) is computed directly on the shape space, and makes use of distances and lengths that are defined strictly on the manifold. It uses the geodesics defined via the exponential map, and iteratively minimizes the average geodesic variance of the collection of shapes. We will adopt the intrinsic approach by computing the Karcher mean for a given set of shapes. The Karcher mean is computed by minimizing the geodesic variance for a given collection of shapes. In other words, given a set of shapes $\{q_i\}, i = 1, \dots, N$, the Karcher mean is given by

$\mu = \operatorname{argmin}_{q_\mu} \sum_{i=1}^N d_e(q_\mu, q_i)^2, i = 1, \dots, N$. This mean is computed by an optimization procedure that involves repeated computations of geodesics from each of the shapes of the population to the current estimate of the mean. Next, we describe the procedure of combining the nonlinear sulcal atlas along with cortical surface registration for mapping brains across populations.

3 Integrating Sulcal Shapes with Cortical Surface Registration

In this section, we introduce our scheme for registering cortical surfaces. There have been several prominent approaches [10,41] for cortical registration in the neuroimaging community. Our method, although based on the same conceptual framework of elastic registration, provides a slightly different model for computing the deformation, with certain improvements in the implementation that increase flexibility and efficiency. A general outline of the process is as follows. The first stage is to establish homology of the curve data, which is accomplished by matching the shapes in a Riemannian shape space framework. Next, the surfaces and curves are conformally mapped to the sphere, establishing a common space where deformation will be defined. Following this, there is a rotational alignment of the surfaces and curves to account for the differences in spherical mapping orientations. Next, the spherical mean of the curves is found to define the atlas curves for the domain of the deformation. Finally, the elastic deformation of the atlas on the sphere is found, which constrained to map the atlas curves to each case's set of curves. The surfaces are reparameterized by this elastic deformation. The resulting surfaces then have homologous coordinate systems, allowing local comparisons across the group. We define a set of N surfaces, $\{M_1, \dots, M_N\}$ where $M_i \subset \mathbb{R}^3$. We represent their mesh geometry using a set of simplicial complexes, $\{(K_1, f_1), \dots, (K_N, f_N)\}$ where K_i is a simplicial complex and $f_i : K \mapsto M_i$. For each surface i , we have a set of M landmarks represented by continuous open curves $\{\beta_{i1}, \dots, \beta_{iM}\}$ where $\beta_{ij} : [0, 2\pi] \mapsto M_i$, and the set of curves $\{\beta_{ij} : i \in [1, N]\}$ represent homologous regions on the set of surfaces. Additionally, the curves are discretized, where the j -th curve has k_j vertices. The first step of the process is to establish correspondence between the homologous landmark curves by computing mappings $\hat{\gamma}_{ij} : [0, 2\pi] \mapsto [0, 2\pi]$ such that for curve j and parameter t , the set $\{\beta_{ij}(\hat{\gamma}_{ij}(t)) : i \in [1, N]\}$ is a set of homologous points on the surfaces. This is accomplished by mapping the curves to a Riemannian manifold shape space, where reparameterizations are defined by geodesics to the Karcher mean of the curves in the shape space.

3.1 Spherical Mapping and Alignment

Next, the meshes are simplified using a QEM-based method, and a set of conformal mappings is found from each surface to the unit sphere, $\{\phi_1, \dots, \phi_N\}$ where $\phi_i : M_i \mapsto S^2$. The spherical mapping of the matched curves is then $\beta_{ij} = \phi_i \circ \beta_{ij} \circ \gamma_{ij}$, which is found using the barycentric coordinates of each curve vertex. A bounded interval hierarchy is used to efficiently search for the coincident face of each curve vertex. Once the data are mapped to the sphere, they are rotationally aligned to enforce a consistent orientation of the spherical mappings. Given an arbitrarily chosen target, each set of curves

is aligned to the target by computing the rotation and reflection that minimizes the least-squared difference between the discretized curve coordinates. This is accomplished by solving the unconstrained orthogonal Procrustes problem using singular value decomposition, allowing reflections to account for the inversion between the hemispheres. For an arbitrary $T \in \{\widehat{\beta}_{ij} : i \in [1, N]\}$, we find an optimal alignment $R_i \in \mathbb{R}^{3 \times 3}$ and then optimally rotate the data as, $\widehat{\phi}_i = R_i \circ \phi_i$, and $\widehat{\beta}_{ij} = R_i \circ \beta_{ij}$.

3.2 Spherical Curve Atlas

Once the data have been aligned on the sphere, the mean curves are computed to serve as the atlas curves in the surface warping. The Karcher mean on the sphere is found for each vertex of each curve. In this method, an initial guess is found by the normalized average of the points. For this point, the tangent space is defined by the gnomonic projection. A new mean is computed in the tangent space and then is mapped back to the sphere, repeating this process until convergence. We can express the curve atlas as the set $\{\overline{\beta}_j : j \in [1, M]\}$, where $\overline{\beta}_j$ is the karcher mean of $\{\widehat{\beta}_{ij} : i \in [1, N]\}$.

3.3 Elastic Surface Warping

For surface i , the deformation of the atlas is $\phi_i : S^2 \mapsto S^2$, where $\phi_i(\overline{\beta}_j(t)) = \widehat{\beta}_{ij}(t)$ for $t \in [0, 2\pi], j \in [1, M]$. Six flattenings of the sphere are defined $\{\varphi_n : S^2 \mapsto [0, 1]^2 : n \in 1, 2, 3, 4, 5, 6\}$. For a point on the sphere, $p \in S^2$, the optimal flattening is chosen as $\varphi_p = \arg \min \varphi_n \|\varphi_n(p) - (\frac{1}{2}, \frac{1}{2})\|$. The displacement field $u_p : S_2 \mapsto \mathbb{R}^2$ is then $u_p(x) = \varphi_p(\phi_i(x)) - \varphi_p(x)$. At non-landmark points, i.e. $x \in S^2, x \notin \cup_j \{\widehat{\beta}_{ij}(t) : t \in [0, 2\pi]\}$, the mapping is constrained to satisfy a small deformation elastic model as proposed by Thompson et al. [10], and is given by

$$\mu \nabla^2 u_p(x) + (\lambda + \mu) \nabla(\nabla \cdot u_p(x)) = 0 \tag{3}$$

The atlas mesh is defined on the sphere by tessellating the sphere with a subdivided octahedron [7]. This representation is advantageous for its multiscale processing and flattening. The flattening can be imagined as follows. First, choose one of the vertices of the octahedron to map the center of the grid. Then, cut the four far edges that do not contain the center vertex. These edges are duplicated and define the boundary of the grid, and the opposite vertex maps to the four corners of the grid. The deformation is computed iteratively using multigrid finite differences, where the octahedral subdivisions and flattenings are used for prolongation and restriction operations. The solver accounts for the spherical topology of the domain by solving the above nonlinear model and resampling the deformed atlas by establishing vertex homology between the meshes.

4 Experimental Results

Our experimental data consisted of 3T MRI acquisitions (GE) for a population of 69 healthy, Chinese, right-handed volunteers (30 men, 39 women; 18-33 yr) using a transverse 3D T1-weighted fast spoiled gradient-echo (FSPGR) sequence (TR/TE = 6.8

ms/2.9 ms; voxel size = 0.47 mm \times 0.47 mm \times 0.70 mm; FOV = 24.0 \times 24.0 cm; matrix size = 512 \times 512; flip angle = 10°, slice thickness = 1.4 mm, and slice gap = 0.7 mm). After preprocessing the raw data, and registering it stereotaxically to a standard atlas space, the cortical surfaces for these subjects were extracted using an automated algorithm [2]. For each of these subjects, a total of 27 landmark curves were manually traced. Figure 1 shows the original 27 landmark curves for each of the 69 subjects for both hemispheres overlaid together. Additionally, Figure 1 also shows the intrinsic sulcal shape averages of the 27 landmark curves, as well as the respective extrinsic Euclidean averages for the same. While computing the extrinsic average, each curve for the same landmark type was mapped to its q representation, thus making it scale and translation invariant. The Euclidean average of all the q functions was then computed after a pairwise rotational alignment. Both the Karcher mean shapes as well as the Euclidean averages were then mapped back to the native space in order to visualize them. It is observed that the intrinsic averages although smooth, have preserved important features along the landmarks, thus representing the average local shape geometry along the sulci and gyri. This also implies that the shape average has not only captured the salient geometric features, but has also reduced the shape variability in the population. In order to demonstrate this property, we plot the variance of the shape deformation for each landmark type as captured by the velocity vector along the geodesic path, both for Euclidean extrinsic, and Riemannian intrinsic averages. This quantity measures the invariant deformation between a pair of shapes, and only depends upon the intrinsic geometry of the shapes. For both of these averages, the tangent vectors were computed using the procedure outline in Section 2.2 and the computations were done using the elastic geodesic method for consistent comparisons. Figure 2 shows a comparison of the plots of $\frac{1}{69} \sum_1^{69} (\dot{\chi}(i), \dot{\chi}(i))^2$ for each of the landmarks, taken along the length of

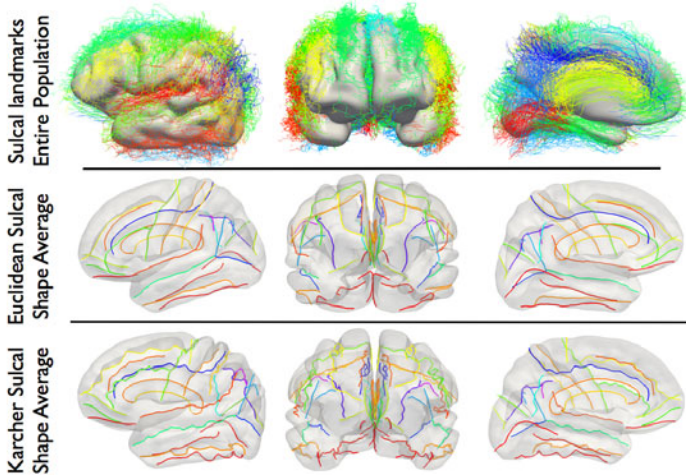


Fig. 1. Lateral, frontal, and medial views of, top row: 27 landmark sulci and gyri for 69 subjects, middle row: Euclidean sulcal shape averages, bottom row: Karcher shape average for each landmark type

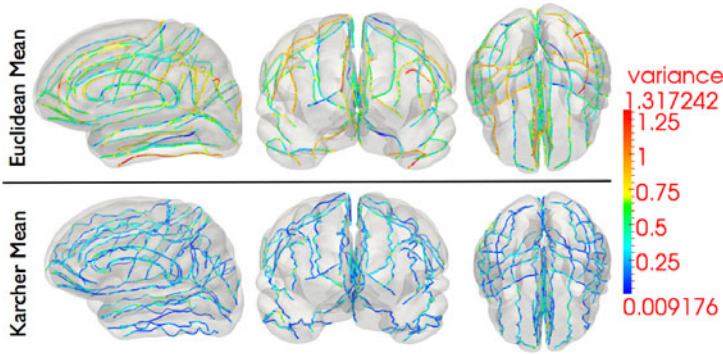


Fig. 2. Comparison of the geodesic variance for the entire sulcal population for each of the 27 landmarks, both for Euclidean shape averages, as well as elastic shape averages, along the length of the curves

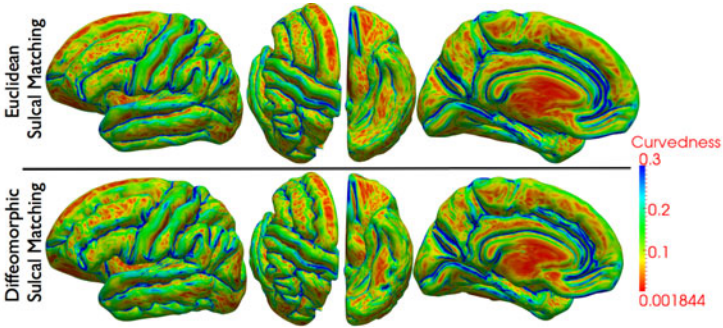


Fig. 3. Lateral, axial, ventral, and medial views of the reconstructed cortical surface with Euclidean sulcal matching (top), and diffeomorphic sulcal matching (bottom). The surfaces are colored according to shape curvedness.

the curve, for both Euclidean shape averages, as well as intrinsic shape averages. Here the $\dot{\chi}(i)$ is the tangent vector from the mean shape to i^{th} shape. From the color-coded map, it is observed that the intrinsic average has reduced the variance in terms of shape geometry deformation, and thus is a better representative of the population.

Next, we demonstrate results of cortical surface registration with and without the incorporation of the above diffeomorphic sulcal atlas in Figure 1. As an initial step, we compute geodesics between the average shape of the landmark, and the set of all sulci belonging to that landmark type, and reparameterize the set of sulci according to inverses of the resulting diffeomorphisms. We then follow the steps outlined in Sec. 3 in order to warp all the surfaces meshes to the atlas. Figure 3 shows three the lateral, axial, ventral, and medial views of the reconstructed cortical surface averages from the flattened representations. The surface is colored by its curvedness in order to highlight the fundi of the sulci as well as the ridges of the gyri. It is observed that the surface with diffeomorphic sulcal mapping shows richer geometric detail than the traditional

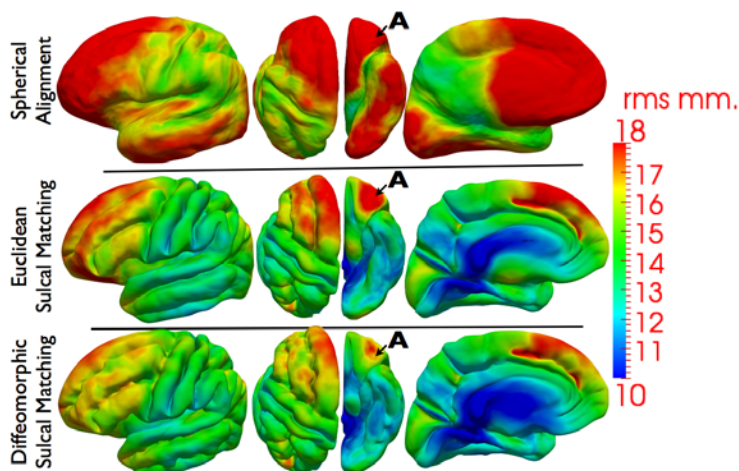


Fig. 4. R.M.S error of the distance from each sample surface to the average reconstructed cortical surface with and without diffeomorphic sulcal matching shown for left hemisphere. There is considerable improvement in registration even where there is an absence of landmarks (labeled as A).

Euclidean reconstruction. As another measure of distortion, we also computed the root mean square (r.m.s.) error of the distance from each sample surface to the average reconstructed cortical surface with and without diffeomorphic sulcal matching. It is observed that the diffeomorphic method yields lower errors throughout the surface as compared to the Euclidean matching. Interesting, the diffeomorphic mapping approach has also shown considerable improvement in the frontal lobe (Labeled as A in Figure 4) even in the absence of sulcal landmarks. The results from spherical alignment are also shown for comparison.

5 Conclusion

We have presented a direct diffeomorphic approach for shape analysis of sulcal and gyral features and demonstrated its application in cortical surface registration. We emphasize that the use of the sulcal atlas is not limited to registration alone, and can be also used to study cortical patterns for developmental, diseased or even normative patterns. The success of our method on deep brain sulci also demonstrates the effectiveness in capturing the intrinsic shape variability of the sulci and gyri. In the future, we intend to perform extensive validation studies for large populations as well as apply the sulcal models for neuroimaging population studies.

References

1. Auzias, G., Glaunès, J., Colliot, O., Perrot, M., Mangin, J.-F., Trounev, A., Baillet, S.: Disco: a coherent diffeomorphic framework for brain registration under exhaustive sulcal constraints. *Med. Image Comput. Assist. Interv.* 12(1), 730–738 (2009)

2. Dale, A.M., Fischl, B., Sereno, M.I.: Cortical surface-based analysis. I. segmentation and surface reconstruction. *NeuroImage* 9(2), 179–194 (1999)
3. Durrleman, S., Pennec, X., Trounevé, A., Ayache, N.: Measuring brain variability via sulcal lines registration: a diffeomorphic approach. *Med. Image Comput. Assist. Interv.* 10(1), 675–682 (2007)
4. Joshi, A., Shattuck, D., Thompson, P., Leahy, R.: Surface-constrained volumetric brain registration using harmonic mappings. *IEEE Tran. Med. Imagng* 26(12), 1657–1669 (2007)
5. Joshi, S., Klassen, E., Srivastava, A., Jermyn, I.: A novel representation for Riemannian analysis of elastic curves in R^n . In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2007)
6. Le, H.: Locating Fréchet means with application to shape spaces. *Advances in Applied Probability* 33(2), 324–338 (2001)
7. Losasso, F., Hoppe, H., Schaefer, S., Warren, J.: Smooth geometry images. In: *ACM SIGGRAPH Symposium on Geometry processing*, vol. 43, pp. 138–145 (2003)
8. Lyttelton, O., Boucher, M., Robbins, S., Evans, A.: An unbiased iterative group registration template for cortical surface analysis. *NeuroImage* 34(4), 1535–1544 (2007)
9. Tao, X., Han, X., Rettmann, M., Prince, J.: *Statistical study on cortical sulci of human brains*. LNCS, pp. 475–487. Springer, Heidelberg (2001)
10. Thompson, P., Toga, A.: A surface-based technique for warping 3-dimensional images of the brain. *IEEE Tran. Med. Imagng* 15(4), 402–417 (1996)
11. Vaillant, M., Davatzikos, C.: Finding parametric representations of the cortical sulci using an active contour model. *Medical Image Analysis* 1(4), 295–315 (1997)

Non-rigid Registration with Missing Correspondences in Preoperative and Postresection Brain Images^{*}

Nicha Chitphakdithai¹ and James S. Duncan^{1,2,3}

Departments of

¹ Biomedical Engineering

² Electrical Engineering

³ Diagnostic Radiology

Yale University, New Haven, CT, USA

nicha.chitphakdithai@yale.edu

Abstract. Registration of preoperative and postresection images is often needed to evaluate the effectiveness of treatment. While several non-rigid registration methods exist, most would be unable to accurately align these types of datasets due to the absence of tissue in one image. Here we present a joint registration and segmentation algorithm which handles the missing correspondence problem. An intensity-based prior is used to aid in the segmentation of the resection region from voxels with valid correspondences in the two images. The problem is posed in a maximum a posteriori (MAP) framework and optimized using the expectation-maximization (EM) algorithm. Results on both synthetic and real data show our method improved image alignment compared to a traditional non-rigid registration algorithm as well as a method using a robust error kernel in the registration similarity metric.

1 Introduction

In localization-related epilepsy, brain resection is often considered the next line of treatment if anticonvulsive medication is unable to control seizures [1]. To better understand epilepsy and the changes that occur in the brain after resection, preoperative and postresection images first need to be aligned. This registration problem is challenging due to the missing correspondences caused by resected tissue coupled with the possibility for highly nonlinear deformations in the postresection brain.

While many traditional non-rigid registration algorithms are available, few handle the problem of missing correspondences. A method for general registration problems was presented in [2] which dealt with partial data by using an EM-style method to simultaneously estimate the registration parameters and the missing data. Each voxel is assumed to have an equal chance of being labeled as missing or valid data. For our application, it is known that resection

^{*} This work was supported by NIH 5R01EB000473-08.

voxels will appear dark in T1-weighted MRI due to cerebrospinal fluid (CSF) taking place of the missing tissue [3], and we will thus incorporate this prior information.

Some application-specific methods have also been proposed to better register images with substantial changes. The correspondence problem in the alignment of dynamic contrast-enhanced breast MRI was handled by “de-enhancing” the contrast images before registration in [4]. Biomechanical models of brain tumor growth and estimation of tissue loss and replacement have been used to improve the registration accuracy around the tumor region [5]. An algorithm to accommodate resection and retraction was presented in [6] for alignment of preoperative to intraoperative brain images. Their method alternated between registering the images using a demons algorithm with an anisotropic diffusion smoother and segmenting the resection by evolving a level set in the region with high intensity errors. We are proposing a probabilistically-grounded approach to the simultaneous registration and resection estimation problem.

In this paper, we present a joint registration and segmentation method in which an indicator function separating the valid and missing correspondence regions is estimated along with the transformation parameters. The addition of the indicator map into the registration estimation allows the incorporation of different models for observing the data under different correspondence assumptions. The general joint registration and segmentation framework follows the method described in [7]. Previously in [8] we used a spatial prior on the valid tissue and resection locations to help estimate these regions. Due to the difficulty in gathering a training set for the spatial prior, we now explore a prior on postresection image intensities. Posing the problem in a MAP framework allows for the natural inclusion of an intensity-based prior.

2 Methods

2.1 Registration and Segmentation Framework

The goal of the registration problem is to determine the optimal transformation T which aligns voxel \mathbf{x} in the preoperative image P with voxel $T(\mathbf{x})$ in the postresection image R . We incorporate a “hidden” indicator map I which gives the segmentation of valid tissue in the postresection image R from the resection region. The registration problem is then posed in a marginalized MAP framework, where we solve for

$$\hat{T} = \arg \max_T \log \sum_I p(T, I | P, R) \quad (1)$$

by using the EM algorithm. In the $(k + 1)$ th iteration, the E-step gives an estimate for the indicator map given the estimate of the transformation T^k from the previous iteration. The M-step then determines the new registration parameters T^{k+1} using the information from the E-step. The algorithm keeps iterating between the E-step and M-step until the registration parameters converge.

In our instance of the EM algorithm, the M-step can be written as

$$T^{k+1} = \arg \max_T E_{I|P,R,T^k} [\log p(P | R, I, T) + \log p(R | I, T) + \log p(I | T) + \log p(T)]. \tag{2}$$

We next make a few simplifications. First we make the common assumption of independence of voxels in the image. We also assume that given the indicator, the postresection image does not depend on T . To simplify (2) further, we assume no prior spatial information for the indicator map. We can then rewrite (2) as

$$T^{k+1} = \arg \max_T \sum_{\mathbf{x} \in R} \sum_{l \in L} p(I(\mathbf{x}) = l | P, R, T^k) [\log p(P(T(\mathbf{x})) | R, I(\mathbf{x}) = l, T) + \log p(R(\mathbf{x}) | I(\mathbf{x}) = l)] + \log p(T), \tag{3}$$

where L is the set of possible labels in the indicator map. In the following we use $L = \{0, 1, 2\}$: $I(\mathbf{x}) = 0$ for the resection, $I(\mathbf{x}) = 1$ for a voxel with a valid correspondence, and $I(\mathbf{x}) = 2$ for a background voxel.

The E-step calculates the probability that a voxel is assigned a certain label. Using Bayes' rule, this probability can be computed from

$$p(I(\mathbf{x}) = l | P, R, T^k) = \frac{p(P(T^k(\mathbf{x})) | R, I(\mathbf{x}) = l, T^k) p(R(\mathbf{x}) | I(\mathbf{x}) = l)}{\sum_{l'} p(P(T^k(\mathbf{x})) | R, I(\mathbf{x}) = l', T^k) p(R(\mathbf{x}) | I(\mathbf{x}) = l')}. \tag{4}$$

After the EM algorithm converges, the final indicator map \hat{I} is estimated from these probabilities by assigning $\hat{I}(\mathbf{x}) = \arg \max_l p(I(\mathbf{x}) = l | P, R, \hat{T})$.

2.2 Probability Models

To calculate (3) and (4), we need to define probability models for the similarity measure $p(P(T(\mathbf{x})) | R, I(\mathbf{x}) = l, T)$, the prior on intensities given the indicator function value $p(R(\mathbf{x}) | I(\mathbf{x}) = l)$, and the transformation prior $p(T)$.

Similarity Term. The probability $p(P(T(\mathbf{x})) | R, I(\mathbf{x}) = l, T)$ acts like the similarity metric of a standard registration algorithm. For different values of the indicator map, we can use a different probability model to describe how we expect the images to match. We assume a voxel labeled as part of the resection in R can match any intensity in P with equal probability and use a uniform distribution. When a voxel is labeled as having a valid correspondence, the intensities should match, so we model $P(T(\mathbf{x})) | R, I(\mathbf{x}) = l, T \sim N(R(x), \sigma_1)$. Finally, for a background voxel, we again employ a normal distribution but assign a larger standard deviation since we are not actually interested in matching the background. To summarize, the models under each indicator label are

$$p(P(T(\mathbf{x})) | R, I(\mathbf{x}) = l, T) = \begin{cases} 1/c, & l = 0 \\ \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{[P(T(\mathbf{x})) - R(\mathbf{x})]^2}{2\sigma_1^2}\right), & l = 1 \\ \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{[P(T(\mathbf{x})) - R(\mathbf{x})]^2}{2\sigma_2^2}\right), & l = 2 \end{cases}, \quad (5)$$

where c is the number of intensity levels in each image and $\sigma_1 < \sigma_2$. For the experiments in this paper, we set $\sigma_2 = 2\sigma_1$.

Intensity Prior. The term $p(R(\mathbf{x}) | I(\mathbf{x}) = l)$ is the intensity-based prior in which given the indicator map label, we assume some knowledge of the intensities in the postresection image. For a voxel labeled as resection, we model its intensity using a normal distribution with a maximum likelihood estimate of the mean and standard deviation from manually segmented resection regions from a training set of postoperative images. If a voxel is labeled as having a valid correspondence, we assume any intensity in R is equally likely to have a matching correspondence in P and use a uniform distribution. Finally, background voxels are assumed to follow a normal distribution with 0 mean and a small standard deviation.

Transformation Prior. We use free form deformations (FFD) based on uniform cubic B-splines to model the non-rigid transformation as in [9]. The transformation parameters are then the B-spline control points \mathbf{t}_i , spaced δ apart. We assume the control points and its components $t_{i,j}$ are independent so that $p(T) = \prod_i \prod_j p(t_{i,j})$, where $p(t_{i,j})$ follows a normal distribution with mean equal to the starting location of the control point component on the uniform grid and standard deviation equal to $\frac{0.4\delta}{3}$ to encourage a smooth transformation. In addition, we enforce a hard constraint to ensure the transformation is injective by restricting the control points to lie within a sphere of radius 0.4δ [10].

2.3 Registration Methods for Comparison

We compare our joint registration and indicator map estimation (RIME) method to a standard non-rigid registration (SNRR) method found in BioImage Suite (BIS) software [11]. In addition, we compare an algorithm using a robust squared error in the similarity measure, i.e. $\rho(R(\mathbf{x}) - P(T(\mathbf{x}))) \sim N(0, \sigma)$, where

$$\rho(x) = \begin{cases} \frac{s^2}{6} \left(1 - \left(1 - \frac{x}{s}\right)^2\right)^3, & |x| \leq s \\ \frac{s^2}{6}, & |x| > s \end{cases}, \quad (6)$$

is the Tukey function with scaling parameter s [12]. The influence of outliers is reduced by mapping errors greater than s to a maximum value. In this way, the intensity disparity in the resection region is given less influence in determining the optimal transformation. We automatically estimate $s = 3\sigma$, where σ is the standard deviation of the intensity differences for the current overlap of images.

Both SNRR and robust Tukey registration (RTR) use a FFD transformation model with the same control point spacing as in RIME. In addition, note that all methods utilize a similarity measure based on direct intensity differences.

3 Results

3.1 Synthetic Images

We used 11 pairs of 2D synthetic MR images. For each pair, a slice from a normal real 3D MRI was taken as the preoperative image. The postoperative image was created by first assigning resection voxels normally distributed random intensities with mean and standard deviation from a CSF sample from the preoperative image. The image was then warped by a physical model. The deformation simulations, while not necessarily mimicking true non-rigid motion after resection, provide a ground truth of displacement vectors at each voxel. The intensity prior for the image to be registered was trained on the remaining postresection images.

Figure 1 shows example results. Difference images were masked by the true valid correspondence region. In Fig. 1(a), traditional registration resulted in large errors, especially near the resection. While RTR in Fig. 1(b) showed great improvement, the RIME difference image in Fig. 1(c) appeared even flatter.

Table 1 lists the minimum, maximum, mean, and standard deviation of displacement errors for voxels in R which had a valid correspondence in P , averaged over the image pairs. We performed one-tailed paired t-tests to evaluate the results and considered $p < 0.05$ significant. Only the mean error for RTR significantly decreased compared to SNRR. On the other hand, all RIME error statistics were significantly reduced compared to SNRR. Furthermore, the minimum, mean, and standard deviation of errors for RIME were significantly smaller than those for RTR. The quantitative analysis confirmed the qualitative observations that RIME produced more accurate results than SNRR and RTR.

The estimated indicator function for valid correspondence pixels is shown in Fig. 2, with the true map outlined in green. Typically, errors occurred in the CSF or near the boundary of the brain and background or resection. The dice coefficient was computed for each pair to measure the overlap between the true and estimated maps for valid correspondences, giving an average of 0.98.

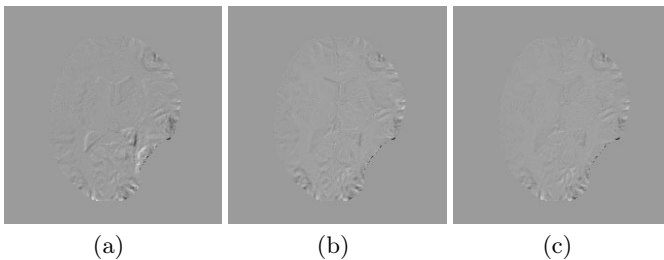


Fig. 1. Sample difference images between a simulated postresection image and the corresponding warped preoperative image using (a) SNRR, (b) RTR, and (c) RIME



Fig. 2. Outline of true indicator map (green) overlaid on estimated map

Table 1. Average displacement field errors (in voxels) after synthetic data registration

Registration Method	Min Error	Max Error	Mean Error	Std Dev Error
SNRR	0.0028	3.1952	0.3479	0.3709
RTR	0.0017	3.1808	0.2757	0.3083
RIME	0.0009	2.9013	0.2272	0.2723

3.2 3D Clinical Data

We applied the registration methods to 6 pairs of preoperative and postresection T1-weighted MR images. First, images were skull stripped using BIS, resampled to 128x128x60, and affinely aligned. Intensities were renormalized so they could be directly compared. Leave-one-out validation was used, resulting in a training set of about 12500 voxels from 5 images for the resection intensity prior.

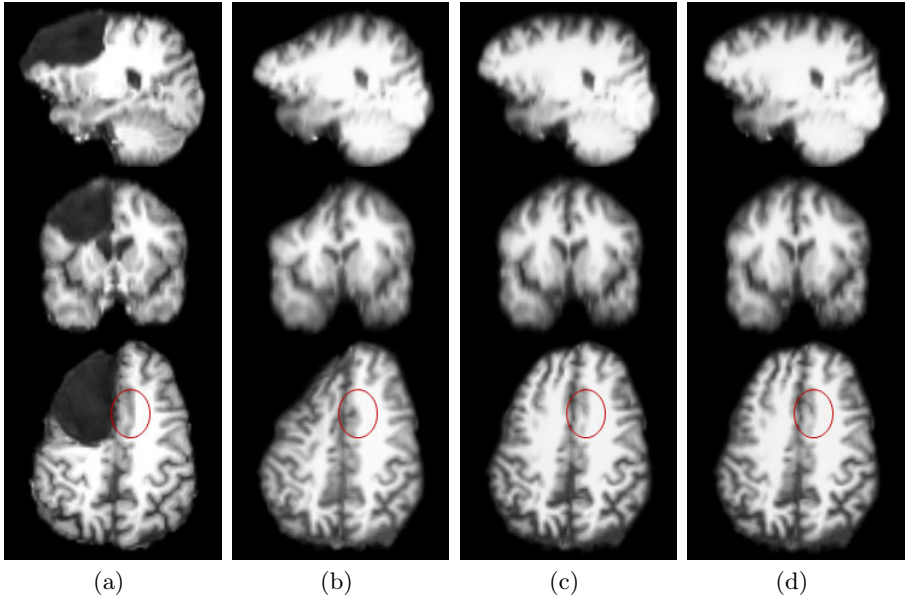


Fig. 3. Registration of real data. Red ring highlights different results. (a) Postresection image. (b) Warped preoperative image using SNRR. (c) RTR result. (d) RIME result.

Figure 3 displays slices from the registration of real 3D images. The patient had a right frontal lobe resection as seen in Fig. 3(a). Warped preoperative images after SNRR, RTR, and RIME are shown in Figs. 3(b), (c), and (d) respectively. SNRR erroneously deformed the right frontal lobe to match the background to the large resection. This caused misalignment of other features, like the right lateral ventricle and sulci left of the longitudinal fissure. RTR resulted

in much better alignment of corresponding structures. RIME results showed further improved alignment, such as in the gray matter left of the longitudinal fissure in the axial slice (red ring) and the lateral ventricle in the sagittal image.

We evaluated the registration error by computing the distance between 8 corresponding landmarks spread throughout the brain in each pair of postresection and registered preoperative images. The average errors were 3.99 mm for affine registration, 2.69 mm for SNRR, 2.16 mm for RTR, and 1.27 mm for RIME. All non-rigid registration methods significantly decreased the error compared to affine (again $p < 0.05$). The reduced error for RTR compared to SNRR missed significance, while the improvement in RIME error over SNRR was significant.

A sample estimated indicator map is overlaid on a brain with a left temporal lobe resection in Fig. 4. The average dice coefficient measuring the overlap between true and estimated indicator maps for valid correspondences was 0.92.

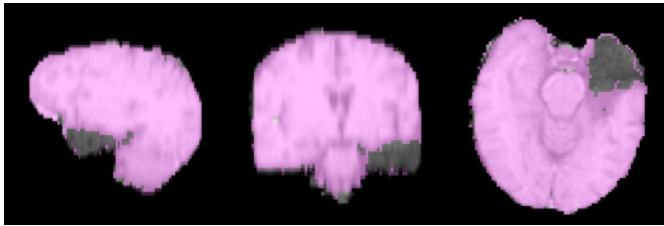


Fig. 4. Estimated indicator map segmenting tissue with correspondences (pink), resection (no color mask), and background (black), overlaid on postresection image

4 Discussion

Our proposed RIME method follows a marginalized MAP framework solved using the EM algorithm to estimate both the registration parameters and a “hidden” map indicating the different correspondence regions. The formulation of the problem allows the inclusion of a prior based on voxel intensities in the postresection images given the correspondence label. Our method resulted in significantly improved accuracy compared to a standard non-rigid registration technique (SNRR) and a method using a robust similarity metric (RTR).

Future work will incorporate probability models based on image histograms for the similarity measure to extend the method to intermodality registration. A greater number of labels in the indicator map, for example to represent gray and white matter, may further increase registration accuracy. We will also work to improve the final estimate of the indicator map, which sometimes mislabeled voxels with valid correspondences as resection. This may be due to dark intensities matching the intensity prior better than the similarity term (such as in the CSF) or due to high dissimilarity in intensities near the brain and background border. To simplify the calculations in this paper, we had assumed no prior spatial information for the indicator map. We plan to reincorporate this

spatial prior to improve the indicator map estimate, which should in turn result in more accurate registration.

References

1. Kwan, P., Sperling, M.R.: Refractory seizures: Try additional antiepileptic drugs (after two have failed) or go directly to early surgery evaluation? *Epilepsia* 50, 57–62 (2009)
2. Periaswamy, S., Farid, H.: Medical image registration with partial data. *Med. Image Anal.* 10, 452–464 (2006)
3. Winterstein, M., Münter, M.W., Burkholder, I., Essig, M., Kauczor, H.U., Weber, M.A.: Partially resected gliomas: Diagnostic performance of fluid-attenuated inversion recovery mr imaging for detection of progression. *Radiology* 254(3), 907–916 (2010)
4. Zheng, Y., Yu, J., Kambhamettu, C., Englander, S., Schnall, M.D., Shen, D.: Enhancing the dynamic contrast-enhanced breast mri for robust registration. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part I. LNCS*, vol. 5241, pp. 933–941. Springer, Heidelberg (2008)
5. Zacharaki, E.I., Hoge, C.S., Shen, D., Biros, G., Davatzikos, C.: Non-diffeomorphic registration of brain tumor images by simulating tissue loss and tumor growth. *Neuroimage* 46(3), 762–774 (2009)
6. Risholm, P., Samset, E., Talos, I.F., Wells, W.: A non-rigid registration framework that accommodates resection and retraction. In: Prince, J.L., Pham, D.L., Myers, K.J. (eds.) *IPMI 2009. LNCS*, vol. 5636, pp. 447–458. Springer, Heidelberg (2009)
7. Pohl, K.M., Fisher, J., Grimson, W.E.L., Kikinis, R., Wells, W.M.: A bayesian model for joint segmentation and registration. *NeuroImage* 31(1), 228–239 (2006)
8. Chitphakdithai, N., Duncan, J.S.: Pairwise registration of images with missing correspondences due to resection. In: 2010 7th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 1025–1028 (2010)
9. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J.: Non-rigid registration using free-form deformations: application to breast mr images. *IEEE Trans Med. Imaging* 18(8), 712–721 (1999)
10. Greene, W.H., Chelikani, S., Purushothaman, K., Knisely, J., Chen, Z., Papademetris, X., Staib, L.H., Duncan, J.S.: Constrained non-rigid registration for use in image-guided adaptive radiotherapy. *Med. Image Anal.* 13(5), 809–817 (2009)
11. Papademetris, X., Jackowski, M., Rajeevan, N., Okuda, H., Constable, R., Staib, L.: BioImage Suite: An integrated medical image analysis suite. Section of Bioimaging Sciences, Dept. of Diagnostic Radiology, Yale School of Medicine, <http://www.bioimagesuite.org>
12. Arya, K.V., Gupta, P., Kalra, P.K., Mitra, P.: Image registration using robust m-estimators. *Pattern Recognition Letters* 28, 1957–1968 (2007)

System to Guide Transcatheter Aortic Valve Implantations Based on Interventional C-Arm CT Imaging

Matthias John¹, Rui Liao², Yefeng Zheng², Alois Nöttling¹, Jan Boese¹,
Uwe Kirschstein¹, Jörg Kempfert³, and Thomas Walther⁴

¹ Siemens AG, Healthcare Sector, Forchheim, Germany

² Siemens Corporate Research, Princeton, USA

³ Department of Cardiac Surgery, Heart Center, University of Leipzig, Germany

⁴ Department of Cardiac Surgery, Kerckhoff Heart Center, Bad Nauheim, Germany
matthias.mj.john@siemens.com

Abstract. Transcatheter aortic valve implantation is an emerging technique to be applied in patients with aortic valve defects. Angiographic and fluoroscopic X-ray imaging with a C-arm system is crucial in these minimally invasive procedures. We describe a prototypical system based on the ability to acquire a 3D C-arm CT image during transcatheter aortic valve implantations. It supports the physician in measuring critical anatomical parameters, finding an optimum C-arm angulation, and guiding the positioning and deployment of the prosthesis by 3D overlay with fluoroscopic images. To yield high acceptance by the physicians in the operating room, our approach is fast, fully integrated into an angiographic C-arm system, and designed to minimize the necessary user interaction. We evaluate the accuracy of our system on 20 clinical cases.

1 Introduction

Aortic valve disease is the most common acquired heart valve disease. Minimally invasive transcatheter aortic valve implantation (TAVI) is clinically being performed in elderly and high risk patients with a severe aortic stenosis. It has the potential to be applied in the future to regular risk patients instead of an open heart surgery including sternotomy, extracorporeal circulation and cardioplegic cardiac arrest. A stent based valve consisting of xenograft leaflet tissue is positioned and deployed in the aortic root of the patient using transcatheter techniques. During transapical TAVI an antegrade access is applied where the valve is inserted via small incisions in the chest and the apex of the left ventricle. During transfemoral TAVI the valve is inserted retrograde via the femoral artery and the aortic arch. Both approaches require X-ray angiographic and fluoroscopic imaging to guide the procedure (see Figure 1). Therefore these procedures are usually performed in operating rooms equipped with a fixed angiographic C-arm system (“Hybrid OR”). In order to visualize the aortic root under X-ray, contrast agent must be injected. Due to frequent renal insufficiency in these patients the amount of contrast agent applied should be minimized.

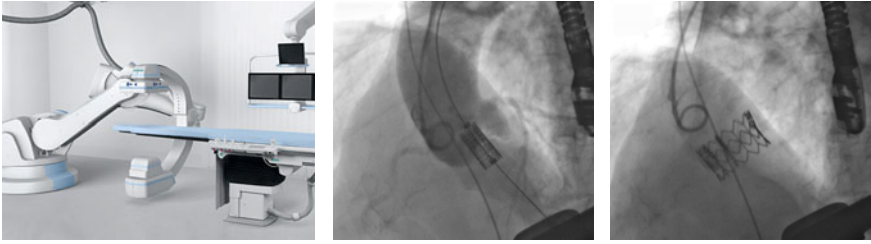


Fig. 1. Transapical aortic valve implantation under X-ray guidance. **Left:** Angiographic C-arm system able to acquire interventional 3D images in an operating room. **Middle:** Contrast injection via pigtail catheter immediately prior to valve deployment. **Right:** Implanted valve.

Prior to implantation it is important to angulate the C-arm with respect to the aortic root anatomy. Rotationally symmetric prostheses require an angulation perpendicular to the aortic root to be placed correctly. Prostheses that model the leaflet anatomy additionally require an angulation that allows for outline of the commissures (attachments of the valvular leaflets). Conventionally, an appropriate angulation is achieved with iterated C-arm angulations, each followed by an angiogram of approximately 15 ml contrast to double-check the aortic root position. Further angiograms are needed later on for correct prosthesis positioning and for functional control after implantation.

Our goal is to provide image guidance based on interventional C-arm CT images to add detailed 3D information to the procedure. The system needs to be set up in the complex environment of an operating room and should be easy to use by a physician during TAVI. Therefore, it is crucial for the acceptance of such a system to be fast, to minimize the user interaction and to allow table-side control.

Previous work about the support of transcatheter aortic valve implantations contains modeling for procedure planning [1,2], guidance by tracking the prosthesis in fluoroscopic images [3], and a robotic system using MRI imaging [4].

2 Method

In this section we describe the design of our system, the steps the system performs and discuss the user interactions. All described components were prototypically integrated into an angiographic system (Artis zee/zeego with syngo X Workplace and syngo DynaCT, Siemens AG, Healthcare Sector, Forchheim, Germany).

Before the implantation, the physician obtains an interventional 3D image of the aortic root by acquiring a rotational 2D image sequence of 200° over 5 seconds on the C-arm system. Via a pigtail catheter 25 ml contrast agent (diluted to 75 ml) is injected over 5 seconds (with a 1-second X-ray delay) into the aortic root. To minimize motion artifacts, patient breathing is suspended and rapid ventricular pacing is applied. This temporarily stops the heart pumping and minimizes blood flow. It allows for a relatively small amount of contrast agent to be used - compared to approximately 80 ml for a conventional CT and approximately 15 ml for a single 2D angiogram [5].

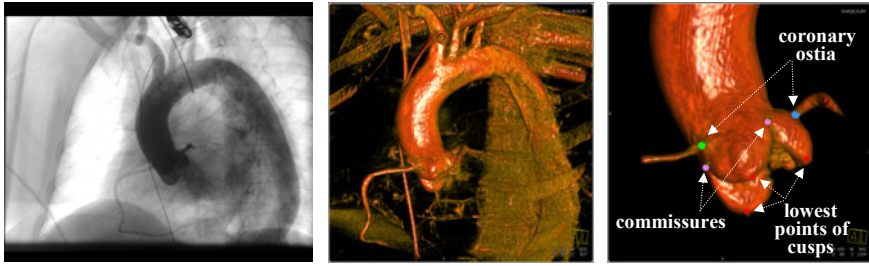


Fig. 2. **Left:** Image from a rotational acquisition scene acquired under rapid ventricular pacing and contrast injection into aortic root. **Middle:** C-arm CT image reconstructed from this scene. **Right:** Segmented aortic root with landmarks.

After the rotational run is finished on the C-arm system, all of the following steps 1–7 are initiated and performed fully automatically.

1. Reconstructing 3D data from acquired rotational image sequence.

The rotational run is reconstructed in about 12 seconds and is based on a software available with the angiographic system (syngo DynaCT [6], see Figure 2).

2. Detecting aortic root shape and landmarks from 3D volume.

Following [5] we detect the aortic root shape and eight landmarks (see Figure 2, right): the lowest point (nadir) of each aortic root cusp (to support finding a C-arm angulation perpendicular to the aortic root), the coronary artery ostia (which need to stay open after prostheses implantation), the commissure points where the cusps meet (to help orient anatomically designed prostheses) and finally the centerline of the aorta.

We use a fast dedicated machine learning based algorithm [7]: first, position, orientation, and scale of the aortic root are estimated by the efficient Marginal Space Learning (MSL) algorithm. Then the mean aortic root shape (calculated from a training set) is aligned with the estimated pose, followed by boundary refinement using a learning-based 3D boundary detector. In addition, the eight landmarks are detected using a discriminative learning based landmark detector. The authors report the following mean detection errors based on a four-fold cross validation on a dataset with 192 volumes: 1.1 mm for aortic root mesh, 2.4 mm for lowest cusp points, 3.5 mm for the aortic commissure points, and 2.7 mm for the coronary ostia.

3. Deriving additional structures from detected landmarks.

We derive a circle parallel to the plane spanned by the three lowest points of the cusps (see red circle in Figure 3). Visually, this perpendicularity circle degenerates to a straight line if and only if the three lowest cusp points are aligned (see Figure 4), which corresponds to an optimal perpendicular angulation for valve implantation.

To estimate how critical the position of the coronary ostia is, the physician will measure their distance to the plane through the three lowest points of the cusps. We create a ruler orthogonal to that plane (see the red ruler in Figure 3). We show a ruler instead of numbers, because this makes the measurement process transparent to the physician and allows for user compensation of the measurement in the case where detected landmarks need to be corrected.

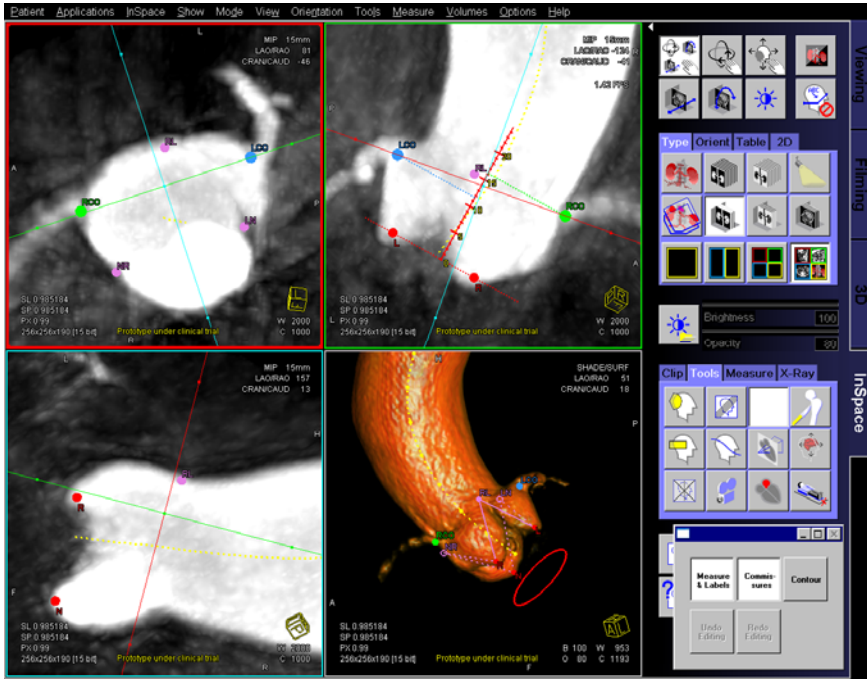


Fig. 3. System appearance after automatically performing steps 1-7. Detected landmarks (coronary ostia in blue and green, commissures in purple, lowest points of cusps in red, centerline in yellow) and derived structures (perpendicularity circle and ruler in red) are shown in 3D volume rendering and three orthogonal intersection planes with 15 mm slice thickness. The panel in the lower right shows the user interface we added to the existing system.

4. Extract interior of detected aortic shape out of 3D volume for volume rendering.

The interior of the detected aortic root is extracted from the volume and visualized with volume rendering. We believe that with volume rendering it is easier for the user to verify the accuracy of the detection step compared to a mesh visualization as volume rendering is still based on the original intensities of the image voxels.

A perfect aortic root detection cannot be guaranteed for all patients. Therefore, we dilate the extracted volumetric shape by 2 mm (determined heuristically) so that the user can visually detect accidentally removed structures. Also visualization of coronary arteries is important, but automatic segmentation is difficult. To make them visible without segmenting them explicitly we add a ball shape of 15 mm (determined heuristically) around each detected coronary ostium. This ensures that coronary artery segments are volume rendered even in cases of slight misdetections.

5. Computing optimized volume rendering transfer function parameters.

We want to avoid having the user manually find the appropriate volume rendering parameters c_{opt} for transfer function center and w_{opt} for transfer function width. Therefore, they are calculated automatically based on the voxel intensities:

$$\begin{aligned} c_{opt} &= f_{c,in} m_{in} + f_{c,out} m_{out} + f_{c,offset} \\ w_{opt} &= f_{w,in} m_{in} + f_{w,out} m_{out} + f_{w,offset} \end{aligned} \quad (1)$$

Here, m_{out} (m_{in}) is a volume specific value and is determined by the mean intensities of all voxels outside (inside) the boundary of the segmented aortic root with a fixed distance to it. The underlying set of voxels is computed using the morphologic dilation and erosion operators with a certain number of iterations. The six parameters $f_{c,in}$, $f_{c,out}$, $f_{w,offset}$, $f_{w,in}$, $f_{w,out}$, $f_{w,offset}$ in Equation (1) are fixed values that are obtained by a trained sample set of segmented volumes: For each of these training volumes we manually adjusted optimal window width and window center values. Together with the calculation of the corresponding values m_{in} and m_{out} we get an over-determined system of linear equations with six unknown parameters, which is then solved by a least-squares fitting method.

6. Computing good volume position and orientations of intersection planes.

To give the user a good initial view, the volume is centered and zoomed based on the position of the two detected coronary ostia.

To allow the user an easy verification of the detected coronary ostia, we show two orthogonal volume intersection planes, both visualizing the two ostia (see Figure 3). One of the planes is chosen to be orthogonal to the plane spanned by the three lowest points of the cusps. This plane contains the ruler discussed in step 3 and therefore allows the user to directly measure distances without any user interactions.

7. Visualizing 3D volume and intersection planes in a 2x2 screen.

According to the computations in the previous steps, the segmented aortic root volume is shown in a 3D volume rendering screen. Three orthogonal intersection planes show the unsegmented volume. All four screens contain the detected landmarks (see Figure 3). The perpendicularity circle is displayed in the 3D volume window and the ruler in one of the intersection planes. To allow for verification of the landmark detection (also in cases the detection went slightly wrong), we show the intersection planes as 15 mm thick maximum intensity projections (MIP).

After the results are presented, the physician is able to adjust landmarks and visualization if necessary. All described interactions can be done immediately at the operating table using a joystick which is part of the angiographic C-arm system.

8. Editing landmarks with automatic adaption of derived structures.

By drag-and-drop functionality the user can edit the detected landmarks in the orthogonal intersection planes and in the volume rendering screen. Undo and redo functionality ensure that editing mistakes can be handled easily. Structures derived from the landmarks (e.g. ruler and perpendicularity circle) are adapted in real-time.

9. Toggling on/off landmarks and structures.

Because displaying too many structures simultaneously can confuse the user, some of them can be toggled on and off (see user interface in Figure 3). The system can be configured in a way that the preferred set of structures is initially displayed.

If landmark positions and visualization are satisfactory, the physician can perform the following steps:

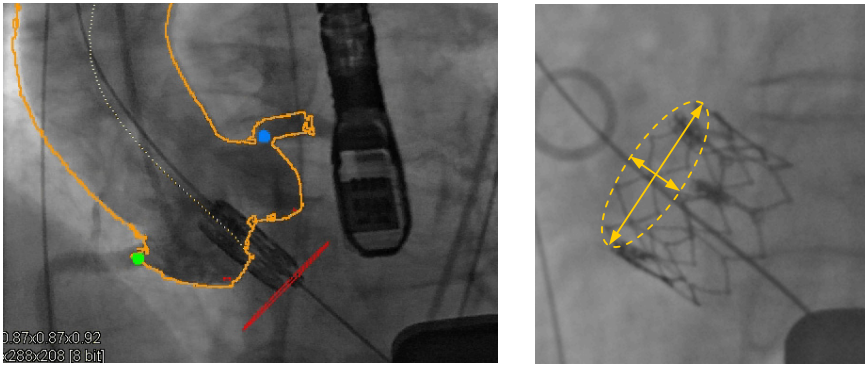


Fig. 4. **Left:** Overlay with perfect matching (measured misalignment is 0 mm). **Right:** Measurement of potential tilting of the prosthesis for evaluation (for better illustration we show an image from an angulation that had not been used for valve deployment).

10. Rotating 3D volume to an appropriate C-arm angulation.

Every rotation of the volume rendered view corresponds to a C-arm angulation (up to in-plane rotation). Therefore, by using the displayed perpendicularity circle or the commissures, the physician can virtually choose a view that corresponds to an appropriate C-arm angulation. The C-arm can then be automatically rotated to that angulation. Alternatively, the volume rendering rotation can be synchronized with any C-arm movement.

11. Overlay of visualized 3D structures onto fluoroscopic images.

The live overlay of the rendered 3D visualization onto fluoroscopic images is based on software available with the angiographic C-arm system (syngo iPilot). The 3D volume is inherently registered to the fluoroscopic images because both images are acquired on the same system. The overlay dynamically adapts to C-arm rotations and table movements. It does not compensate for patient and heart motions, but can be corrected manually in these cases. This mode can be used to adjust or fine-tune the C-arm angulation needed for implantation (see Figure 4, left).

12. Switching to contour view.

When overlaying the 3D volume onto a live fluoroscopic scene, we offer a contour view of the aortic root which shows only the essential information needed for prosthesis positioning and deployment. Therefore, there is less of the live fluoroscopic image space being hidden compared to volume rendering which yields better fluoroscopic image quality in the overlay for a safer implantation (see Figure 4, left).

The contours are computed by efficient edge detection. First, a volume rendered aortic root image is segmented by simple image intensity thresholding. Second, gradients of the segmented pixels are calculated. The maximum gradient value is also assigned to border pixels (pixels with at least one neighboring background pixel). Third, edge pixels are detected by a hysteresis thresholding (similar to the Canny edge detector) on the gradient values. Finally, connected components with a small number of detected edge pixels are removed.

3 Evaluation

We evaluated the accuracy of our system in patients who received an Edwards Sapien valve prosthesis (Edwards Lifesciences, Irvine, USA). These valves are deployed under rapid ventricular pacing. Retrospectively we analyzed the first 20 cases that were supported by our system and where the overlay image scene was documented.

First, we were interested in how well the system can help to position the valve in the aortic root with a correctly tilted angle. Assuming that a C-arm angulation adjusted with our system influences the final valve tilt angle, it would be interesting to evaluate the valve position in relation to the individual patient anatomy. Unfortunately this would require the effort of a post-op 3D image scan including additional contrast agent and X-ray dose. Instead we determined for each patient the tilting angle of the implanted prosthesis in the 2D fluoroscopic image under the chosen angulation. We therefore assume that ideally an optimal angulation would result in a valve image not showing any tilting. With this assumption we ignore other factors like the complex interaction of operators, devices, and patient anatomy. We measured the ellipsoid diameters of the upper prosthesis ring in the image in pixels (see Figure 4, right) and derived the tilting by $\arcsin(diam_{min}/diam_{max})$. This value demonstrates how perpendicular the valve prosthesis was imaged immediately after implantation. For the 20 evaluated patients, we measured a tilt angle of $5.7^\circ \pm 5.2^\circ$ (mean \pm standard deviation). Clinically, a tilt angle of $< 5^\circ$ can be stated as very good (obtained in 60% of the patients in our study), 5° - 10° as good (30%), 10° - 15° as acceptable (5%) and $> 15^\circ$ as inappropriate (5%). The values show that a procedure with 3D image support by our system yields overall positive results. Reasons for suboptimal angulation estimation in some cases might be misdetections due to C-arm CT image artifacts, resulting from asymmetric position of the injection pigtail catheter in the aortic root and severe aortic regurgitation.

Second, we evaluated the accuracy of the overlay of the 3D image and the X-ray images. We assume that a misalignment corresponds to a shift parallel to the projection plane, which simplifies the evaluation but ignores rotation and zoom. For each patient we took an image from the recorded overlay scene that showed a contrast injection with rapid pacing right before deployment of the prosthesis (see Figure 4, left). We then measured the shift error as the distance (in pixels) of a landmark point that could be identified in X-ray image and 3D overlay, e.g. lowest cusp point or coronary ostium. Furthermore we measured the shift along the aortic root centerline, which is the most important direction for guiding the implantation. Because of the projective geometry of the images, a measurement in pixels must be scaled using the known length of an object in approximately the same distance to the X-ray detector. For this we used the known length of the implanted prosthesis (in mm) divided by its measured height (in pixels). For the 20 evaluated patients we measured a shift error of $3.1 \text{ mm} \pm 1.9 \text{ mm}$ (mean \pm standard deviation) and in centerline direction a shift error of $1.9 \text{ mm} \pm 1.5 \text{ mm}$. The reasons for pronounced deviation in a few patients could be the dislocation of the aortic root by sheath-manipulation and accidentally movements of the patients caused by the physicians. The measured accuracy is only valid under repeat rapid ventricular pacing, which duplicates the heart position that we had during the 3D imaging and which minimizes heart motion.

4 Conclusion

Our system design enables efficient integration of interventional 3D imaging in transcatheter aortic valve implantations and provides valuable guidance tools. It is prototypically used in parallel to conventional fluoroscopy and angiography. The main benefits are the adjustment of an optimal C-arm angulation based on 3D information requiring a low amount of contrast agent only, the measurement of critical coronary ostia distances and the additional anatomical orientation by the fluoroscopic overlay when implanting the valve.

The system is very fast: post-processing of 3D data (steps 2-7 in section 2) takes about 4 seconds. In the cases evaluated up to this point, all described algorithm steps are stable. In practice, landmark adjustments were only rarely done by the user. Under such rare circumstances the table sided joystick was an appropriate tool which eliminates the need to operate a mouse in the non-sterile environment of the control room.

Valuable extensions of our system would be an automated motion compensation of the 3D overlay on fluoroscopic images and the integration of pre-operative CT images. For evaluation purposes it would be interesting to compare our proposed system with the conventional approach to implant aortic valves.

References

1. Gessat, M., Merk, D.R., Falk, V., Walther, T., Jacobs, S., Nöttling, A., Burgert, O.: A planning system for transapical aortic valve implantation. In: *Medical Imaging: Visualization, Image-Guided Procedures, and Modeling*, vol. 7261. SPIE (2009)
2. Ionasec, R., Georgescu, B., Gassner, E., Vogt, S., Kutter, O., Scheuering, M., Navab, N., Comaniciu, D.: Dynamic model-driven quantification and visual evaluation of the aortic valve from 4D CT. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part I. LNCS*, vol. 5241, pp. 686–694. Springer, Heidelberg (2008)
3. Karar, M.E., Chalopin, C., Merk, D.R., Jacobs, S., Walther, T., Falk, V., Burgert, O.: Localization and tracking of aortic valve prosthesis in 2D fluoroscopic image sequences. In: *Medical Imaging: Visualization, Image-Guided Procedures, and Modeling*, vol. 7261. SPIE (2009)
4. Li, M., Mazilu, D., Horvath, K.A.: Robotic System for Transapical Aortic Valve Replacement with MRI Guidance. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part II. LNCS*, vol. 5242, pp. 476–484. Springer, Heidelberg (2008)
5. Kempfert, J., Falk, V., Schuler, G., Linke, A., Merk, D.R., Mohr, F.W., Walther, T.: Dyna-CT during minimally invasive off-pump transapical aortic valve implantation. *Ann. Thorac. Surg.* 88(6), 2041 (2009)
6. Zellerhoff, M., Scholz, B., Ruehrnschopf, E.-P., Brunner, T.: Low contrast 3D reconstruction from C-arm data. In: *Medical Imaging: Physics of Medical Imaging*, vol. 5745. SPIE (2005)
7. Zheng, Y., John, M., Liao, R., Boese, J., Kirschstein, U., Georgescu, B., Zhou, S.K., Kempfert, J., Walther, T., Brockmann, G., Comaniciu, D.: Automatic Aorta Segmentation and Valve Landmark Detection in C-Arm CT: Application to Aortic Valve Implantation. In: *Medical Image Computing and Computer Assisted Intervention, MICCAI (2010)*

Cardiac Anchoring in MRI through Context Modeling

Xiaoguang Lu^{1,*}, Bogdan Georgescu¹, Marie-Pierre Jolly¹, Jens Guehring¹,
Alistair Young², Brett Cowan², Arne Littmann³, and Dorin Comaniciu¹

¹ Siemens Corporate Research, Princeton, NJ, USA
xiaoguang.lu@siemens.com

² Auckland MRI Research Group, University of Auckland, Auckland, New Zealand

³ Magnetic Resonance, Siemens Healthcare, Erlangen, Germany

Abstract. Cardiac magnetic resonance imaging (MRI) has advanced to become a powerful diagnostic tool in clinical practice. Robust and fast cardiac modeling is important for structural and functional analysis of the heart. Cardiac anchors provide strong cues to extract morphological and functional features for diagnosis and disease monitoring. We present a fully automatic method and system that is able to detect these cues. The proposed approach explores expert knowledge embedded in a large annotated database. Exemplar cues in our experiments include left ventricle (LV) base plane and LV apex from long-axis images, and right ventricle (RV) insertion points from short-axis images. We evaluate the proposed approach on 8304 long-axis images from 188 patients and 891 short-axis images from 338 patients that are acquired from different vendors. In addition, another evaluation is conducted on an independent 7140 images from 87 patient studies. Experimental results show promise of the proposed approach.

1 Introduction

In cardiology, precise information on both the dimensions and functions of the heart chambers is essential in clinical applications for diagnosis, prognostic, and therapeutic decisions. The precision on the measures extracted from MR images has been demonstrated and makes MR imagery a standard for left ventricle (LV) analysis [1]. Although cardiac MR imaging technologies have rapidly advanced [2,3], due to considerable amount of available data, analysis of cardiac images for quantification is time consuming and error-prone for human operators.

Typical cardiac MR studies contain both long-axis and short-axis slices. Long axis slices are not only used as scout images for acquisition planning, but also are complementary to the short axis stack [4]. Long axis slices capture heart chamber shape information and can also be used to correct mis-registration of the short axis stack. Anchoring is helpful for accurate and efficient cardiac modeling, such as initialization of deformable model based approaches [5], accelerating

* Corresponding author.

acquisition time by facilitating fully automatic planning of cardiac MR examinations, and accurate assessment of mass and volume [6], where demarcation of the base is important [7]. For example, anchoring the base plane and the apex in long-axis images facilitates and accelerates the LV segmentation of the short stack. Anchoring RV insertion (intersection between RV outer boundary and the LV epicardium) helps analyze LV functions according to AHA myocardial segmentation models [8]. Each anchoring component can be of different geometric representations, e.g., a line segment for the base plane and a point for the apex.

We propose a unified approach to detecting anchoring components. Anchoring components are converted into parameterized bounding box representations, which fit into an object detection framework. Such representation embeds not only individual anchoring components but also their context, which contains rich information to distinguish the anchoring components from its background and other anatomical structures. We apply a learning-based method to train detectors on expert annotations in order to handle complex appearance and heterogeneous characteristics of anatomical features in medical images, as the complex prior knowledge is implicitly encoded. Learning based object detection approaches have been demonstrated successful in many applications [9,10].

The proposed approach provides a large flexibility to be applied to a wide range of anatomical structures. We apply our framework to detect LV base mitral valve plane and LV apex in long-axis images, and RV insertions and RV lateral (a point where the RV outer boundary changes directions significantly within the image) in short-axis images, as shown in Fig. 1. Our approach is fully automated.

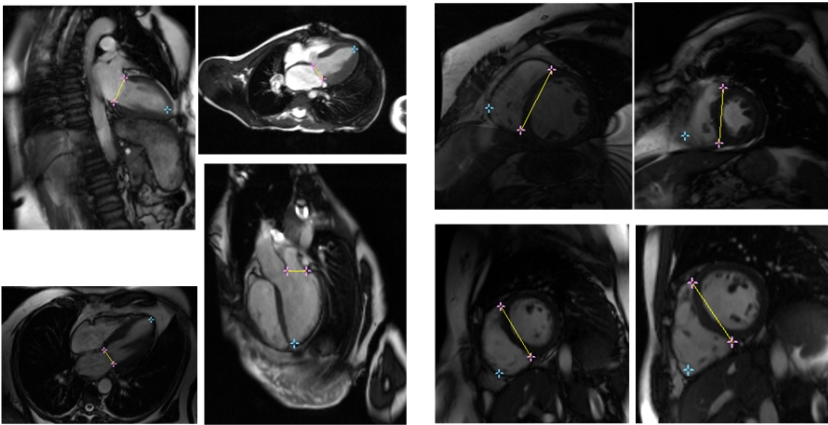


Fig. 1. Examples of cardiac images and associated anchoring components of interest

2 Methodology

We propose a unified framework to anchor the anatomy of interest. We convert detection of different types of anchoring components, such as base plane (a line)

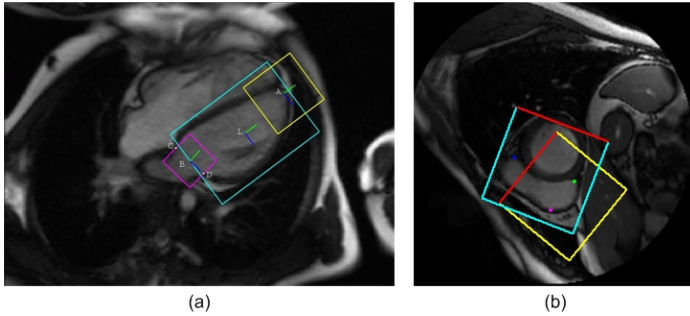


Fig. 2. Context construction and conversion from anchoring components to objects represented by a parameterized bounding box. Base plane (pink) and apex (yellow) are shown in a long-axis image (a). RV insertion (cyan) and RV lateral (yellow) are illustrated in a short-axis image (b). Notice that each bounding box is parameterized by its positions, orientation (green axis as local x-axis in (a), and red edge indicating orientation in (b)), and scales. By adjusting these parameters, a different context around the anchor component is selected.

and apex (a point) for LV in each long-axis image, into the same object detection framework by designing a contextual representation for each anchoring component. Each anchoring component is represented by a bounding box as an object with 5 parameters (2 translations, 1 orientation, and 2 scales), as shown in Fig. 2. A probabilistic learning approach [11] is applied to solve a two-class classification task, i.e., object vs. background. In order to reduce computational cost of searching through a large 5-dimensional parameter space, we adopt the marginal space search strategy proposed in [12].

2.1 Context Learning

A 2D object (bounding box) parameter set consists of five degrees of freedom. Exhaustively searching in this 5-dimensional space is prohibitive for online applications. Therefore, we adopt the marginal space search strategy, where we design a series of detectors that estimate plane parameters at a number of sequential stages in the order of complexity, i.e., translation, orientation, and scale, as the parameter degrees of freedom increase [12]. Different stages utilize different image features. Multiple hypotheses are maintained between algorithm stages, which quickly removes false hypotheses at the earlier stages while propagating the right hypotheses to the final stage. Only one hypothesis is consolidated as the final detection result.

We use a probabilistic boosting tree [11] for each detector to achieve a discriminative task between the object and background. The classifier is a tree-based structure with which the posterior probabilities of the presence of the object of interest are calculated from given image data. Therefore, each detector not only provides a binary decision for a given sample, but also a confidence value associated with the

decision. The nodes in the tree are constructed by a combination of simple classifiers using boosting techniques [11].

Each detector selects a set of discriminative features that are used to distinguish the object from background from a large pool of features. For the classifiers at the translation stage, we choose Haar wavelet-like features [9], which are efficiently calculated using integral image-based techniques. For the classifiers at the orientation and scale stages, steerable features [12] are applied, because their computation does not require image rotation and re-scaling, which are computationally expensive, especially when the hypothesis search space is large.

2.2 Context Modeling in Long-Axis Slices

For each long-axis image, two anchoring components are targets of interest, the base plane and the apex. We associate a two-dimensional bounding box with each contextual object around the anchoring component. Each bounding box is specified by a five-parameter set θ , containing two positions $\langle x, y \rangle$, one orientation $\langle \phi \rangle$, and two scales $\langle s_x, s_y \rangle$. Although only positions are mostly in use, orientation and scales are useful in encoding proper and consistent context learned during offline training process, where a set of contextual models/classifiers are obtained.

We collect a set of cardiac long-axis images and annotate the anchoring components to learn contextual models. Based on this annotated training set, we build a contextual model for each target object. For long-axis images, a joint contextual model [13] for the pair of $\langle \text{apex, base plane} \rangle$ is also constructed as shown by the cyan box in Fig. 2, which is used for inter-anchoring validation. Let $\langle x_a, y_a \rangle$, $\langle x_{b1}, y_{b1} \rangle$, and $\langle x_{b2}, y_{b2} \rangle$ denote the positions of the apex, and two basal annulus points, respectively. The contextual parameter set for the base plane is: positions $\{(x_{b1} + x_{b2})/2, (y_{b1} + y_{b2})/2\}$; orientation $\{\text{orthogonal to the line segment connecting the two basal annulus points, and pointing toward the apex side}\}$; and scales $\{s_b, s_b\}$, where $s_b = \sqrt{(y_{b2} - y_{b1})^2 + (x_{b2} - x_{b1})^2} * \alpha$, where α is a factor that can be used to adjust the contextual range set to 2.4 in our experiments. Selection of α is a tradeoff between rich context and noise. For the apex, the context parameters are constructed as: positions $\{x_a, y_a\}$; orientation $\{\arctan((y_a - (y_{b1} + y_{b2})/2)/(x_a - (x_{b1} + x_{b2})/2))\}$; and scales $\{s_a, s_a\}$, where $s_a = \sqrt{((y_{b2} - y_{b1})^2 + (x_{b2} - x_{b1})^2)} * \alpha$.

2.3 Context Modeling in Short-Axis Slices

For each short-axis image, two anchoring components are targets of interest, the RV insertion and the RV lateral. For RV insertion, the anterior and posterior anchors are identified through the following context modeling. Let $\langle x_{an}, y_{an} \rangle$, $\langle x_{po}, y_{po} \rangle$, and $\langle x_{La}, y_{La} \rangle$ denote the positions of the RV insertion anterior, RV insertion posterior, and RV lateral, respectively. The contextual parameter set for RV insertion is: positions $\{(x_{an} + x_{po})/2, (y_{an} + y_{po})/2\}$; orientation $\{\text{orthogonal to the line segment connecting RV insertion anterior and RV insertion posterior, and pointing toward the LV}\}$; and scales $\{s_{ins}, s_{ins}\}$,

where $s_{ins} = \sqrt{(x_{an} - x_{po})^2 + (y_{an} - y_{po})^2} * \beta$. β is set to 1.5 in our experiments. For RV lateral, in order to utilize its RV context, we constructed the context parameters as: positions $\{(x_{La} + x_{po})/2, (y_{La} + y_{po})/2\}$; orientation {orthogonal to the line segment connecting RV Lateral and RV insertion posterior, and pointing toward RV insertion anterior side}; and scales $\{s_{La}, s_{La}\}$, where $s_{La} = \sqrt{((x_{La} - x_{po})^2 + (y_{La} - y_{po})^2)} * \gamma$. γ is set to 2.4 in our experiments.

3 Experiments

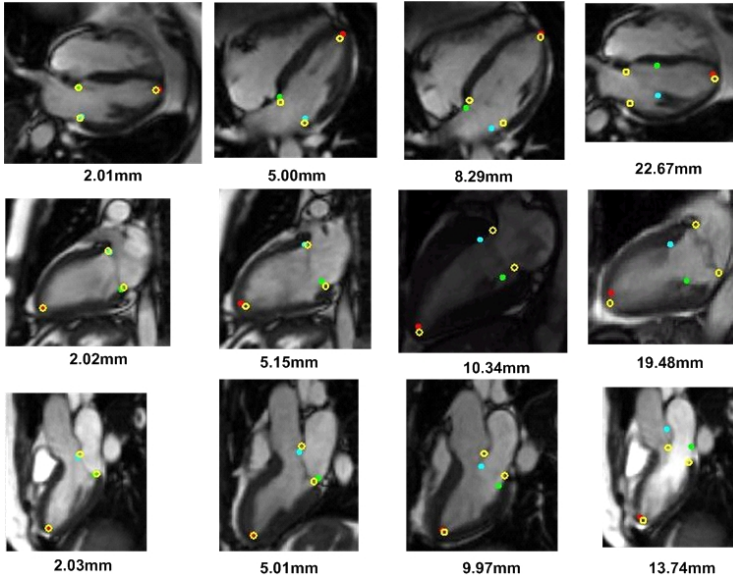
We collected 490 long-axis sequences from 188 patients, whose ages ranged from 11 to 72 years old. In total, 8304 images were used to construct our long-axis image database. Long-axis image can contain different chamber views, namely, 4-chamber, 3-chamber, and 2-chamber views. For each image, the base plane (two annulus anchors) and the apex of the LV were manually annotated by experts and used as ground truth for evaluation. Our short-axis database contains 891 images from 756 sequences of 338 patients, which were provided by two different vendors, each providing 296 and 42 patient studies, respectively. For short-axis images, the ground truth positions of the two RV insertion anchors and the RV lateral anchor were annotated for evaluation purposes.

We applied our context learning algorithm to detect the objects of ‘base plane’, ‘apex’, ‘RV insertion’, and ‘RV lateral’ on respective long-axis and short-axis images for anchoring purposes. The long-axis and short-axis images can be distinguished using orientation information captured during acquisition. With each object detected, corresponding anchors are inferred based on reverse object/context modeling process, i.e., calculating anchor positions from the detected parameterized bounding box. We computed Euclidean distance between the detected anchor position and its corresponding ground truth as the detection error for each anchor. The average distance of all anchors in each image was used as the metric to evaluate the overall system performance.

A 4-fold cross-validation scheme was applied for evaluation. The entire database was randomly partitioned into four quarters. For each fold evaluation, three quarters were combined for training and the remaining one quarter was used as unseen data for testing. This procedure was repeated four times so that each image was used once for testing. Performance is summarized based on all 4 folds and provided in Table 1 and Figs. 4(a) and 4(b). Fig. 3 shows examples of the detection results along with the calculated distance metrics to provide a visual correlation between the distance and quality of detection. In addition to the large parameter search space, cardiac MR images in a large population present a large variation of appearance intensities along with the anatomy shape changes across the heart beat cycle, leading to difficulties for accurate identification. The significant performance difference between RV insertion and RV lateral is due to lack of consistent definition of RV lateral. On the average, it took about 1.5 seconds to detect the base plane and the apex on a 400×400 long-axis image on a duo core 2.8GHz CPU, and 0.5 seconds to detect RV insertion and RV lateral anchors on a 256×256 short-axis image.

Table 1. Average distance of all detected anchors from ground truth positions by a 4-fold cross validation. Distances are in unit of *mm*.

(a) Long-axis				(b) Short-axis			
	Mean	Std	Median		Mean	Std	Median
Overall	4.9	7.0	3.8	Overall	6.7	15.9	4.8
Baseplane	5.1	6.8	3.7	RV insertion	5.9	16.0	3.9
Apex	4.5	6.7	3.3	RV lateral	8.4	16.5	5.9

**Fig. 3.** Examples of anchoring results along with annotated ground truth (yellow). The average detection distance is provided below each image. Our mean and median distances are 5.5mm and 3.8mm, respectively.

In addition, we were able to conduct an independent evaluation on another large expert validated database, which is collected as a mixture of multiple vendors and called ONTARGET database. Patients with vascular disease at high risk of cardiac events were imaged as part of the ONTARGET MRI substudy, described in [14]. Standardized cardiac MRI exams were performed in six countries around the world using Siemens, Philips and GE scanners. Either prospectively or retrospectively gated steady state free precession (SSFP) CMR cines were acquired in six equally spaced short axis (SA) locations from apex to base. Typical imaging parameters were TR/TE/flip/FOV = 30ms/1.6ms/60°/360mm, slice thickness 6mm, image matrix 256×208. There were typically 25 temporal cine frames per slice, depending on the heart rate. All cines were acquired during breath-holding of 815 seconds duration. Most patients had coronary heart disease (87%), myocardial infarction (58%), and/or hypertension (61%). 23% were

Table 2. Average distance of the detected two base plane anchors from ground truth positions from the independent evaluation. Distances are in unit of *mm*.

	Mean	Std	Median
Long-axis	5.2	7.8	3.9

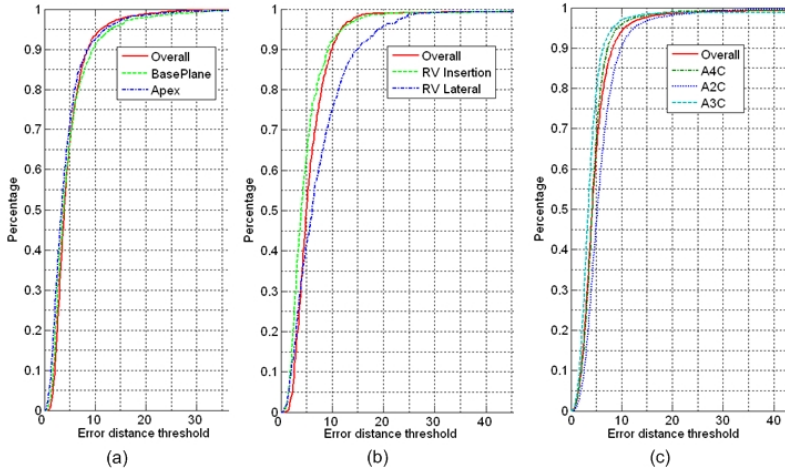


Fig. 4. Percentile evaluation results of 4-fold cross-validation. Each curve shows the percentage of cases whose distance of the automatic detection results from ground truth is less than an error distance threshold. (a) Long-axis anchoring evaluation. (b) Short-axis anchoring evaluation. (c) ONTARGET baseplane evaluation along with performance breakdown based on the chamber views.

female and 36% were Asian. The entire database contains 7140 long-axis images from 87 patient studies. Only base plane anchoring is evaluated. The experimental results are presented in Table 2 and Fig. 4(c). A performance breakdown based on the chamber views is reported in Fig. 4(c). For this evaluation, all original 8304 images were used for training. Then the system was evaluated on the independent 7104 images. Overall, the independent evaluation results are consistent with the internal cross-validation results, both showing promise and robustness of the proposed approach.

4 Conclusions

We have proposed a unified approach to explore contextual information and integrated it with a learning-based object detection framework. We have developed a fully automatic system for cardiac anchoring in both MR long-axis and short-axis images. The principle of the proposed approach is generic. With concrete design, which is application-specific, the proposed approach is able to be applied to a wide range of applications in addition to acquisition planning and cardiac segmentation.

References

1. Cousty, J., Najman, L., Couprie, M., Clement-Guinaudeau, S., Goissen, T., Garot, J.: Automated, accurate and fast segmentation of 4D cardiac MR images. In: Sachse, F.B., Seemann, G. (eds.) FIHM 2007. LNCS, vol. 4466, pp. 474–483. Springer, Heidelberg (2007)
2. Frangi, A., Niessen, W., Viergever, M.: Three-dimensional modeling for functional analysis of cardiac images: A review. *IEEE Trans. on Medical Imaging* 20(1), 2–25 (2001)
3. Finn, J.P., Nael, K., Deshpande, V., Ratib, O., Laub, G.: Cardiac MR imaging: State of the technology. *Radiology* 241(2), 338–354 (2006)
4. Koikkalainen, J., Pollari, M., Lotjonen, J., Kivisto, S., Lauerma, K.: Segmentation of cardiac structures simultaneously from short- and long-axis MR images. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) MICCAI 2004. LNCS, vol. 3216, pp. 427–434. Springer, Heidelberg (2004)
5. Jolly, M.P.: Automatic segmentation of the left ventricle in cardiac MR and CT images. *International Journal of Computer Vision* 70(2), 151–163 (2006)
6. Young, A., Cowan, B., Thrupp, S., Hedley, W., Dell’Italia, L.: Left ventricular mass and volume: Fast calculation with guide-point modeling on MR images. *Radiology* 216(2), 597–602 (2000)
7. Weaver, A., Jones, H., Jelaco, G., Cha, J., Hall, J., Walker, K., Blatter, D., Anderson, J.: Magnitude and causes of interobserver discrepancies in CMR volume measurements: critical importance of choice of the basal slice. *Journal of Cardiovascular Magnetic Resonance* 8, 82–83 (2006)
8. Cerqueira, M., Weissman, N., Dilsizian, V., Jacobs, A., Kaul, S., Laskey, W., Pennell, D., Rumberger, J., Ryan, T., Verani, M.: Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. *Circulation* 2002, 105–539 (2002)
9. Viola, P., Jones, M.J.: Robust real-time face detection. *International Journal of Computer Vision* 57(2), 137–154 (2004)
10. Georgescu, B., Zhou, X., Comaniciu, D., Gupta, A.: Database-guided segmentation of anatomical structures with complex appearance. In: Proc. IEEE CVPR (2005)
11. Tu, Z.: Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering. In: Proc. ICCV, pp. 1589–1596 (2005)
12. Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., Comaniciu, D.: Fast automatic heart chamber segmentation from 3D CT data using marginal space learning and steerable features. In: Proc. ICCV (2007)
13. Lu, X., Georgescu, B., Littmann, A., Mueller, E., Comaniciu, D.: Discriminative joint context for automatic landmark set detection from a single cardiac MR long axis slice. In: Ayache, N., Delingette, H., Sermesant, M. (eds.) FIMH 2009. LNCS, vol. 5528, pp. 457–465. Springer, Heidelberg (2009)
14. Cowan, B., Young, A., Anderson, C., Doughty, R., Krittayaphong, R., Lonn, E., Marwick, T., Reid, C., Sanderson, J., Schmieder, R., Teo, K., Wadham, A., Worthley, S., Yu, C., Yusuf, S., Jennings, G.: Left ventricular mass and volume with telmisartan, ramipril, or combination in patients with previous atherosclerotic events or with diabetes mellitus (from the ONgoing Telmisartan Alone and in combination with Ramipril Global Endpoint Trial [ontarget]). *American Journal of Cardiology* 104(11), 1484–1489 (2009)

Real-Time Respiratory Motion Correction for Cardiac Electrophysiology Procedures Using Image-Based Coronary Sinus Catheter Tracking

YingLiang Ma¹, Andy P. King¹, Nicolas Gogin², C. Aldo Rinaldi³,
Jaswinder Gill³, Reza Razavi¹, and Kawal S. Rhode¹

¹ Division of Imaging Sciences, King's College London, SE1 7EH, UK

² Medisys Research Laboratory, Philips Healthcare, Paris, France

³ Department of Cardiology, Guy's & St. Thomas' Hospitals NHS Foundation Trust, London,
SE1 7EH, UK

y.ma@kcl.ac.uk

Abstract. X-ray fluoroscopically guided cardiac electrophysiological procedures are routinely carried out for diagnosis and treatment of cardiac arrhythmias. X-ray images have poor soft tissue contrast and, for this reason, overlay of static 3D roadmaps derived from pre-procedural volumetric data can be used to add anatomical information. However, the registration between the 3D roadmap and the 2D X-ray data can be compromised by patient respiratory motion. We propose a novel method to correct for respiratory motion using real-time image-based coronary sinus (CS) catheter tracking. The first step of the proposed technique is to use a blob detection method to detect all possible catheter electrodes in the X-ray data. We then compute a cost function to select one CS catheter from all catheter-like objects. For correcting respiratory motion, we apply a low pass filter to the 2D motion of the CS catheter and update the 3D roadmap using this filtered motion. We tested our CS catheter tracking method on 1048 fluoroscopy frames from 15 patients and achieved a success rate of 99.3% and an average 2D tracking error of $0.4 \text{ mm} \pm 0.2 \text{ mm}$. We also validated our respiratory motion correction strategy by computing the 2D target registration error (TRE) at the pulmonary veins and achieved a TRE of $1.6 \text{ mm} \pm 0.9 \text{ mm}$.

1 Introduction

Cardiac electrophysiological (EP) procedures are traditionally carried out under X-ray fluoroscopic guidance to diagnose and treat cardiac arrhythmias. However, X-ray images have poor soft tissue contrast and it is difficult to interpret the anatomical context directly from these images. To overcome the lack of soft tissue contrast, a three-dimensional (3D) roadmap can be generated from 3D high-resolution computed tomography (CT)/ magnetic resonance images (MRI), registered and overlaid in real-time with X-ray fluoroscopy images [1]. Currently, the 3D roadmap remains static and does not move with the patient's respiratory motion. In some cases, respiratory motion can cause a two-dimensional (2D) registration error of over 14 mm [2], which

is a significant compromise in the accuracy of guidance. A number of groups have previously addressed the issue of respiratory motion correction for cardiac interventions. Motion-compensated navigation for coronary interventions based on magnetic tracking was suggested in [3], but it required additional special hardware. Several image-based approaches have been developed that use only information from the X-ray fluoroscopic images themselves. Shechter et al. [4] constructed a model of cardiac and respiratory motion of the coronary arteries from biplane contrast-enhanced X-ray image sequences. The model was applied by tracking the motion of the diaphragm in subsequent (non-enhanced) X-ray images. However, forming the model from X-ray images under contrast injection means that it will be constructed from a limited amount of data. Furthermore, the diaphragm is not always in the X-ray field of view, particularly for obese patients. Brost et al. [5] developed an image-based respiratory motion correction method for EP procedures by tracking the 3D position of a lasso catheter from biplane X-ray images. Unlike tracking the diaphragm, this method directly tracks an instrument very close to the target region of the EP procedure. However, it also has some limitations. Firstly, the lasso catheter is particular for only a subset of EP procedures and it does not always remain stationary inside the heart. Secondly, the majority of X-ray systems are monoplane systems. Finally, the maximum frame rate of the lasso catheter tracking was only 3 frames per second and the tracking method required manual initialization.

In this paper, we present a novel method to track the coronary sinus (CS) catheter from X-ray images in real-time. The reason we choose the CS catheter is that it is almost ubiquitously present during EP procedures. The CS catheter has several electrodes which are highly visible in normal dose and low dose X-ray images. Furthermore, the CS catheter remains in place throughout the procedure, its position is not routinely altered and it is normally not close to other catheters. Our technique works by first using a fast multi-scale blob detection method to detect all possible electrode-like objects in the X-ray image. The main novelty of our approach is that we then use prior knowledge of the CS catheter geometry to identify a single CS catheter from all of these candidate objects. This allows us to quickly and reliably distinguish the CS catheter from other catheters and instruments in the X-ray image. Our tracking method uses the whole image as a region of interest so that we are able to track the CS catheter even in the presence of large sudden motion. Furthermore, the proposed method does not require any user interaction. We apply the technique to correct respiratory motion during cardiac EP procedures by applying filtered 2D in-plane translational motion to the 3D roadmap.

2 Method

CS catheters are often used as diagnostic catheters and one of most commonly used CS catheters is the 10-electrode catheter. The arrangement of electrodes has two variations. One is evenly distributed and the other is paired (figure 1a). The presence of 10 electrodes is the unique feature which we use to separate the CS catheter from other catheters, instruments and ECG leads. The first step of our proposed method is to run a blob detection method to detect all possible electrode-like objects.

2.1 Blob Detection

We use a multi-scale blob detector based on the determinant of the Hessian matrix [6]. We consider a continuous function of two variables such that the value of the function at (x, y) is given by $f(x, y)$. Then the determinant of the Hessian matrix is defined as

$$\det H(f(x, y)) = f_{xx}f_{yy} - f_{xy}^2 \tag{1}$$

where f_{xx} , f_{yy} and f_{xy} are the second-order partial derivatives of the function $f(x, y)$. Blobs can be defined as local extrema of the determinant of the Hessian matrix. To translate this theory to work with 2D images, we first replace the function $f(x, y)$ with discrete image pixel intensities $I(x, y)$. Then we calculate the derivatives by convolving the image with second order Gaussian derivative masks (figure 1b).

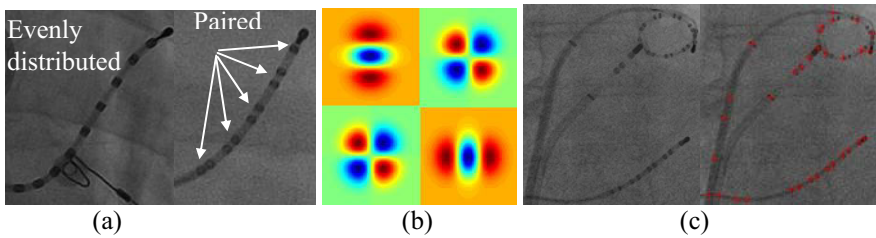


Fig. 1. a) Two variations of 10 electrode CS catheter. b) Gaussian derivative masks. c) The result from the blob detection method. Red crosses are the positions of electrode-like blobs.

To introduce a scale factor t , we first introduce the scale-space representation $L(x, y; t)$ of the image $I(x, y)$ by $L(x, y; t) = I(x, y) * g(x, y; t)$, where $*$ is the convolution operator and the Gaussian filter $g(x, y; t)$ is defined as $g(x, y; t) = \frac{1}{2\pi} e^{-(x^2+y^2)/2t}$.

Here scale factor $t = \sigma^2$ where σ is the standard deviation of the Gaussian function. Now we can derive the scale-normalized determinant of the Hessian from eq. (1) and we get

$$\det H(L(x, y; t)) = t^2(L_{xx}L_{yy} - L_{xy}^2) \tag{2}$$

Here $L_{xx} = I(x, y) * g_{xx}$ where

$$g_{xx}(x, y; t) = -\frac{1}{2\pi^2} \left(1 - \frac{x^2}{t}\right) e^{-(x^2+y^2)/2t} \tag{3}$$

Similar definitions apply for L_{yy} and L_{xy} . These Gaussian derivatives (g_{xx} , g_{yy} and g_{xy}) are often known as Laplacian of Gaussians (LoG). In practice, we just pre-compute the masks of these Gaussian derivatives (figure 1b), convolve with the input image and calculate the determinant of the Hessian matrix using eq. (2). Blobs

are detected as regional maxima of the determinant of the Hessian matrix and we also define the strength of the blob as the normalized value of the determinant of the Hessian matrix.

Finally, we need to decide the scale factors for detecting all electrodes on the CS catheter. As shown in figure 1a, there are only two sizes of electrodes. One is a big catheter tip electrode and the others are small electrodes. Therefore, two fixed scale factors are sufficient which could lead to a more efficient and robust implementation than other automatic scale selection methods. To calculate the scale factor t , we use $t = ((s-1)/3)^2$ where s is the size of blob. This equation is motivated by the “ 3σ ” ($t = \sigma^2$) rule that 99% of energy of the Gaussian is within three standard deviations and was empirically tested for our application. Figure 1c demonstrates the result of blob detection in a low-dose X-ray image. The strength of the blob is dependent on the blob size. The value of the blob size, s , was set to 6 and 3 pixels for the catheter tip electrode and the smaller electrodes, respectively.

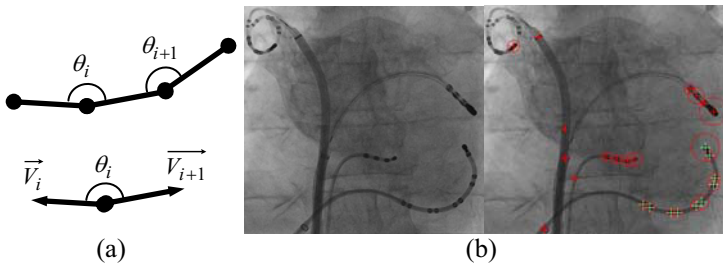


Fig. 2. a) Definition and calculation of the deviation angle. b) The result from the CS catheter detection method. Green crosses are the positions of CS catheter electrodes. Red crosses are the positions of other catheter electrodes. The size of the red circles represents the strength of the blobs.

2.2 Catheter Detection

The next step involves deciding which combination of blobs represents the CS catheter. We select the 50 highest strength blobs from the blob detection method. Each blob is connected to its nearest neighbour. Based on the shape of the 10-electrode CS catheter we design a cost function to estimate the likelihood of candidate catheter-like objects. Catheter-like objects should have a large blob as the catheter tip and several smaller blobs after it. As we know there are two variations of the arrangement of electrodes (figure 1a). We designed a unified cost function. The cost function consists of two parts. One is the smooth variation of deviation angle between two consecutive edges (figure 2a). The other is the smooth variation of the strength difference between two consecutive blobs except the catheter tip electrode. We exclude the catheter tip electrode from the calculation of the cost function because the strength of the response of the blob detector is dependent on the blob size. Therefore, the catheter tip electrode would have a higher weight in the cost function. The cost function is defined as follows

$$CSCathCost = \frac{\sum_{i=1}^{N-2} |\cos\theta_i - \overline{\cos\theta}|}{N-2} + \frac{\sum_{i=2}^N |Blob_i - \overline{Blob}|}{N-1} \quad (4)$$

where $\overline{\cos\theta}$ is the mean of cosines of all deviation angles and \overline{Blob} is the mean of blob strengths of the 9 electrodes (excluding the catheter tip electrode). $Blob_i$ is normalized to a range of zero to one which is the same value range as $\cos\theta_i$. N is the total number of electrodes ($N=10$ in our case). $\cos\theta_i$ can be computed efficiently using $\cos\theta_i = \vec{V}_i \cdot \vec{V}_{i+1} / (\|\vec{V}_i\| \cdot \|\vec{V}_{i+1}\|)$, where \cdot is the dot product and $\|\cdot\|$ is the vector length. \vec{V}_i and \vec{V}_{i+1} are defined in figure 2a.

To reduce the number of catheter candidates, we introduce two constraints to quickly remove unwanted candidates. The first constraint is the maximum electrode gap. Currently, we define it as twice the length of the maximum distance Max_{dist} between two neighbouring CS catheter electrodes (we obtained Max_{dist} by physical measurement of a CS catheter). 10-electrode CS catheters have two variations of electrode distributions and we choose whichever maximum distance between two neighbour electrodes is larger. The second constraint is the minimum deviation angle which must be larger than 90 degrees as catheters in human vessels could not have sharp turns. So, this condition can be translated into $\vec{V}_i \cdot \vec{V}_{i+1} < 0$. The overall algorithm for detecting one 10-electrode CS catheter is presented in Algorithm 1. Figure 2b gives an example CS catheter detection result in one clinical X-ray image.

Algorithm 1. 10-electrode CS catheter detection Algorithm

- 1) Run blob detection method to find 50 highest strength blobs, put them into a Global Blob List and select 10 highest strength blobs from the Global Blob List.
 - 2) For each blob B in 10 highest strength blobs
 - 3) Remove blob B from Global Blob List.
 - 4) Create an empty Catheter Blob List and store blob B into the list.
 - 5) Find the nearest neighbour blob NB from the remaining blobs in the Global Blob List.
 - 6) If $Dist_{B \rightarrow NB} > 2 * Max_{dist}$ Goto Step 2 EndIf
 - 7) If the number of blobs in Catheter Blob List larger than 2.
 - 8) If $\vec{V}_i \cdot \vec{V}_{i+1} > 0$, Goto Step 2. EndIf
 - 9) EndIf
 - 10) Store NB into Catheter Blob List, Remove NB from Global Blob List and $B=NB$
 - 11) If the number of blobs in Catheter Blob List less than 10 Goto Step 5.
 - 12) End of for loop
 - 13) Compute the cost function using eq. (4) for all candidate catheters and select the highest one as the CS Catheter.
-

2.3 Motion Correction

We apply our CS detection method to correct for respiratory motion in 2D fluoroscopic image sequences during cardiac EP procedures. CS catheters have been previously used in computing the initial registration for image-guided cardiac EP procedures [7] because they are known to be relatively stable in position with respect to the anatomy. Similarly, our approach tracks the position of the proximal electrode of the 10-electrode CS catheter to use for respiratory motion correction. The reason for choosing this electrode (the distal electrode is the catheter tip) is that the proximal electrode has less cardiac cycle motion and has lower tracking errors than the distal electrode (see section 3.1). For a monoplane X-ray system, we first apply a low pass filter to the 2D translational motion of the proximal electrode to reduce cardiac motion further and then apply the filtered 2D translational motion directly to the 3D roadmap.

3 Results

We tested the proposed CS catheter tracking method on 1048 clinical X-ray images. There were a total of 18 different clinical fluoroscopy sequences which came from 15 clinical cardiac catheterization cases. 28% of the clinical X-ray images that we tested were low dose and contained high frequency noise. 2 clinical cases (136 images) used evenly distributed 10-electrode CS catheters and the other 13 cases (912 images) used paired 10-electrode CS catheters. Typically, the clinical X-ray images contained one CS catheter, one ablation catheter and sometimes one lasso catheter. Some X-ray fluoroscopy images, which were acquired during pace-maker implant cases, contained ECG electrodes and wires, and also sternal wire loops.

3.1 Accuracy, Robustness and Efficiency Test

Table 1. 2D electrodes detection errors in mm (excluding partial failure images and electrode 1 is the distal electrode)

Electrode	1	2	3	4	5	6	7	8	9	10
Low dose images	0.4	0.6	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
Normal dose images	0.4	0.8	0.3	0.4	0.3	0.3	0.4	0.4	0.3	0.3

To test the accuracy of our method, we asked two clinical experts to manually pick the centre of all electrodes of the CS catheter on all 1048 images. Then we calculated the 2D distance between the manually defined centres and the centres detected by our method for all 10 electrodes. Measurements were made in pixel space and then converted to mm space using the magnification factor available from the DICOM header of the X-ray data. Overall, we achieved a 2D detection error of $0.39 \text{ mm} \pm 0.22 \text{ mm}$ for all electrodes in all images and we obtained a very similar error of $0.36 \text{ mm} \pm 0.21 \text{ mm}$ for low dose X-ray images only. The errors of individual electrodes are given in table 1. It is noticeable that electrode No 2 has larger errors than others. This is

because the second electrode is very close to the catheter tip (the distal electrode) and it could be detected together with the catheter tip as a single large blob.

Our method is robust in low dose images as the errors in low dose images were very similar to the errors in normal dose images. We only observed the partial failure of detection in 8 frames of normal dose images because one of electrodes was occluded by ECG electrodes. We classify a partial failure when the detection method correctly detected the CS catheter but any of electrodes had an error of more than 5 mm. The choice of 5mm is motivated by the size of the smallest target structures for EP procedures (the pulmonary veins, approximately 5mm in radius). There were no complete failure cases which did not detect the correct CS catheter. The overall success rate of detection was 99.3%.

The major computational load of our method is the blob detection algorithm. We implemented it on the Intel Integrated Performance Primitives Library for fast image convolution calculation. Our method currently achieves a frame rate of 21 frames-per-second using a single-threaded CPU implementation. The performance was evaluated on an Intel Core 2 Duo 2.0GHz laptop with an nVidia Quadro FX 350M graphics card.

3.2 Motion Correction Results

The intended application of our CS catheter tracking technique is to update the position of a 3D roadmap. Therefore, as further validation we computed the target registration error (TRE) at the pulmonary veins (PVs), which are the main structures of interest in many EP procedures. For the cases in which the lasso catheter was used for validation of accuracy, it remained stable in one of the pulmonary veins for all the X-ray frames evaluated. We applied the computed 2D translation determined from the CS catheter tracking to the position of the lasso catheter which acts as a surrogate for the position of the PVs since it is rigidly placed within these structures during the procedure. The TRE was then computed as the distance error between this predicted position and the actual position of the lasso catheter in the X-ray data. We calculated the TRE at the PVs on 565 fluoro images (7 patients) and achieved a 2D accuracy of $1.6 \text{ mm} \pm 0.9 \text{ mm}$. The TRE before motion correction was $6.1 \pm 1.8 \text{ mm}$. Figure 3b

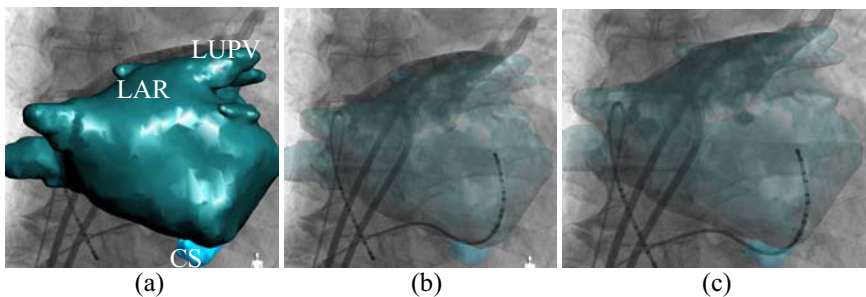


Fig. 3. Illustrates an MRI overlay on to x-ray fluoroscopy. The initial registration was done using multiple view X-ray acquisitions followed by manual alignment. a) b). Frame of a non-motion compensated sequence with fluoroscopy overlay. c) Motion compensated fluoroscopy overlay.

presents an example of a fluoroscopy overlay with respiratory motion correction. In the 3D roadmap, LAR is the left atrium roof and LUPV is the left upper pulmonary vein. From figure 3a and 3b, there is a noticeable mis-alignment between the 3D roadmap and the X-ray angiography, particularly in the area of LAR and LUPV. The roadmap and X-ray angiography are well matched and the CS is aligned with the CS catheter in figure 3c.

4 Conclusion and Discussions

We have developed an accurate and robust method for real-time CS catheter tracking in X-ray fluoroscopy images and applied it to correct respiratory motion in X-ray fluoroscopy image-guided EP procedures. Our tracking method remains robust and accurate even in low dose fluoroscopy images as CS catheter electrodes remain highly visible. Because the majority of normal dose images we tested had fast-changing background contrast (contrast agent injection), our results showed less tracking errors in low dose images than normal dose ones as the accuracy of blob detection is influenced by contrast agent injection. We achieved a sub-millimeter accuracy of our method as the average tracking error for all CS catheter electrodes is 0.4 mm and maximum error is 0.8 mm. Updating the 3D roadmap by the filtered 2D motion of the CS catheter can significantly improve the accuracy of fluoroscopy overlays for cardiac EP procedures. Our CS catheter tracking method has several advantages. First, it is real-time so that as well as being used to detect respiratory motion it could potentially also be applied to the detection of the much faster cardiac cycle motion. Secondly, it does not require any user interaction and can detect the CS catheter position without defining a region of interest in the X-ray image. Finally, it has potential application in more types of cardiac catheterization procedure, rather than only for EP procedures. Other potential applications of our CS catheter detection method that we plan to investigate include gating cardiac and respiratory motion in 3D C-Arm CT sequences to improve the quality of the reconstructed 3D volumes. Also, we plan to investigate its use for 3D real-time registration between the CS catheter and the main branch of the CS geometry in the 3D roadmap. This could lead to a real-time solution to the problem of bulk patient motion. Because our CS catheter tracking method is based on blob detection, it can be extended to simultaneously track a CS catheter and an ablation catheter in real-time. This opens the possibility to design a system to automatically tag ablation positions with motion correction in real-time.

References

1. Rhode, K.S., Hill, D.L.G., Edwards, P.J., Hipwell, J., Rueckert, D., Sanchez-Ortiz, G., Hegde, S., Rahunathan, V., Razavi, R.: Registration and tracking to integrate X-ray and MR images in an XMR facility. *IEEE Transactions on Medical Imaging* 24(11), 810–815 (2003)
2. King, A.P., Boubertakh, R., Rhode, K.S., Ma, Y.L., Chinchapatnam, P., Gao, G., Tangcharoen, T., Ginks, M., Cooklin, M., Gill, J.S., Hawkes, D.J., Razavi, R.S., Schaeffter, T.: A subject-specific technique for respiratory motion correction in image-guided cardiac catheterisation procedures. *Med. Image Anal.* 13(3), 419–431 (2009)

3. Timinger, H., Krueger, S., Dietmayer, K., Borgert, J.: Motion Compensated Coronary Interventional Navigation by Means of Diaphragm Tracking and Elastic Motion Models. *Phys. Med. Biol.* 50(3), 491–503 (2005)
4. Shechter, G., Shechter, B., Resar, J.R., Beyar, R.: Prospective motion correction of X-ray images for coronary interventions. *IEEE Transactions on Medical Imaging* 24(4), 441–450 (2005)
5. Brost, A., Liao, R., Hornegger, J., Strobel, N.: 3-D respiratory motion compensation during EP procedures by image-based 3-D lasso catheter model generation and tracking. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009*. LNCS, vol. 5761, pp. 394–401. Springer, Heidelberg (2009)
6. Tony, L.: Detecting salient blob-like image structures and their scales with a scale-space primal sketch: a Method for focus-of-attention. *International Journal of Computer Vision* 11(3), 283–318 (1993)
7. Sra, J., Krum, D., Belanger, B., Vaillant, R.: Registration of three dimensional left atrial computed tomographic images with fluoroscopy. *Heart Rhythm* 2(9), 1020 (2005)

Accurate Segmentation of the Left Ventricle in Computed Tomography Images for Local Wall Thickness Assessment

J. Peters^{1,*}, J. Lessick², R. Kneser¹, I. Wächter², M. Vembar³,
O. Ecabert¹, and J. Weese¹

¹ Philips Research Europe – Aachen, 52066 Aachen, Germany

jochen.peters@philips.com

² Philips Healthcare – Advanced Technology Centre, Haifa 31004, Israel

³ Philips Healthcare – CT Clinical Science, Cleveland, OH 44143, USA

Abstract. In recent years, the fully automatic segmentation of the whole heart from three-dimensional (3D) CT or MR images has become feasible with mean surface accuracies in the order of 1mm. The assessment of local myocardial motion and wall thickness for different heart phases requires highly consistent delineation of the involved surfaces. Papillary muscles and misleading pericardial structures lead to challenges that are not easily resolved. This paper presents a framework to train boundary detection functions to explicitly avoid unwanted structures. A two-pass deformable adaptation process allows to reduce false boundary detections in the first pass while detecting most wanted boundaries in a second pass refinement. Cross-validation tests were performed for 67 cardiac datasets from 33 patients. Mean surface accuracies for the left ventricular endo- and epicardium are 0.76mm and 0.68mm, respectively. The percentage of local outliers with segmentation errors $> 2\text{mm}$ is reduced by a factor of 3 as compared to a previously published approach. Wall thickness measurements in full 3D demonstrate that artifacts due to irregular endo- and epicardial contours are drastically reduced.

Keywords: cardiac wall thickness, papillary muscles, boundary detection, image segmentation, multi-slice computed tomography.

1 Introduction

Fully automatic segmentation of three-dimensional (3D) cardiac images enables efficient assessment of global and local functional cardiac parameters such as ejection fraction or wall thickness. Model-based segmentation approaches [1,2,3] enable full-heart segmentation with surface-to-surface accuracies in the order of 1mm in both computed tomography (CT) and magnetic resonance (MR) image volumes (see [4]). Shape-constrained deformable models for whole heart segmentation have been described in detail in [5]. There, a triangulated mesh

* Corresponding author.

models the endocardial surfaces of the four cardiac chambers as well as the left ventricular (LV) epicardium and short trunks of the major vasculature. The mesh is adapted towards corresponding image boundaries. External image forces are balanced by internal forces that roughly maintain the model shape. This approach reaches an average accuracy of 0.82mm on 28 CT datasets. Clinically, global parameters such as cardiac chamber volumes and myocardial mass are sufficiently accurate and automation results in appreciable time savings and excellent reproducibility [5,6,7,8].

To characterize irregularities of the heart motion, a local assessment of wall motion and wall thickness for different cardiac phases is desirable. For this task, a highly consistent segmentation of the involved surfaces is needed. For the endocardium, papillary muscles and trabeculations lead to an ambiguity that must be resolved. To measure blood volume, the segmentation should consistently follow the bloodpool border. Wall thickness measurements should be based on a convex hull around the bloodpool that excludes the papillaries from the myocardium. For the epicardium, it is important to avoid local confusions with the nearby pericardium or lung transition.

In [9], manual contours from one cardiac phase are propagated to other phases. Thereby, user-defined deviations between the wanted contours and the visible bloodpool-papillary transitions are propagated. In [10], a convex hull is established around classified LV bloodpool voxels. Similarly, in [11], papillaries and trabeculations at the endocardium are addressed by post-processing of a model-based segmentation: If contrast suffices to classify the LV bloodpool voxels near the initial segmentation, a convex mesh is established around the classified voxels. [12] presents a model-based segmentation with locally disabled image forces near the papillary muscles without contour post-processing.

This paper presents a method to change the boundary detection for model-based segmentation such that (1) papillaries and trabeculations are excluded from the myocardium and (2) the epicardial segmentation is no longer misled by false boundaries. Apart from quantitative improvements, we demonstrate that obvious errors for local wall thickness estimation are eliminated.

2 Improved Boundary Detection Training

A crucial point for robust and accurate segmentation is the detection of corresponding image boundaries for each model surface element. Discriminative boundary detectors can be trained using the “Simulated Search” described in detail in [4]. This method evaluates the performance of a set of boundary detection functions on a set of representative images with corresponding reference segmentations. The set of tested boundary detectors is generated from some templates with parameters estimated from the training examples. The evaluation metric to rank the proposed detectors simulates the geometric segmentation errors without shape constraints: For each triangle, boundary detection is performed from systematically modified triangle poses to simulate the mesh adaptation. Each detected boundary point is compared to the reference segmentation

and the geometric distance to the reference surface is averaged over all simulated mesh states and all training images. Finally, the boundary detector with minimal simulated root-mean-square (RMS) error is assigned to the triangle.

For the task of accurate LV endo- and epicardium segmentation, two problems motivate algorithmic extensions of the above optimization metric. First, image boundaries are *missing* for some surface regions. Bloodpool boundaries at the papillaries shall *not* attract the mesh if a convex hull around the bloodpool is required. Furthermore, the epicardium often touches the diaphragm and no epicardial edges are visible there. Second, the epicardium is often parallel to nearby boundaries such as the pericardium or the lung transition (see Fig. 1). In all these cases, the image contains boundaries close to the wanted contours. These, however, shall *not* be detected.

We can summarize our training requirements: (1) At papillary muscles and trabeculations, bloodpool contours *inside* the reference endocardium shall *not* be detected. (2) For the epicardium, nearby competing boundaries shall be rejected. The simulated geometric error alone does not fully address these goals. E.g., a detector that often locates the weak epicardial transition but sometimes detects false points or rejects all edges may have a larger error than another detector that reliably locates the pronounced lung transition if both boundaries are adjacent. At the papillaries, learning to detect bloodpool contours is evidently good if geometric proximity is optimized and if no better edges can be found. Rejecting all edges produces simulated errors equal to the simulated mesh modifications, which are typically larger than the papillaries.

To include the above explicit goals (1) and (2) into the training, the optimization criterion of the Simulated Search was extended. First, regions containing *unwanted* boundaries were explicitly annotated in the image volume (see Fig. 2). All detected boundaries in these “forbidden” regions that would pull the mesh towards unwanted boundaries are counted. The resulting percentage p of “forbidden” detections is linearly combined with the geometric RMS error d from the standard training procedure to arrive at a new optimization criterion:

$$\text{Cost}(F, i) = (1 - \alpha) \cdot d(F, i) + \alpha \cdot D \cdot p(F, i) \quad (1)$$

Here, F is the tested boundary detector for the given triangle i . d and p include all simulated detections in all images. D converts the percentage p into a distance

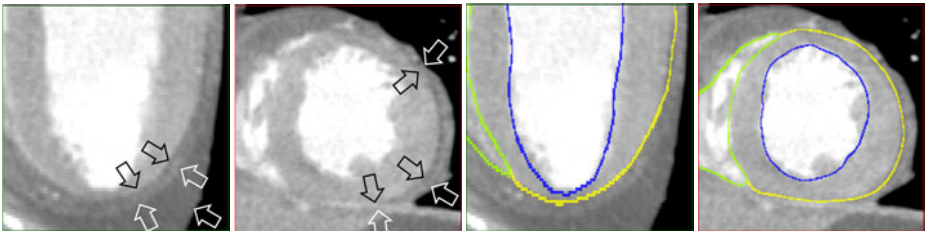


Fig. 1. Closeup of endo-, epi-, and pericardium. Black outlined arrows indicate the wanted epicardial boundary. White arrows indicate competing edges (pericardium, lung, diaphragm). Endo- and epicardial reference segmentations are shown on the right.

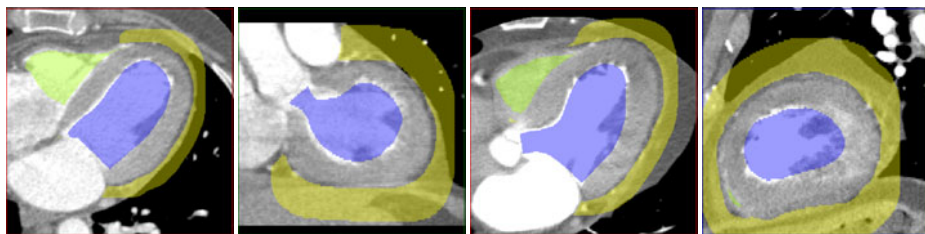


Fig. 2. Masks describing “forbidden” regions for the endocardium (blue, including the papillaries) and for the epicardium (yellow, covering various false edges; the green mask covers trabeculations in the right ventricle). Note that the masks do *not* include the immediate neighborhood of the reference segmentations.

comparable to typical simulated values of d . $\alpha = 0$ restores the standard training, while $\alpha = 1$ would focus only on avoiding forbidden edges and ignore the geometric accuracy of all detected boundaries.

3 Model for Accurate LV Segmentation

3.1 Reference Segmentations and Shape Model

Segmentations of 67 image volumes from 33 patients at different cardiac phases were carefully inspected and corrected in short and long axis views as well as in axial, coronal, and sagittal views. For the endocardial contour, a convex hull tightly enclosing the visible bloodpool was created. In some images, however, faint contrast and image artifacts made the proper definition of this contour difficult. Epicardial contours needed corrections where the pericardium or a nearby transition to the lung or beyond the diaphragm had been segmented. A smoothly interpolating mesh was used in regions of missing visible edges at the diaphragm. From the resulting 67 meshes, a new shape model was trained that better represents the convex shape of the endocardium.

3.2 Segmentation Chain with Two-Pass Deformable Adaptation

We follow the segmentation framework from [11]: The heart is first roughly localized by a Generalized Hough Transform [13]. The mesh is then adapted towards corresponding image boundaries. First, a global similarity transformation followed by a piecewise affine transformation initializes the mesh close to the desired boundaries. Finally, the mesh is deformed elastically but image forces are balanced by shape-preserving forces. For this deformable adaptation, we adopt a two-pass scheme from [12] that allows to reduce false positive boundary detections in the first pass while improving true positive detections in the second pass. In the first pass, the whole heart model is adapted with a search range of 10mm. Here, boundary detectors trained with large α in (1) can be used to minimize detection in forbidden regions. The shape model will smoothly interpolate the

surface across missing boundaries. In the second pass, only LV endo- and epicardial mesh parts are further adapted using a short search range of 2mm. The mesh can thus adapt to nearby boundaries missed in the first pass. Many forbidden edges are beyond the short search range and boundary detectors optimized for this range with small α can focus on local accuracy.

3.3 Boundary Detection Functions

Our boundary detectors combine edge detection with additional rejection criteria [14]: We either project the image gradient onto the triangle normal (to suppress edges with false orientation) as in [14] or we calculate the difference between inside and outside gray values (averaged over several points). To reject false boundaries, one or several local image features are evaluated and compared to trained acceptance intervals. If any feature value violates its interval, the edge is rejected. Image features offered to the training covered the gray values on both sides of the triangle (possibly averaged over several points) and Taylor coefficients of gray values along the triangle normal direction as in [14]. Here, variances of gray values parallel and normal to the mesh have been added to the feature set.

3.4 Detector Training and Selection

Using the extended training framework of Sec. 2, we first studied the trade-off between simulated geometric errors and forbidden boundary detections. Optimizing all $\{F_i\}$ for increasing α in (1), d increases and p decreases as expected. Plotting the mesh average of both quantities versus α , p shrinks most rapidly for small α , while d increases most for large α . Increasing d and shrinking p may indicate that no edge was detected for some triangles. (This contributes to d with the simulated displacements.) Under the influence of shape constraints, local errors may remain small if some neighboring edges are still found. After model adaptation, we typically find minima at medium values of α for both the geometric errors and the percentages of triangles ending up in forbidden regions.

To select the best combination of first and second pass boundary detectors, we decided to train and test endo- and epicardial detectors *separately*. This has two advantages: First, it allows to efficiently explore a multitude of training setups with different image features for boundary rejection since the “problems” at endo- and epicardium are of different nature. Second, it allows the use of different values of α for both mesh regions. To select the final boundary detectors from the different tested training setups, separate shape-constrained endo- and epicardial models were adapted to all training images in the two-pass scheme. Amongst several detector combinations with close-to-minimal surface-to-surface errors the percentage of triangles ending up in forbidden regions was minimized. This resulted in one detector combination for both the endocardium and the epicardium. These detectors were finally integrated into the full heart model.

4 Results

4.1 Surface-to-Surface Errors

To evaluate the accuracy of the segmentation, we measure constrained surface-to-surface distances: For each triangle, the distance to the closest surface point in a corresponding patch of the compared mesh is determined. Corresponding patches around a triangle are constrained by anatomical labels and limited to a geodesic radius of 10mm over the mesh surface. The distances are measured from the automatic mesh to the reference mesh and vice versa since they are not symmetric. From each distance pair, the *maximum* is retained. These distances over all test images and all mesh triangles per anatomical region are finally evaluated statistically.

Experiments were conducted on 67 retrospectively ECG-gated cardiac multi-slice CT image volumes from 33 patients (1–3 cardiac phases per patient). These had been obtained using 16-, 40-, and 64-channel CT scanners (Brilliance CT, Philips Healthcare, Cleveland, Ohio, USA). Bloodpool contrast and image quality varied considerably. The mean in-plane voxel resolution was 0.45mm (range 0.30–0.49mm) using a 512×512 matrix. The mean slice thickness was 0.72mm (range 0.33–2.0mm). Our baseline model had been generated and trained as described in [1] on 28 datasets from 13 patients using old reference segmentations that had *not* been corrected at the LV endo- and epicardium. For the improved LV segmentation, we performed a 3-fold cross-validation on all 67 datasets. Here, a new shape model as well as the LV endo- and epicardial boundary detectors were trained on 22 patients and the datasets of the left-out 11 patients were segmented with the resulting cross-validation model. The boundary detectors of all other structures were taken from the baseline model.

Table 1 summarizes the resulting error statistics for the LV surfaces. Mean errors for epi- and endocardium are reduced by 30% and 40%, resp. In view of our carefully corrected reference segmentations we consider this reduction as a clear improvement. More important, however, is the distribution of errors over the mesh and across images. Local outliers with large errors exceeding 2mm that are beyond typical inter-observer uncertainties are reduced by a factor of 3.

Table 1. Mean surface-to-surface errors [in mm] with standard deviations covering variations both across mesh triangles and across images. The error distribution describes the percentages [in %] of triangles within the listed error bins [in mm] in all images (NumTriangles × NumImages data points). The LV epicardium includes the septal boundary towards the right ventricle.

	Baseline model					New model				
	Mean ± SDev	[0, 1)	[1, 2)	[2, 3)	≥ 3	Mean ± SDev	[0, 1)	[1, 2)	[2, 3)	≥ 3
LV endo	1.26 ± 1.35	58.5	22.3	8.9	10.2	0.76 ± 0.88	77.0	16.6	3.6	2.9
LV epi	0.96 ± 1.02	67.1	21.6	6.8	4.4	0.68 ± 0.62	79.1	16.9	3.1	1.0

4.2 Qualitative Results and Wall Thickness Measurements

Fig. 3 shows some segmentation examples illustrating the achieved improvements. The consistent segmentation of the endocardium and the improvements for the epicardium enable us to measure the LV wall thickness. To account for the bending 3D geometry of the myocardium towards the apex, we do *not* measure wall thickness in short-axial slices. Rather, the normal of the endocardial wall is estimated for endocardial triangles from a regression plane fitted through the triangle vertices and their neighbors. We use 1st and 2nd order vertex neighbors for robust direction estimation. A ray is then cast from the triangle center outwards, and the distance to the epicardial wall is taken as wall thickness. We restrict our measurements to the myocardial region assigned to the 17 AHA segments proposed in [14].

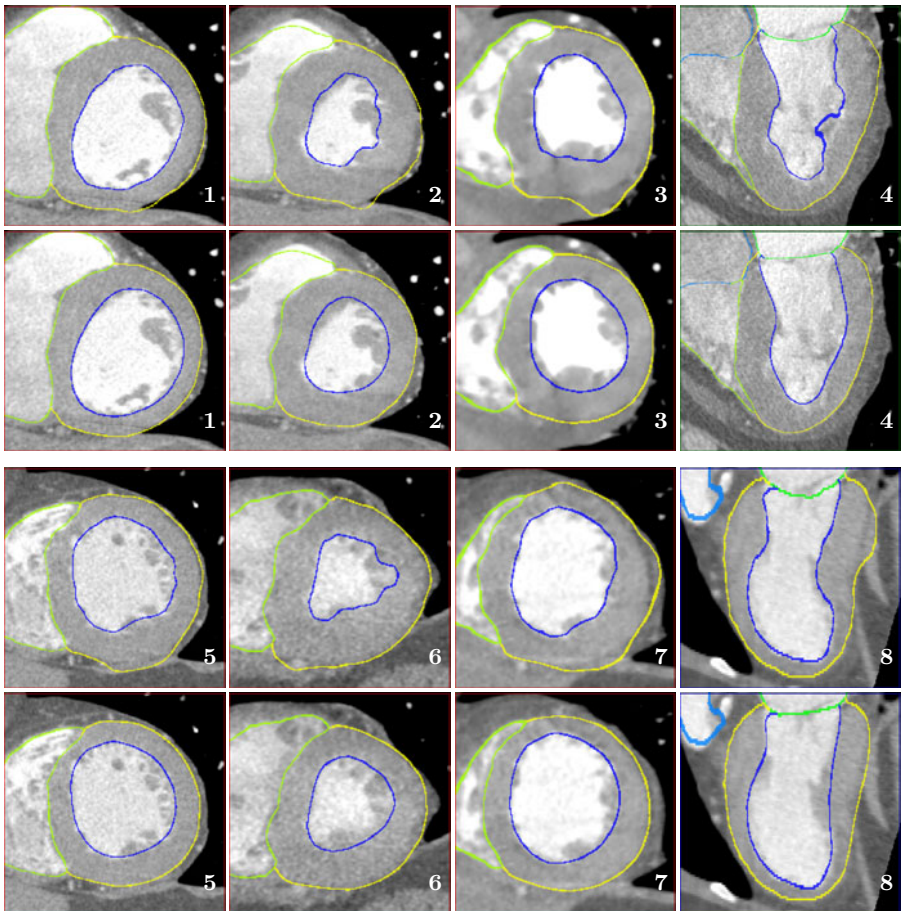


Fig. 3. Segmentations obtained with the baseline model (rows 1 and 3) and with the cross-validation models of Sec. 4.1 (rows 2 and 4). The first two examples stem from the same patient. The first column is at mid-diastole, all other examples are end-systolic.

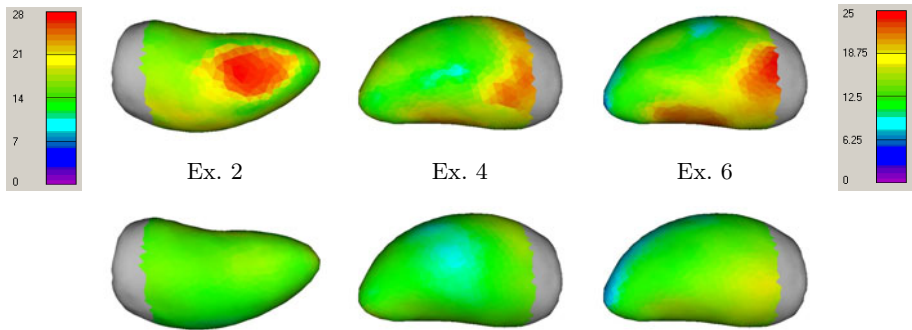


Fig. 4. Wall thickness measurements for examples 2, 4, and 6 from Fig. 3. Example 2: inferior view (color bar 0–28mm). Examples 4 and 6: lateral view (color bar 0–25mm, all measurements projected onto the mean mesh). Local maxima due to irregular endocardial and epicardial surfaces with the baseline model (top row) are eliminated with the cross-validation models (bottom).

This region excludes the LV outflow tract towards the aorta and the transition towards left atrium. Results are shown in Fig. 4.

5 Discussion and Conclusion

The quantitative and qualitative results demonstrate clear improvements of the LV segmentation. Using the extended optimization criterion of the Simulated Search and carefully corrected reference segmentations, boundary detectors could be trained that effectively avoid unwanted edges. Locally visible structures are well segmented in the two-pass adaptation and outliers with large errors are drastically reduced. The consistent exclusion of papillary muscles from the myocardium and the accurate segmentation of the epicardial surface are the basis for local wall thickness measurements.

References

1. Ecabert, O., et al.: Automatic model-based segmentation of the heart in CT images. *IEEE Trans. Medical Imaging* 27(9), 1189–1201 (2008)
2. van Assen, H.C., et al.: A 3-D active shape model driven by fuzzy inference: Application to cardiac CT and MR. *IEEE Trans. Information Tech. Biomed.* 12(5), 595–506 (2008)
3. Zheng, Y., et al.: Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features. *IEEE Trans. Medical Imaging* 27(11), 1668–1681 (2008)
4. Peters, J., et al.: Optimizing boundary detection via simulated search with applications to multi-modal heart segmentation. *Medical Image Analysis* 14(1), 70–84 (2010)

5. Plumhans, C., et al.: Comparison of manual, semi- and fully automated heart segmentation for assessing global left ventricular function in multidetector computed tomography. *Investigative Rad.* 44(8), 476–482 (2009)
6. Ghersin, E., et al.: Clinical evaluation of a fully automated model-based algorithm to calculate left ventricular volumes and ejection fraction using multidetector computed tomography. *Acute Cardiac Care* 11, 47–55 (2009)
7. Coche, E., et al.: Quantitative right and left ventricular functional analysis during gated whole-chest MDCT: A feasibility study comparing automatic segmentation to semi-manual contouring. *European Journal of Radiology* (in Press), <http://dx.doi.org/10.1016/j.ejrad.2009.05.037>
8. Abadi, S., et al.: Feasibility of automatic assessment of four-chamber cardiac function with MDCT: Initial clinical application and validation. *European Journal of Radiology* (in Press), <http://dx.doi.org/10.1016/j.ejrad.2009.01.035>
9. Hautvast, G., et al.: Automatic contour propagation in cine cardiac magnetic resonance images. *IEEE Trans. Medical Imaging* 25(11), 1472–1482 (2006)
10. Cocosco, C.A., et al.: Automatic image-driven segmentation of the ventricles in cardiac cine MRI. *J. Magn. Reson. Imaging* 28(2), 366–374 (2008)
11. Zheng, Y., et al.: Left ventricle endocardium segmentation for cardiac CT volumes using an optimal smooth surface. In: *Proc. SPIE Medical Imaging*, vol. 7259, 72593V–1–11 (2009)
12. Lehmann, H., et al.: Integrating viability information into a cardiac model for interventional guidance. In: Ayache, N., Delingette, H., Sermesant, M. (eds.) *FIMH 2009*. LNCS, vol. 5528, pp. 312–320. Springer, Heidelberg (2009)
13. Ballard, D.H.: Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recogn.* 13(2), 111–122 (1981)
14. Cerqueira, M.D., et al.: Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. *Circulation* 105, 539–542 (2002)

Regional Heart Motion Abnormality Detection via Information Measures and Unscented Kalman Filtering

Kumaradevan Punithakumar¹, Ismail Ben Ayed¹, Ali Islam²,
Ian G. Ross³, and Shuo Li^{1,4}

¹ GE Healthcare, London, Ontario, Canada

² St. Joseph's Health Care, London, Ontario, Canada

³ London Health Science Center, London, Ontario, Canada

⁴ University of Western Ontario

Abstract. This study investigates regional heart motion abnormality detection using various classifier features with Shannon's Differential Entropy (SDE). Rather than relying on elementary measurements or a fixed set of moments, the SDE measures global distribution information and, as such, has more discriminative power in classifying distributions. Based on functional images, which are subject to noise and segmentation inaccuracies, heart wall motion analysis is acknowledged as a difficult problem and, therefore, incorporation of prior knowledge is desirable to enhance the accuracy. Given noisy data and nonlinear dynamic model to describe the myocardial motion, unscented Kalman filter, a recursive nonlinear Bayesian filter, is devised in this study so as to estimate LV cavity points. Subsequently, a naive Bayes classifier algorithm is constructed from the SDEs of different features in order to automatically detect abnormal functional regions of the myocardium. Using 90×20 segmented LV cavities of short-axis magnetic resonance images obtained from 30 subjects, the experimental analysis carried over 480 myocardial segments demonstrates that the proposed method perform significantly better than other recent methods, and can lead to a promising diagnostic support tool to assist clinicians.

1 Introduction

Assessment of left ventricular function is of utmost importance in the diagnosis of coronary heart disease, the leading cause of death worldwide. It primarily relies on the visual analysis and interpretation of wall motion and, thus, subject to high interobserver variability. Alternatively, computer-aided detection systems have been attempted in recent years in order to automatically analyze the wall motion quantitatively [1,2,3], and to classify hearts into normal or abnormal [4,5,6]. However, in clinical practice, the wall motion is commonly scored by following a standard issued by American Heart Association (AHA) [7], where the myocardium is divided into 17 segments. Therefore, regional abnormality analysis is more desirable for clinical purposes. Existing heart motion regional

analysis methods are based on, among others, shape models with localized variations [8], a tensor-based classification to conserve the spatiotemporal structure of the myocardium deformation [9], a hidden Markov model for local wall motion classification based on stress echocardiography [10], a pattern recognition method based on intra-segment correlation [11], and an independent component analysis classifier to detect and localize abnormally contracting regions [12]. Most of existing methods either suffer from poor accuracy, use data that are not available in regular clinical routine, or require extensive user interaction to define myocardial boundaries.

Nevertheless, automating abnormality analysis would reduce inter- and intra-observer variability and, therefore, subjectivity in the analysis. However, it requires precise estimation of the myocardial points. As such, accurate characterization of dynamic behavior of Left Ventricle (LV) is essential in order to enhance the performance of motion estimation. In this connection, linear dynamic models may not be sufficient to describe the LV dynamics and, therefore, a *nonlinear cyclic model* is proposed in this study. The proposed model consists of a time-varying parameter, angular frequency, to be estimated from the data along with other state elements. This is achieved by augmenting the angular frequency with other state elements, and by estimating the augmented state using Unscented Kalman Filter (UKF), a recursive nonlinear Bayesian filter, given initial LV cavity points and a nonlinear cyclic model.

Cine Magnetic Resonance (MR) sequences are widely used for cardiac functional analysis, and provide a large number of images¹. Therefore, tracking based on manual delineation of the LV boundary in all these images is prohibitively time consuming. Alternatively, automating the process can be of great interest [13,14]. However, it is subject to the difficulties due to the low contrast and photometric similarities between the connected cardiac regions - for instance, the papillary muscles within the cavity and myocardial wall have approximately the same intensity. To tackle the problem of obtaining initial LV segmentations, this study adopts a non-rigid image registration method that minimizes the sum of squared intensity differences [15], given an initial segmentation of the first frame. Rather than using a segmentation algorithm to obtain initial LV cavity points, using registration is advantageous in our study as it provides the sequence of corresponding points over time, an essential attribute to analyze wall motion regionally.

Classification of regional abnormality is a difficult problem due to the similarity between the statistical information associated with normal and abnormal heart motions. In this study, we investigate the problem with a global measure based on the Shannon's Differential Entropy (SDE) of the distributions of normalized radial distance, radial velocity, segment area and circumferential length of a segment. Rather than relying on elementary measurements or a fixed set of moments, the SDE measures global distribution information and, as such, has more discriminative power in classifying distributions. The individual classification ability of elements were measured using Receiver Operating Characteristic (ROC) curves with

¹ Typically, the number of images per subject is equal to 200.

the corresponding Area Under the Curves (AUCs), and the Bhattacharyya distance metric [16]. Subsequently, a naive Bayes classifier algorithm is constructed from the SDEs of classifier elements with the best performance in order to automatically detect abnormal functional regions of the myocardium.

Using 90 image datasets, each consisting of 20 segmented LV cavities of short-axis MR functional images obtained from 30 subjects, the proposed method is quantitatively evaluated by comparison with ground truth classifications by radiologists over 480 myocardial segments. The proposed method performed significantly better than other recent methods with an overall classification accuracy of 91.0%, and can lead to a promising diagnostic support tool to assists clinicians.

2 Temporal Smoothing of the Dataset

Dynamic Model for Temporal Periodicity: Let (x, y) be a point on the boundary of the LV cavity and $\mathbf{s} = [\bar{x} \times \dot{x} \bar{y} y \dot{y} \omega]^T$ be the state vector that describes the corresponding dynamics. The elements \dot{x} , \bar{x} , \dot{y} and \bar{y} denote, respectively, velocity and the mean position over a cardiac cycle in x and y coordinate directions, and ω the angular frequency. We assume that the heart motion is periodic. The discrete-time dynamic model that describes the cyclic motion of the point is given by

$$\mathbf{s}_{k+1} = F_k \mathbf{s}_k + v_k \tag{1}$$

where

$$F_k = \begin{bmatrix} F_{cy}(k) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & F_{cy}(k) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 1 \end{bmatrix}, F_{cy}(k) = \begin{bmatrix} 1 & 0 & 0 \\ 1 - \cos(\omega_k T) & \cos(\omega_k T) & \frac{1}{\omega_k} \sin(\omega_k T) \\ \omega_k \sin(\omega_k T) & -\omega_k \sin(\omega_k T) & \cos(\omega_k T) \end{bmatrix},$$

and T is the time interval between two subsequent image frames. The process noise sequence $\{v_k\}$ is Gaussian with zero-mean and covariance Q_k that accommodates the unpredictable errors due to modeling uncertainties. The measurement equation is given by

$$z_k = H_k \mathbf{s}_k + \eta_k \tag{2}$$

where $H_k = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$, and $\{\eta_k\}$ is a zero-mean Gaussian noise sequence with covariance $R_k = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}$.

Unscented Kalman Filter: The dynamic model in (1) is nonlinear and, therefore, we adopt a recursive nonlinear Bayesian filter, the UKF [17], to estimate the state s_k at each time step. The UKF uses unscented transformation, a more direct and explicit mechanism for transforming mean and covariance information, that addresses the deficiencies of linearization inherent to other adaptive filters such as Extended Kalman Filter (EKF). UKF propagates the second order properties of the distribution with only a small amount of statistical information by choosing sample points deterministically. Therefore, it provides sufficient accuracy for nonlinear filtering applications with computational cost of order similar to the EKF.

3 Information Theoretic Measures and Classifier

The state estimates from UKF are subsequently processed towards the classification of normal or abnormal motions. In order to measure the information associated with regional LV function, the SDEs of a set of classifier elements were evaluated.

Classifier elements: Let $\hat{\mathbf{s}}_{k,i} = [\hat{x}_{k,i} \ \hat{x}_{k,i} \ \hat{x}_{k,i} \ \hat{y}_{k,i} \ \hat{y}_{k,i} \ \hat{y}_{k,i} \ \hat{\omega}_{k,i}]^T$ be the estimated state of i^{th} point by UKF at time step k . Let $I_N = \{1, 2, \dots, N\}$ and $I_S = \{i_s, i_s + 1, \dots, i_s + N_s - 1\} \subset I_N$ be the set of points on the LV cavity and segment S , respectively. The center $(c_{x,k}, c_{y,k})$ of the LV cavity at time step k is given by $c_{x,k} = \frac{1}{N} \sum_{i \in I_N} \hat{x}_{k,i}$ and $c_{y,k} = \frac{1}{N} \sum_{i \in I_N} \hat{y}_{k,i}$. We propose to use the following classifier elements.

1. Normalized radial distance

$$r_{k,i} = \frac{\sqrt{(\hat{x}_{k,i} - c_{x,k})^2 + (\hat{y}_{k,i} - c_{y,k})^2}}{\max_{i \in I_N} \sqrt{(\hat{x}_{k,i} - c_{x,k})^2 + (\hat{y}_{k,i} - c_{y,k})^2}} \quad \forall i \in I_S \quad (3)$$

2. Radial velocity

$$v_{k,i} = \left\langle \mathbf{v}_{k,i} \cdot \frac{\mathbf{r}_{k,i}}{\|\mathbf{r}_{k,i}\|} \right\rangle \quad \forall i \in I_S \quad (4)$$

where $\mathbf{v}_{k,i} = [\hat{x}_{k,i}, \hat{y}_{k,i}]^T$ and $\mathbf{r}_{k,i} = [\hat{x}_{k,i} - c_{x,k}, \hat{y}_{k,i} - c_{y,k}]^T$

3. Circumferential length

$$l_k = \sum_{i=i_s}^{i_s+N_s-2} \sqrt{(\hat{x}_{k,i+1} - \hat{x}_{k,i})^2 + (\hat{y}_{k,i+1} - \hat{y}_{k,i})^2} \quad (5)$$

4. Segment area

$$a_k = \frac{1}{2} \left[(c_{x,k} \hat{y}_{k,i_s} - c_{y,k} \hat{x}_{k,i_s}) + (\hat{x}_{k,i_s+N_s} c_{y,k} - \hat{y}_{k,i_s+N_s} c_{x,k}) + \sum_{i=i_s}^{i_s+N_s-2} (\hat{x}_{k,i} \hat{y}_{k,i+1} - \hat{x}_{k,i+1} \hat{y}_{k,i}) \right] \quad (6)$$

Shannon’s differential entropy and Bayesian classifier: The kernel density estimate of a classifier element $\chi_n \in \{r_{k,i}, v_{k,i}, l_k, a_k\}$ is given by

$$f(\xi) = \frac{\sum_{n=1}^{n_\chi} \mathcal{K}(\chi_n - \xi)}{n_\chi} \quad (7)$$

where $n_\chi = KN_s$ for $\chi_n \in \{r_{k,i}, v_{k,i}\}$ and $n_\chi = K$ for $\chi_n \in \{l_k, a_k\}$. Typical choices of $\mathcal{K}(\cdot)$ are the Dirac function and the Gaussian kernel. In this study, the SDE is derived as follows

$$S_f = - \int_{\xi \in \mathbb{R}} \frac{\sum \mathcal{K}(\chi_n - \xi)}{n_\chi} \left(\ln \sum \mathcal{K}(\chi_n - \xi) - \ln n_\chi \right) d\xi \quad (8)$$

It is advantageous to use multiple SDE measures towards the classification as they measure different information associated with the myocardial function. A naive Bayesian classifier that provides a quadratic decision boundary is constructed from SDEs of classifier elements.

4 Experiments

The data contains 90 short-axis image datasets, each consisting of 20 functional 2D images acquired from 20 normal and 10 abnormal hearts. The data were acquired on 1.5T MRI scanners with fast imaging employing steady state acquisition (FIESTA) mode. The results of 480 myocardial segments from apical, mid-cavity and basal frames were compared with ground truth classification of the cine MRI datasets by radiologists². The dynamic model and UKF parameters are chosen as $q_1 = 0.01$, $q_2 = 0.1$ $q_3 = 1$ and $r = 0.01$ to accommodate noise that account for modeling uncertainties. We used Dirac function for $\mathcal{K}(\cdot)$ in the kernel density estimation.

In Fig. 1, we give a representative sample of the segmentation results for apical, mid-cavity and basal frames. The frames were automatically segmented following the standard issued by the AHA [7], given *anatomical landmarks*³ on the first frame.

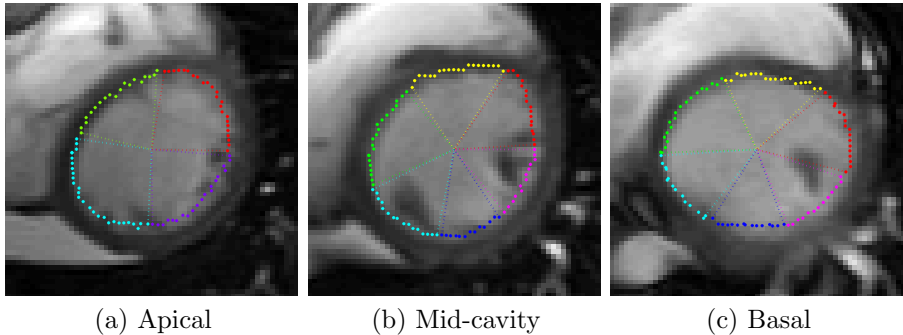


Fig. 1. Representative examples of segmented myocardium using the proposed approach. Apical, mid-cavity and basal frames were segmented, respectively, into 4, 6 and 6 segments following the standard in [7].

We used two criteria to measure the performance of each classifier elements, namely, the ROC curves with corresponding AUCs, and Bhattacharyya measure [16] to assess the discriminative power of each classifier elements. Furthermore, we assessed the performance of the proposed approach via *leave-one-subject-out* method.

² Among the 480 myocardial segments, 389 segments were marked as normal and 91 as abnormal.

³ As suggested by [7], the attachment of the right ventricular wall to the LV is used to identify and separate the septum from the LV anterior and inferior free walls.

ROC, AUC and Bhattacharyya measure: The ROC curves for classifier elements is shown in Fig. 2. The figure shows that SDEs of segment area and normalized radial distance have better classifying ability than other classifier elements. The AUCs corresponding to the ROC curves in Fig. 2 are reported in Table 1. We used the Bhattacharyya distance metric to evaluate the overlap between the distributions of classifier elements over normal and abnormal motions. The SDEs of segment area and normalized radial distance yielded the higher \mathcal{B} as reported in Table 1 and, therefore, have the best discriminative ability. This is consistent with the previous findings based on ROC/AUC evaluations.

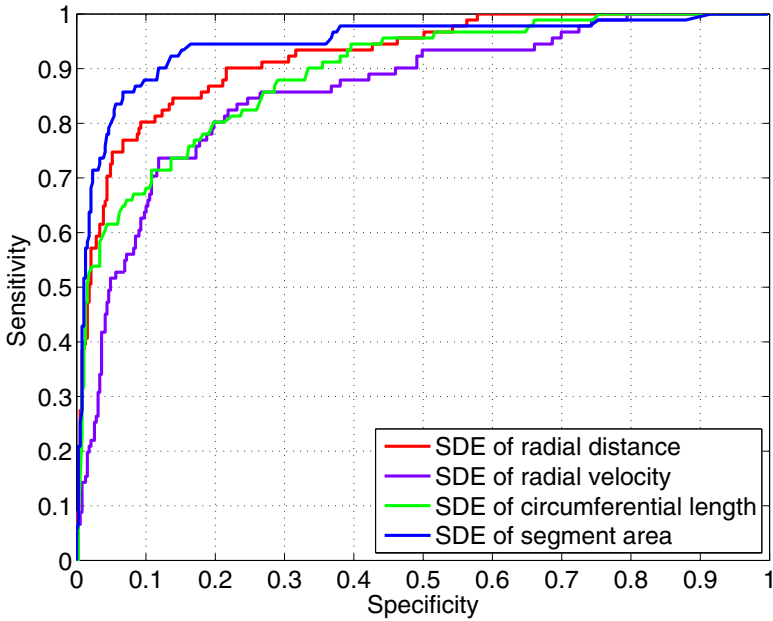


Fig. 2. Receiver operating characteristics of classifier elements. The closer the curve to the left hand top corner, the better the classification performance.

Table 1. The area under the curve corresponding to Fig. 2 and the Bhattacharyya distance metric (\mathcal{B}) of normal/abnormal distributions. The higher the values the better the discriminative ability of the classifier.

Classifier element	AUC (%)	Bhattacharyya distance metric (\mathcal{B})
SDE of segment area	94.3	0.66
SDE of normalized radial distance	92.1	0.61
SDE of circumferential shrink	89.0	0.53
SDE of radial velocity	85.8	0.48

Table 2. The percentage of classification accuracy using leaving-one-subject-out method. The proposed method achieved an overall classification accuracy of 90.8%.

	Accuracy (%)	Sensitivity (%)	Specificity (%)
Apex	92.5	90.0	93.3
Mid-cavity	93.3	93.1	94.1
Base	87.2	100.0	84.9

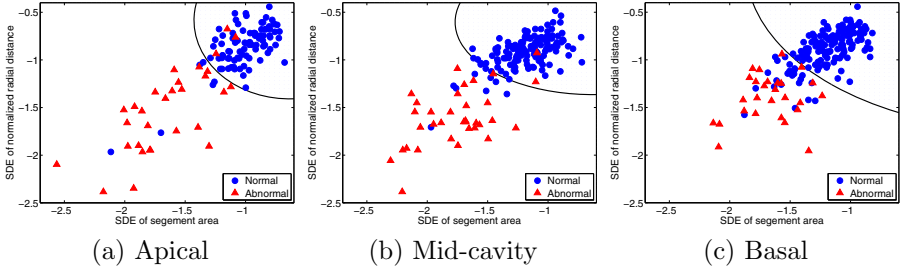


Fig. 3. Decision boundary for normal and abnormal regional myocardial functions using a Bayesian classifier

Classification performance: Table 2 reports the classification performance of correctly classified hearts using a *leaving-one-subject-out* method. A naive Bayes classifier algorithm is constructed from the SDEs of the segment area and normalized radial distance, the elements with better classifier ability. Fig. 3 shows the quadratic decision boundary for normal/abnormal classification, where blue circles represent the normal function and red triangles the abnormal. The overall classification accuracy is equal to 90.8%, with a sensitivity of 94.5% and specificity of 90.0%. The highest performance was achieved for mid-cavity frames with average of 93.3% for accuracy, 93.1% for sensitivity and 94.1% for specificity.

5 Conclusions

We presented a regional heart motion abnormality detection method using the SDE and UKF. A non-rigid image registration method was adopted to initialize the LV cavity points, and the results were subsequently processed using UKF, given a nonlinear dynamic model. The myocardial segments were identified following the standard issued by AHA. The SDE of normalized radial distance, radial velocity, segment area and circumferential segment length were evaluated for each myocardial segments and a naive Bayesian classifier is constructed from the SDEs. The experimental analysis carried over 90×20 segmented LV cavities of short-axis MR images obtained from 30 subjects demonstrates that the proposed method perform significantly better than other recent methods, and can lead to a promising diagnostic support tool to assists clinicians.

References

1. McEachen, J., Nehorai, A., Duncan, J.: Multiframe temporal estimation of cardiac nonrigid motion. *IEEE Trans. Image Process.* 9(4), 651–665 (2000)
2. Spottiswoode, B., Zhong, X., Hess, A., Kramer, C., Meintjes, E., Mayosi, B., Epstein, F.: Tracking myocardial motion from cine dense images using spatiotemporal phase unwrapping and temporal fitting. *IEEE Trans. Med. Imag.* 26(1), 15–30 (2007)
3. Sundar, H., Davatzikos, C., Biros, G.: Biomechanically-constrained 4D estimation of myocardial motion. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009. LNCS*, vol. 5762, pp. 257–265. Springer, Heidelberg (2009)
4. Punithakumar, K., Li, S., Ayed, I.B., Ross, I., Islam, A., Chong, J.: Heart motion abnormality detection via an information measure and bayesian filtering. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009. LNCS*, vol. 5762, pp. 373–380. Springer, Heidelberg (2009)
5. Fung, G., Qazi, M., Krishnan, S., Bi, J., Rao, B., Katz, A.: Sparse classifiers for automated heartwall motion abnormality detection. In: *Proc. IEEE 4th Int. Conf. Machine Learning and Applications*, pp. 194–200 (2005)
6. Qazi, M., Fung, G., Krishnan, S., Bi, J., Bharat Rao, R., Katz, A.: Automated heart abnormality detection using sparse linear classifiers. *IEEE Eng. Med. Biol. Mag.* 26(2), 56–63 (2007)
7. Cerqueira, M.D., Weissman, N.J., Dilsizian, V., Jacobs, A.K., Kaul, S., Laskey, W.K., Pennell, D.J., Rumberger, J.A., Ryan, T., Verani, M.S.: Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: A statement for healthcare professionals from the cardiac imaging committee of the council on clinical cardiology of the American Heart Association. *Circulation* 105(4), 539–542 (2002)
8. Leung, K.Y., Bosch, J.G.: Localized shape variations for classifying wall motion in echocardiograms. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part I. LNCS*, vol. 4791, pp. 52–59. Springer, Heidelberg (2007)
9. Mansor, S., Noble, J.: Local wall motion classification of stress echocardiography using a hidden Markov model approach. In: *The 5th IEEE Int. Symp. Biomedical Imaging: From Nano to Macro*, pp. 1295–1298 (2008)
10. Qian, Z., Liu, Q., Metaxas, D.N., Axel, L.: Identifying regional cardiac abnormalities from myocardial strains using spatio-temporal tensor analysis. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part I. LNCS*, vol. 5241, pp. 789–797. Springer, Heidelberg (2008)
11. Lu, Y., Radau, P., Connelly, K., Dick, A., Wright, G.: Pattern recognition of abnormal left ventricle wall motion in cardiac MR. In: Yang, G.G., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009. LNCS*, vol. 5762, pp. 750–758. Springer, Heidelberg (2009)
12. Suinesiaputra, A., Frangi, A., Kaandorp, T., Lamb, H., Bax, J., Reiber, J., Lelieveldt, B.: Automated detection of regional wall motion abnormalities based on a statistical model applied to multislice short-axis cardiac MR images. *IEEE Trans. Med. Imag.* 28(4), 595–607 (2009)
13. Ben Ayed, I., Punithakumar, K., Li, S., Islam, A., Chong, J.: Left ventricle segmentation via graph cut distribution matching. In: Yang, G.Z., et al. (eds.) *MICCAI 2009. LNCS*, vol. 5762, pp. 901–909. Springer, Heidelberg (2009)

14. Ben Ayed, I., Lu, Y., Li, S., Ross, I.: Left ventricle tracking using overlap priors. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, pp. 1025–1033. Springer, Heidelberg (2008)
15. Chen, H.M., Ben Ayed, I., Garvin, G., Goela, A., Li, S.: A parameterization of deformation fields for diffeomorphic image registration and its application to myocardial delineation. In: MICCAI 2010 (2010) (accepted for publication)
16. Comaniciu, D., Ramesh, V., Meer, P.: Kernel-based object tracking. *IEEE Trans. Pattern Anal. Mach. Intell.* 25(5), 564–577 (2003)
17. Julier, S., Uhlmann, J.: Unscented filtering and nonlinear estimation. *Proc. IEEE* 92(3), 401–422 (2004)

In vivo Human 3D Cardiac Fibre Architecture: Reconstruction Using Curvilinear Interpolation of Diffusion Tensor Images

Nicolas Toussaint^{1,2}, Maxime Sermesant^{1,2}, Christian T. Stoeck³,
Sebastian Kozerke^{1,3}, and Philip G. Batchelor¹

¹ King's College London, Imaging Sciences, London, UK
nicolas.toussaint@kcl.ac.uk

² INRIA, Asclepios Research Group, Sophia Antipolis, France

³ ETH Zürich, Institute for Biomedical Engineering, Switzerland

Abstract. *In vivo* imaging of the cardiac 3D fibre architecture is still a challenge, but it would have many clinical applications, for instance to better understand pathologies and to follow up remodelling after therapy. Recently, cardiac MRI enabled the acquisition of Diffusion Tensor images (DTI) of 2D slices. We propose a method for the complete 3D reconstruction of cardiac fibre architecture in the left ventricular myocardium from sparse *in vivo* DTI slices. This is achieved in two steps. First we map non-linearly the left ventricular geometry to a truncated ellipsoid. Second, we express coordinates and tensor components in Prolate Spheroidal System, where an anisotropic Gaussian kernel regression interpolation is performed. The framework is initially applied to a statistical cardiac DTI atlas in order to estimate the optimal anisotropic bandwidths. Then, it is applied to *in vivo* beating heart DTI data sparsely acquired on a healthy subject. Resulting *in vivo* tensor field shows good correlation with literature, especially the elevation (helix) angle transmural variation. To our knowledge, this is the first reconstruction of *in vivo* human 3D cardiac fibre structure. Such approach opens up possibilities in terms of analysis of the fibre architecture in patients.

1 Introduction

Cardiac fibre architecture has a crucial role in the cardiac function, as it influences heavily the muscle electrophysiological and mechanical behaviours. For instance electrical propagation is three times faster in the fibre direction than in the orthogonal plan [8]. For such reasons its study can have an important impact on clinical decisions, as several cardiac pathologies – such as myocardial infarction, cardiomyopathy, hypertension, or valvular heart diseases – involve a rearrangement of myocardial fibres [16].

Diffusion Tensor MRI (DTI) can depict non-invasively the fibre orientation distribution of the myocardium [6,14,17]. Moreover, taking into account the full information given by the tensors gives an insight on the intrinsic laminar sheet

structure [12]. However, using this technique *in-vivo* is challenging: complex motion patterns due to heartbeat creates important distortions of the signal. Furthermore, the muscle tissue has relatively low anisotropy. Recent studies however raised a new hope [25], but the difficulty and sensitivity of these techniques limit the data that could be acquired in clinical time, resulting in sparsely distributed acquisitions.

We set up a method to interpolate the *in vivo* Left Ventricle (LV) sparse tensor data in a shape-adapted curvilinear coordinate system (i.e. Prolate Spheroidal System, PSS). We explain how we non-linearly register the LV to an ellipsoid to express a Kernel regression scheme in PSS. Spatial coherence of a statistical cardiac DTI atlas is used as a prior for our curved interpolation. By mapping back the resulting tensor information to the LV, we are able to reconstruct the complete tensor structure of the LV, and compare it with those obtained with Cartesian interpolation.

2 Curvilinear Interpolation of Tensors

2.1 Truncated Prolate Spheroid 3D Diffeomorphic Mapping

The shape of the LV has been assimilated to a truncated ellipsoid [19], on which we can use the shape adapted curvilinear coordinate system called Prolate Spheroidal System [15] (PSS). Note that this coordinate system is already used to describe mechanical behaviour of the heart [119].

A 3D binary mask of the left ventricle is manually delineated on a 3D MRI of the heart, and registered to a 3D synthetic binary mask of a perfect truncated prolate spheroid (see Fig. 1c). In order to have an invertible and symmetric mapping of positions, we chose a symmetric version of the Diffeomorphic Demons registration algorithm [3]. In addition, an elasticity constraint, as introduced in [10] has been incorporated to ensure a smooth displacement of the middle of the ventricle wall. The advantage of this method is that it provides us with smooth and elastic displacement fields, from the LV to the spheroid and vice versa.

In this setting, any acquired tensor D_x lying within the LV wall is warped using the forward transformation Φ . We reorient it using the Finite Strain strategy as it seems well suited to preserve geometric features of cardiac diffusion tensor fields [12], with the Jacobian of the backward transformation Φ^{-1} .

PSS is well adapted to the LV shape [9] as local coordinates have physiological meaning (see Fig. 1(b)): ξ_1 is the positive transmural abscissa, ξ_2 is the apex-base abscissa from 0 to $\pi/2$, and ξ_3 is the circumferential abscissa from 0 to 2π . Thus, once the tensor data is mapped on the prolate spheroid, we can compute PSS coordinate position vectors (ξ_1, ξ_2, ξ_3) of each measure point, and describe tensors components in local contravariant basis $\mathcal{B}_\xi = (g_1, g_2, g_3)$. The transformation between basis are then written: $D_\xi = \mathcal{B}_\xi^T \cdot \Phi(D_x) \cdot \mathcal{B}_\xi$ and $D_x = \Phi^{-1}(\mathcal{B}_\xi \cdot D_\xi \cdot \mathcal{B}_\xi^T)$.

2.2 Anisotropic Kernel Estimator in Curvilinear Coordinates

Let us consider the situation where we have a set of N known tensors $(D_{\xi_i})_1^N$ over the domain of interest Ω' , which is the image of the LV: $\Omega' = \Phi(\Omega)$.

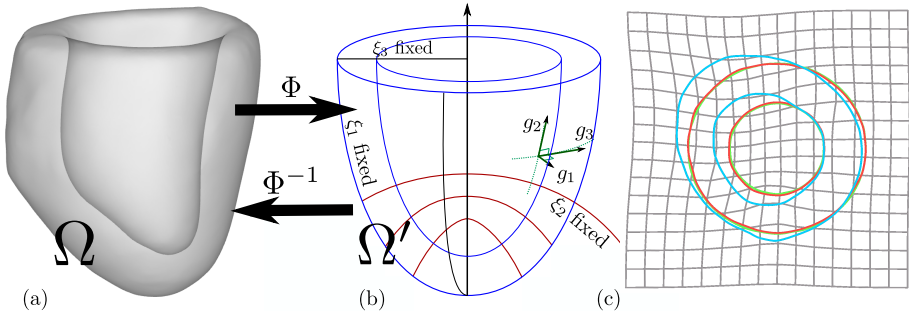


Fig. 1. 3D Mapping of a binary mask of the LV (a) to a synthetic prolate spheroid (b) using adapted symmetric diffeomorphic demons. Transformations Φ and Φ^{-1} are smooth and inverse of each other. (b) shows level-sets of PSS coordinates, as well as the local contravariant basis vectors (g_1, g_2, g_3) . (c) demonstrates the quality of the 3D registration. The LV (blue) is registered towards the synthetic spheroid (green). The registered LV outline (red) fits with its target with a smooth displacement (grey).

We want to estimate the missing tensor D_ξ at position ξ . Assuming that the tensor field is continuous and relatively homogeneous in the myocardium, we can claim that D_ξ is directly influenced by its neighbours. In this context we use a Gaussian kernel estimator K_σ of bandwidth σ . The kernel estimate $\hat{m}_\sigma(D_\xi)$ is taken as the Log-Euclidean [4] weighted mean of surrounding tensors, and thus written:

$$\hat{m}_\sigma(D_\xi) = \exp \left(\frac{\sum_{i=1}^N K_\sigma(\xi - \xi_i) \log(D_{\xi_i})}{\sum_{i=1}^N K_\sigma(\xi - \xi_i)} \right)$$

PSS coordinates are highly non-homogeneous, therefore an anisotropic interpolation with respect to each of the coordinates is needed. We thus replace the scalar σ by a positive definite matrix H . Let X be the position in a given coordinate system as a column vector $(X_1, X_2, X_3)^T$ and let $dX = (X - X_i)$. The kernel becomes [7]:

$$K_H : dX \rightarrow K_H(dX) = \det(H)^{-1} K \left(\sqrt{dX^T H^{-2} dX} \right)$$

where K is the univariate kernel function (we use Normal Gaussian function). H is invertible as positive definite. The difficulty remains in the choice over the matrix H . We follow ideas from Härdle in [7] and define a rule to find optimal bandwidth H_{opt} to be used. It consists in a Leave-One-Out approach where we minimize residual norm (or Least Square LS). In the next section we estimate appropriate bandwidth matrices, in both PSS and Cartesian coordinate systems, using a cardiac DTI atlas.

3 Parameter Estimation Using a Cardiac DTI Atlas

In order to estimate the interpolation parameters, we took advantage of a publicly available cardiac DTI atlas of canine hearts [12]. Our study is limited to the

left ventricle in order to match with the *in vivo* data studied in the next section. If we only know partial data $(D_{\xi_i})_1^N$ on the atlas, let α be the data coverage ratio between the known data N and the total number of voxels of the atlas LV M : $\alpha = N/M$. To reproduce *in-vivo* acquisition environment, The atlas is divided slice by slice in the long axis direction (i.e. imitating sparsely distributed short axis slices). Normal distribution noise (of variance $V = 0.01$) was added over the tensor’s corresponding DWIs. We propose to find the optimal bandwidth H_{opt} that minimizes the LS criterion over the entire ventricle domain Ω . We are looking for a trade-off between fit quality and smoothness, so we minimize the criterion:

$$CV(H) = \underbrace{\sum_{x_j \in \Omega} dist(D_{x_j}, \hat{m}_H(D_{x_j}))^2}_{\text{data fit } LS(H)} + \lambda \cdot \underbrace{\sum_{x_j \in \Omega} \|\nabla \hat{m}_H(D_{x_j})\|^2}_{\text{smoothness } Reg(H)}$$

where the distance *dist* is taken in Log-Euclidean metrics, i.e. the Frobenius norm of the log difference $dist(A, B) = \|\log(A) - \log(B)\|$. The criterion $CV(H)$ to be minimized thus includes a scalar λ to control the regularization. We chose to use a gradient-free multivariate optimization scheme derived from Powell & al. [13] to minimize $CV(H)$. Furthermore, H was constrained to be diagonal. Note that optimal bandwidth matrices are different in PSS interpolation (H_{opt}^p) and in Cartesian interpolation (H_{opt}^c).

We ran the minimization process for several values of λ to find a suitable regularization ratio. Fig. 2.a shows a log-log scatter graph of the regularization term vs. the data fit term as result of the L-curve in PSS coordinates. We see an inflexion point of maximal curvature at $\lambda_{opt} = 1.5$ for a coverage of $\alpha = 60\%$. We used this value of λ for the remaining of this paper.

The long axis data coverage ratio α has a physical meaning: in practice, a full coverage of the ventricle is of great difficulty and *in vivo* data is often limited to short axis views around the equatorial region. We thus reproduced the optimization experiment varying α . The evolution of $CV(H_{opt})$ against α is shown in Fig. 2.b. At low data coverage the quality of data fit using PSS interpolation is more than 100% better compared to Cartesian interpolation. Moreover we achieve a very good fit after 50% of coverage, while it is only above 80% that we have similar fit with both coordinate systems.

The eigenvalues of H_{opt} (bandwidths in each axis direction) are of great importance as they will be used as spatial coherence prior for the *in vivo* study. They depend on the noise level as well as the data coverage ratio of the ventricle. Fig. 3 summarizes the variation of these optimal bandwidths against the α ratio, using respectively Cartesian and PSS coordinate systems for interpolation. As the ventricle is aligned to the z-axis, a decrease of its corresponding H_{opt} eigenvalue in Cartesian coordinates can be seen (Fig. 3.a). Cartesian bandwidths at full coverage converge to a value of 0.9 mm. In the PSS coordinate case (3.b), optimal bandwidths in both g_2 and g_3 directions are decreasing while the one corresponding to g_1 is stable. They converge respectively to 4 % of wall thickness, 3.4 deg. in apex-base direction, and 6.3 deg. in circumferential direction. Now that we

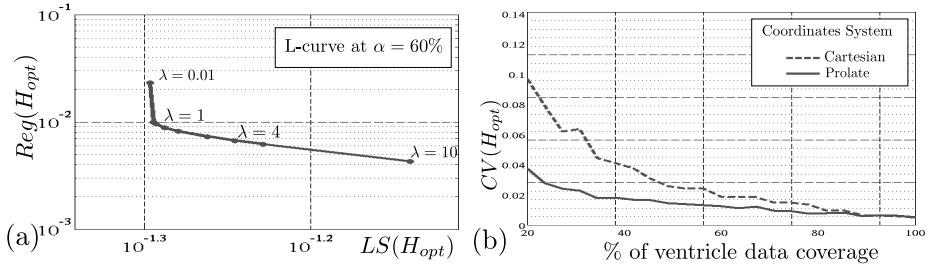


Fig. 2. (a): L-curve of the regularization parameter λ in the PSS case for a data coverage ratio $\alpha = 60\%$. The curve shows an inflexion point at $\lambda_{opt} = 1.5$. (b): Evolution of the criterion $CV(H_{opt})$, using $\lambda = \lambda_{opt}$, against the data coverage ratio α . We compare the PSS interpolation performance (solid) versus Cartesian interpolation one (dashed). We see better fit to data in PSS interpolation, especially for low values of α .

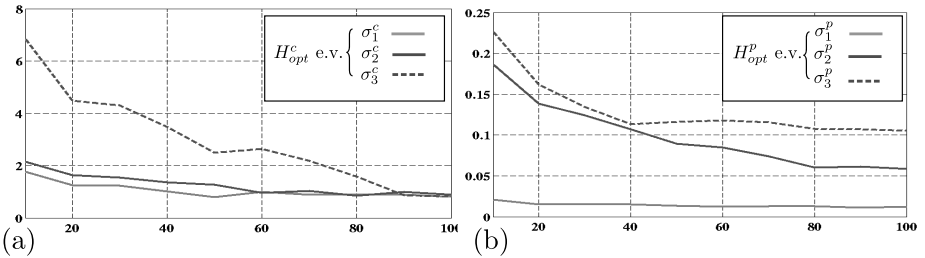


Fig. 3. H_{opt} eigenvalues obtained function of the data coverage α , using Cartesian interpolation (a), or PSS interpolation (b). While σ_1^c and σ_2^c (corresponding to x and y directions) stay stable, σ_3^c (z direction) decreases as the long-axis data resolution raises. Note also the stability of σ_1^p which corresponds to the transmural direction.

estimated optimal bandwidth parameters for our curvilinear interpolation, we can apply this framework to *in-vivo* data.

4 Application to *in-vivo* Human Data

Multi-slice cardiac DTI data were acquired on a healthy volunteer using a single-shot spin echo sequence on a 3T Philips Achieva System (Philips Healthcare, Best, The Netherlands) with reduced field-of-view excitation and flow compensated diffusion encoding gradients [5]. A total of thirty-one LV short-axis DTI views were obtained around the equatorial level of the left ventricle. The data was acquired in seven separate sessions on the same subject, totalling one hour and thirty minutes of effective scanning. Image acquisition was triggered to peak systole to account for strain effects and utilise maximum myocardial thickness [5]. To account for inconsistencies in volunteer and breath-hold positions among the separate scans, 3D misalignment was corrected using the Insight ToolKit (<http://www.itk.org>). An additional anatomical 3D dataset was acquired at the same trigger delay (TD

= 300ms) as the DTI scans. This anatomical dataset was used to define the left ventricle binary mask to be mapped on the prolate spheroid.

Fig. 4a shows the input data distribution over the ventricle. Most of the acquisitions have been done in the equatorial part of the LV, because cardiac motion (especially rotation) is of less intensity in this region. We applied the full method detailed above to interpolate tensor data over the entire LV area both using Cartesian coordinates and PSS coordinates. We used the optimal bandwidths obtained at the same coverage level in the atlas study. In that sense we are using the spatial coherence of the atlas as a prior information rather than the atlas tensor data itself. Fig. 4b shows the evolution of fit to data over the number of slices taken into account. The graph demonstrates better performance of PSS interpolation.

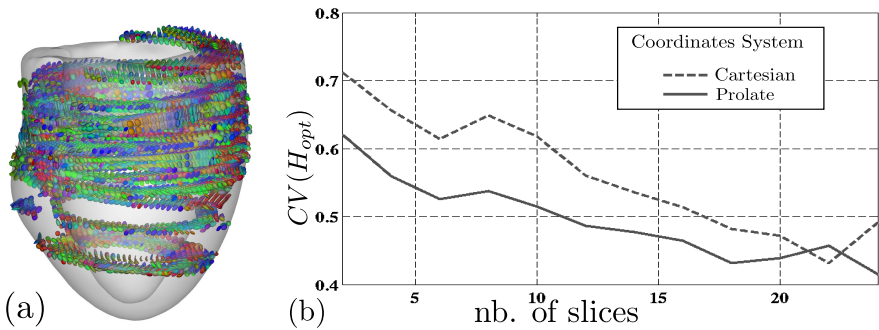


Fig. 4. The *in vivo* DTI 2D slices were acquired at different occasions (they are not all parallel to each other), and registered to the anatomical left ventricle. We show tensors color-coded by their first eigenvector direction superimposed with the ventricle outline in (a): The lack of data around the apex is visible. (b) compares performances of interpolation in Cartesian (dashed) and PSS (solid) case against the number of slices.

The helix angle is the angle that the tensor first eigenvector makes with the (g_1, g_3) plane (see Fig. 1). The variation of this angle from endocardium to epicardium is an important characteristic of the cardiac function, clearly seen in the DTI atlas and crucial in heart modelling. We show in Fig. 5 joint histograms of the angle against the normalized transmural distance from endocardium to epicardium. The pattern of this variation is reproduced when PSS Coordinates are used, whereas it is less visible in the Cartesian case. Fig. 5 shows a selected slice resulting from *in vivo* DTI data using Cartesian (Fig. 5.a) and PSS (Fig. 5.b). Tensors are color-coded with the main eigenvector direction. The regularization follows the ventricle shape in the PSS case and thus shows more plausible fibre directions. Finally we were able to track fibres (Fig. 5.c) from these tensors using streamline algorithm. The *in vivo* fibres have a general circumferential pattern and we can recognize the main helical layers of the endocardium and the epicardium.

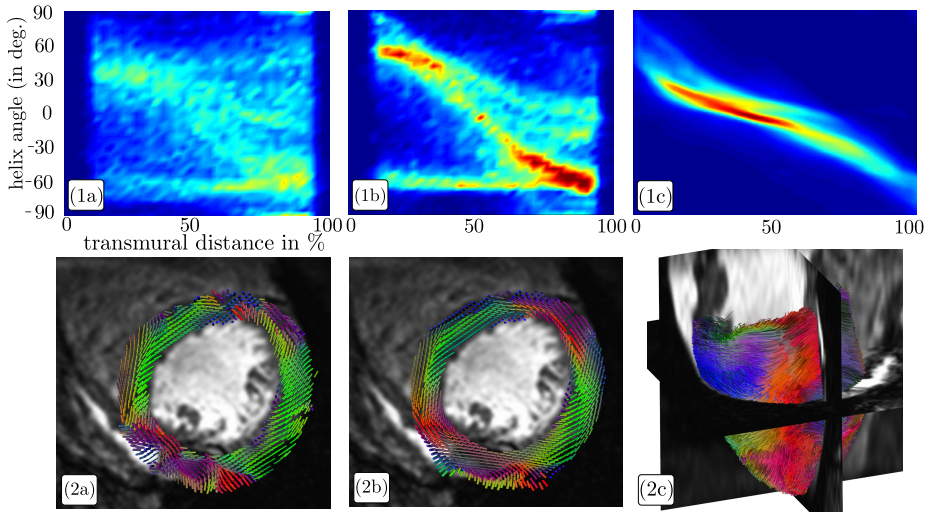


Fig. 5. Top: Joint histograms of the elevation (or helix) angle and the normalized transmural distance from endo to epi. (1a) *in-vivo* interpolated results using Cartesian coordinates, (1b) *in-vivo* interpolated results using PSS coordinates, and (1c) as a reference, the fully sampled LV statistical atlas. The correlation is visible using PSS coordinates. **Bottom:** Interpolated DTI slice color coded by eigenvector direction, using Cartesian Coordinates (2a) and PSS (2b). (2c) is a streamline fibre tractography result from the PSS interpolated tensor field.

5 Conclusions

In this paper, we demonstrated that shape adapted curvilinear coordinates – Prolate Spheroidal – are appropriate for the tensor reconstruction over the left ventricle wall volume. We set up a mathematical framework for the kernel regression of tensor data in those coordinates, using anisotropic kernel regression with an optimised bandwidth matrix. We have shown that the resulting interpolated tensors better fit the physiomy of the heart. As the left ventricle has a very characteristic ellipsoidal shape, its fibre architecture (and thus the underlying tensor field) has an important spatial coherence in PSS coordinates, whereas it is less sensible in Cartesian coordinates. We have shown that using the PSS anisotropic spatial coherence of a statistical cardiac DTI atlas as a prior information for *in vivo* tensor interpolation and regularization helps us to reconstruct full tensor information. We applied our method to reconstruct the fibre architecture of the left ventricle of a healthy volunteer, and, to the best of our knowledge, it is the first time that the *in vivo* human 3D structure of the heart has been reconstructed. The *in vivo* results show a good correlation with literature values of *ex vivo* human studies. We were able to reproduce the typical pattern of transmural variation of the helix angle. The presented approach opens up possibilities in terms of analysis of the fibre architecture in patients.

References

1. Dieudonne, J.M.: The left ventricle as confocal prolate spheroids. *Bulletin of Mathematical Biology* 31(3), 433–439 (1967)
2. Dou, J., Tseng, W.Y.I., Reese, T.G., Wedeen, V.J.: Combined diffusion and strain MRI reveals structure and function of human myocardial laminar sheets in vivo. *Magnetic Resonance in Medicine* 50, 107–113 (2003)
3. Dru, F., Vercauteren, T.: An ITK implementation of the symmetric log-domain diffeomorphic demons algorithm. *Insight Journal* –2009 January - June 2009 (May 2009)
4. Fillard, P., Arsigny, V., Pennec, X., Ayache, N.: Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics. *IEEE Transactions on Medical Imaging* 26(11), 1472–1482 (2007); pMID: 18041263
5. Gamper, U., Boesiger, P., Kozerke, S.: Diffusion imaging of the in vivo heart using spin echoes - considerations on bulk motion sensitivity. *Magnetic Resonance in Medicine* 57, 331–337 (2007)
6. Garrido, L., Wedeen, V.J., Kwong, K.K., Spencer, U.M., Kantor, H.L.: Anisotropy of water diffusion in the myocardium of the rat. *Circulation Research* 74, 789–793 (1994)
7. Härdle, W.K., Marron, J.S.: Optimal bandwidth selection in nonparametric regression function estimation. *Annals of Statistics* 13(4), 1465–1481 (1985)
8. Kanai, A., Salama, G.: Optical mapping reveals that repolarization spreads anisotropically and is guided by fiber orientation in guinea pig hearts. *Circulation Research* 77, 784–802 (1995)
9. LeGrice, I.J., Takayama, Y., Covell, J.W.: The architecture of the heart: a data-based model. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* 359, 1217–1232 (2001)
10. Mansi, T., Peyrat, J.M., Sermesant, M., Delingette, H., Blanc, J., Boudjemline, Y., Ayache, N.: Physically-constrained diffeomorphic demons for the estimation of 3d myocardium strain from cine-MRI. In: Ayache, N., Delingette, H., Sermesant, M. (eds.) *FIMH 2009*. LNCS, vol. 5528, pp. 201–210. Springer, Heidelberg (2009)
11. Nielsen, P.M., Grice, I.J.L., Smaill, B.H., Hunter, P.J.: Mathematical model of geometry and fibrous structure of the heart. *Am. J. Physiol. Heart Circ. Physiol.* 260, 1365–1378 (1995)
12. Peyrat, J.M., Sermesant, M., Pennec, X., Hervé Delingette, C.X., McVeigh, E.R., Ayache, N.: A computational framework for the statistical analysis of cardiac diffusion tensors: Application to a small database of canine hearts. *IEEE Transactions in Medical Imaging* 25(5), 612–625 (2006)
13. Powell, M.J.D.: Developments of newuoa for minimization without derivatives. *Journal of Numerical Analysis*, 1–16 (February 2008)
14. Reese, T.G., Weisskoff, R.M., Smith, R.N., Rosen, B.R., Dinsmore, R.E., Wedeen, V.J.: Imaging myocardial fiber architecture in vivo with magnetic resonance. *Magnetic Resonance in Medicine* 34, 786–791 (1995)
15. Rohmer, D., Gullberg, G.T.: A bloch-torrey equation for diffusion in a deforming media. Tech. rep., University of California (2006)
16. Sutton, M.G.S.J., Sharpe, N.: Left ventricular remodeling after myocardial infarction. *Circulation: Cardiovascular Imaging* 101, 2981–2988 (2000)
17. Tseng, W.Y.I., Wedeen, V.J., Reese, T.G., Smith, R.N., Halpern, E.F.: Diffusion tensor MRI of myocardial fibers and sheets: Correspondence with visible cut-face texture. *Magnetic Resonance in Medicine* 42, 17–31 (2003)

Physics-Based Modeling of Aortic Wall Motion from ECG-Gated 4D Computed Tomography

Guanglei Xiong¹ and Charles A. Taylor²

¹ Biomedical Informatics Program, Stanford University, CA, USA

² Departments of Bioengineering and Surgery, Stanford University, CA, USA
{glxiong, taylorca}@stanford.edu

Abstract. Recent advances in electrocardiogram (ECG)-gated Computed Tomography (CT) technology provide 4D (3D+T) information of aortic wall motion in high spatial and temporal resolution. However, imaging artifacts, e.g. noise, partial volume effect, misregistration and/or motion blurring may preclude its usability in many applications where accuracy and reliability are concerns. Although it is possible to find correspondence through tagged MRI or echo or image registration, it may be either inconsistent to the physics or difficult to utilize data from all frames. In this paper, we propose a physics-based filtering approach to construct a dynamic model from these 4D images. It includes a state filter that corrects simulated displacements from an elastic finite element model to match observed motion from images. In the meantime, the model parameters are refined to improve the model quality by applying a parameter filter based on ensemble Kalman filtering. We evaluated the performance of our method on synthetic data where ground-truths are available. Finally, we successfully applied the method to a real data set.

Keywords: physics-based modeling, dynamic model, aorta, wall motion, Kalman filtering, 4DCT.

1 Introduction

CT angiography (CTA) is now considered as a preferred technique to conventional X-ray angiogram for assessing of diseases of heart, coronary artery, and thoracic aorta [1]. The development of CTA with ECG gating has been driven by imaging of coronary artery but it is also quite beneficial to the imaging of the thoracic aorta [2]. Although the frame rate has not yet reached typical video rates, it is already adequate to characterize the aortic wall motion. Visualization using real-time rendering has been routinely used to qualitatively examine functional abnormalities of the aorta that are not obvious on static images. In contrast, only a limited number of preliminary quantitative studies were carried out in healthy [3] or disease-related cases [4].

Although these studies have provided useful insights of aortic wall motion, there are several drawbacks that need to be improved, which is the purpose of this paper. First, imaging artifacts, such as noise, partial volume effect, misregistration, and/or motion blurring, should be explicitly considered since the goal is to quantify the motion. A physics-based approach may be helpful to overcome the effects of these artifacts, because the

motion of the aorta should obey general principles of elasticity. Another advantage of this approach is the ability to obtain not only kinematic quantities, e.g. strain, but also the dynamic quantities, e.g. stress. Second, all data, both spatial and temporal, should be utilized in the analysis. It is beneficial to make use of full volumes of images instead of 2D slices and all frames in a cycle rather than merely two single systolic and diastolic phases. Third, it is helpful to establish correspondences or trajectories of material points across all frames. Although tagged MRI or echo or image registration is traditionally used for this task, the problems are inconsistency to the physics and difficulty to utilize data from all frames. A compromise between the image data and the physical model may be necessary in this respect. Lastly, the model could be improved by refining its parameters, i.e. the physical properties of the wall. This step is necessary and important because the dynamic model only becomes valid and useful once the parameters are correct. In other words, the model with reasonable parameters explains better what are observed in images.

A relevant area is to extract heart (or mostly left ventricle) motion in cardiac-gated SPECT or CT. Most recently, temporal relationship is explicitly employed in some studies. To ensure robustness of the segmentation, temporally smoothing regularization in deformable models is naturally taken into account [5, 6]. However, these models are empirical, not necessarily physical. Until recently, Moireau *et al.* proposed filtering techniques to estimate heart motion from 4D position measurements [7], although results were only demonstrated on synthetic data. In contrast of heart motion, studies of quantification of thoracic aorta motion are lacking. While pulsation in the thoracic aorta is not as large as that in the heart, the cyclic strain is still quite significant, up to 10% [8], which may have critical implications on surgical interventions, e.g. endovascular aneurysm repair [4]. Therefore, it is demanding of a systematic method to model dynamics of thoracic aorta with the goal of facilitating quantification of thoracic aorta motion. Instead of using image registration approaches, we seek a physics-based filtering approach for this task.

2 Segmentation and Surface Mesh

Figure 1 shows a volume-rendered data set of healthy thoracic aorta from ECG-gated 4D CT, which has 10 frames per cardiac cycle. The size for each frame is $512 \times 512 \times 482$. The spatial resolution is $0.62 \times 0.62 \times 1.0$ mm. The anatomy of interest includes ascending thoracic aorta, aortic arch, descending thoracic aorta, as well as three main branches: brachiocephalic trunk, left common carotid artery, and left subclavian artery. Prominent motion artifacts are observed in systolic phase (frames 0-3), especially at ascending aorta.

To consider thoracic aorta itself, we need to separate it from other objects, e.g. heart, veins, and bones, etc. We chose to segment thoracic aorta for each frame (shown in Fig. 2) by geodesic active contour method [9]. Only the first frame needs to be segmented from scratch. Segmentations of the following frames are simpler if initialized by that of the last frame noting the difference between neighboring frames is relatively small.

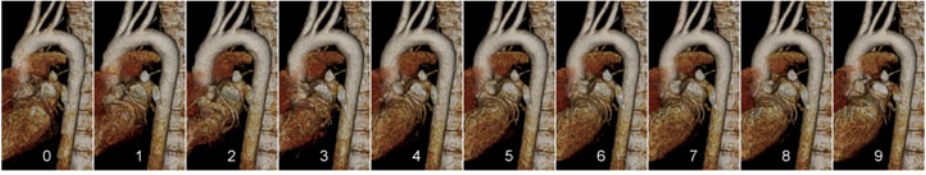


Fig. 1. A volume-rendered data set of thoracic aorta from ECG-gated 4D CT

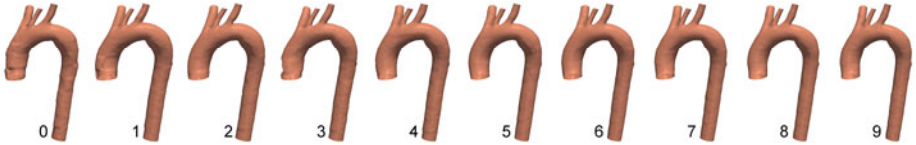


Fig. 2. Segmentations of thoracic aorta in Fig. 1

Since the end-diastole frame (frame 9) often has least noise, the segmentation of this frame is used to extract the triangular mesh (nodes: 1622, triangles: 3144) of the aorta as in Fig. 3(a), upon which will be solved by the finite element method. The mesh is trimmed perpendicular to the vessel wall at all the outlets to get planar boundaries. Note we also include three main branches coming off the aorta arch in the geometric model in Fig. 3(b).

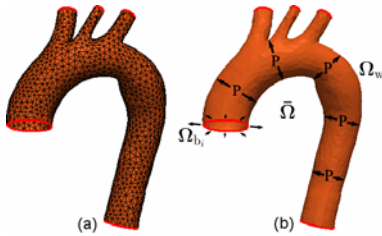


Fig. 3. Surface mesh (a) and geometric model (b)

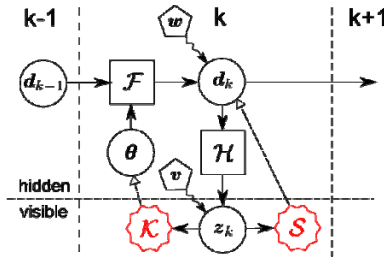


Fig. 4. Overview of our approach (details in text)

3 Methods

Figure 4 provides a graphical overview of the proposed physics-based filtering approach to construct a dynamic model of thoracic aorta. The dynamics of the aorta is modeled by an operator \mathcal{F} (defined in Section 3.1), which relates the displacements \mathbf{d} (named as states in the sequel) at current time step k from $k-1$, with the mechanical properties θ , i.e. Young’s modulus (named as parameters), and some additive model error w :

$$\mathbf{d}_k = \mathcal{F}(\mathbf{d}_{k-1}, \theta) + w \tag{1}$$

The observed displacements \mathbf{z} (named as measurements) are led by an operator \mathcal{H} , from the actual displacements \mathbf{d} , subject to some measurement error v :

$$\mathbf{z}_k = \mathcal{H}(\mathbf{d}_k) + \mathbf{v} \quad (2)$$

Our goals are

(a) To develop a state filter \mathcal{S} , which corrects simulated displacements with measurements (Section 3.2):

$$\hat{\mathbf{d}}_k = \mathcal{S}(\mathbf{d}_k, \boldsymbol{\theta}, \mathbf{z}_k) \quad (3)$$

(b) To develop a parameter filter \mathcal{K} , which improves the model by refining the parameters (Section 3.3):

$$\hat{\boldsymbol{\theta}} = \mathcal{K}(\mathbf{d}_k, \boldsymbol{\theta}, \mathbf{z}_k) \quad (4)$$

Note we separate entities hidden from visible in Fig. 4.

3.1 Physical Model

The dynamics of aorta is modeled as a linear elastic membrane. The linearity is justified by the aortic deformation is generally less than 10% and we do not account for the visco-elastic effects for our scope. Let $\bar{\Omega} = \Omega_w \cup \Omega_{b_1} \cup \Omega_{b_2} \cup \dots \cup \Omega_{b_n}$ be the entire aorta as in Fig. 3(b), where Ω_w is the aortic wall and Ω_{b_i} is the i -th boundary contour. The displacement $\mathbf{u} = [u_x, u_y, u_z]$ satisfies:

$$\rho \ddot{\mathbf{u}} = \nabla \cdot \boldsymbol{\sigma} + \mathbf{f} \quad (5)$$

where the stress $\boldsymbol{\sigma} = [\sigma_{xx}, \sigma_{yy}, \tau_{xy}, \tau_{xz}, \tau_{yz}]$ is [10]:

$$\frac{E}{1-\nu^2} \begin{bmatrix} 1 & \nu & 0 & 0 & 0 \\ \nu & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1-\nu}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{5(1-\nu)}{12} & 0 \\ 0 & 0 & 0 & 0 & \frac{5(1-\nu)}{12} \end{bmatrix} \begin{bmatrix} \frac{\partial u_x}{\partial x} \\ \frac{\partial u_y}{\partial y} \\ \frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x} \\ \frac{\partial u_x}{\partial z} \\ \frac{\partial u_z}{\partial y} \end{bmatrix} \quad (6)$$

where E is Young's modulus and $\nu = 0.5$ is the Poisson's ratio. The body force \mathbf{f} is related to driving internal pressure (see Fig. 3(b)) and is periodic.

Using finite element method [10], equation (5) reduces to the matrix form:

$$\mathbf{M} \ddot{\mathbf{d}}(t) + \mathbf{C} \dot{\mathbf{d}}(t) + \mathbf{K} \mathbf{d}(t) = \mathbf{F}(t) \quad (7)$$

where \mathbf{M} , \mathbf{C} , and \mathbf{K} are mass, damping, and stiffness matrices, respectively. To eliminate acceleration $\ddot{\mathbf{d}}$ and velocity $\dot{\mathbf{d}}$, we use a two-step Newmark's time stepping method [11] to solve for displacements \mathbf{d} , which leads to the operator \mathcal{F} in Eq. (1):

$$\begin{aligned} & \left[\mathbf{M} + \gamma \Delta t \mathbf{C} + \beta \Delta t^2 \mathbf{K} \right] \mathbf{d}_{k+1} + \left[-2\mathbf{M} + (1-2\gamma)\Delta t \mathbf{C} + \left(\frac{1}{2} - 2\beta + \gamma \right) \Delta t^2 \mathbf{K} \right] \mathbf{d}_k \\ & + \left[\mathbf{M} - (1-\gamma)\Delta t \mathbf{C} + \left(\frac{1}{2} + \beta - \gamma \right) \Delta t^2 \mathbf{K} \right] \mathbf{d}_{k-1} = \Delta t^2 \left[\beta \mathbf{F}_{k+1} + \left(\frac{1}{2} - 2\beta + \gamma \right) \mathbf{F}_k + \left(\frac{1}{2} + \beta - \gamma \right) \mathbf{F}_{k-1} \right] \end{aligned} \quad (8)$$

where Δt is the time step size. $\beta = 0.25$ and $\gamma = 0.5$ retain second-order accuracy and unconditional stability. For the nodes on the boundary contours Ω_{b_i} , we allow them to deform freely in the radial direction, while keeping them fixed longitudinally by

specifying Dirichlet boundary condition $\mathbf{u} \cdot \mathbf{n}_{b_i} = 0$ for nodes on Ω_{b_i} in Eq. (5), where \mathbf{n}_{b_i} is the normal of the corresponding trimming plane. The operator \mathcal{H} in Eq. (2) reduces to an observation matrix \mathbf{H} where its element is 1 if the corresponding displacement component has measurement, otherwise 0.

3.2 State Filter

To correct simulated displacements from the physical model, we propose to use a state filter to weigh the trajectories between the model and the measurement as shown in Fig. 5. Our approach is justified from the fact that either the model or the measurement is not perfect, i.e. subject to errors as discussed before.

We define a joint strain energy $\Pi(\bar{\mathbf{d}})$ from both contributions:

$$\underbrace{(\bar{\mathbf{d}} - \mathbf{d})^T \mathbf{K} (\bar{\mathbf{d}} - \mathbf{d})}_{\text{from model}} + \varepsilon \underbrace{(\mathbf{z} - \mathbf{H}\bar{\mathbf{d}})^T \mathbf{W}_{vv}^{-1} (\mathbf{z} - \mathbf{H}\bar{\mathbf{d}})}_{\text{from measurement}} \tag{9}$$

where ε controls the relative weighting between model and measurement. Clearly, small ε pulls the trajectory closer to the model, while large ε favors measurement. The matrix \mathbf{W}_{vv} is the covariance of the measurement errors, which is $\text{var}(\mathbf{v})\mathbf{I}$ if they are the same. The filtered trajectory is then $\hat{\mathbf{d}}$ that minimizes $\Pi(\bar{\mathbf{d}})$, i.e. $\hat{\mathbf{d}} = \text{argmin}_{\bar{\mathbf{d}}} \Pi(\bar{\mathbf{d}})$. A simple differentiation leads to:

$$\hat{\mathbf{d}} = (\mathbf{K} + \varepsilon \mathbf{H}^T \mathbf{W}_{vv}^{-1} \mathbf{H})^{-1} (\mathbf{K} \mathbf{d} + \varepsilon \mathbf{H}^T \mathbf{W}_{vv}^{-1} \mathbf{z}) \tag{10}$$

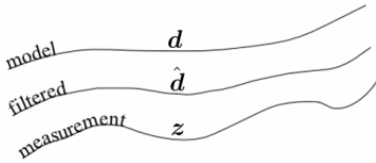


Fig. 5. The state filter weighs between model and measurement

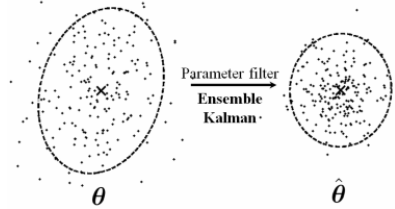


Fig. 6. The parameter filter based on ensemble Kalman filtering

3.3 Parameter Filter

To improve the physical model, we employ ensemble Kalman filtering [12] (EnKF) to refine the parameters. EnKF was originated from well-known Kalman filter but specialized for solving large problems as in our problem. It approximates probability densities of system variables as ensembles, i.e. collections of particles (Fig. 6).

In our case, an ensemble for displacements or states $\mathbf{D} = [\mathbf{d}^1, \mathbf{d}^2, \dots, \mathbf{d}^N]$ and an ensemble for parameters $\Theta = [\theta^1, \theta^2, \dots, \theta^N]$ are defined, where N is the number of particles. In general, N should be large enough to represent the underlying distribution. We choose $N = 64$ considering the computational cost. Also by subtracting the

means, we have the state and parameter perturbations: $\tilde{D} = D - \sum_{n=1}^N d_n \mathbf{1}_N$ and $\tilde{\Theta} = \Theta - \sum_{n=1}^N \theta_n \mathbf{1}_N$, where $\mathbf{1}_N$ is a matrix with all elements equal to $1/N$. The measurement ensemble is $Z = [z^1, z^2, \dots, z^N]$, where each particle $z^n = z + v^n$ is generated by perturbing random variables v^n sampled from the distribution of measurement error v . Using the Kalman gain

$$K = (H\tilde{D})^T \left[(H\tilde{D})(H\tilde{D})^T + (N-1)W_{vv} \right]^{-1} \quad (11)$$

The refined parameter ensemble can be written as

$$\hat{\Theta} = \Theta + \tilde{\Theta}K(Z - HD) \quad (12)$$

Finally, the filtered parameter $\hat{\theta}$ is the ensemble mean of $\hat{\Theta}$. The uncertainty of $\hat{\theta}$ can be estimated by calculating the ensemble variance of $\hat{\Theta}$. The uncertainty generally decreases after applying the parameter filter as in Fig. 6.

4 Results

4.1 Synthetic Data

The geometric model was divided into 8 regions, 5 of which are on the aortic wall. Each region was assigned different Young's modulus as shown in Fig.7. The synthetic data were generated with this set of parameters by simulating the dynamics using Eq. (8). The time step size is 1/200 of the cardiac cycle, which is 1.0 s. The model and measurement errors are Gaussian with standard deviations 0.4 mm and 0.5 mm, respectively.

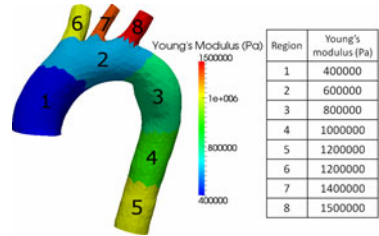


Fig. 7. Division of the geometric model. Each region has a different Young's modulus

The state filter was applied to correct the displacements using measurement. Note the measured displacements were subject to measurement error and the true displacements were blinded to the filter. Figure 8 presents the mean distances from the filtered model to the true model and the effects of different weighting ε . The distance for the non-filtered model is served as the baseline. As ε increases, the filtered model gets closer to the true model. But it cannot coincide with the true model considering the measurement error. The distance becomes flat after a short period.

Next, we consider refine the parameters for 5 regions on the aortic wall. The parameter ensembles were assigned by randomly sampling from a uniform distribution with minimum 0.5×10^6 and maximum 2.0×10^6 . The parameter filter was applied while using state filter with $\varepsilon = 10^3$. This weighting was selected due to the fact that we need to leave enough differences between simulated displacements and measurements in Eq. (12) to update parameters while correcting the displacement. This is justified because the parameter refinement will make the model trajectory closer to the measurement.

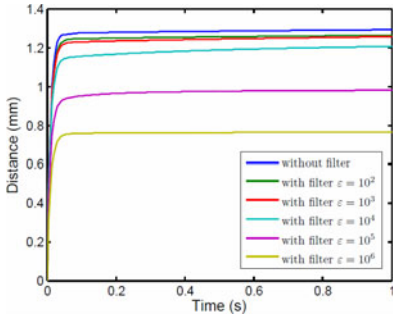


Fig. 8. Mean distance of the filtered model to the true model

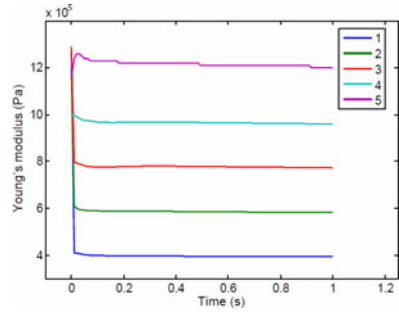


Fig. 9. Parameter refinement on synthetic data

Figure 9 shows the parameter refinement process. We can see that the parameter filter was very efficient to push the parameters to correct values, in less than 1/5 of the cardiac cycle.

4.2 Real Data

We consider the real data in Fig. 1. Results with and without the state filter are shown in Fig. 10. The non-filtered result is just from the segmentation where we notice the irregular trajectory of mesh nodes. For the filtered case, the model and measurement

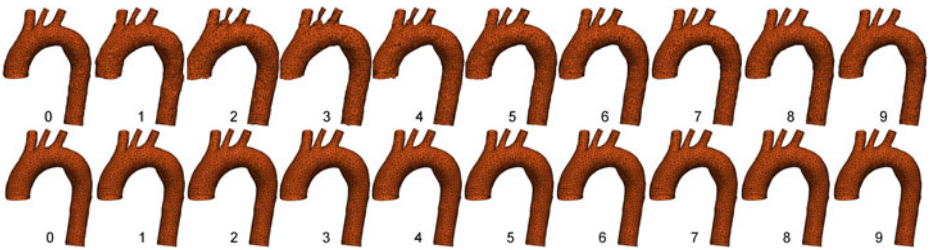


Fig. 10. Results on real data without state filter (top) and with state filter (bottom)

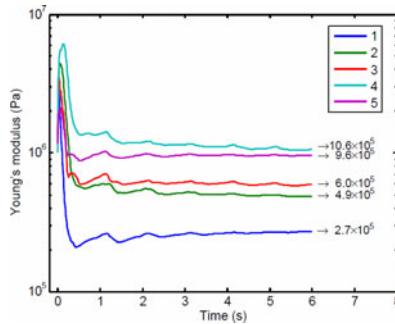


Fig. 11. Parameter refinement on real data

errors are set to be 0.1 mm and 0.5 mm. The weighting ε is 10^3 . After the state filter, the trajectories are much smoother. More results with other weightings are provided in supplementary materials. Next, parameter was refined (as shown in Fig. 11) using the proposed parameter filter with the same divisions of the geometric model in Fig. 7. We arrive at a set of meaningful parameters: “*distal is stiffer than proximal*”.

5 Discussion and Future Work

We have proposed physics-based filtering techniques that can be used to model the dynamics of thoracic aorta from ECG-gated 4D CT data. The proposed approach features physical model of thoracic aorta, a state filter to correct simulated displacements, and a parameter filter to refine the model parameters. We have tested our method on the synthetic data where the ground truth is available and successfully applied it to a real data set with no known parameters. Our results demonstrate that the approach may be useful to obtain subject-specific kinematic information from 4D CT and to quantify the dynamics of the thoracic aorta. We believe it is generalizable to other situation involved in 4D data.

In the future, we will apply the method to more real data sets, especially to some diseased cases, e.g. aneurysms where the wall stiffens. In addition, since we are using a temporally varying but spatially uniform pressure internal to the aortic wall, minor pressure loss and shear force are not currently taken into account. A possibility is to substitute a more sophisticated model of fluid structure interaction. However, this approach needs to be justified since the computational cost dramatically increases and may make parameter filtering intractable in practice. Finally, the parameters obtained on the real data need to be validated before becoming more physically meaningful.

Acknowledgements

We are grateful to Nan Xiao, Dr. Philippe Moireau and Dr. C. Alberto Figueroa for helpful discussions.

References

1. Rubin, G.D.: Helical CT Angiography of the Thoracic Aorta. *Journal of Thoracic Imaging* 12, 128–149 (1997)
2. Marten, K., Funke, M., Rummeny, E.J., Engelke, C.: Electrocardiographic Assistance in Multidetector CT of Thoracic Disorders. *Clinical Radiology* 60, 8–21 (2005)
3. Lin, F.Y., Devereux, R.B., Roman, M.J., Meng, J., Jow, V.M., Jacobs, A., Weinsaft, J.W., Shaw, L.J., Berman, D.S., Gilmore, A., Callister, T.Q., Min, J.K.: Assessment of the Thoracic Aorta by Multidetector Computed Tomography: Age- and Sex-Specific Reference Values in Adults without Evident Cardiovascular Disease. *Journal of Cardiovascular Computed Tomography* 2, 298–308 (2008)
4. van Prehn, J., Bartels, L.W., Mestres, G., Vincken, K.L., Prokop, M., Verhagen, H.J., Moll, F.L., van Herwaarden, J.A.: Dynamic Aortic Changes in Patients with Thoracic Aortic Aneurysms Evaluated with Electrocardiography-Triggered Computed Tomographic Angiography before and after Thoracic Endovascular Aneurysm Repair: Preliminary Results. *Annals of Vascular Surgery* 23, 291–297 (2009)

5. Montagnat, J., Delingette, H.: 4D Deformable Models with Temporal Constraints: Application to 4D Cardiac Image Segmentation. *Medical Image Analysis* 9, 87–100 (2005)
6. Brankov, J.G., Yang, Y., Wernick, M.N.: Spatiotemporal Processing of Gated Cardiac Spect Images Using Deformable Mesh Modeling. *Medical Physics* 32, 2839–2849 (2005)
7. Moireau, P., Chapelle, D., Le Tallec, P.: Filtering for Distributed Mechanical Systems Using Position Measurements: Perspectives in Medical Imaging. *Inverse Problems* 25, 035010 (2009)
8. Morrison, T.M., Choi, G., Zarins, C.K., Taylor, C.A.: Circumferential and Longitudinal Cyclic Strain of the Human Thoracic Aorta: Age-Related Changes. *Journal of Vascular Surgery* 49, 1029–1036 (2009)
9. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic Active Contours. *International Journal of Computer Vision* 22, 61–79 (1997)
10. Hughes, T.J.R.: *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Prentice-Hall, Englewood Cliffs (1987)
11. Zienkiewicz, O.C.: New Look at Newmark, Houbolt and Other Time Stepping Formulas - Weighted Residual Approach. *Earthquake Engineering & Structural Dynamics* 5, 413–418 (1977)
12. Evensen, G.: *Data Assimilation: The Ensemble Kalman Filter*. Springer, New York (2009)

Whole Heart Segmentation of Cardiac MRI Using Multiple Path Propagation Strategy

X. Zhuang^{1,*}, K. Leung^{1,2}, K. Rhode³, R. Razavi³, D. Hawkes¹,
and S. Ourselin^{1,2}

¹ CMIC, Medical Physics Department, University College London
x.zhuang@ucl.ac.uk

<http://www.cs.ucl.ac.uk/staff/x.zhuang/>

² Dementia Research Centre, Institute of Neurology, University College London

³ Division of Imaging Science, King's College London

Abstract. Automatic segmentation of cardiac MRI is an important but challenging task in clinical study of cardiac morphology. Recently, fusing segmentations from multiple classifiers has been shown to achieve more accurate results than a single classifier. In this work, we propose a new strategy, MULTiple Path Propagation and Segmentation (MUPPS), in contrast with the currently widely used multi-atlas propagation and segmentation (MAPS) scheme. We showed that MUPPS outperformed the standard MAPS in the experiment using twenty-one *in vivo* cardiac MR images. Furthermore, we studied and compared different path selection strategies for the MUPPS, to pursue an efficient implementation of the segmentation framework. We showed that the path ranking scheme using the image similarity after an affine registration converged faster and only needed eleven classifiers from the atlas repository. The fusion of eleven propagation results using the proposed path ranking scheme achieved a mean Dice score of 0.911 in the whole heart segmentation and the highest gain of accuracy was obtained from myocardium segmentation.

1 Introduction

Magnetic resonance imaging (MRI) has become a routine modality for the determination of patient cardiac morphology. The extraction of this morphological information can be important for the development of new clinical applications as well as the planning and guidance of cardiac interventional procedures. Manual delineation is labor intensive and subject to inter- and intra-observer variability. Therefore, it is highly desirable to develop an automatic technique for whole heart segmentation of cardiac magnetic resonance (MR) images. However, automating this process is complicated by the limited quality of acquired images and large shape variation of the heart between subjects.

Many works have shown the applicability of registration-based atlas propagation for the automatic segmentation, particularly in brain MR segmentation,

* This work has been funded by EPSRC grant EP/H02025X/1.

where the multi-atlas propagation and segmentation (MAPS) has gained its wide popularity in recent years [1,2,3,4,5]. In MAPS, each atlas is a combination of an intensity image and a label image. The intensity image is used for the registration process to map the space of the atlas to the unseen image, while the label image contains the segmentation information for propagation to the unseen image. The label image can also have prior knowledge such as the shape of the heart to assist the registration process [6].

In MAPS, good quality images, such as MR data with high signal-to-noise ratios and minimal artifacts, are selected as atlases for propagation to maximize the registration accuracy [2,3,4,5]. However, clinical cardiac MR images often contain strong artifacts, in particular the high resolution volumetric data, due to the effects from complex heart motions and long acquisition. Therefore, the registration of these images may have large errors or even fail. Furthermore, constructing an atlas with good image quality may need a large amount of training data and may be practically expensive in terms of manual image processing and data acquisition. Therefore, in this work we propose to build a good quality atlas from limited quality images and use the Multiple Path Propagation and Segmentation (MUPPS) strategy to achieve a result of multiple classifiers.

In MUPPS, the propagation of an atlas to unseen image can be done through a number of different paths to resemble the multiple classifier strategy. A set of propagation results are then available for fusion using the available methods. However, the number of available paths from the training subjects can be large, leading to the problem of scale. Three path selection strategies are hence investigated and an efficient method is proposed.

The main contribution of this work includes: (1) propose MUPPS and demonstrate its superiority over standard MAPS; (2) investigate propagation path selection and the effect on accuracy and efficiency.

2 Data and Method

2.1 Data

The cardiac MRI sequence used in the experiment was the balanced steady state free precession (b-SSFP) for whole heart imaging, at the end diastolic phase. The sequence was implemented on a 1.5T clinical scanner (Philips Healthcare, Best, The Netherlands) equipped with 32 independent receive channels. A fat saturation and T2 preparation pulses were used to null fat and to increase the contrast between blood and cardiac muscle. A free breathing scan was realized by enabling one navigator beam before data acquisition for each cardiac phase.

A *test dataset* of twenty-one cardiac MR volumes, all with voxel size $2 \times 2 \times 2$ (mm), was used in our experiments. The blood cavities of the four chambers and the myocardium of the left ventricle of all the images in the test set were separately segmented as gold standard for accuracy assessment. In addition, the pulmonary artery, ascending aorta and aortic arch, and descending aorta were also separately delineated for the atlas construction registration. They were done by fitting a deformable mesh model [7] with manual corrections. The manual



Fig. 1. Left: an MR image. Middle: the corresponding label image of the MR image. Right: an average intensity image.

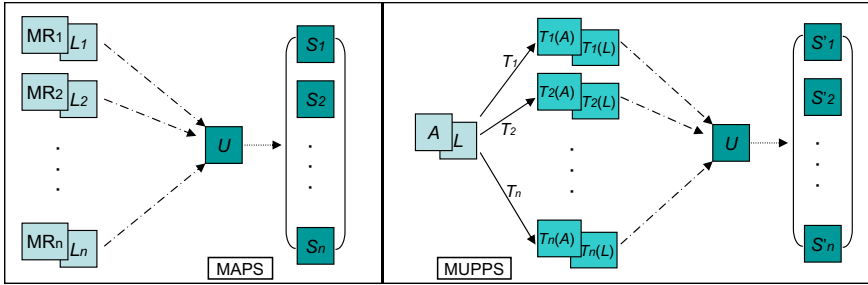


Fig. 2. The two propagation and segmentation schemes: $\{MR_i\}$ are the MR images of each subject and $\{L_i\}$ are the corresponding label images, U indicates the unseen image, and $\{S_i\}$ are the resultant propagation segmentations; A and L are the atlas intensity image and corresponding label image, $\{P_i\}$ and $\{T_i\}$ are the propagation paths and corresponding transformations.

segmentation was completed with the agreement by at least one cardiologist. Fig. 1 (left and middle) shows an MR image and its segmentation label image.

2.2 MULTIPLE Path Propagation and Segmentation (MUPPS)

In MAPS, different segmentations are obtained by registering the MR images, $\{MR_i\}$, in the atlas repository to the unseen image and propagating the corresponding labels, $\{L_i\}$, as Fig. 2 (left) shows. The clinical cardiac MR images often contain strong artifacts which may lead to large errors or even failures in the *propagation registration*. Therefore, we propose to construct a common atlas, consisting of an *average intensity image* A and the corresponding label image L , and register it to the unseen image though multiple paths to obtain multiple segmentations for label fusion, as Fig. 2 (right) illustrates.

The average intensity image A is computed from the set of training data, $\{MR_i\}$, by transforming all of them into a user-defined common space and averaging the intensity values [6]. These transformations, $\{T_i\}$, are computed from the registration of their corresponding binary label images $\{L_i\}$ and L , using the following three registration steps [6]: global affine registration, locally affine registration method (LARM) [8], and the fluid registration. The resultant average intensity image has better image quality, such as high signal and contrast-to-noise ratio and minimal artifacts, than the original atlas images in MAPS, and

the propagation registration using A is expected to achieve better results for segmentation. Fig. 1 (right) shows an example of A .

The paths in MUPPS are defined by $\{T_i\}$. For each path, we replace the original atlas images in MAPS, MR_i and L_i , by the images $T_i(A)$ and $T_i(L)$, transformed from A and L using T_i , for the propagation registration. This is equivalent to register A to the unseen image via the different paths in the shape manifold [9] defined by the spaces of $\{MR_i\}$, and hence the multiple path propagation and segmentation.

2.3 Registration and Path Ranking

Two registration procedures are used in MUPPS: propagation registration and selection registration.

Propagation registration is used to register $T_i(A)$ in MUPPS or MR_i in MAPS to the unseen image to achieve the segmentation. This process includes three steps [6]: the global affine registration, LARM using the local regions defined by the labels in the segmentation, and the FFD registration.

Selection registration is used to register the unseen image to the atlas intensity image and then compute the similarity between them to rank the path or atlas for propagation selection. In the multiple classifier strategy, the number of available atlases is normally greater than that of atlases needed to achieve the best result [3]. Therefore, the selection registration is also required to be computationally efficient compared to the propagation registration. Three registration schemes will be investigated in this work:

- the affine registration, referred to as *After affine*;
- the affine plus LARM registration, referred to as *After LARM*;
- the registration using global affine, LARM, and FFD registration, referred to as *After nonrigid*.

Path or atlas ranks are then calculated using the similarity values of the registered atlas intensity image and unseen image. We assess the similarity using the normalized mutual information (NMI) [10] with a mask on the interested region, similar to the work for brain MR MAPS [3]. The mask region is the endo- and epi-cardial surfaces of the transformed atlas label image and with an operation of morphological dilation of 5 mm.

2.4 Fusion and Accuracy Assessment

We evaluate three fusion strategies in MUPPS: the ‘vote rule’ [2], referred to as VOTE, shape-based average (SBA) [11], and the simultaneous truth and performance level estimation (STAPLE) [12]. The accuracy of segmentation results are assessed using Dice coefficient ($\frac{2|V_s \cap V_g|}{|V_s| + |V_g|}$) [13] between the segmentation and gold standard. The Dice score is calculated on the blood pool of the left ventricle (LV), left atrium (LA), right ventricle (RV), right atrium (RA), and myocardium (MYO). The mean of the volume size adjusted Dice of the five local regions is computed as the error of whole heart (Whole):

$$\text{MeanDice} = \frac{2 \sum_{i=1}^5 |V_{s_i} \cap V_{g_i}|}{\sum_{i=1}^5 (|V_{s_i}| + |V_{g_i}|)}. \quad (1)$$

We compare the mean Dice scores of MUPPS and MAPS using the two tailed, paired t-test.

3 Results and Discussions

MAPS vs MUPPS. Fig. 3 and Table 1 present the results of the standard MAPS, using the MR image of each subject as the atlas intensity image, and the proposed MUPPS, using the transformed mean atlas intensity image through different paths. In both schemes, we employed the leave-one-out to test the segmentation propagation, resulting in 420 cases in total for each scheme. Fig. 3 (left) presents the mean Dice scores of the two segmentation schemes. The mean Dice score of MUPPS was 0.059 ($P < 0.001$ and 95% confidence interval [0.0564, 0.0630]) greater than that of MAPS. Fig. 3 (right) gives the mean Dice scores of the VOTE fusion results. Both MAPS and MUPPS improved the accuracy after using the fusion. However, the best accuracy of MAPS was lower than that of MUPPS. The results of each local regions are provided in Table 1.

Path selection. Fig. 4 plots the mean Dice scores of the VOTE, STAPLE, SBA fusion results of MUPPS using the three different registration schemes for path ranking: the *After affine*, *After LARM*, and *After nonrigid*. All the three fusion methods agreed that the path selection strategy using the *After affine* was better than the other two. The best results of the *After affine* scheme were achieved when using the best 11-13 ranked paths for propagation and segmentation, while the other two both needed more, about 20, cases. Hence, we needed N_{sr} , which was 20 in our experiment, affine registration for path selection and ranking and 11-13 propagation registration to achieve the best result using the *After affine*. Also, affine registration was the fast one among the three selection registration schemes. It could be achieved within one minute while the nonrigid registration, LARM plus FFDs, might need over one hour in our experiment.

It may seem counterintuitive that the *After affine* performed better than the *After nonrigid*, as the image similarity after nonrigid registration should be more accurate in evaluating the segmentation accuracy and thus in ranking the paths. This could be an inconclusive result from the specific test dataset, as the difference between the *After nonrigid* and the other two on 11-13 of Fig. 3 (right) was small. This could also be the indication that the “good” segmentations ranked by *After nonrigid* might tend to have similar bias, resulting in the gain of accuracy by fusing them not being significant, compared to the fusion results of “bad” segmentations which nevertheless have unrelated random errors.

Accuracy of MUPPS. Fig. 5 and Table 1 provide the mean Dice scores of local regions and the whole heart. The $VOTE_U$, $STAPLE_U$, and SBA_U in Table 1 are the fusion results of MUPPS using the *After affine* registration to select eleven highest ranked paths for propagation. The accuracy improvement using fusion was evident, as all the three fusion techniques gained at least 0.016,

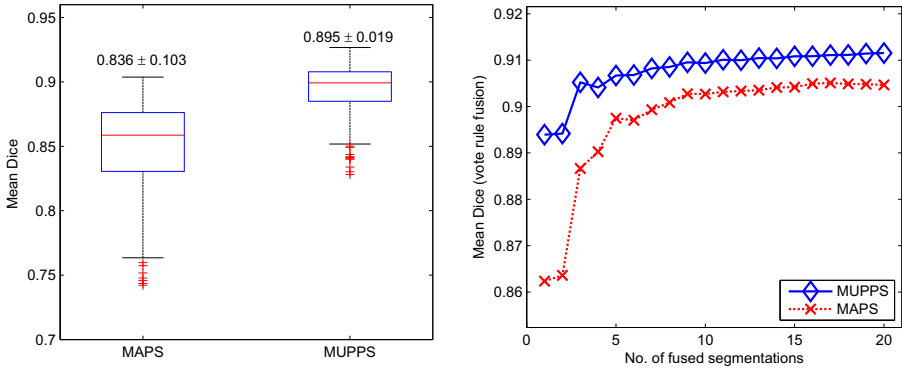


Fig. 3. Left: Box-and-Whisker diagram of the Dice scores using MAPS and MUPPS. Right: the Dice scores from the VOTE fusion results of them. Note that the difference of MAPS and MUPPS in local regions and the fusion results using STAPLE and SBA are almost identical to the results presented here.

Table 1. The mean Dice score of segmentation using standard MAPS and its best fusion results $VOTE_A$, and using MUPPS and the best fusion results, $VOTE_U$, $STAPLE_U$, and SBA_U . Note that the fusion results of MAPS using STAPLE and SBA are almost identical to $VOTE_A$.

Dice	Whole	MYO	LV	LA	RV	RA
MAPS	$.836 \pm .103$	$.722 \pm .116$	$.879 \pm .114$	$.788 \pm .084$	$.852 \pm .075$	$.783 \pm .105$
MUPPS	$.895 \pm .019$	$.836 \pm .027$	$.940 \pm .016$	$.869 \pm .028$	$.915 \pm .024$	$.871 \pm .050$
$VOTE_A$	$.905 \pm .018$	$.858 \pm .025$	$.946 \pm .016$	$.881 \pm .034$	$.921 \pm .021$	$.879 \pm .041$
$VOTE_U$	$.911 \pm .017$	$.866 \pm .026$	$.949 \pm .014$	$.883 \pm .025$	$.930 \pm .023$	$.892 \pm .041$
$STAPLE_U$	$.911 \pm .017$	$.866 \pm .025$	$.949 \pm .014$	$.880 \pm .026$	$.930 \pm .021$	$.888 \pm .044$
SBA_U	$.912 \pm .016$	$.887 \pm .025$	$.949 \pm .014$	$.883 \pm .026$	$.930 \pm .021$	$.891 \pm .038$

about one standard deviation, in Whole category. The gains in local regions were different. The myocardium, which resembles a thin cup-shaped sheet and normally reports a worst segmentation result in the five regions, achieved the most significant improvement, about 0.03 after using the fusion. By contrast, the left ventricle had the least improvement, less 0.01, and other categories had less than 0.02. We therefore conclude that it can be more beneficial to use MUPPS and fusion techniques in the segmentation of thin, sheet-shaped objects where one propagation path tends to have certain random errors in the result. Note that the conclusion from the surface distance as error measure was similar, as the mean root mean squared surface distance of MUPPS without fusion, and the fusion results using VOTE, STAPLE, and SBA were 1.29 ± 0.319 , 1.13 ± 0.275 , 1.25 ± 0.346 , and 1.13 ± 0.248 (mm), respectively. Only that STAPLE fusion seemed to be disfavored in this error measure.

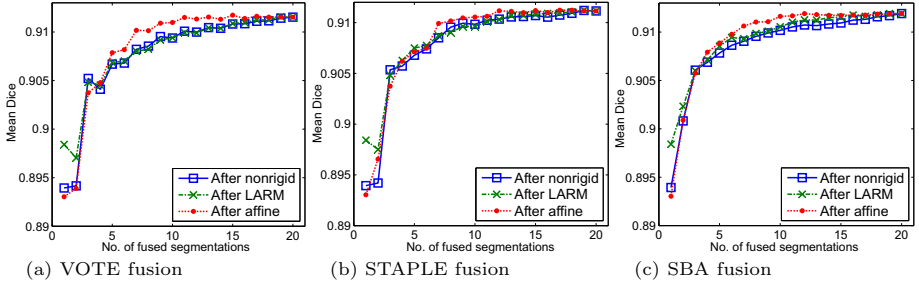


Fig. 4. The mean Dice scores using different path selection strategies: *After affine*, *After LARM*, and *After nonrigid*

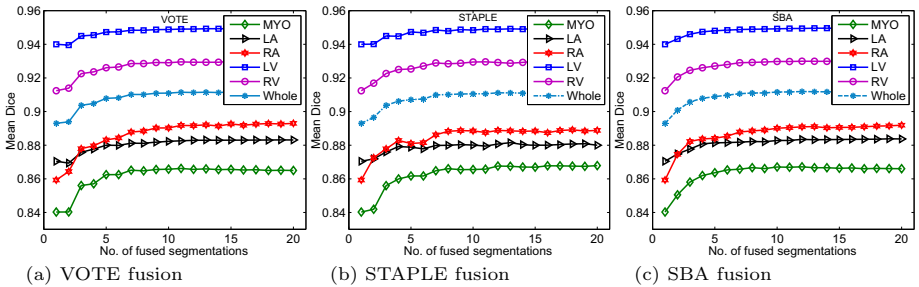


Fig. 5. The mean Dice scores of MUPPS for whole heart and each local region segmentation using the *After affine* path selection

4 Conclusion

In this paper, we propose a new multi-classifier method, the Multiple Path Propagation and Segmentation (MUPPS), for the whole heart segmentation of cardiac MRI. We showed that MUPPS achieved better mean Dice scores than the standard MAPS, 0.895 VS 0.836 before the fusion and 0.911 VS 0.905 after the fusion. To minimize the number of propagation paths in MUPPS for a segmentation case, we studied three path ranking schemes. The best scheme was to rank the path using the image similarity of the transformed atlas and unseen image after an affine registration. The segmentation framework of MUPPS achieved 0.911 ± 0.016 mean Dice score, a 0.015 gain of accuracy, using eleven highest ranked paths and the fusion techniques. In particular for the segmentation of myocardium, a gain of 0.03 was reported. The whole heart segmentation error of RMS surface distance was 1.13 ± 0.248 (mm).

Note that the shape manifold in MUPPS is different from the linear space formed from the principal components in the statistical shape model [14]. Also, it only needs about 10 images to construct A with good quality [6]. Hence, for a new image MR_{n+1} , only one off-line registration is needed to compute the new path T_{n+1} and no reconstruction of A is required.

A potential advantage of MUPPS is to further discretize the shape manifold to increase the number of searching paths such as these between two existing paths T_i and T_j . Furthermore, we used a mask on the whole heart for the path ranking. It may be more accurate to rank the path based on the similarity within certain local region such as for the myocardium. The segmentation of the whole heart can then be considered as several segmentations of a number of local regions and different rankings can be used for different regions [5]. Finally, the results of MUPPS contain a set of deformation fields which are the transformations between the atlas and unseen image plus certain random errors. By averaging these fields, we should obtain a better estimation of the registration ground truth for the applications where the spatial mapping is of the interest. We will address these in future work.

References

1. Rohlfing, T., Maurer, J.C.R.: Multi-classifier framework for atlas-based image segmentation. *Pattern Recognition Letters* 26, 2070–2079 (2005)
2. Heckemann, R.A., Hajnal, J.V., Aljabar, P., Rueckert, D., Hammers, A.: Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *NeuroImage* 33(1), 115–126 (2006)
3. Aljabar, P., Heckemann, R., Hammers, A., Hajnal, J., Rueckert, D.: Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. *NeuroImage* 46(3), 726–738 (2009)
4. Leung, K., Barnes, J., Ridgway, G., Bartlett, J., Clarkson, M., Macdonald, K., Schuff, N., Fox, N., Ourselin, S.: Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer’s disease. *Neuroimage* 51(4), 1345–1359 (2010)
5. van Rikxoort, E., Isgum, I., Arzhaeva, Y., Staring, M., Klein, S., Viergever, M., Pluim, J., van Ginneken, B.: Adaptive local multi-atlas segmentation application to the heart and the caudate nucleus. *Medical Image Analysis* (14), 39–49 (2010)
6. Zhuang, X., Rhode, K., Razavi, R., Hawkes, D.J., Ourselin, S.: A registration-based propagation framework for automatic whole heart segmentation of cardiac MRI. *IEEE Transactions on Medical Imaging* (2010), doi:10.1109/TMI.2010.2047112
7. Peters, J., Ecabert, O., Meyer, C., Kneser, R., Weese, J.: Optimizing boundary detection via simulated search with applications to multi-modal heart segmentation. *Medical Image Analysis* 14, 70–84 (2009)
8. Zhuang, X., Rhode, K., Arridge, S., Razavi, R., Hill, D., Hawkes, D., Ourselin, S.: An atlas-based segmentation propagation framework using locally affine registration – application to automatic whole heart segmentation. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part II. LNCS*, vol. 5242, pp. 425–433. Springer, Heidelberg (2008)
9. Hamm, J., Davatzikos, C., Verma, R.: Efficient large deformation registration via geodesics on a learned manifold of images. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009. LNCS*, vol. 5761, pp. 680–687. Springer, Heidelberg (2009)
10. Studholme, C., Hill, D.L.G., Hawkes, D.J.: An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition* 32(1), 71–86 (1999)

11. Rohlfing, T., Maurer, C.R.: Shape-based averaging. *IEEE Trans. Image Processing* 16(1), 153–161 (2007)
12. Warfield, S.K., Zou, K.H., Wells III, W.M.: Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation. *IEEE Transaction on Medical Imaging* 23(7), 903–921 (2004)
13. Dice, L.R.: Measures of the amount of ecologic association between species. *Ecology* 26(3), 297–302 (1945)
14. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models: Their training and application. *Computer Vision and Image Understanding* 61(1), 38–59 (1995)

Combined Model-Based Segmentation and Elastic Registration for Accurate Quantification of the Aortic Arch

Andreas Biesdorf¹, Karl Rohr¹, Hendrik von Tengg-Kobligh²,
and Stefan Würz¹

¹ University of Heidelberg, BIOQUANT, IPMB, and DKFZ Heidelberg
Dept. Bioinformatics and Functional Genomics, Biomedical Computer Vision Group

² German Cancer Research Center (DKFZ) Heidelberg, Dept. of Radiology and
University Hospital Heidelberg, Dept. of Diagnostic and Interventional Radiology

Abstract. Accurate quantification of the morphology of vessels is important for diagnosis and treatment of cardiovascular diseases. We introduce a new approach for the quantification of the aortic arch morphology that combines 3D model-based segmentation with elastic image registration. The performance of the approach has been evaluated using 3D synthetic images and clinically relevant 3D CTA images including pathologies. We also performed a comparison with a previous approach.

1 Introduction

Diseases of the aortic arch are a major cause of death in the western world and can be diagnosed using, for example, computed tomography angiography (CTA), and treated by minimally-invasive placement using an endovascular graft. For this task, individual morphological parameters such as the centerline position and the vessel diameters have to be quantified. The geometry of the aortic arch can be automatically determined from radiological images by *segmentation* approaches. A wide spectrum of different approaches exists, for example, approaches based on differential measures, deformable models, or parametric intensity models. Alternatively, segmentation may also be achieved by *registration*, where an atlas or a model is registered with an image of a patient (e.g., [23]).

In recent years, increased attention has been paid to *combined* approaches for vessel analysis that integrate both segmentation and registration (e.g., [14, 5, 8]). Existing approaches can be classified according to the transformation model (e.g., rigid, affine, elastic), the type of information used for registration (e.g., point sets, surfaces, intensities), and whether prior information is employed (e.g., an atlas or a template image). Most approaches use rigid transformation models (e.g., [1]), or splines on a regular grid such as B-Splines (e.g., [4]) and Cardinal-Splines (e.g., [5]). In addition, registration is performed based on, for instance, surfaces (e.g., [5]), point sets (e.g., [8]), centerlines (e.g., [1]), or binarized image volumes (e.g., [4]). Thus, the image intensities are not directly exploited. Moreover, none of the combined approaches uses an incremental tracking scheme or employs a parametric intensity model for vessel segmentation.

In this contribution, we introduce a novel approach for the quantification of the aortic arch morphology from 3D tomographic images. Our approach combines 3D fitting of a parametric intensity model with intensity-based elastic image registration. The image intensities are exploited directly, thus the full intensity information is incorporated. In comparison to a pure model fitting approach (e.g., [6]), the combined approach can cope with a larger spectrum of vessel shapes and even with shapes that deviate significantly from the model used for segmentation. At the same time, our approach is constrained to meaningful shapes by employing a physically-based deformation model for registration. Moreover, our approach can cope with very different shapes of vessel centerlines since an incremental tracking scheme is used. In contrast to atlas- or model-based segmentation (e.g., [23]), our approach does not require a prior segmentation. Instead, the intensity template is automatically generated by model-based segmentation. We have successfully applied our approach to 3D synthetic images and clinically relevant 3D CTA images, and the performance has been evaluated.

2 Model-Based Segmentation and Elastic Registration

Our approach for the segmentation of vessels in 3D tomographic images combines model-based segmentation with elastic image registration. The approach is based on an energy-minimizing functional J_k corresponding to a vessel segment k :

$$J_k(\mathbf{p}_k, \mathbf{u}_k) = J_M(g_M, g_{I,k}^{roi}, \mathbf{p}_k) + J_R(g_{I,k}^{roi}, g_{M,k}^{roi}, \mathbf{u}_k) \quad (1)$$

The first term J_M denotes an intensity similarity measure between a 3D cylindrical intensity model g_M with parameters \mathbf{p}_k and the intensities $g_{I,k}^{roi}$ within a region-of-interest (ROI) of a 3D tomographic image g_I . The second term J_R denotes an energy-minimizing functional for elastic registration of $g_{I,k}^{roi}$ with an image $g_{M,k}^{roi}$ generated from the 3D intensity model g_M . The result of elastic registration is described by the deformation field \mathbf{u}_k . By minimizing the overall functional J_k , the segmentation result from 3D model fitting is refined using elastic image registration, while at the same time the registration result is used to improve 3D model fitting. For segmentation of an entire vessel such as the aortic arch, we incrementally minimize J_k along the vessel using a tracking approach based on a Kalman filter assuming a linear motion model. Typically, a segmentation result of the aorta consists of several hundreds overlapping vessel segments. Branches can be detected based on a connected components analysis.

The 3D parametric intensity model used in J_M represents an ideal sharp 3D cylinder convolved with a 3D Gaussian. The model includes parameters for the width R of the tubular structure and the image blur σ , and is well-suited to describe the plateau-like intensity structure of thick vessels such as the aorta:

$$g_{Cyl}(\mathbf{x}, R, \sigma) = \Phi\left(\frac{c_2 - 1}{c_1} + c_1\right), \quad c_1 = \frac{2}{3}\sigma \frac{\sqrt{\sigma^2 + r^2}}{2\sigma^2 + r^2}, \quad c_2 = \sqrt[3]{\frac{R^2}{2\sigma^2 + r^2}} \quad (2)$$

where $\Phi(x) = \int_{-\infty}^x (2\pi)^{-1/2} e^{-\xi^2/2} d\xi$ denotes the Gaussian error function, $\mathbf{x} = (x, y, z)^T$, and $r = \sqrt{x^2 + y^2}$. The complete model also incorporates intensity

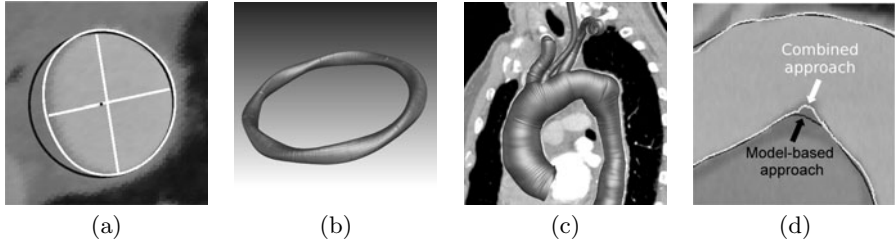


Fig. 1. (a) Cross-section of a 3D CTA image of an aorta and overlaid result of model-based segmentation (black) and combined model-based segmentation and elastic registration (white). (b) Segmentation result of the 2D combined approach for a 3D synthetic image of a twisted torus. (c) Segmentation result of the 2D combined approach for a 3D CTA image. (d) Computed vessel contours using the model-based approach (black) and the 2D combined approach (white) for a section of a 3D CTA image.

levels a_0 (surrounding tissue) and a_1 (vessel) as well as a 3D rigid transform \mathcal{R} with rotation $\boldsymbol{\alpha} = (\alpha, \beta, \gamma)^T$ and translation $\mathbf{x}_0 = (x_0, y_0, z_0)^T$, which yields

$$g_M(\mathbf{x}, \mathbf{p}) = a_0 + (a_1 - a_0) g_{Cyl}(\mathcal{R}(\mathbf{x}, \boldsymbol{\alpha}, \mathbf{x}_0), R, \sigma) \quad (3)$$

with 10 parameters $\mathbf{p} = (R, a_0, a_1, \sigma, \alpha, \beta, \gamma, x_0, y_0, z_0)^T$ [6].

The cylindrical model g_M can accurately represent a vessel segment if the vessel has circular cross-sections. However, the model may be inaccurate in the case of non-circular cross-sections (e.g., Fig. 1b, black contour). To improve the accuracy between the model and the true vessel shape in this case, we suggest using elastic registration of an image $g_{M,k}^{roi}$ generated from the 3D intensity model g_M with a ROI of the original image $g_{I,k}^{roi}$. The result of elastic registration is a deformation field \mathbf{u}_k which can be used to compute a refined vessel contour and centerline position (e.g., Fig. 1a, white contour). To limit the final segmentation result in (1) to physically meaningful shapes, the deformations \mathbf{u}_k are computed based on Gaussian elastic body splines (e.g., [7]).

2.1 Optimization of the Energy-Minimizing Functional J_k

The functional in (1) is optimized by an iterative scheme which alternately minimizes J_M and J_R for each vessel segment k to obtain estimates for the model parameters \mathbf{p}_k and the deformation field \mathbf{u}_k (see Fig. 2). For a vessel segment k , we estimate \mathbf{p}_k by least-squares model fitting of g_M to the image intensities $g_{I,k}^{roi}$ by minimizing

$$J_M(\mathbf{p}_k) = \sum_{\mathbf{x} \in g_{I,k}^{roi}} (g_M(\mathbf{x}, \mathbf{p}_k) - g_{I,k}^{roi}(\mathbf{x}))^2 \quad (4)$$

using the method of Levenberg-Marquardt. To compute the deformation field \mathbf{u}_k , we generate an image $g_{M,k}^{roi}$ from the fitted intensity model g_M and perform intensity-based registration with $g_{I,k}^{roi}$ by minimizing

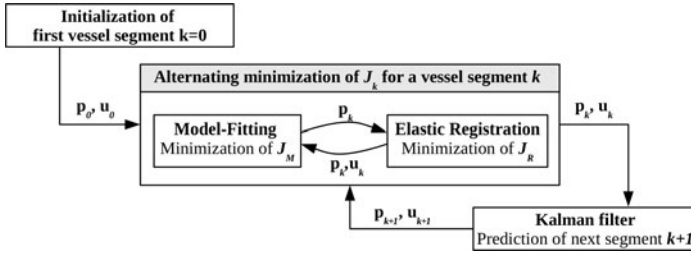


Fig. 2. Incremental combined vessel segmentation approach

$$J_R(\mathbf{u}_k) = J_{Data,I}(g_{I,k}^{roi}, g_{M,k}^{roi}, \mathbf{u}_k^I) + \lambda_I J_I(\mathbf{u}_k, \mathbf{u}_k^I) + \lambda_E J_{Elastic}(\mathbf{u}_k) \quad (5)$$

where λ_I and λ_E are scalar weights. We use $\lambda_E = 0.1$ while λ_I is estimated automatically. The first term $J_{Data,I}$ describes the intensity-based similarity measure between $g_{I,k}^{roi}$ and $g_{M,k}^{roi}$. Since g_M incorporates intensity levels a_0 and a_1 , minimization of $J_{Data,I}$ can be considered a monomodal registration problem. Therefore, we use the sum-of-squared intensity differences for $J_{Data,I}$ (cf. (4)). With the second term J_I , the intensity-based deformation field \mathbf{u}_k^I is coupled with the final deformation field \mathbf{u}_k using a weighted Euclidean distance. The third term $J_{Elastic}$ represents the regularization of the deformation field according to the Navier equation of linear elasticity. Optimization of J_R is performed alternately w.r.t. \mathbf{u}_k^I and \mathbf{u}_k . Note that the functional (5) has been formulated such that for the minimization of $\lambda_I J_I + \lambda_E J_{Elastic}$ w.r.t. \mathbf{u}_k an analytic solution can be derived [7]. For the minimization w.r.t. \mathbf{u}_k^I , $J_{Data,I} + J_I$ has to be minimized for which we use the method of Levenberg-Marquardt.

The result of elastic registration is used to improve the result of model fitting by re-estimating the model parameters \mathbf{p}_k including the radius R , the orientation $\boldsymbol{\alpha}$, as well as the translation \mathbf{x}_0 . To update the radius R , we compute the mean radius along the vessel contour based on \mathbf{u}_k . For the translation \mathbf{x}_0 we obtain new estimates based on the deformation $\mathbf{u}_k(\mathbf{x}_0)$. The orientation $\boldsymbol{\alpha}$ can be re-estimated using finite differences of the positions of two points close to \mathbf{x}_0 .

Based on the updated parameter vector \mathbf{p}_k and the deformation field \mathbf{u}_k , we again perform model-based segmentation with subsequent elastic registration for minimizing J . Note that now, model fitting is performed on a deformed image which is obtained by applying \mathbf{u}_k to $g_{I,k}^{roi}$. Performing model fitting on the deformed image has the advantage that the model parameters can be estimated more accurately since the deformed image is more similar to the underlying model than the original image. This alternating optimization is repeated until the results of model fitting and elastic registration converge for a vessel segment k . After convergence and having estimated the parameters for the current vessel segment, a new parameter vector \mathbf{p}_{k+1} is predicted based on a Kalman filter and used as initialization for the next vessel segment along a vessel. For initialization of \mathbf{u}_{k+1} in the next iteration, the current deformation field \mathbf{u}_k is used.

2.2 Exploiting 2D and 3D Image Information

With our combined approach for vessel segmentation there are different possibilities to exploit the intensity information. We have developed two approaches, a 3D and a 2D approach, which differ w.r.t. computational efficiency and segmentation accuracy. The first approach performs model fitting within a 3D ROI and uses 3D image registration of the 3D ROI. Since with this approach all image information within a 3D ROI is exploited, the accuracy as well as the computational complexity is expected to be relatively high. The second variant uses 3D model fitting only for estimating the initial 3D orientation α , while J_k in (II) is minimized based on model fitting and image registration of 2D image cross-sections orthogonal to the vessel centerline. Compared to the 3D approach, the computational complexity of the 2D approach is significantly lower.

3 Experimental Evaluation

We have applied our approach to 240 3D synthetic images and 17 clinically relevant 3D CTA images of the human thorax. To quantify the segmentation accuracy, we have computed the mean errors for clinically relevant measures comprising the minimum, mean, and maximum vessel diameters, $\bar{e}_{D,min}$, $\bar{e}_{D,mean}$, and $\bar{e}_{D,max}$, respectively, as well as the mean error for the centerline position $\bar{e}_{\mathbf{x}_0}$. Note that, in general, vessel cross-sections are noncircular. To compute $\bar{e}_{D,min}$, $\bar{e}_{D,mean}$, and $\bar{e}_{D,max}$, we have defined a diameter by the length of a straight line that connects two points on the vessel boundary and that passes through the center \mathbf{x}_0 . The error measures for the diameters were determined by evaluating a sample of 1000 different diameters of a cross-section where the directions of the diameters are equiangularly distributed. $e_{\mathbf{x}_0}$ is defined as Euclidean distance between the true center of the vessel and the estimated position \mathbf{x}_0 . For all measures, we have computed mean errors by averaging over all vessel segments.

3.1 3D Synthetic Images

In a first experiment, we have generated two different sets of 3D synthetic images. The first set contains 120 different images of *straight twisted cylinders* with *elliptical* cross-sections that differ in radii and the level of Gaussian image noise ($\sigma_n = 0, 1, 3, 5, 20$), and have a size of $200 \times 200 \times 200$ voxels. The second image set contains 120 different images of *twisted tori* with *elliptical* cross-sections that also differ in radii and the level of Gaussian image noise, and have a size of $400 \times 400 \times 200$ voxels (see Fig. IIb for an example). Table II shows the results of the segmentation accuracy of the two new combined approaches (2D and 3D) in comparison to a previous model-based approach [6].

Segmentation results: 120 twisted cylinders. For $\bar{e}_{\mathbf{x}_0}$ we obtain similar very good results for the new approaches and the previous approach with sub-voxel accuracy of $\bar{e}_{\mathbf{x}_0} \leq 0.02$ voxels. For $\bar{e}_{D,mean}$, the results of the approaches are comparable, while the 2D combined approach yields the best result. For

$\bar{e}_{D,min}$ and $\bar{e}_{D,max}$, the accuracy of the new approaches is significantly better than that of the previous approach. In comparison to the previous approach, the 2D and 3D combined approaches yield improvements of 61% - 66% for $\bar{e}_{D,min}$ and $\bar{e}_{D,max}$.

Segmentation results: 120 twisted tori. For $\bar{e}_{D,min}$, $\bar{e}_{D,mean}$, and $\bar{e}_{D,max}$, the previous approach yields similar results as for the first image set with $\bar{e}_{D,min} = 3.14$ voxels, $\bar{e}_{D,mean} = 0.09$ voxels, and $\bar{e}_{D,max} = 3.83$ voxels. However, for the 2D combined approach we obtain a significant improvement for all diameter measures with $\bar{e}_{D,min} = 0.61$ voxels, $\bar{e}_{D,mean} = 0.05$ voxels, and $\bar{e}_{D,max} = 0.88$ voxels, which is an improvement of 77% - 80% for $\bar{e}_{D,min}$ and $\bar{e}_{D,max}$ as well as 44% for $\bar{e}_{D,mean}$ compared to the previous approach. The 3D combined approach, however, yields the best result for all diameter measures with $\bar{e}_{D,min} = 0.37$ voxels, $\bar{e}_{D,mean} = 0.04$ voxels, and $\bar{e}_{D,max} = 0.52$ voxels, which, in comparison to the previous approach, are improvements of 88%, 56%, and 86%, respectively. For \bar{e}_{x_0} , all approaches yield very good results with sub-voxel accuracy, however the 3D combined approach yields the best result with $\bar{e}_{x_0} = 0.08$ voxels, which is an improvement of 50% compared to the previous approach and 38% compared to the 2D combined approach.

3.2 3D CTA Images

In a second experiment, we applied our approach to two different sets of 3D CTA images of the thorax. The first set of images contains ten 3D CTA images of patients with only slight pathologies. The second set of images contains seven 3D CTA images of patients with severe pathologies such as aneurysms or highly curved vessel centerlines. The CTA images comprise 619 to 829 slices with a size of 512×512 voxels. For evaluation, manual segmentation by a radiologist was performed for 15 of the 3D CTA images, while two of the images were segmented by a trained observer. Table 2 shows the results for the two new combined approaches and a previous model-based approach [6] for both sets of images.

Segmentation results: Ten 3D CTA images. For the first set of 3D CTA images and for $\bar{e}_{D,min}$ as well as $\bar{e}_{D,max}$ the accuracy of our new approaches is significantly better than that of the previous approach, while the best result is obtained for the 3D combined approach. For $\bar{e}_{D,mean}$ and \bar{e}_{x_0} we obtain similar good results for the new approaches and the previous approach while for $\bar{e}_{D,mean}$ the best result is obtained for the 3D combined approach and for \bar{e}_{x_0}

Table 1. Mean errors for the diameters $\bar{e}_{D,min}$, $\bar{e}_{D,mean}$, and $\bar{e}_{D,max}$, and the centerline position \bar{e}_{x_0} for different approaches

Accuracy \ Approach	120 twisted cylinders				120 twisted tori			
	$\bar{e}_{D,min}$	$\bar{e}_{D,mean}$	$\bar{e}_{D,max}$	\bar{e}_{x_0}	$\bar{e}_{D,min}$	$\bar{e}_{D,mean}$	$\bar{e}_{D,max}$	\bar{e}_{x_0}
Model-based approach	3.01	0.09	3.67	0.01	3.14	0.09	3.83	0.16
2D combined approach	1.04	0.07	1.42	0.02	0.61	0.05	0.88	0.13
3D combined approach	1.02	0.09	1.42	0.01	0.37	0.04	0.52	0.08

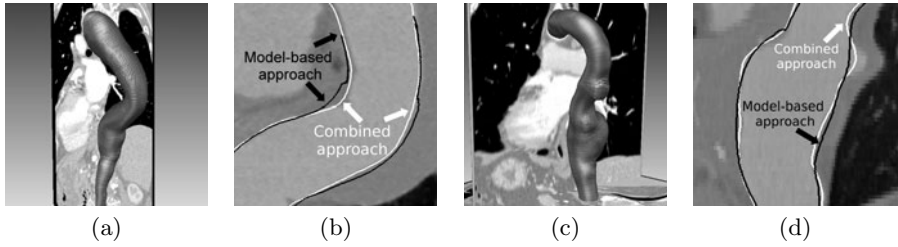


Fig. 3. (a), (c) Segmentation results of the 3D combined approach for two 3D CTA images showing a pathology. (b), (d) Vessel contours using the model-based approach (black) and the 3D combined approach (white) for a section of the 3D CTA images.

the best result is obtained for the 2D combined approach. In Fig. 11c,d we show the segmentation result of the 2D combined approach for a 3D CTA image as well as a section of the same image and the result of the 2D combined approach in comparison to a previous approach. It can be seen that the new approach yields a significant improvement in highly curved regions.

Segmentation results: Seven 3D CTA images with pathologies. For the second set of 3D CTA images, we consistently obtain more accurate results for the new combined approaches. For $\bar{e}_{D,min}$, $\bar{e}_{D,mean}$, and $\bar{e}_{D,max}$ the accuracy of the new approaches is significantly better than that of the previous approach, while the best result is obtained for the 3D combined approach with improvements of 21% to 57% compared to the previous approach as well as improvements of 9% to 15% compared to the 2D combined approach. For \bar{e}_{x_0} , the 2D combined approach consistently yields the most accurate result. In Fig. 3 we show segmentation results of the 3D combined approach for two 3D CTA images as well as sections of the same images and the result of the new approach in comparison to the previous approach. It can be seen that the new approach yields a significant improvement. Note that for the image in Fig. 3b, only the new approach succeeds to fully segment the pathology, while the previous approach fails to segment the last part of the pathology.

Overall, it turns out that for $\bar{e}_{D,min}$ and $\bar{e}_{D,max}$ the new 2D and 3D combined approaches yield more accurate results than the previous approach. For \bar{e}_{x_0} , the 2D combined approach yields the best result, while for $\bar{e}_{D,mean}$, the 3D combined approach yields the best result. The computation time for a vessel segment on a

Table 2. Mean errors for the diameters $\bar{e}_{D,min}$, $\bar{e}_{D,mean}$, and $\bar{e}_{D,max}$, and the centerline position \bar{e}_{x_0} for different approaches

Accuracy Approach	Ten 3D CTA images				Seven pathologies			
	$\bar{e}_{D,min}$	$\bar{e}_{D,mean}$	$\bar{e}_{D,max}$	\bar{e}_{x_0}	$\bar{e}_{D,min}$	$\bar{e}_{D,mean}$	$\bar{e}_{D,max}$	\bar{e}_{x_0}
Model-based approach	5.40	2.27	1.67	0.65	7.42	2.32	6.15	1.00
2D combined approach	4.67	2.38	1.58	0.47	5.19	1.89	3.14	0.63
3D combined approach	4.24	2.19	1.46	0.83	4.70	1.83	2.66	0.96

2.40 GHz Intel Core 2 Quad CPU is about 1-2 seconds for the 2D approach and about 50-60 seconds for the 3D approach.

4 Discussion

We have introduced a new approach for the quantification of the aortic arch from 3D CTA images that combines fitting of a parametric intensity model with intensity-based elastic image registration. We have demonstrated the applicability of our approach using 3D synthetic images and clinically relevant 3D CTA images. From the experiments it turned out that the new combined approach consistently yields more accurate segmentation results than a previous segmentation approach for the minimum and maximum diameters, which are the most relevant clinical measures. For real 3D CTA images, the 2D combined approach is most accurate for estimating the centerline position, while the 3D combined approach is most accurate for estimating the diameters. It also turned out that for the new approach significant improvements are obtained for difficult segmentation tasks, in particular, for pathologies and highly curved vessel centerlines.

Acknowledgment. Support of the Deutsche Forschungsgemeinschaft (DFG) within the project QuantVessel (RO 2471/6) is gratefully acknowledged.

References

1. Aylward, S.R., Jomier, J., Weeks, S., Bullitt, E.: Registration and Analysis of Vascular Images. *Internat. Journal of Computer Vision* 55(2/3), 123–138 (2003)
2. Barber, D., Oubel, E., Frangi, A., Hose, D.: Efficient computational fluid dynamics mesh generation by image registration. *Med. Imag. Anal.* 11(5), 648–662 (2007)
3. Isgum, I., Staring, M., Rutten, A., Prokop, M., Viergever, M., van Ginneken, B.: Multi-Atlas-Based Segmentation With Local Decision Fusion – Application to Cardiac and Aortic Segmentation in CT Scans. *IEEE Trans. Med. Imaging* 28(7), 1000–1010 (2009)
4. Kang, D.G., Suh, D., Ra, J.: Three-Dimensional Blood Vessel Quantification via Centerline Deformation. *IEEE Trans. Med. Imaging* 28(3), 405–414 (2009)
5. Wong, W., Chung, A.: Augmented Vessels for Quantitative Analysis of Vascular Abnormalities and Endovascular Treatment Planning. *IEEE Trans. Med. Imaging* 25(6), 665–684 (2006)
6. Wörz, S., Rohr, K.: Segmentation and Quantification of Human Vessels Using a 3-D Cylindrical Intensity Model. *IEEE Trans. on Image Processing* 16(8), 1994–2004 (2007)
7. Wörz, S., Rohr, K.: Hybrid Physics-Based Elastic Image Registration Using Approximating Splines. In: Reinhardt, J., Pluim, J. (eds.) *Proc SPIE Medical Imaging 2008: Image Processing*, San Diego, CA/USA, pp. 6914–6930 (February 2008)
8. Zhao, F., Zhang, H., Wahle, A., Matthew, T.T., Stolpen, A.H., Scholz, T.D., Sonka, M.: Congenital aortic disease: 4D magnetic resonance segmentation and quantitative analysis. *Medical Image Analysis* 13(3), 483–493 (2009)

Conditional Shape Models for Cardiac Motion Estimation

Coert Metz^{1,*}, Nora Baka^{1,2,*}, Hortense Kirisli^{1,2}, Michiel Schaap¹,
Theo van Walsum¹, Stefan Klein¹, Lisan Neeffjes³, Nico Mollet³,
Boudewijn Lelieveldt^{2,4}, Marleen de Bruijne^{1,5}, and Wiro Niessen^{1,4,**}

¹ Dept. of Rad. and Med. Informatics, Erasmus MC, Rotterdam, The Netherlands

² Leiden University Medical Center, Leiden, Netherlands

³ Dept. of Radiology and Cardiology Erasmus MC, Rotterdam, The Netherlands

⁴ Delft University of Technology, The Netherlands

⁵ University of Copenhagen, Denmark

Abstract. We propose a conditional statistical shape model to predict patient specific cardiac motion from the 3D end-diastolic CTA scan. The model is built from 4D CTA sequences by combining atlas based segmentation and 4D registration. Cardiac motion estimation is, for example, relevant in the dynamic alignment of pre-operative CTA data with intra-operative X-ray imaging. Due to a trend towards prospective electrocardiogram gating techniques, 4D imaging data, from which motion information could be extracted, is not commonly available. The prediction of motion from shape information is thus relevant for this purpose. Evaluation of the accuracy of the predicted motion was performed using CTA scans of 50 patients, showing an average accuracy of 1.1 mm.

1 Introduction

Coronary angioplasty is an often applied procedure to reopen narrowed or occluded coronary arteries. Real time X-ray visualization guides these interventions, providing information about the morphology of the patent vessels. For difficult interventions, such as reopening chronic total occlusions, integration of a pre-operative CTA acquisition is expected to improve the intervention result. Hereto, a correct alignment of the CTA image with the 2D X-ray projection images is required, which consists of both a rigid alignment to determine the pose and orientation of the data and a non-rigid alignment to compensate for cardiac motion. A patient specific cardiac motion prior can be derived from 4D retrospectively gated CTA data [1]. However, the trend in cardiac CT acquisition is towards prospective electrocardiogram (ECG) gating techniques to decrease the effective patient dose. In these situations, the derivation of dynamic information from CTA data is often not possible, as only one phase of the cardiac

* Both authors contributed equally to this work.

** This work was supported by SenterNovem, project IGIT4Health (HTT09011), NWO, STW/EW and The Danish Council for Strategic Research.

cycle is imaged. Moreover, it is often not known before the CTA acquisition if a patient will undergo coronary angioplasty, making selective dynamic imaging for a subset of patients difficult.

The purpose of this work is to develop and evaluate a method to predict patient specific cardiac motion from a single 3D CTA image. To this end a 4D statistical shape model is built from 50 ECG-gated dynamic CTA images by combining 3D multi-atlas registration and 4D registration. Motion is estimated for an unseen patient by first segmenting the cardiac structures in the acquired 3D CTA scan. Subsequently we derive the most probable motion given the segmentation and the statistical model using a conditional Gaussian distribution.

Statistical shape models have been frequently applied for 3D image segmentation [2]. The use of active shape models for dynamic image segmentation has been investigated, for instance in the work of Ordas et al., who built a single 3D model with training data from multiple ECG phases [3]. However, this method cannot be used for the cardiac motion prediction problem, as a single 3D shape model of the heart is built, without distinguishing between inter-patient and intra-patient variability. Models built from 4D landmark positions can overcome this problem. Such a 4D statistical model was, for instance, built by Perperidis et al. for segmenting the left ventricle, right ventricle, and myocardium, but was not used for motion prediction [4]. Hoogendoorn et al. built a bilinear model for the extrapolation of cardiac motion, assuming that the motion of the heart is independent of its shape [5]. We, in contrast, build statistical shape models of the shape and motion of the heart without assuming their independence and evaluate its applicability for the prediction of cardiac motion by conditioning the motion model on single time point shape information.

2 Methods

The statistical model is built using multi-atlas based segmentation and 4D registration. The structures of interest are the aorta (Ao), endocardium left ventricle (endoLV), epicardium left ventricle (epiLV), right ventricle, left atrium (LA) and right atrium (RA). An atlas landmarking procedure ensures anatomical landmark correspondence. Cardiac motion is predicted by conditioning the motion model on the landmark points of the shape in the 3D end-diastolic image. These steps are explained in more detail in the following sections.

2.1 Statistical Shape Model

The K 4D CTA sequences that are used to build the statistical models [6] consist of T time points, each of which contains 3D shapes represented by n landmark points. Landmark correspondence is ensured both along the sequence and across the training set (see Sect. 2.2). First, a base time point b is determined, which is defined as the time point in the cardiac cycle for which the 3D reconstructions are available. In this work, we use the end-diastolic time point for this purpose. Subsequently, we represent the 4D sequence as a 3D shape in time point b , and

a motion sequence given by the vectors describing the displacements from this time point to all time points in the sequence. Accordingly, two separate models are built: a 3D statistical shape model (at time point b) and a 4D statistical motion model. Because the models are created for prediction purposes, the alignment (translation, rotation and isotropic scaling) of the shapes in the training sequences is performed using the landmarks of the shape in the base time point. This is in agreement with the alignment of the segmented end-diastolic target shape to the model. Assuming a Gaussian shape and motion distribution, PCA analysis is performed to extract the main modes of variation resulting in the following shape and motion model

$$\mathbf{s}_i \approx \bar{\mathbf{s}}_b + \Phi \mathbf{p}, \quad \text{and} \quad \mathbf{m} \approx \begin{bmatrix} \mathbf{s}_1 - \mathbf{s}_b \\ \vdots \\ \mathbf{s}_T - \mathbf{s}_b \end{bmatrix} = \bar{\mathbf{m}} + \Psi \mathbf{q} \quad (1)$$

where \mathbf{s}_i is the shape of the i -th time point of a sequence, b denotes the base time point, and \mathbf{m} is the motion vector. The mean shape and motion is denoted with $\bar{\mathbf{s}}_b$ and $\bar{\mathbf{m}}$, while \mathbf{p} and \mathbf{q} are the shape and motion parameter vectors respectively. The models were reduced to explain 95% of the total variance.

2.2 Model Construction

Atlases and Atlas Point Correspondence. 3D segmentation of the cardiac structures at one time point in the cardiac cycle is achieved by multi-atlas registration using $J = 8$ atlases [7]. We used `elastix` for all registration procedures in this work [8]. The following procedure was applied to the atlases to obtain automatic anatomical landmark correspondence:

1. Non-rigid B-spline registration of all atlas pairs to each other, resulting in $J * J$ transformations denoted as \mathbf{T}_{ij} where i and j indicate the atlas numbers.
2. Determination of the mean transformation for every atlas: $\bar{\mathbf{T}}_i = \frac{1}{J} \sum_j \mathbf{T}_{ij}$ and subsequent transformation of the manual annotations to the mean space.
3. Creation of a signed distance map SDM_{i_s} for every atlas i and structure s in the atlas (a) and mean (m) space and averaging of the signed distance maps for all structures in the mean space: $\overline{\text{SDM}}_s^m = \frac{1}{J} \sum_i \text{SDM}_{i_s}^m$.
4. Creation of a mean surface with approximately equally distributed landmark locations for every structure by extracting the surface at the zero level set of $\overline{\text{SDM}}_s^m$. The number of landmarks used was 15835 (Ao: 1316, endoLV: 2723, epiLV: 4106, LA: 2034, RA: 2079, RV: 3577).
5. Transformation of the obtained mean surface landmarks back to the atlas spaces using $\bar{\mathbf{T}}_i^{-1}$ and projection of the landmark points of the mean surfaces onto the annotated surfaces by non-rigidly registering their SDMs.

Landmark Determination in 4D CTA Sequences. 3D segmentation of the heart structures in the end-diastolic time point of the K 4D CTA sequences is performed by:

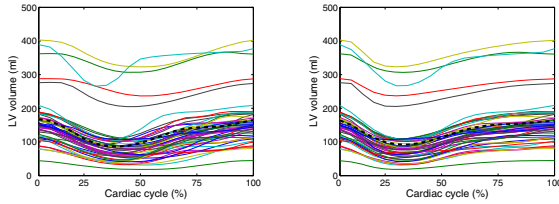


Fig. 1. Left ventricular volume during the cardiac cycle before and after alignment of the volume curves. Each colored line represents a patient and the dashed black line represents the reference curve.

1. Non-rigid B-spline registration of all atlases to the end-diastolic CTA image.
2. Transformation of the surface landmarks from the J atlases to the target image using the resulting transformations.
3. Combination of the J atlas landmark sets into one landmark set by application of the mean shift algorithm to the landmark coordinates [9]. This ensures a robust average of the corresponding landmark points and reduces the influence of possible registration errors.

This procedure results in a set of landmarks describing the cardiac structures at end-diastole. Subsequently, the segmentation of the cardiac structures in all time phases is achieved by determining the cardiac motion for each patient using 4D B-spline registration on the 4D CTA image. During registration, the variance of intensity values at corresponding spatial locations is minimized. Only deformations in the spatial domain are allowed and the deformation is forced to be periodic.

By combining the segmented shapes and the deformations resulting from the 4D registration procedure we determine 4D landmark sets for every patient. To ensure temporal correspondence between the 4D landmark sets, an additional alignment step is performed to align the moment of contraction and relaxation during the cardiac cycle between patients. To this end, time curves of left ventricular volume are obtained from the K 4D segmentations. These curves are subsequently aligned by a 1D non-rigid group-wise registration approach in which the cross-correlation between the curves is optimized. After this alignment, new landmark positions at twenty regular intervals of the cardiac cycle are determined by applying spline interpolation between corresponding landmark points in time. Volume graphs before and after alignment are shown in Fig. 1.

2.3 Conditional Model

We derive the motion given the shape of the cardiac structures at one time point in the cardiac cycle by assuming a Gaussian distribution of the shape vectors $\mathbf{s}_b \in S$ and motion vectors $\mathbf{m} \in M$ respectively. Their combined distribution $P(M, S)$ is a normal distribution with mean

$$\mu = \begin{bmatrix} \bar{\mathbf{m}} \\ \bar{\mathbf{s}}_b \end{bmatrix} \quad \text{and variance} \quad \Sigma = \begin{bmatrix} \Sigma_{MM} & \Sigma_{MS} \\ \Sigma_{SM} & \Sigma_{SS} \end{bmatrix}, \quad (2)$$

with Σ_{SS} and Σ_{MM} the covariance matrix of shape and motion respectively, and Σ_{SM} the covariance matrix between shape and motion. The most likely motion $\hat{\mathbf{m}}$ given a single shape \mathbf{s}_b^* is estimated as the mean of the conditional probability function $P(M|\mathbf{s}_b^*)$ [10]:

$$\hat{\mathbf{m}} = \bar{\mathbf{m}} + \Sigma_{MS}\Sigma_{SS}^{-1}(\mathbf{s}_b^* - \bar{\mathbf{s}}_b) \quad (3)$$

However, as the dimensionality of the 3D shape vectors is larger than the number of training shapes, the covariance matrix Σ_{SS} becomes singular, and cannot be inverted. Therefore, conditioning is performed on the shape and motion model parameters after applying PCA rather than on the shape and motion vectors themselves. As the mean parameter vectors are zero, Equation 3 simplifies to:

$$\hat{\mathbf{q}} = \Sigma_{qp}\Sigma_{pp}^{-1}\mathbf{p}^*, \quad \text{where } \mathbf{p}^* = \Phi^T(\mathbf{s}_b^* - \bar{\mathbf{s}}_b), \quad (4)$$

with $\hat{\mathbf{q}}$ being the parameter representation of the estimated motion $\hat{\mathbf{m}}$. The generation of plausible motion is ensured by scaling $\hat{\mathbf{q}}$ such that it always lies within ± 3 standard deviations of the motion model. The final conditional sequence is computed from the known 3D shape \mathbf{s}_b^* and the derived motion $\hat{\mathbf{m}} = \bar{\mathbf{m}} + \Psi\hat{\mathbf{q}}$.

3 Experiments and Results

3.1 Imaging Data

Eight 3D diastolic CTA reconstructions were used as the atlas images and 50 retrospectively gated 4D CTA images were used for building and evaluating the 4D model. All scans were acquired with Siemens CT scanners and 20 time points were reconstructed for the 4D CTA scans. Note that the variety of the data is large as healthy subjects are not imaged because of radiation regulations.

3.2 Cardiac Motion Prediction

The accuracy of the motion prediction was evaluated in leave-one-out experiments by applying the conditional model described in Sect. 2.3 on end-diastolic segmentations, which are most often available in clinical practice. The accuracy of the resulting 4D shapes was determined by computation of the average root mean squared point-to-surface distance per structure and time point. The results are compared to a projection of the complete sequence onto the model space. This projection represents the best possible reconstruction of an entirely known 4D input sequence given the training data. Note, that this represents a practical limit for accuracy and that the 4D information is not available for the motion estimation. Also, a replication of the end-diastolic segmentations for all time points was used as a worst-case estimate for benchmarking the errors of the conditional shape estimates. Graphs of the accuracy over time are shown in Fig. 2. Table 1 lists the root mean squared point-to-surface distance for both the replication and the conditioning averaged over all time points and over the

Table 1. Mean and standard deviation of root mean squared point-to-surface distance in mm. for replication (assuming no motion) and the proposed method. Values are listed for all time points and for the minimal left ventricular volume time point (30%).

Structure	Ao	endoLV	epiLV	RV	LA	RA
Replication (all)	1.6 (1.1)	2.2 (1.8)	1.7 (1.4)	1.7 (1.3)	1.6 (1.0)	1.2 (0.9)
Conditioning (all)	0.9 (0.6)	1.1 (0.8)	0.9 (0.6)	0.9 (0.7)	1.0 (0.6)	0.8 (0.5)
Replication (at 30%)	3.2 (1.0)	5.1 (2.0)	4.0 (0.9)	3.7 (0.8)	2.8 (0.8)	2.5 (0.6)
Conditioning (at 30%)	1.4 (0.6)	2.0 (0.7)	1.7 (0.5)	1.5 (0.5)	1.4 (0.5)	1.2 (0.5)

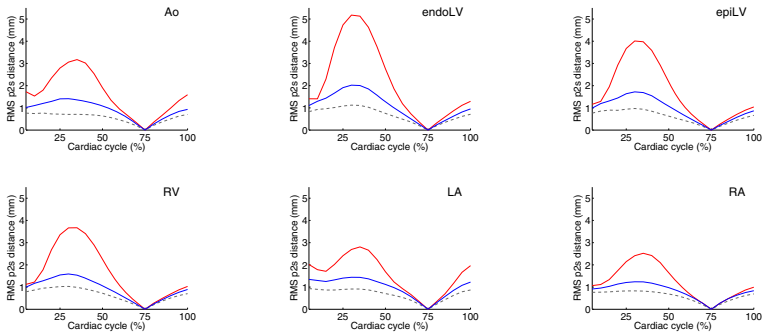


Fig. 2. Average root mean squared (RMS) point-to-surface (p2s) distance over the cardiac cycle. RMS values are averaged over all patients. The red line indicates the replication error. The grey dotted line shows the error for projecting the entire sequence onto the model space and the blue line shows the error when conditioning is used.

time point at which the left ventricular volume was minimal (30%). We expect the shapes at this time point to differ the most from the end-diastolic shapes used for the motion estimation. Fig. 3 shows a color coding of the distances for the left and right ventricle, and atrium at end-systole and end-diastole for two randomly selected patients.

3.3 Model Generalization Ability and Training Set Size

An experiment was conducted to assess the dependency between the accuracy of the model and the training set size. The available training sequences were grouped randomly in subgroups, each containing 5 sequences. Subsequently, for every tested training set size all possible combinations of groups were taken as training set, and a random left out group was used for testing. Fitting was performed by projecting the motion of the 4D segmentation of a CTA sequence on the subspace of the motion model, and back to the original space to derive the motion vectors for these parameters. The resulting motion was applied to

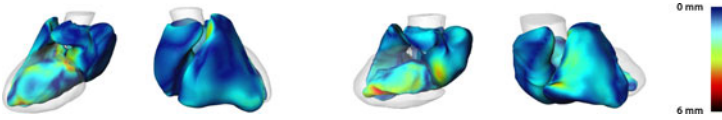


Fig. 3. Two views of end-systolic predicted surfaces for two randomly selected patients. Color coded are the point-to-surface distances from the predicted shape to the segmented shape. The Ao and epiLV are for visualization purposes not color coded.

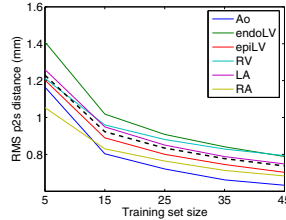


Fig. 4. Root mean squared point-to-surface (p2s) distance in mm. for every structure with respect to training set size

the atlas segmented shape, and compared with the actual 4D segmentations. Fig. 4 shows the results.

4 Discussion and Conclusion

The experiments show that the conditional shape model is able to predict the motion of the heart with an average accuracy of around 1.0 mm (see Table 1). This represents a large improvement compared to shape replication, which is the only option when only a 3D CTA scan is available (see Fig. 2). Accuracy is measured by the point-to-surface error, but depending on the application of interest, the point-to-point error can be used as well. For the here presented experiments, point-to-point errors were on average larger, but showed very similar trends.

Furthermore, the remaining errors when projecting the entire sequence onto the model subspace and the results of the generalization ability experiment show that the size of the training set is most probably not sufficient. Therefore, the effect of adding more training shapes on the accuracy will be investigated in the future. The errors at the left of the graph are related to the magnitude of the motion found by the registration approach. Errors induced by landmark propagation are currently not taken into account and subject to future work.

In the current work no assumptions were made about the relation between cardiac shape and motion, in contrast to the method proposed by Hoogendoorn et al., in which shape and motion variations are explicitly decoupled [5]. Although the motion of the heart depends for a large amount on its electrophysiology, which is not apparent from its shape, we think that shape can still predict cardiac

motion when both the shape and motion of the heart are affected by disease (e.g. due to cardiac remodelling after myocardial infarction). Motion changes related to the shape at other time points in the cardiac cycle are not taken into account.

The conditioning framework described in this paper exploits any available correlation between shape and motion given in the training sequences, and naturally reduces to the mean motion in case no correlation is present. Due to the small training set size we were not able to draw any conclusions with respect to this correlation. We plan to investigate this topic in detail as part of future work. For the conditioning 95% of the statistical shape model variance was retained, however the effect of the value of this parameter on the prediction results should still be investigated.

In conclusion, we presented and evaluated a method that uses conditional shape models to predict cardiac motion from 3D cardiac shape. The results suggest that it can reliably predict the motion of the heart. Our goal is to use this model for the dynamic alignment of pre-operatively acquired CTA images with intra-operative X-ray imaging, which should be applied and evaluated in future work.

References

1. Metz, C., et al.: Patient specific 4D coronary models from ECG-gated CTA data for intra-operative dynamic alignment of CTA with X-ray images. In: *Med. Image Comput. Assist. Interv. Int.* (2009)
2. Heimann, T., Meinzer, H.P.: Statistical shape models for 3D medical image segmentation: a review. *Med. Image Anal.* 13(4), 543–563 (2009)
3. Ordas, S., Oubel, E., Leta, R., Carreras, F., Frangi, A.: A statistical shape model of the heart and its application to model-based segmentation. In: *Proc. SPIE Med. Imaging*, vol. 6511 (2007)
4. Perperidis, D., Mohiaddin, R., Rueckert, D.: Construction of a 4D statistical atlas of the cardiac anatomy and its use in classification. In: *Med. Image Comput. Assist. Interv. Int.*, vol. 8, pp. 402–410 (2005)
5. Hoogendoorn, C., Sukno, F., Ordás, S., Frangi, A.: Bilinear models for spatio-temporal point distribution analysis. *Int. J. of Comput. Vision* 85(3), 237–252 (2009)
6. Cootes, T., Cooper, D., Taylor, C., Graham, J.: Active shape models - their training and application. *Comput. Vision and Image Understanding* 61(1), 38–59 (1995)
7. Kirisli, H., et al.: Fully automatic cardiac segmentation from 3D CTA data: a multiatlas based approach. In: *Proc. SPIE Med. Imaging* (2010)
8. Klein, S., Staring, M., et al.: *elastix*: a toolbox for intensity based medical image registration. *IEEE T. Med. Imaging* 29(1), 195–205 (2010)
9. Comaniciu, D., Meer, P.: Mean Shift: A robust approach toward feature space analysis. *IEEE T. Pattern Anal.* 24(5), 603–619 (2002)
10. de Bruijne, M., Lund, M., Tankó, L., Pettersen, P., Nielsen, M.: Quantitative vertebral morphometry using neighbor-conditional shape models. *Med. Image Anal.* 11(5), 503–512 (2007)

Cross-Modality Assessment and Planning for Pulmonary Trunk Treatment Using CT and MRI Imaging

Dime Vitanovski^{1,2}, Alexey Tsymbal¹, Razvan Ioan Ionasec¹, Bogdan Georgescu¹, Martin Huber¹, Andrew Taylor³, Silvia Schievano³, Shaohua Kevin Zhou¹, Joachim Hornegger², and Dorin Comaniciu¹

¹ Integrated Data Systems, Siemens Corporate Research, Princeton, USA

² Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg, Germany

³ Great Ormond Street Hospital for Children, London, England

Abstract. Congenital heart defect is the primary cause of death in newborns, due to typically complex malformation of the cardiac system. The pulmonary valve and trunk are often affected and require complex clinical management and in most cases surgical or interventional treatment. While minimal invasive methods are emerging, non-invasive imaging-based assessment tools become crucial components in the clinical setting. For advanced evaluation and therapy planning purposes, cardiac Computed Tomography (CT) and cardiac Magnetic Resonance Imaging (cMRI) are important non-invasive investigation techniques with complementary properties. Although, characterized by high temporal resolution, cMRI does not cover the full motion of the pulmonary trunk. The sparse cMRI data acquired in this context include only one 3D scan of the heart in the end-diastolic phase and two 2D planes (long and short axes) over the whole cardiac cycle. In this paper we present a cross-modality framework for the evaluation of the pulmonary trunk, which combines the advantages of both, cardiac CT and cMRI. A patient-specific model is estimated from both modalities using hierarchical learning-based techniques. The pulmonary trunk model is exploited within a novel dynamic regression-based reconstruction to infer the incomplete cMRI temporal information. Extensive experiments performed on 72 cardiac CT and 74 cMRI sequences demonstrated the average speed of 110 seconds and accuracy of 1.4mm for the proposed approach. To the best of our knowledge this is the first dynamic model of the pulmonary trunk and right ventricle outflow track estimated from sparse 4D cMRI data.

1 Introduction

Congenital Heart Defect (CHD) is the primary cause of death in newborns characterized by complex malformations of the heart and great vessels. Often, the right side of the heart is affected and especially the pulmonary trunk, as in Tetralogy of Fallot (TOF) and pulmonary artesia or stenosis. The clinical management

of such conditions is confronted with complex treatment decisions, which include pulmonary valve procedures in the majority of the cases.

Percutaneous interventions for pulmonary valve replacement are emerging as feasible treatment alternatives to the classical cardiac surgery with important benefits: less invasive, reduced risks associated with cardiopulmonary bypass, bleeding, infections and reduced expenses for postoperative intensive care [1]. Nevertheless, comprehensive investigation, based on non-invasive imaging techniques, is still mandatory for clinical decision making and treatment success.

For therapy planning purposes, the pulmonary trunk is increasingly imaged using either cardiac Computer Tomography (CT) or cardiac Magnetic Resonance Imaging (cMRI) [2]. While CT has a high spatial resolution, fast acquisition times without anesthesia, it has the disadvantages of poor temporal resolution and ionizing radiation. Contrary, cMRI has high temporal resolution without X-ray radiation, but long acquisition times and usually does not cover the full 4D information. The regular protocol, so called sparse 4D cMRI, involves an end-diastolic (ED) 3D heart image and two orthogonal cine projections 2D+t, short axis (SA) and long axis (LA). LA passes through the main pulmonary artery and the descending aorta, while SA is aligned with pulmonary valve, perpendicular to the LA (see Fig. 1). Ideally, clinicians would be provided with a accurate morphological and functional quantification of the pulmonary trunk, independent of the employed imaging technique.

In this paper we present a cross-modality framework for the evaluation of the pulmonary trunk, which combines the advantages of both, cardiac CT and cMRI, non-invasive imaging techniques. A physiological model, which captures complex morphological, dynamic and pathologic variations of the pulmonary trunk is presented in Sec. 2. In Sec. 4, the patient-specific model parameters are estimated from both modalities within a hierarchical learning-based framework, which involves three-stages: landmark detection, center line detection and dynamics estimation. A novel dynamic regression-based reconstruction is proposed in Sec. 3 to infer the incomplete temporal information characteristic to the sparse cMRI protocols is presented.

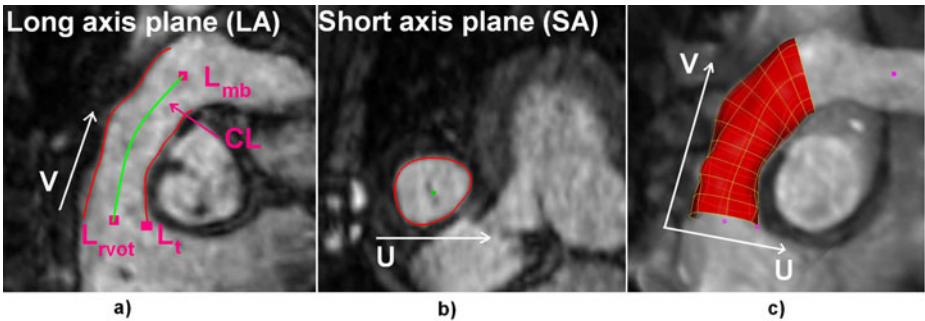


Fig. 1. 3D cMRI scan of the whole heart in the ED phase (a). 2D long axis (LA) plane (b) and short axis (SA) plane (c) of the pulmonary artery over the cardiac cycle.

2 Physiological Pulmonary Trunk Modelling

In this section we introduce our physiological model of the RVOT and pulmonary trunk, which represents both morphological and dynamic variations. Similar to [3], the anatomical complexity is reduced by employing a coarse to fine parameterization which includes: anatomical landmarks, pulmonary artery center line and the full surface model of the pulmonary trunk. As illustrated in Fig. 1(a), the considered anatomical landmarks include, Trigone (L_t), RVOT (L_{rvot}) and Main-Bifucation (L_{mb}), each represented in the Euclidean 3D space, $L_x \in R^3$. The centre line CL passes through the pulmonary artery center and is parameterized by 12 points, $CL = CL_0 \dots CL_{11}$. The surface model S is represented by a structured grid, spanned along two anatomical directions, u – circumferential and v – longitudinal, using 50×40 vertices (see Fig. 1(c)). Point correspondence in time and across patients is enforced by intrinsic resampling of S , using a set of anatomical-driven cutting-planes, described by center line points CL_x and corresponding tangential directions. Given the different modalities supported and characteristic imaging protocols, we differentiate among two dynamic extensions of the proposed physiological model. The definition of a full 4D model, which can be directly estimated from 4D cardiac CT data, is rather straightforward and realized by plain concatenating a time variable t :

$$Model_{full4D} = \{L_t, L_{rvot}, L_{mb}, CL, S\}_t \tag{1}$$

However, given the sparse 4D acquisition, common to cMRI, the extension to a temporal model includes two additional representations: LA and SA . LA which describes the contour of S intersected with the plane with the origin in L_{rvot} and the normal obtained from the cross-product between $L_{trigone}$ and L_{rvot} , and the center line tangent at CL_0 represents a specific 2D+time long axis acquisition (Fig. 1(a)). SA describes the contour of S intersected with the plane center in CL_{middle} and the corresponding tangent as normal (Fig. 1(b)). Hence, the sparse dynamic model is parameterized as follows:

$$Model_{sparse4D} = \{L_t, L_{rvot}, L_{mb}, CL, S\}_{ED} + \{LA, SA\}_t \tag{2}$$

3 Dynamic Regression Based 4D Model Reconstruction

As describe above, the dynamic information in case of 4D cMRI is incomplete but rather available only in two orthogonal projections, as opposed to 4D cardiac CT, which provides full dynamics over the cardiac cycle.

In *regression* a solution to the following optimization problem is normally sought [4]:

$$\hat{\mathcal{R}}(\mathbf{x}) = argmin_{\mathcal{R} \in \mathfrak{S}} \sum_{n=1}^N L(y(\mathbf{x}_n), \mathcal{R}(\mathbf{x}_n)) / N \tag{3}$$

where \mathfrak{S} is the set of possible regression functions, $L(\circ, \circ)$ is a loss function that penalizes the deviation of the regressor output $\mathcal{R}(\mathbf{x}_n)$ from the true output, and

N is the number of available training examples. In our case the reconstruction task is defined as a regression problem between the full dynamic model of the pulmonary trunk extracted from 4D CT data and the sparse one extracted from the sparse cMRI data:

$$Y(Model_{full4D}) = \hat{\mathcal{R}}(X(Model_{sparse4D})) + \epsilon \quad (4)$$

In our regression problem we focus on shape information and completely neglect volume data. As descriptors both for the input $X(Model_{sparse4D})$ and output elements $Y(Model_{full4D})$ of the models we choose *coordinates* of mesh vertices normalized with the generalized procrustes analysis. This representation has a uniform representation of the input and the output data. The training set T used to generate the regression model thus includes feature vectors T_i as follows:

$$T_i = \langle (S_i^{ED}, LA_i^t, SA_i^t)_{MRI}, (S_i^t)_{CT} \rangle, \quad (5)$$

where t is a time step within the cardiac cycle, S_i^{ED} is a set of 3D coordinates representing each point of the end-diastolic model (2000 3D points), LA_i^t and SA_i^t are point sets (80 and 50 3D points respectively) representing the model curves extracted from the cMRI's long axis stack and short axis stack respectively, for the current time step t , and $(S_i^t)_{CT}$ are the corresponding point coordinates for the point set to be reconstructed (238 3D points). Due to the dense representation of our model (2000 3D points) we reconstruct only the most significant 238 3D points from the associated CT model. The rest of the points are interpolated and projected onto the PCA shape space from which the complete model is then obtained.

The formulated regression problem is solved by learning the regression function \mathcal{R} with two different methods: *boosting-based additive regression* [5] and *random forest* [6]. Two main reasons motivate our choice. First, these techniques were shown to be robust to high-dimensional data with many irrelevant, redundant and noisy features, without the need for additional data pre-processing and feature selection. This was shown both for classification [7, 8] and regression [4] tasks. Second, both boosting-based and random forest-based models are relatively fast to train and to evaluate comparing for example with Support Vector Regression. In the spirit of [7, 4], we use *simple 1D linear regression* as the base learner for boosting-based regression. Each weak learner is a simple linear regressor of the form:

$$y = \beta_0 x + \beta_1 \quad (6)$$

where x is the selected scalar input coordinate and y is a scalar output coordinate. Using more sophisticated weak learners such as CART decision trees and multiple linear regression with greedy forward feature inclusion, has proven to always result in a worse or no better performance while the resulting model gets significantly more complicated. Using simple 1D binary decision stumps as in [4] has also proven to lead to suboptimal accuracy; the reason for this is perhaps the nature of the data, as it is rather impossible to generate as many candidate decision stumps with the coordinate - based features as it is possible with

the Haar-like features. For each boosting-based model, we generate 200 weak learners. The accuracy plateaus with this number of component models, and the further accuracy increase is always insignificant with this data.

For random forests, we always generate 25 component trees. The accuracy usually remains same or even decreases with the addition of more trees to the model. The minimum leaf size is set to 1; the trees are thus generated to the full with no pruning. The number of features considered at each node is set to the value recommended by Breiman [6], which is one third of the total number of features for regression. Using other parameter settings was shown to lead to worse or no better accuracy in our preliminary experiments.

In boosting-based regression the output function is assumed to take a linear form as follows [4]:

$$\hat{\mathcal{R}}(\mathbf{x}) = \sum_{t=1}^T \alpha_t h_t(x) \in H \quad (7)$$

where $h_t(x)$ is a base (weak) learner and T is the number of boosting iterations. Having a *linear* base learner (simple linear regression), a *linear* final solution is thus also found. In contrast to this, random forests seek for a non-linear function approximation, recursively splitting the feature space in the nodes of component decision trees.

In contrast to [4], we use *naive decoupling* of the regression problem into a number of single output problems. While multi-output regression solutions do exist both for boosting [4], for our task multi-output optimization was not shown to lead to error decrease and time savings were rather insignificant.

4 Estimating Patient-Specific Model Parameters

The patient-specific model parameters described in Section 2 are estimated from cardiac acquisition using a learning-based approach. Detectors are learned separately for both modalities, CT and cMRI, and applied to estimate model parameters in a hierarchical three-stage approach: Anatomical Landmarks Estimation, Center Line Estimation and Full Surface Model Estimation.

Anatomical Landmarks Estimation. By defining the localization as a classification problem, the anatomical landmarks, L_t, L_{rvot}, L_{mb} , are estimated within the Marginal Space Learning (MSL) framework [9]. Separate detectors $D_t^L, D_{rvot}^L, D_{mb}^L$, are learned using the Probabilistic Boosting Tree (PBT) [8] in combination with Haar-like feature from a training dataset annotated by experts.

$$p(L_x|x, y, z) = D_x^L(x, y, z), (x, y, z) \in \sigma^x \quad (8)$$

The trained detectors D_x^L models the target posteriori distribution $p(L_x|x, y, z)$ for a specific search space σ^x formed by the training set. MSL is applied to conduct guided search in the parameter space using the learned detectors and obtained the location of the anatomical landmarks. Note that in case of 4D cardiac CT anatomical landmark are detected in each volume to obtain the dynamic parameters $\{L_t, L_{rvot}, L_{mb}\}_t$, while in sparse cMRI only a static detection in the end-diastolic volume is performed: $\{L_t, L_{rvot}, L_{mb}\}_{ED}$.

Center Line Estimation. CL passes through the centre of the pulmonary trunk and is initialized by the previously detected landmarks L_t and L_{rvot} . A robust detector D^{CL} is learned using the same MSL framework to detect circular structures, parameterized by center line points CL_x , corresponding tangent and having fixed radius $r = 20mm$ obtained from the average value in the training set. An incremental approach is used to search circles on a series of successively updating planes. Please note, that for the Anatomical Landmarks Estimation, a temporal center-line model CL_t is detected in CT and a static CL in cMRI.

Full Surface Model Estimation. The full model of the pulmonary trunk S is initialized in the end-diastolic frame using the estimated landmarks and center-lines, using a piecewise affine transformation along the center line [3]. Robust boundary detectors D^s , trained using the PBT and steerable feature [9] are applied to locally refine the surface by moving it along normal directions towards the position with the highest boundary probability. To obtain spatially smooth delineation, the final results are obtained by projecting S to a previously learned shape space model.

In case of CT, the above described algorithm is applied in each time step to obtain the full temporal model $\{L_t, L_{rvot}, L_{mb}, CL, S\}_t$. In case of MR, the estimated surface in the end-diastolic frame S_{ED} is used to initialize the contours LA and SA . These are refined using a trained D^c contour detector as described above. A full dynamic 4D model is then estimated by using a learned regression model (see Eq. 4) to predict the missing temporal information.

5 Results

5.1 Patient-Specific Model Parameters Estimation

The proposed framework for detecting a personalized pulmonary trunk model in 4D CT and sparse cMRI data was evaluated on 50 4D CT (500 volumes) and 74 sparse cMRI (74 ED Volumes associated with 4736 LA/SA planes) studies from patients with different CHD. Each volume in the data set is associated with annotation, manually generated by experts, which is considered as ground truth. Three-fold cross validation was used to divide the data set into training and test data.

Table 1 summarizes the detection performance on both modalities (CT and sparse cMRI), from the test data. Point-to-mesh measurement error was used to evaluate the detection accuracy between the ground truth and the detected model for both modalities. Average speed of 10sec per frame was achieved for both modalities on a standard 2.0GHz Dual Core PC.

Table 1. Detection accuracy

CT/MRI	Mean Error(mm)	Median(mm)	Std.Dev(mm)
Landmarks	3.5/4.3	5.1/6.4	2.7/3.0
Center Line	3.0/3.3	2.3/2.3	1.7/2.0
Full Surface	1.6/1.9	1.2/1.3	0.2/0.2

5.2 Intra-modality Comparison between CT and MRI

The inter-modality consistency of the model was demonstrated on a subset of 10 patients which underwent both imaging investigations, 4D CT and sparse cMRI (see Fig. 2). Ground truth and detected pulmonary trunk models from both modalities were compared using the point-to-mesh measurement and clinically relevant diameter measurements: RVOT, hinges and commissures. Results are summarized in Table 2. A strong inter-modality correlation, $r = 0.992$, $p < 0.0001$ and confidence of 98%, was obtained for CT and cMRI for the pulmonary trunk models.

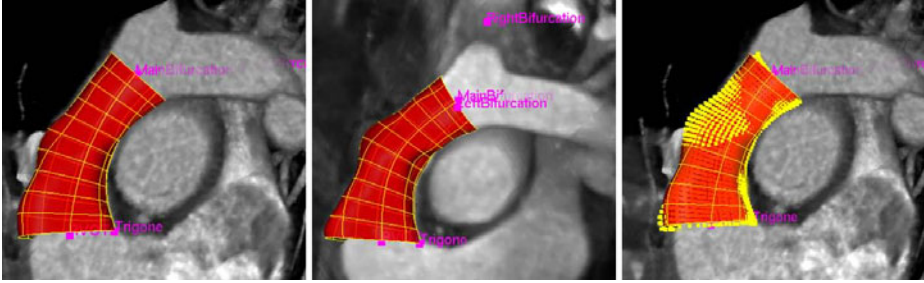


Fig. 2. Pulmonary trunk model in CT (left) and cMRI (middle) data for the ED phase. Inter-modality consistency by projecting the cMRI model (yellow points) into the CT data (right).

Table 2. Model based intra-modality comparison between CT and MRI

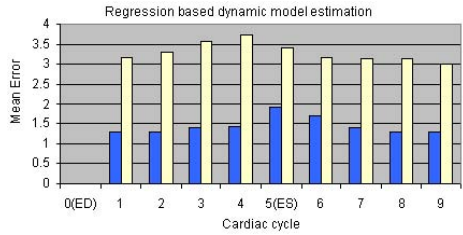
(mm)	Ground truth	Estimation
RVOT	0.7 ± 0.5	3.8 ± 1.5
Hinges	1.2 ± 1.4	2.6 ± 4.7
Commissures	1.5 ± 1.2	3.2 ± 1.7
Point-to-mesh	1.4 ± 0.1	2.5 ± 0.7

5.3 Regression Based Dynamic Model Reconstruction

As described in Sec. 1 the sparse cMRI protocol is able to capture the full anatomy of the pulmonary trunk only in the ED phase (3D volume) of the heart and parts of the pulmonary trunk in 2D planes (LA and SA) over the cardiac cycle. However, a full 4D model of the pulmonary trunk can be still computed from the available sparse data by learning the full motion from 4D CT data. For this purpose we learned a regression model as presented in Sec. 3 on a training data set of 72 4D CT (720 Volumes) studies. Two different machine-learning techniques (boosting and random forest) are used to train the regression model and to evaluate the reconstruction error. Table 3 presents reconstruction results obtained by applying the learned regression model on the sparse cMRI images and evaluating the reconstructed model on 4D CT, for a set of 10 patients, which underwent both imaging modalities. Figure 3 illustrates the reconstruction error distributed over the cardiac cycle.

Table 3. Reconstruction error for Random Forest and Boosting

	Boosting	Random Forest
Mean Err.	1.44(mm)	3.2(mm)
Std. Dev	0.21(mm)	0.23(mm)
Speed	3.07 (ms)	6.21 ms



6 Conclusion

In this paper we propose a cross-modality detection framework for estimating a dynamic personalized model of the pulmonary trunk from the available data, 4D CT and sparse cMRI. A novel regression based reconstruction method is presented and used to infer the incomplete temporal information characteristic to the sparse cMRI protocols. The estimated model from both modalities can be utilized to extract morphological and functional information of the pulmonary trunk and dynamics over the cardiac cycle. Extensive experiments performed on a large heterogeneous data set demonstrated a precision of 1.44mm data at a speed of 11 seconds per volume. The proposed method has the potential to significantly advance the pulmonary trunk treatment.

References

1. Bonhoeffer, P., et al.: Percutaneous insertion of the pulmonary valve. *Journal of the American College of Cardiology* 39(10), 1664–1669 (2002)
2. Taylor, A.: Cardiac imaging: Mr or ct? which to use when. *Pediatr. Radiology* 38 (2008)
3. Vitanovski, D., Ionasec, R., Georgescu, B., Huber, M., Taylor, A., Hornegger, J., Comaniciu, D.: Personalized pulmonary trunk modeling for intervention planning and valve assessment estimated from CT data. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009*. LNCS, vol. 5761, pp. 17–25. Springer, Heidelberg (2009)
4. Zhou, S.K., Georgescu, B., Zhou, X.S., Comaniciu, D.: Image based regression using boosting method. In: *ICCV* (2005)
5. Friedman, J.H.: Greedy function approximation: A gradient boosting machine. *Annals of Statistics* 29, 1189–1232 (2000)
6. Breiman, L.: Random forests. In: *Machine Learning*, pp. 5–32 (2001)
7. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: *CVPR* (2001)
8. Tu, Z.: Probabilistic boosting-tree: Learning discriminativemethods for classification, recognition, and clustering. In: *ICCV* (2005)
9. Zheng, Y., Barbu, A., et al.: Fast automatic heart chamber segmentation from 3d ct data using marginal space learning and steerable features. In: *ICCV* (2007)

Automatic Cardiac MRI Segmentation Using a Biventricular Deformable Medial Model

Hui Sun¹, Alejandro F. Frangi^{2,3}, Hongzhi Wang¹, Federico M. Sukno^{2,3},
Catalina Tobon-Gomez^{2,3}, and Paul A. Yushkevich¹

¹ Penn Image Computing and Science Laboratory (PICSL), Department of Radiology, University of Pennsylvania, Philadelphia, USA

² Center for Computational Imaging & Simulation Technologies in Biomedicine, Universitat Pompeu Fabra, Barcelona, Spain

³ Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Zaragoza, Spain

Abstract. We present a novel approach for automatic segmentation of the myocardium in short-axis MRI using deformable medial models with an explicit representation of thickness. Segmentation is constrained by a Markov prior on myocardial thickness. Best practices from Active Shape Modeling (global PCA shape prior, statistical appearance model, local search) are adapted to the medial model. Segmentation performance is evaluated by comparing to manual segmentation in a heterogeneous adult MRI dataset. Average boundary displacement error is under 1.4 mm for left and right ventricles, comparing favorably with published work.

1 Introduction

Segmentation of the human myocardium in *in vivo* MRI is a necessary first step for various computational analyses of heart structure and function. Automatic segmentation is a challenge because intensity characteristics of cardiac MRI are complex, and because the myocardium is thin relative to typical MRI voxel size, particularly in the right ventricle (RV). Most existing techniques only segment the left ventricle (LV) (e.g. [11,12,116]), or segment either the inner surface (endocardium) or outer surface (epicardium) of the RV [9,17,113]. When characterizing pathology such as myocardial hypertrophy, myocardial infarction, and ventricular arrhythmias, a segmentation of both surfaces in LV and RV is desirable. Examples of dual-surface RV and LV segmentation exist in the literature [10], but there remains substantial room for improvement in accuracy.

When segmenting structures known to be thin *a priori*, it is natural to build thickness constraints into the segmentation algorithm. Standard active shape model (ASM), level set, and registration-based approaches, used widely for cardiac segmentation, do not provide an explicit way of doing so. However, coupled-surface approaches incorporated into the level set framework [16,12] can be used to constrain the distance between endocardial and epicardial surfaces. In this paper, we explore an alternative approach, where a Markovian prior on model thickness is introduced by representing the myocardium using a deformable *medial*

model. The medial model explicitly represents the skeleton of the myocardium as well as myocardial thickness. The boundary of the model (epicardial and endocardial) is derived from the skeleton and thickness data using a simple analytical expression. A parameterization of the volumetric region enclosed by the model is also easily derived. Thus, the model can explicitly represent constraints and prior knowledge regarding myocardial thickness, and there is no possibility of endocardial and epicardial surfaces crossing during model deformation.

Our approach builds on earlier work [14] that demonstrated the feasibility of capturing and analyzing myocardium shape using the *continuous medial representation (cm-rep)* with branching skeletons (Fig. 1). This paper makes the step from shape modeling to segmentation by developing, from a large training set, a shape and thickness prior, which combines standard PCA for shape with a Markov model for thickness; and by incorporating into the cm-rep framework a statistical learning-based appearance model derived from the Active Shape Model (ASM) framework [43]. This leads to excellent segmentation accuracy on short-axis cardiac MRI from a heterogeneous patient population.

2 Materials and Methods

2.1 Subjects, Imaging and Manual Segmentation

The primary dataset in this study, used for appearance learning and evaluation, consists of data from 40 subjects in four clinical cohorts: normal controls (n=10), myocardial infarction (MI) (n=10), hypertrophy (n=10) and dilation (n=10). Short-axis cardiac MRI (TR/TE=2.9/1.2 ms, FA=45°, in-slice resolution = 1.56 mm × 1.56 mm, slice thickness= 8 mm, no gap, FOV = 400 mm × 340 mm) were acquired at CETIR Sant Jordi Centre (Barcelona, Spain) using a GE Signa CVi-HDx 1.5T scanner (GE Healthcare, Milwaukee, USA). Author XX manually outlined the endocardial and epicardial surfaces of the LV and RV at end diastole, forming a volumetric two-chamber heart representation that typically spans 8-12 slices from the base to the apex. A secondary dataset of manual segmentations derived from similar short axis MRI of 81 subjects with a variety of common cardiovascular pathologies was used in the construction of the shape and thickness priors.

2.2 Medial Models

The medial model is defined as a pair (\mathbf{m}, R) , where \mathbf{m} is the skeleton, i.e., a collection of adjoining co-dimension 1 manifolds in \mathbb{R}^3 , and R is a positive scalar field defined on \mathbf{m} , called the radial field. In practice, a triangular mesh $(\mathbf{m}, R; \mathcal{E})$ is defined by applying Loop [8] subdivision rules (modified to handle junctions between manifolds) to a triangular mesh of control points $(\mathbf{m}_{\text{ctl}}, R_{\text{ctl}}; \mathcal{E}_{\text{ctl}})$, where \mathcal{E} , \mathcal{E}_{ctl} denote the edges in these meshes. Edges in the medial mesh (either control or subdivided) may belong to one, two or three triangles. The set of edges with three adjacent triangles is called the *medial seam*, and the set of edges adjacent

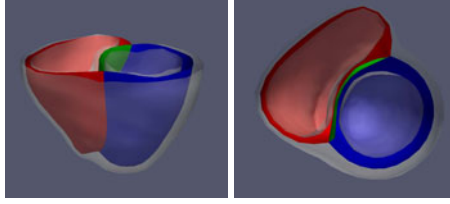


Fig. 1. The medial model of the myocardium that is used as a deformable template in our method. The skeleton of the model is partitioned into red, blue, and green surfaces corresponding to the RV wall, LV wall, and interventricular septum. The translucent gray surface is the boundary of the model, i.e. epicardium and endocardium.

to just one triangle is called the *medial edge*. Below, we treat (\mathbf{m}, R) as smooth manifolds, although in practice they are approximated by discrete meshes.

The *boundary* corresponding to the pair (\mathbf{m}, R) is defined as a set of points \mathcal{B} , such that the Blum medial axis transform (MAT) [2] of \mathcal{B} is (\mathbf{m}, R) . Generally, \mathcal{B} cannot be described analytically as a function of (\mathbf{m}, R) , as MAT may map whole patches of \mathcal{B} to a single point on \mathbf{m} . However, if (\mathbf{m}, R) satisfy a set of conditions given in [15], such an analytical description is possible. It is given by $\mathcal{B} = \mathbf{b}^+ \cup \mathbf{b}^-$, where \mathbf{b}^+ , \mathbf{b}^- are surfaces that lie on the opposite sides of \mathbf{m} :

$$\mathbf{b}^\pm = \mathbf{m} + R \left[-\nabla_{\mathbf{m}} R \pm \sqrt{1 - \|\nabla_{\mathbf{m}} R\|^2} \cdot \mathbf{N}_{\mathbf{m}} \right], \quad (1)$$

where $\nabla_{\mathbf{m}} R$ is the Riemannian gradient of R on \mathbf{m} , and $\mathbf{N}_{\mathbf{m}}$ is the unit normal vector to \mathbf{m} . The sufficient conditions for $\mathcal{B} = \mathbf{b}^+ \cup \mathbf{b}^-$ given in [15] include equality conditions that must hold along medial edges and seams, as well as inequality conditions that hold everywhere on the skeleton. For brevity, we do not discuss them here in detail. It suffices to say that the constraints ensure that surfaces \mathbf{b}^+ , \mathbf{b}^- do not self-intersect and that \mathbf{b}^+ , \mathbf{b}^- join together along medial seams and edges to form a single closed surface \mathcal{B} .

The medial model of the myocardium is shown in Fig. 1. Surfaces \mathbf{b}^+ and \mathbf{b}^- describe the the epicardium and endocardium. The skeleton \mathbf{m} has three branches, corresponding to the LV wall, the RV wall, and the ventricular septum.

2.3 Statistical Shape Model

Training data are used to construct statistical models for myocardial shape and thickness, as follows. The myocardial model (Fig. 1) is deformed to binary segmentations of the myocardium in training data. The likelihood and prior terms used in this deformation are specified in [14]. In brief, the likelihood measures volume overlap between the binary segmentation and the model’s interior, and the prior ensures model validity by penalizing violation of constraints in [15]. Rudimentary correspondence between training data is achieved by penalizing the distortion in area element of \mathbf{m} during deformation (this is roughly equivalent to equal arc length parameterization for 2D models). The Generalized

Procrustes algorithm is applied to points in the skeleton mesh to remove pose variability from fitted medial models.

Variability in myocardial shape is modeled using a multivariate Gaussian distribution: principal component analysis (PCA) is applied to the x, y, z coordinates of vertices on the skeleton \mathbf{m} . Thickness variability is modeled by a more localized Gibbs distribution (i.e., thickness is treated as a Markov random field):

$$-\log p(R_1, \dots, R_N; \mu_\bullet, \sigma_\bullet) \sim \lambda_1 \sum_{i=1}^N \frac{(R_i - \mu_i)^2}{\sigma_i^2} + \lambda_2 \sum_{(j,k) \in \mathcal{E}} \frac{(|R_j - R_k| - \mu_{jk})^2}{\sigma_{jk}^2}, \quad (2)$$

where $i = 1 \dots N$ indexes vertices in the skeleton, and $\{\mu_i, \sigma_i, \mu_{jk}, \sigma_{jk}\}$ are parameters estimated from the training data. The reason we use different priors for \mathbf{m} and R stems from the fact that thickness varies much more smoothly than the x, y, z coordinates of the skeleton. Thus, the MRF is an appropriate model for thickness, but it is too restrictive for shape. We performed experiments that show that using PCA for both \mathbf{m} and R leads to worse segmentation performance than using PCA for \mathbf{m} and MRF for R .

2.4 Appearance Model

At each vertex in the medial model, we build a model of local appearance, which is subsequently used to drive image segmentation. Recall that [\(II\)](#) associates each vertex \mathbf{m}_i on the skeleton with two boundary vertices (BV), \mathbf{b}_i^+ and \mathbf{b}_i^- . At each BV, we train a classifier to discriminate between a “well-placed” BV and a “misplaced” BV. A well-placed BV lies on the corresponding anatomical boundary in the training image, and a misplaced BV lies some distance away from the anatomical boundary. Well-placed BVs are obtained by fitting models to manual segmentations of the myocardium in the training data. Misplaced BVs are obtained by applying displacements to the well-placed BVs along the direction from \mathbf{b}_i^+ to \mathbf{b}_i^- (called the *chord direction* because $\mathbf{b}_i^+ \mathbf{b}_i^-$ is a chord of the sphere (\mathbf{m}_i, R_i)). So training exemplars for each classifier include well-placed and misplaced versions of a given BV across all subjects included in the training subset. To further increase the number of training exemplars and make classifiers less sensitive to location, we include, as training exemplars for each classifier, misplaced and well-placed versions of the BVs in the two-ring neighborhood of the BV associated with the classifier.

A rich set of features is used to build these classifiers. First, for each MRI slice, we compute a set of rotation-invariant texture descriptors at different scales [\[7\]](#). Using Einstein notation, with L_α denoting image derivative in direction $\alpha \in \{x, y\}$, these descriptors are given by $L, L_\alpha L_\alpha, L_{\alpha\alpha}, L_\alpha L_{\alpha\beta} L_\beta, L_{\alpha\beta} L_{\beta\alpha}$. These texture descriptors are sampled around each BV using a cylindrical sampling grid oriented along the chord direction. Linear interpolation is used to sample texture descriptors between slices. Thousands of features are obtained for each BV. For each feature, a simple threshold-based *weak classifier* is constructed. AdaBoost [\[5\]](#) is used to combine these weak classifiers into a single *strong classifier*.

2.5 Automatic Segmentation

The automatic segmentation is initialized manually by identifying five landmarks in the most basal slice of the MRI image and one landmark at the apex. The template medial model is affinely aligned to the six landmarks. Subsequently, the model is deformed by iteratively applying two steps: *local search* and *Bayesian deformation*. Both of these steps are adaptations of the segmentation algorithm used in active shape models [43], and we only summarize them here. During local search, we displace each BV along the chord direction, so as to maximize the probability of it being a “well-placed” BV; this probability is estimated using the AdaBoost classifier trained for that BV. Given two displaced BVs $\hat{\mathbf{b}}_i^+$ and $\hat{\mathbf{b}}_i^-$, we compute a new position $\hat{\mathbf{m}}_i$ such that the triangles $\hat{\mathbf{b}}_i^+ \hat{\mathbf{m}}_i \hat{\mathbf{b}}_i^-$ and $\mathbf{b}_i^+ \mathbf{m}_i \mathbf{b}_i^-$ lie in the same plane and have the same shape; the new radius value \hat{R}_i is equal to the radius of the circumscribed circle of $\hat{\mathbf{b}}_i^+ \hat{\mathbf{m}}_i \hat{\mathbf{b}}_i^-$. The mesh $(\hat{\mathbf{m}}, \hat{R}, \mathcal{E})$ resulting from the local search step is not guaranteed to satisfy any of the constraints for a “valid” medial model; nor does it respect statistical shape models computed from training data. During the Bayesian deformation step, we deform the template myocardial model to optimize a log posterior energy that is decomposed into log likelihood and log prior terms. The log likelihood simply measures similarity between (\mathbf{m}, R) in the model and $(\hat{\mathbf{m}}, \hat{R})$ computed by the local search:

$$\log p(\hat{\mathbf{m}}, \hat{R} | \mathbf{m}, R) = \sum_{i=1}^N [\|\mathbf{m}_i - \hat{\mathbf{m}}_i\|^2 + (R_i - \hat{R}_i)^2].$$

This likelihood treats \mathbf{m} and R as having the same units, which has not been a problem in practice. A likelihood term based on Mahalanobis distance could also be used as an alternative. Prior terms include (a) validity priors that ensure constraints on the medial model are satisfied; (b) PCA-based statistical shape prior on \mathbf{m} ; (c) Markovian prior on R in (2); (d) a penalty on distortion in area element, used to maintain rudimentary correspondence between subjects.

3 Results

Segmentation performance is evaluated using a cross-validation strategy. We perform ten experiments in which the primary dataset (n=40) is divided into a training subset (n=24, 6 images from each cohort) and test subset (n=16, 4 images from each cohort). In each experiment, we build shape and thickness priors and train AdaBoost classifiers for appearance modeling using the training subset. We then apply the segmentation method to the test subset. We measure the difference between automatic segmentation results and manual segmentations and report average errors over the ten cross-validation experiments. Errors are reported in terms of the widely used point-to-mesh distance metric (e.g., [1]). For each point on the mesh, the closest point (not necessarily a vertex) on the other mesh is located and the Euclidean distance between these two points is

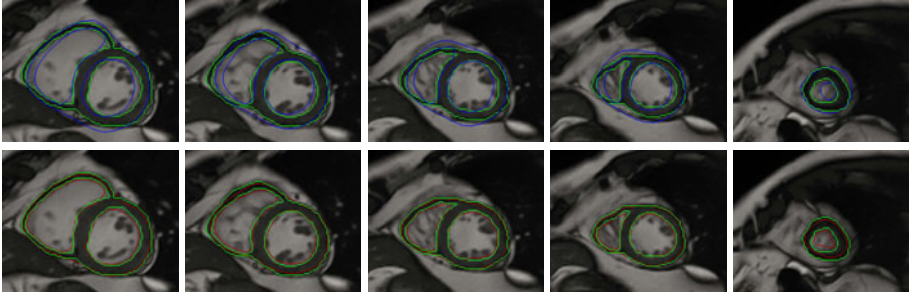


Fig. 2. Example of automatic segmentation in a single subject. The top row shows the manual segmentation in green and the model initialized by landmarks in blue. The bottom row shows the segmentation result in red, with manual segmentation in green. From left to right, slices progress from most basal slice to the apex.

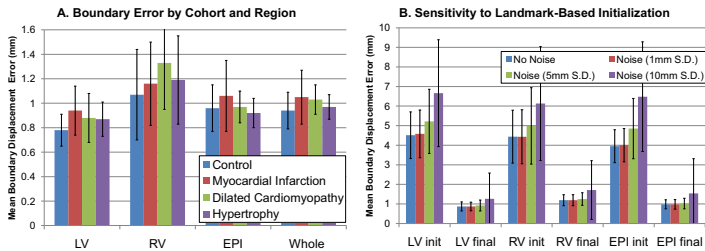


Fig. 3. Accuracy of the proposed method, expressed in terms of average mesh-to-point boundary displacement error. **A.** Average error in four cohorts. Error is computed separately for endocardial LV and RV surfaces, epicardium, and whole myocardium. **B.** Sensitivity to initialization. Random displacement with standard deviation of 1, 5 or 10mm is applied to the landmarks used during initialization. Boundary displacement error before (Init) and after segmentation (Final) is plotted for endocardial LV and RV surfaces and the epicardium. Error bars indicate standard deviation in both plots.

calculated. This distance is computed for each point on the mesh and the weighted average (according to the area) defines the point-to-mesh distance to the other mesh. The distance is calculated from model-based segmentation mesh to the ground-truth and vice versa to make the measurement symmetric.

Fig. 2 shows an example of automatic segmentation, compared with initial landmark-based placement of the model and manual segmentation. The model correctly finds boundaries of the myocardium, separating it even from other structures with the same intensity, such as papillary muscles. Fig. 3A plots the average segmentation error in different regions of the myocardium for each clinical cohort. Overall, the error is close to 1 mm, slightly less in the LV and slightly higher in the RV. Errors are slightly higher in presence of pathology than in controls. Sensitivity to initialization is illustrated in Fig. 3B, showing that the method tolerates small errors in initial model placement.

Table 1. Errors (\pm s.d. if known) reported in prior cardiac MR segmentation work

Method	Resolution (mm ³)	LV (mm)	RV (mm)	EPI (mm)
Mitchell '02 [11] ^{p,c}	$1.56 \times 1.56 \times 9$	$2.75 \pm 0.86^*$	-	$2.63 \pm 0.76^\dagger$
Lotjonen '04 [10] ^c	$1.0 \times 1.0 \times \text{variable}^{**}$	2.01 ± 0.31	2.37 ± 0.50	$2.77 \pm 0.49^\ddagger$
van Assen '08 [1] ^c	$1.5 \times 1.5 \times 10$	1.72	-	1.55^\dagger
Zhuang '08 [17] ^c	$2 \times 2 \times 2$	2.4 ± 1.1	2.6 ± 1.5	$1.3 \pm 0.21^\dagger$
Jolly '09 [6]	$1.25 \times 1.25 \times 8$	2.26	-	1.97^\dagger
Peters '10 [13] ^p	$0.6 \times 0.6 \times 0.8$	0.69^{***}	0.74	0.83^\dagger
Our approach ^{p,c}	$1.56 \times 1.56 \times 8$	0.87 ± 0.23	1.19 ± 0.28	$0.98 \pm 0.23^\ddagger$

^p Evaluation in patient data. ^c Evaluation in control data. ^{*}Distance measured in 2D. ^{**}Short and long axis slices. ^{***}Error measured as surface to surface distance. [†]LV epicardium only. [‡]LV+RV epicardium.

4 Discussion and Conclusions

Our results show that the automatic segmentation method is rather accurate relative to manual segmentation, with errors on the order of 1 mm. Table 1 shows that the accuracy of our method compares favorably with recently published boundary-based segmentation cardiac techniques on similar data, although one should use caution when comparing techniques that use different MRI sequences, different segmentation protocols and different error criteria. Even when the landmarks used for initialization are randomly displaced, our method achieves very good accuracy. Additional experiments not detailed here show that the MRF thickness prior is largely responsible for the strong performance of our method, as without this prior, the accuracy degrades substantially. Such a prior is made possible by the adoption of medial modeling.

As any complex Bayesian segmentation approach, our method requires many parameters to be set, including the relative weights of different prior terms, numbers and types of features used to train the appearance model, range of the local search, etc. In future work, it will be essential to measure sensitivity of the method to these parameters. Furthermore, it is critical to estimate how well the method generalizes to cardiac MRI from different MRI scanners and cohorts.

In conclusion, we have presented and evaluated a novel technique for biventricular myocardial segmentation. To our knowledge, it is the first such technique to use medial models and to leverage thickness priors that such models enable. The thickness prior, combined with established techniques from the boundary-based cardiac segmentation literature result in excellent accuracy relative to manual segmentation.

Acknowledgments. The project described was supported by the US National Institutes of Health (K25 AG027785, R21 NS061111), by the Spanish Ministry of Innovation and Science (CEN-20091044, TIN2009-14536-C02-01), and the European Commission Seventh Framework Programme (FP7/2007-2013) under grants agreement n.224495 (euHeart Project) and n.223920 (VPH NoE).

References

1. van Assen, H.C., Danilouchkine, M.G., Dirksen, M.S., Reiber, J.C., Lelieveldt, B.F.: A 3-D active shape model driven by fuzzy inference: Application to cardiac CT and MR. *IEEE T. Inf. Technol. B* 12(5), 595–605 (2008)
2. Blum, H., Nagel, R.: Shape description using weighted symmetric axis features. *Pattern Recognit.* 10(3), 167–180 (1978)
3. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models - their training and application. *Comput. Vis. Image Underst.* 1(61), 38–59 (1994)
4. Cristinacce, D., Cootes, T.: Facial feature detection using adaboost with shape constraints. In: *British Machine Vision Conference*, vol. 1, pp. 231–240 (2003)
5. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* 55(1), 119–139 (1997)
6. Jolly, M.P., Xue, H., Grady, L., Guehring, J.: Combining registration and minimum surfaces for the segmentation of the left ventricle in cardiac cine MR images. *Med. Image Comput. Comput. Assist Interv.* 12(Pt 2), 910–918 (2009)
7. Koenderink, J.J., van Doorn, A.J.: Representation of local geometry in the visual system. *Biol. Cybern.* 55(6), 367–375 (1987)
8. Loop, C., DeRose, T.: Generalized b-spline surfaces of arbitrary topology. In: *Computer Graphics (ACM SIGGRAPH Proceedings)*, pp. 347–356 (1990)
9. Lorenzo-Valdés, M., Sanchez-Ortiz, G.I., Elkington, A.G., Mohiaddin, R.H., Rueckert, D.: Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm. *Med. Image Anal.* 8(3), 255–265 (2004)
10. Lotjonen, J., Kivisto, S., Koikkalainen, J., Smutek, D., Lauerma, K.: Statistical shape model of atria, ventricles and epicardium from short- and long-axis MR images. *Med. Image Anal.* 8(3), 371–386 (2004)
11. Mitchell, S.C., Bosch, J.G., Lelieveldt, B.P.F., van der Geest, R.J., Reiber, J.H.C., Sonka, M.: 3-D active appearance models: segmentation of cardiac MR and ultrasound images. *IEEE T. Med. Imaging* 21(9), 1167–1178 (2002)
12. Paragios, N.: A Variational Approach for the Segmentation of the Left Ventricle in Cardiac Image Analysis. *Int. J. Comput Vision* 50(3), 345–362 (2002)
13. Peters, J., Ecabert, O., Meyer, C., Kneser, R., Weese, J.: Optimizing boundary detection via simulated search with applications to multi-modal heart segmentation. *Med. Image Anal.* 14(1), 70–84 (2010)
14. Sun, H., Avants, B.B., Frangi, A.F., Sukno, F., Gee, J.C., Yushkevich, P.A.: Cardiac medial modeling and time-course heart wall thickness analysis. *Med. Image Comput. Comput. Assist Interv.* 11(Pt. 2), 766–773 (2008)
15. Yushkevich, P.A., Zhang, H., Gee, J.: Continuous medial representation for anatomical structures. *IEEE T. Med. Imaging* 25(2), 1547–1564 (2006)
16. Zeng, X., Staib, L., Schultz, R., Duncan, J.: Volumetric layer segmentation using coupled surfaces propagation. In: *Proc. CVPR IEEE*, pp. 708–715 (1998)
17. Zhuang, X., Rhode, K., Arridge, S., Razavi, R., Hill, D., Hawkes, D., Ourselin, S.: An atlas-based segmentation propagation framework locally affine registration–application to automatic whole heart segmentation. *Med. Image Comput. Comput. Assist Interv.* 11(Pt 2), 425–433 (2008)

Automatic Aorta Segmentation and Valve Landmark Detection in C-Arm CT: Application to Aortic Valve Implantation

Yefeng Zheng¹, Matthias John², Rui Liao¹, Jan Boese², Uwe Kirschstein²,
Bogdan Georgescu¹, S. Kevin Zhou¹, Jörg Kempfert³,
Thomas Walther⁴, Gernot Brockmann⁵, and Dorin Comaniciu¹

¹ Siemens Corporate Research, Princeton, USA

² Siemens AG, Healthcare Sector, Forchheim, Germany

³ Department of Cardiac Surgery, Heart Center, University of Leipzig, Germany

⁴ Department of Cardiac Surgery, Kerckhoff Heart Center, Bad Nauheim, Germany

⁵ Department of Cardiovascular Surgery, German Heart Center, Munich, Germany

yefeng.zheng@siemens.com

Abstract. C-arm CT is an emerging imaging technique in transcatheter aortic valve implantation (TAVI) surgery. Automatic aorta segmentation and valve landmark detection in a C-arm CT volume has important applications in TAVI by providing valuable 3D measurements for surgery planning. Overlaying 3D segmentation onto 2D real time fluoroscopic images also provides critical visual guidance during the surgery. In this paper, we present a part-based aorta segmentation approach, which can handle aorta structure variation in case that the aortic arch and descending aorta are missing in the volume. The whole aorta model is split into four parts: aortic root, ascending aorta, aortic arch, and descending aorta. Discriminative learning is applied to train a detector for each part separately to exploit the rich domain knowledge embedded in an expert-annotated dataset. Eight important aortic valve landmarks (three aortic hinge points, three commissure points, and two coronary ostia) are also detected automatically in our system. Under the guidance of the detected landmarks, the physicians can deploy the prosthetic valve properly. Our approach is robust under variations of contrast agent. Taking about 1.4 seconds to process one volume, it is also computationally efficient.

1 Introduction

Affecting 1.8% of the global population and 10.7% of persons older than 65, aortic valve disease is the most common valvular disease in developed countries [1]. Implantation of a prosthetic aortic valve is often necessary to replace the severely damaged native valve. Though open-chest valve surgery is a well established procedure, minimally invasive transcatheter aortic valve implantation is an emerging technique, especially for high-risk patients, to minimize the surgical trauma. Before the surgery, several important parameters of the aortic valve (as shown in Fig. 1) need to be extracted for surgery planning. For example, the diameter of aortic valve annulus needs to be measured to select a

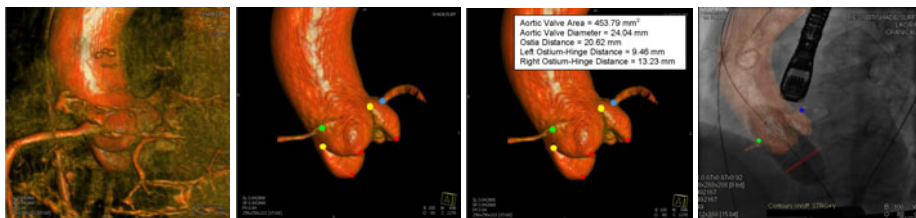


Fig. 1. Applications of C-arm CT to transcatheter aortic valve implantation. **Left:** A C-arm CT volume. **Left Middle:** Automatically segmented aorta together with the detected valve landmarks. **Right Middle:** 3D geometric measurements of the valve. **Right:** Overlay of the segmented aorta onto a 2D fluoroscopic image for visual guidance during surgery.

prosthetic valve with an appropriate size. During the surgery, 2D fluoroscopic images are captured in real time in a C-arm system to provide guidance to physicians [2]. The aortic root structure is distinguishable from the background only during a short period when the contrast agent is applied. However, the contrast agent is toxic and its usage should be minimized. Computed tomography (CT) is often used to provide the necessary 3D geometric measurements in surgery planning. However, CT is rarely used during valve implantation surgery because 2D/3D overlay (or registration) of data captured from different imaging devices is quite difficult. Recently, C-arm CT emerges as a new imaging technique with the following advantages, compared to conventional CT. Since both the 3D volume and 2D fluoroscopic images are captured on the same device within a short time interval, overlay of the 3D patient-specific aorta model onto a 2D image is straightforward and accurate (except patient motion). Besides providing visual guidance, the extracted aortic root can predict the best C-arm angulation (the optimal orientation of the imaging plane) to mitigate the foreshortening effect. For more details on the clinical applications of C-arm CT, please refer to [2,3].

A fully automatic system of aorta segmentation and valve landmark detection pays a key role in seamlessly integrating C-arm CT into the TAVI workflow. There are only a few methods proposed in literature to segment the aorta in a 3D volume. Zhao *et al.* [4] proposed a semi-automatic method to segment aorta in MR images. A user needs to manually select a seed point to initialize the fast marching method, which generates a rough segmentation result. Graph cut is exploited for final boundary refinement. Rueckert *et al.* [5] presented a tracking based segmentation method for the ascending and descending aortas. The intersection of the ascending/descending aorta with the image slice is roughly a circle, which is easy to track along slices. However, the curved aortic arch cannot be handled elegantly. Since the aorta is a tubular structure, many generic tubular structure detection approaches [6,7,8] can be extended to detect and segment it. Automatic aortic valve landmark detection is a new topic with very few publications in literature. Ionasec *et al.* [9] presented a comprehensive aortic valve model, which included the important valve landmarks, *e.g.*, hinge points, commissure points, and coronary ostia. All the previous work focuses on a

relatively consistent imaging protocol with much fewer variations than ours. For example, a roughly same portion of the aorta is captured in the volume and the usage of contrast agent is consistent, resulting in stable image characteristics. Furthermore, most of the previous approaches are semi-automatic (a user needs to click at least one point [4]) and very slow (taking up to 450 seconds to process one volume [7]).

All of the previous methods work on the well established imaging modalities, such as MR and CT. However, automatic segmentation of the aorta in a C-arm CT volume is far more challenging. First, the image quality from different clinical sites varies quite a lot since C-arm CT is too new to have a well accepted scanning protocol. We also observed significant variations inside the same clinical site since physicians were testing different scanning parameters (*e.g.*, the amount of contrast agent and timing of the image acquisition). Conventional image processing techniques, *e.g.*, intensity-based thresholding, region growing, and the watershed method, are not robust under such large variations. We propose to use machine learning techniques to exploit the rich information embedded in an expert-annotated dataset. Second, the field of view varies quite a lot for a C-arm CT volume. For example, the aortic arch and descending aorta may be captured in some volumes (see the first two examples in Fig. 2), but missing in others (see the last two examples in Fig. 2). To address this challenge, we propose a part-based aorta model. As shown in Fig. 3, the whole aorta is split into four parts: aortic root, ascending aorta, aortic arch, and descending aorta. Using the part-based model, the whole aorta does not need to be fully present.

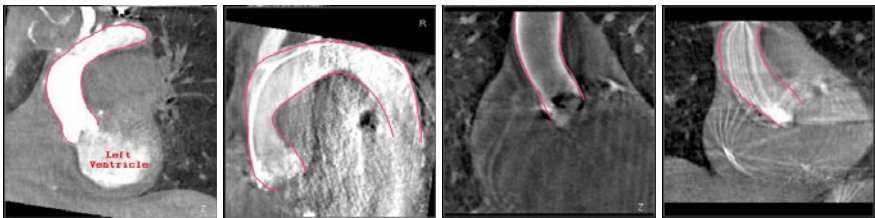


Fig. 2. Automatic aorta segmentation on a few example volumes. **Left:** Good contrast, however, with severe valve regurgitation. **Left Middle:** Fair image quality. **Right Middle:** Contrast agent is almost washed out due to bad timing. **Right:** Streak artifacts generated by the catheters.

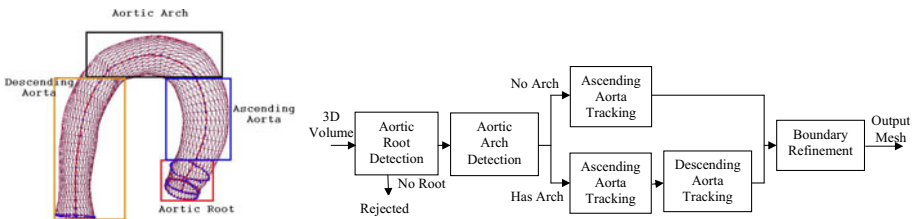


Fig. 3. Part-based aorta model (left) and automatic segmentation workflow (right)

Depending on the structure that can be detected, different workflows can be exploited, therefore, a large structure variation can be handled elegantly.

2 Part-Based Aorta Modeling and Segmentation

Due to the variation in the field of view, the aorta captured in a C-arm CT volume changes a lot in its structure. In this paper, we present a part-based aorta model (as shown in Fig. 3) by splitting the whole aorta into four parts: aortic root, ascending aorta, aortic arch, and descending aorta. The aortic root is required to be present in this application, therefore, it is detected and segmented as the first step. To be specific, we use the recently proposed marginal space learning (MSL) method [10] to segment the aortic root. MSL is an efficient method to detect and segment a 3D anatomical structure in medical images based on a discriminative machine learning technique. It is robust and works for different imaging modalities. Due to the space limit, we would like to refer readers to [10] for more details of MSL. As shown in the system diagram in Fig. 3, the aortic root is detected first. If no aortic root is detected, the input volume is rejected. We then detect the aortic arch. Similarly, MSL is exploited to train a separate detector for the aortic arch. For about half of the volumes in our dataset, the aortic arch may be out of the field of view. If no aortic arch is present, normally the descending aorta is also missing in the volume (see the last two cases in Fig. 2).

The length of the ascending and descending aortas captured in a volume varies significantly. It is difficult to detect them as whole objects. We propose to use a tracking technique to deal with this variation. Since the intersection of the ascending and descending aortas with an image slice is close to a circle, we train a 2D circle detector using Haar wavelet features and the boosting learning algorithm [10] to detect aortic circles as primitive structures for tracking. Starting from the aortic root, we detect an aortic circle on the next slice (toward the patient's head). The detector outputs multiple circle candidates around the true position. We pick the one closest to the circle on the current slice. If the aortic arch is detected in the volume, the tracking procedure stops on the slice touching the aortic arch. Otherwise, it stops when no aortic circle is detected or it reaches the top volume border. Tracking of the descending aorta is similar except that it starts from the aortic arch and moves toward the patient's toe. It stops on the slice with no aortic circle detected.

Assembling all the aortic parts together (the tracked aortic circles, aortic root, and aortic arch if it is present), we get an initial surface mesh of the aorta. The initialization is close to the true aorta boundary, however, a circle does not fit the boundary exactly. A learning based boundary detector is applied for final boundary delineation. Specifically, a two-step iterative approach is used. 1) Use the learning-based boundary detector to adjust each mesh point along the surface normal to the optimal position where the response of the boundary detector is the largest. 2) Apply generic mesh smoothing [11] to get a smooth surface. The above two steps repeat a few iterations to improve the boundary delineation accuracy.

MSL [10] can efficiently detect an object as a whole. However, it cannot deal with structural variations. Therefore, almost all previous work uses bottom-up approaches [4,5,6,7,8] to track the aorta centerline to handle the variations. They are neither automatic nor robust on noisy images. In comparison, we use MSL to detect the aortic root and arch, and use bottom-up tracking to detect ascending/descending aortas that have large variations in length. Our system is a nice combination of both approaches.

3 Aortic Valve Landmark Detection

Besides segmenting the aorta, we also detect eight aortic valve landmarks: three aortic hinge points, three aortic commissure points, and left and right coronary ostia since they are important in both surgery planning and providing visual guidance during surgery [3]. Though it is possible to detect each landmark independently, the detection results may be inconsistent in geometry. It also wastes computation power by ignoring the strong geometric constraint among the landmarks. We propose to use a hierarchical approach by first detecting a global object comprised with all eight valve landmarks. From the position, orientation, and scale of this global object, we can infer the rough position of individual landmarks. Each landmark is then refined in a small region (*e.g.*, a cube of 20 mm centered on the initial position) under the guidance of its own specific landmark detector.

Similar to the aortic root detection, we use marginal space learning (MSL) [10] to efficiently detect the position, orientation, and scales of the global landmark object. For a learning based method, we need to specify the ground truth of object pose for each training volume, therefore, a learning algorithm can learn the implicit relationship to infer the correct pose from an unseen volume. However, there is no standard way to define the pose of the global object containing eight landmarks. After detecting the global landmark object, we align the mean shape (which is the average shape of the training set after global transformation has been compensated) to the global pose to get an initial estimate of each individual landmark's position (see Eq. (2)). In this paper, we propose a method to search for an optimal shape which can represent the whole shape population accurately, therefore improving the initialization accuracy of the landmarks. Given a group of shapes, M_1, M_2, \dots, M_N , we want to find an optimal shape \bar{m} to represent the whole population such that it can minimize the residual errors after alignment,

$$\bar{m} = \arg \min_m \sum_{i=1}^N \|\mathcal{T}_i(m) - M_i\|^2. \quad (1)$$

The optimal shape \bar{m} is called the mean shape in this paper. \mathcal{T}_i is the corresponding transformation from the mean shape \bar{m} to each individual shape M_i . This procedure is called generalized Procrustes analysis in statistical shape analysis [12].

Previously, the generalized Procrustes analysis is only performed under the similarity transformation (*i.e.*, \mathcal{T} is a similarity transformation). MSL can estimate anisotropic scales quite efficiently. With more deformation compensated, the mean shape is more accurate to represent the whole shape population. Therefore, in our approach \mathcal{T} represents translation ($T = [X, Y, Z]'$), rotation (represented as a rotation matrix R), and anisotropic scaling (S_x, S_y, S_z). The transformation of a 3D point P is

$$\mathcal{T}(P) = R \begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & S_z \end{bmatrix} P + T. \quad (2)$$

To the best of our knowledge, there are no closed-form solutions for estimating the anisotropic similarity transformation. We propose an iterative algorithm. We first estimate the similarity transformation (translation, rotation, and isotropic scaling), which has closed-form solutions [12]. After compensating the similarity transformation, we estimate the three anisotropic scaling parameters (S_x, S_y, S_z), for which we derive a closed-form solution. With a module solving the anisotropic similarity transformation between two shapes, we can plug it into the generalized Procrustes analysis method to search for the optimal mean shape \bar{m} . Besides the optimal mean shape, the transformation \mathcal{T}_i of the mean shape to each example shape M_i is also calculated as a by-product, which provides the pose ground truth that MSL can learn to estimate.

4 Experiments

A dataset of 192 C-arm CT volumes from 152 patients were collected from two clinical sites to evaluate our method. The size of each slice in a volume is 256×256 or 512×512 pixels. A volume contains around 200-300 slices. The image resolution is isotropic and varies from volume to volume in the range of $[0.70, 0.84]$ mm.

A four-fold cross-validation is performed to evaluate our algorithm. The aorta segmentation accuracy is measured using the symmetric point-to-mesh distance E_{p2m} [10]. The mean segmentation error of the aorta is 1.1 mm, with a standard deviation of 0.41 mm. We cannot compare our error with those reported in the literature directly because they used different datasets captured from different imaging modalities. Roughly, our accuracy is comparable to (or better than) the state-of-the-art, *e.g.*, 1.55 mm mean error reported in [4] on 21 MR datasets and 1.4 mm mean error reported in [7] on 23 CT datasets. Fig. 2 shows aorta segmentation results on a few volumes. The first example shows a volume with good image quality. However, due to severe valve regurgitation the contrast leaks into the left ventricle. The third example shows a case with bad image acquisition timing where the contrast agent has almost been washed out. Severe streak artifacts generated by catheters are clearly visible in the last volume.

In the following experiment, we evaluate the valve landmark detection accuracy. There are a total of 28 volumes with extremely poor image quality that

the landmarks cannot be identified even by an expert (though the aorta can be successfully segmented from these volumes). So these images are excluded. A four-fold cross-validation is performed on the remaining 164 volumes for aortic valve landmark detection. The landmark detection accuracy is measured using the Euclidean distance from the detected landmark to the ground truth. Table 1 shows the detection errors. After global landmark object pose estimation, we can get a good initial estimate of the landmark position. The mean errors range from 5.40 to 5.81 mm for different landmarks. After local refinement for each landmark, the error is further reduced. For example, the mean error for the aortic hinges reduces from 5.40 mm to 2.41 mm. Fig. 4 shows the detected valve landmarks in two typical volumes. Our approach is computationally efficient, taking about 1.4 seconds to process a volume on a computer with 3.2 GHz duo-core processors and 3 GB memory. It is at least 10 times faster than the previous methods [5,7].

Table 1. Aortic valve landmark detection errors based on a four-fold cross-validation on 164 volumes. The mean, standard deviation (STD), and median of the errors (measured in millimeters) are reported.

	Aortic Hinges			Coronary Ostia			Aortic Commissures		
	Mean	STD	Median	Mean	STD	Median	Mean	STD	Median
After Global Pose Estimation	5.40	2.51	4.83	5.81	2.45	5.18	5.43	2.26	5.20
After Local Refinement	2.41	1.50	1.90	2.74	2.43	1.77	3.46	1.78	3.11

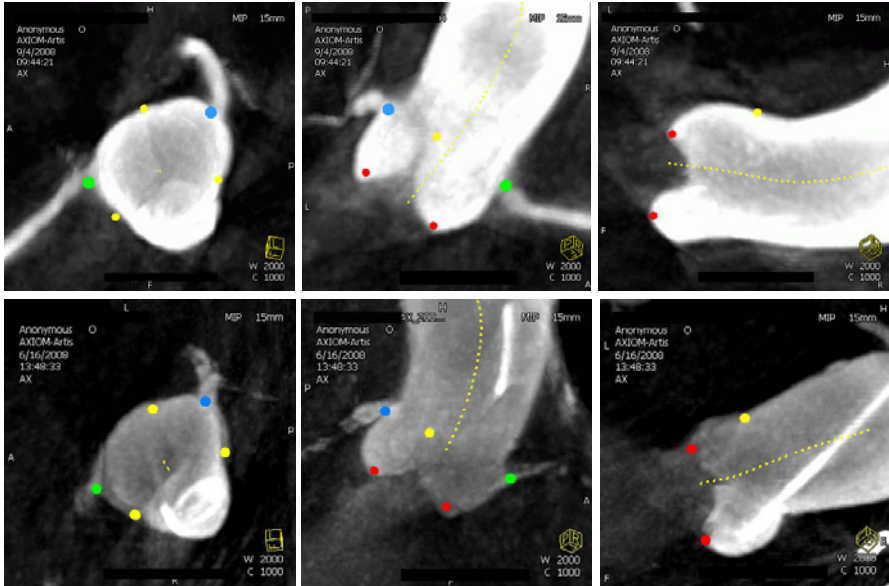


Fig. 4. The aortic valve landmark detection results on two example datasets with red dots for the hinge points, yellow for the commissure points, blue for the left coronary ostium, and green for the right coronary ostium. Each row shows three orthogonal cuts of a volume.

5 Conclusion

In this paper, we presented a fully automatic aorta segmentation and valve landmark detection system in C-arm CT with applications to transcatheter aortic valve implantation (TAVI). The initial clinical trial demonstrated the usefulness of our system in the TAVI workflow, *e.g.*, providing a proper angulation to avoid large tilting of a prosthetic valve after deployment [3]. Our approach is generic, therefore can be extended easily to other imaging modalities by simple retraining, without any manual parameter tuning. A similar system has been built on cardiac CT datasets to use conventional CT for surgery planning.

References

1. Nkomo, V.T., Gardin, J.M., Skelton, T.N., Gottdiener, J.S., Scott, C.G., Enriquez-Sarano, M.: Burden of valvular heart diseases: a population-based study. *Lancet* 368(10), 1005–1011 (2006)
2. Gessat, M., Merk, D.R., Falk, V., Walther, T., Jacobs, S., Nöttling, A., Burgert, O.: A planning system for transapical aortic valve implantation. In: *Proc. of SPIE Medical Imaging* (2009)
3. John, M., Liao, R., Zheng, Y., Nöttling, A., Boese, J., Kirschstein, U., Kempfert, J., Walther, T.: System to guide transcatheter aortic valve implantations based on interventional 3D C-arm CT imaging. In: *Proc. Int'l. Conf. Medical Image Computing and Computer Assisted Intervention* (2010)
4. Zhao, F., Zhang, H., Wahle, A., Scholz, T.D., Sonka, M.: Automated 4D segmentation of aortic magnetic resonance images. In: *Proc. British Machine Vision Conference* (2006)
5. Rueckert, D., Burger, P., Forbat, S.M., Mohiaddin, R.D., Yang, G.Z.: Automatic tracking of the aorta in cardiovascular MR images using deformable models. *IEEE Trans. Medical Imaging* 16(5), 581–590 (1997)
6. Behrens, T., Rohr, K., Stiehl, H.H.: Robust segmentation of tubular structures in 3-D medical images by parametric object detection and tracking. *IEEE Trans. Syst., Man, Cybern. B* 33(14), 554–561 (2003)
7. de Bruijne, M., van Ginneken, B., Viergever, M.A., Niessen, W.J.: Adapting active shape models for 3D segmentation of tubular structures in medical images. In: *Proc. Information Processing in Medical Imaging*, pp. 136–147 (2003)
8. Saur, S.C., Kühnel, C., Boskamp, T., Székely, G., Cattin, P.: Automatic ascending aorta detection in CTA datasets. In: *Bildverarbeitung für der Medizin 2008*, pp. 323–327 (2008)
9. Ionasec, R., Georgescu, B., Gassner, E., Vogt, S., Kutter, O., Scheuring, M., Navab, N., Comaniciu, D.: Dynamic model-driven quantification and visual evaluation of the aortic valve from 4D CT. In: *Proc. Int'l. Conf. Medical Image Computing and Computer Assisted Intervention*, pp. 686–694 (2008)
10. Zheng, Y., Barbu, A., Georgescu, B., Scheuring, M., Comaniciu, D.: Four-chamber heart modeling and automatic segmentation for 3D cardiac CT volumes using marginal space learning and steerable features. *IEEE Trans. Medical Imaging* 27(11), 1668–1681 (2008)
11. Taubin, G.: Optimal surface smoothing as filter design. In: *Proc. European Conf. Computer Vision*, pp. 283–292 (1996)
12. Dryden, I.L., Mardia, K.V.: *Statistical Shape Analysis*. John Wiley, Chichester (1998)

Lesion-Specific Coronary Artery Calcium Quantification for Predicting Cardiac Event with Multiple Instance Support Vector Machines

Qingshan Liu¹, Zhen Qian², Idean Marvasty², Sarah Rinehart², Szilard Voros²,
and Dimitris N. Metaxas¹

¹ Rutgers University, Piscataway, NJ, 08854 USA

² Piedmont Heart Institute, Atlanta, GA, 30309, USA

Abstract. Conventional whole-heart CAC quantification has been demonstrated to be insufficient in predicting coronary events, especially in accurately predicting near-term coronary events in high-risk adults [1]. In this paper, we propose a lesion-specific CAC quantification framework to improve CAC's near term predictive value in intermediate to high-risk populations with a novel multiple instance support vector machines (MISVM) approach. Our method works on data sets acquired with clinical imaging protocols on conventional CT scanners without modifying the CT hardware or updating the imaging protocol. The calcific lesions are quantified by geometric information, density, and some clinical measurements. A MISVM model is built to predict cardiac events, and moreover, to give a better insight of the characterization of vulnerable or culprit lesions in CAC. Experimental results on 31 patients showed significant improvement of the predictive value with the ROC analysis, the net reclassification improvement evaluation, and the leave-one-out validation against the conventional methods.

1 Introduction

Atherosclerosis is the leading cause of morbidity and mortality worldwide. Later stages of atherosclerosis are characterized by progressive deposition of calcium in the coronary arterial vessel-wall. CT-based CAC scanning is a three dimensional imaging technique that efficiently quantifies calcium in the coronary vasculature in a non-invasive, low-radiation way. CT-based CAC measurements are widely used clinically for cardiac disease diagnosis and treatment planning.

Presently the popular CAC measurements are the whole-heart Agatston score (AS) [2] and the whole-heart volume score (VS) [3]. The AS is calculated by multiplying the area of calcification by an arbitrary weighted density score based on the maximum Hounsfield unit (HU) value. The VS is defined as the total volume of the calcification in the major epicardial arteries, where calcification is identified as voxels with $HU \geq 130$. The whole-heart AS/VS have been validated to be independent of, and additive to, the Framingham Risk Score (FRS) in predicting major cardiovascular events [4]. Recently, a coronary calcium coverage score

was proposed to quantify the percentage of coronary arteries affected by calcific plaque, but it was not superior to the whole-heart AS/VS in prediction of hard cardiac events [5]. Some standard statistical methods, such as Cox regression and Kaplan-Meier estimator, are widely used to evaluate predictive values of coronary events on relatively long term data sets [6,7], and a very significant conclusion has been made: the absolute CAC measure is more predictive of coronary events than other regular clinical measurements. However, the information obtained by the relatively long-term studies is limited to individual patient treatment. In clinics, near-term event prediction is more desirable for treatment planning in intermediate to high-risk patients. In [1], a comparison study has shown that the whole heart AS/VS CAC measurements does not accurately predict near-term future coronary events in high-risk adults.

In this paper, we aim at improving CAC's near-term predicative value of major adverse cardiac events (MACE) including death, nonfatal myocardial infarction, and revascularization, in intermediate to high-risk population by using lesion-specific CAC quantification. Besides the whole heart AS/VS, a three dimensional CAC volume contains other clinically significant relevant information, such as the number of plaque lesions and geometric characteristics of lesions. These lesion-specific CAC evaluations will be more helpful for near-term prediction of MACE than the conventional whole heart CAC. For instance, the decision making of revascularization is largely based on the severity and location of the arterial stenosis. [8] showed that distance to corresponding ostium was a good indicator in analyzing cardiac events. However, different patients have different numbers of calcific lesions with different lesion distributions. MACE is generally caused or induced by one or a few culprit or vulnerable lesions, but it is hard to discriminate such risky lesions without further examinations. Thus, conventional pattern-level supervised learning approaches can not work well.

To handle these issues, we proposed to use multiple instance support vector machines (MISVM) [9] to estimate the predictive value of CAC scan for the incidences of MACE. The idea of MISVM is to perform a soft-margin based SVM on the unobserved instances, subjected to constrains defined by the labeled bags. The bag is labeled as positive if at least one instance in a bag is positive. Otherwise a bag is negative. We model each patient's lesion-specific CAC quantification results as a bag of instances, in which each lesion is taken as an instance. While each instance may possess an associated true label, i.e., leading to a event (positive) or not leading to a event (negative), the labels of the instances are only implicitly accessible through the labels associated with the bags. The MISVM model can not only predict the event risk efficiently, but also it can evaluate the culprit or vulnerable lesions. We assessed our method on 31 patients' data including 11 patients with MACE and 20 patients without MACE all within 2 years after the CAC scan. Experimental results showed it significantly improved the prediction accuracy with the ROC analysis, the net reclassification improvement evaluation, and the leave-one-out testing compared to the conventional methods including the whole AS/VS [2,3] and FRS [4].

2 Methodology

2.1 Lesion-Specific CAC Quantification

Although the whole-heart CAC scores can indicate the risk of cardiac disease, the lesion-specific details of the calcification have a potential to play an additive and indispensable role in determining the extent and severity of the coronary disease. In this paper, we aim to extract lesion-specific detailed information for CAC event analysis. Our method works on data sets acquired with typical clinical imaging protocols on conventional CT scanners without modifying the CT hardware or updating the imaging protocol.

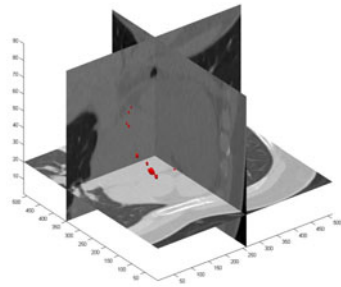
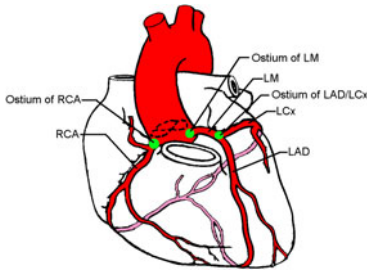


Fig. 1. The major coronary arteries and the arterial ostia **Fig. 2.** Calcific lesion segmentation result

We only take account of 4 major coronary arteries: the left main coronary artery (LM), the left anterior descending artery (LAD), the left circumflex artery (LCx) and the right coronary artery (RCA), as shown in Figure 1. A simple but efficient 3D 6-connectedness flood-fill operation is adopted to segment the lesions, which is defined as a group of calcific voxels that are 6-connected to each other, but not connected to any other calcific voxel in the binary calcium volume. The definition of 6-connectiveness is shown in Figure 3: (a) the blue voxel in the center is 6-connected to its 6 neighboring green voxels; (b) Voxel 1 is 6-connected with voxel 2 by a 6-connected path. Starting from the user-specified seed point in an epicardial arterial calcified lesion, the 3D lesion is easily segmented by propagating the 3D 6-connectedness flood-fill operations. Figure 2 shows the segmentation result of one sample. The locations of the major arterial ostia are identified manually in the 3D CAC volume.

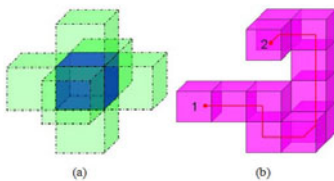


Fig. 3. 6-connectiveness in 3D volumes

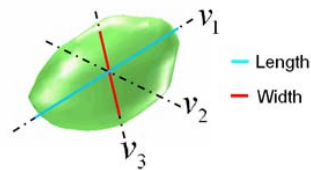


Fig. 4. The lesion's length and width

Suppose lesion \mathbf{L} has N voxels with physical coordinates p_i , $i = 1, 2, \dots, N$ and the same resolution of $r = (r_x, r_y, r_z)'$; The physical coordinates of the corresponding artery ostium is p_o . We quantify \mathbf{L} with a 10 dimensional feature covering geometrical information, density, and clinical measurement as:

Length and Width: Given the covariance matrix of the voxels in the lesion, we calculate its eigenvectors v_1, v_2, v_3 and corresponding eigenvalues $\lambda_1, \lambda_2, \lambda_3$. Assuming $\lambda_1 \geq \lambda_2 \geq \lambda_3$, we define the length as the maximum distance between any two lesion voxels along the direction of v_1 , and the width is the maximum distance between any two lesion voxels along the direction of v_3 , $L = \max_{i,j=1,2,\dots,N,i \neq j} |(p_i - p_j)' \cdot v_1|$ and $W = \max_{i,j=1,2,\dots,N,i \neq j} |(p_i - p_j)' \cdot v_3|$, respectively. Figure 4 shows an example.

VS: The lesion's volume score is calculated by, $VS = N \cdot r_x \cdot r_y \cdot r_z$. If the volume score is smaller than 1, we will treat it as noise and discard it.

Distance to Ostium: We also take account of the distance of the lesion to the ostium, and it is calculated by, $D = \|p_o - \bar{p}\|_2$, where $\bar{p} = (\sum_{i=1}^N p_i)/N$.

Neighborhood Relationship: A group of adjacent calcific lesions with spotty appearance are usually associated with higher event risks [10]. To describe such geometric inter-dependence between the lesions, we introduce a relationship feature for each lesion, defined as: $R_i = \frac{1}{\sigma\sqrt{2\pi}} \sum_j \exp(-(d_{ij})^2/(2\sigma^2))$, $j \neq i$, where the j -lesion is the neighbor of the i -lesion locating in the same artery; σ determines the neighborhood size; d_{ij} is the distance between two lesions.

Maximum HU and Average HU Values: HU value is widely to measure the density of CAC. Besides the maximum HU (MHU) value in the lesion, we also calculate the average HU (AHU) value to describe the lesion's density.

2D/3D AS: 2D AS is the popular clinical measurement used by the doctors. We also define 3D AS to measure the lesions, which is scored by the whole 3D volume and the highest HU value in the whole 3D lesion like the 2D AS scoring.

Artery Index: We should also indicate the artery where the lesion exists in. We label LM, LAD, LCx, and RCA by $Ind = 1, 2, 3, 4$, respectively.

2.2 MISVM Based Cardiac Event Prediction Model

As mentioned in the introduction, different patient's CAC scans contain different number of calcific lesions, which will result in different length of the patient-scan-based feature vector. MACE is generally caused or induced by one or a few culprit or vulnerable lesions, but it is hard to label culprit or vulnerable ones without further examinations. All these issues make conventional pattern-level supervised learning methods difficult to predict cardiac event.

In this paper, we developed a novel MISVM model to predict MACE using lesion-specific CAC quantifications. MISVM is an extension of the conventional SVM classifier, in which training class labels are associated with sets of samples (or bags), instead of individual samples (or instances) [9]. The labels of instances

are only indirectly associated with the labels of bags. If a bag is labeled as positive, it means that at least one instance inside the bag is positive. A bag is labeled as negative if and only if all the instances in the bag are negative. Besides MISVM does not need the label information of instances, the size of instances in each bag can be different, so MISVM can deal with the above issues well. We take each patient as a bag, and each calcific lesion as an instance. We label the patient with MACE as a positive bag and one without MACE as a negative bag.

For convenience, we define some notations first. The patient bag B_i has n_i instances (lesions), $B_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,n_i}\}$, and its label $Y_i = 1$ or -1 . The labels of the instances in the bag B_i as $\{y_{i,j}\} \in \{1, -1\}_{j=1,2,\dots,n_i}$, but they are unknown. $\{y_{i,j}\}$ is indirectly associated to Y_i following this constrain:

$$\sum_{j=1}^{n_i} \frac{y_{i,j} + 1}{2} \geq 1 : Y_i = 1, \text{ and } \forall j, y_{i,j} = -1 : Y_i = -1. \quad (1)$$

Similar to the conventional SVM, the goal of MISVM is to find a separating hyperplane (w, b) to make maximum margin between the positive instances and the negative instances, so as to efficiently separate the positive bags and the negative bags. The MISVM classifier can be defined as:

$$\begin{aligned} & \min_{\{y_{i,j}\}} \min_{w,b,\xi} \frac{1}{2} \|w\|^2 + C \sum_{i,j} \xi_{i,j} \\ \text{s.t. } & \forall i, y_{i,j} (< w, x_{i,j} > + b) \geq 1 - \xi_{i,j}, y_{i,j} \in \{1, -1\} \text{ and } \xi_{i,j} \geq 0 \\ & \forall i, Y_i = 1 : \sum_{j=1}^{n_i} \frac{y_{i,j} + 1}{2} \geq 1, Y_i = -1 : \forall j, y_{i,j} = -1 \end{aligned} \quad (2)$$

where $\xi_{i,j}$ are slack factors. This is a typical mixed integer programming problem, which is difficult to be solved directly, for it aims to find the optimal labeling of instances and the optimal hyperplane based on the bags. [9] presented a heuristic optimization to solve this problem as: (1) given the labels of the instances, solve the associated quadratic program (QP) as the conventional SVM; (2) use the solution of the QP to label the instances; (3) update some labels of the instances to follow the definition of bags. The iteration continues until the labels no longer change. Initially, the labels of the instances are set according to their associated bags' labels.

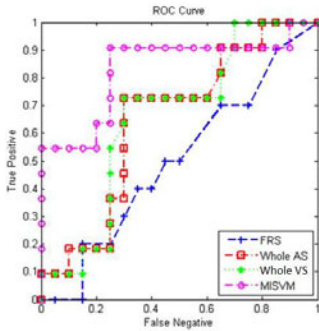
3 Experiments

The experimental data consisted of 31 patients including 11 patients with MACE and 20 patients without MACE in 2 years after the CAC scans. 10 MACE patients underwent invasive revascularization procedures, and another MACE patient died from a myocardium infarction. CAC examinations were performed on a 64 MDCT system (Siemens Somatom), and the images were acquired using 3mm collimation with 2mm inter-slice gap with non-contrast enhanced scans. Acquisition parameters included a gantry rotation of 330-375ms, pitch 0.24, tube voltage 120kV, and tube current of 250mAs.

3.1 MACE Prediction Result Analysis

We adopted three testing analysis methods to evaluate the performance of the proposed MISVM model: ROC analysis, the net reclassification improvement evaluation [11], and leave-one-out validation. We compared the proposed MISVM model with the whole AS/VS [2,3] and FRS [4]. Different from Cox Regression and Kaplan-Meier estimation that require longer follow-up [6,7], we focus on near-term estimation, so we do not consider them for comparison. The Gaussian kernel is adopted for the MISVM model, $k(x, y) = \exp(-\gamma \|x - y\|^2)$, and $\gamma = 0.65$. For two lesions in the same vessel, if their distance is less than $5mm$, we take them as neighbors and set $\sigma = 3$.

ROC Analysis: We first evaluate the proposed method by the ROC analysis. The ROC curves of the whole AS, the whole VS, and FRS were directly generated by an adaptive decision threshold search on the original values, and the ROC curve of MISVM was obtained by adaptively adjusting the decision values of SVM [12]. Figure 5 shows the comparison results, where the left is the ROC curves and the corresponding accuracy (AUC), standard error (S.E.), and 95% confidence interval (C.I.) are listed in the right. We can see MISVM obtains a better performance than the other three methods. The AUC of MISVM is 0.8318, while the whole-heart AS/VS and FRS are 0.65, 0.6636, and 0.5075 respectively. At the sensitivity level of 80%, which is commonly used as the clinical cut-off point for choosing preventive treatment, the specificity of MISVM is improved to over 75% compared to the specificities of about 35% by using the whole-heart AS/VS respectively. Because FRS is basically equal to guess from Figure 5, we will not discuss it in the following experiments.



	AUC	S. E.	95% C. I.
MISVM	0.8318	0.0839	0.6673-0.9962
Whole AS	0.65	0.107	0.4402-0.8598
Whole VS	0.6636	0.1061	0.4557-0.8716
FRS	0.5075	0.1138	0.2844-0.7306

Fig. 5. Comparison of the ROC analysis

The Net Reclassification Improvement Evaluation: The goal of MACE risk prediction is not to simply give an individual’s risk, but to help further diagnostic and therapeutic decision-making for improving clinical outcomes and lower cost. As suggested by the American Heart Association guideline, the net reclassification improvement (NRI) quantitatively assesses the improvement in risk prediction offered by the new marker by quantifying the correct movement

Table 1. Comparison of NRI

	MISVM vs. Whole AS	MISVM vs. Whole VS
NRI	76.36%	90.45%
p	0.001	0.0005

Table 2. Leave-One-Out Validation

	MISVM	Whole AS	Whole VS
α_{determ}	65.38%	66.7%	61.5%
η	16.13%	61.5%	58.1%
$\alpha_{overall}$	54.84%	25.8%	25.8%

in risk categories [13], in which the risky is divided into three categories: a safe category that is below 90% sensitivity cut-off point for conservative treatment, a high-risk category that is above the 90% specificity cut-off point for aggressive treatment, and an indeterminate category in between the two cut-off points for further diagnostic testing. The NRI-based experimental results are reported in Table 1. NRI of MISVM over both the whole AS/VS are 76.36% with $p = 0.001$ and 90.45% with $p = 0.0005$ respectively, which shows the significant improvement of MISVM to CAC's event predication accuracy.

Leave-One-Out Validation: Furthermore, we cross validated our method on a leave-one-out basis with the above safe, high-risk, and indeterminate categories. For diagnostic purposes, physicians will be interested in improving the predictive accuracy in the safe and high-risk categories, and reducing the proportion of patients in the indeterminate category. Thus, we define three measurements for evaluation: (1). the predictive accuracy in the determinate safe and high-risk categories, $\alpha_{determ} = (N_n^s + N_p^r)/(N_p^s + N_n^s + N_p^r + N_n^r)$; (2). the patient percentage in the indeterminate zone, $\eta = (N_t - N_p^s - N_n^s - N_p^r - N_n^r)/(N_t)$; (3). the overall predictive accuracy, $\alpha_{overall} = (N_n^s + N_p^r)/(N_t)$, where N_p^s and N_p^r are the number of safe and high-risky categories in the positives respectively, N_n^s and N_n^r are the numbers of safe and high-risky categories in the negatives, and N_t is the number of the total patients.

Table 2 lists the results of the leave-one-out validation. We can see that the proposed MISVM achieves similar α_{determ} accuracy level compared to the whole AS/VS (65.38%, 66.7%, and 61.5% respectively) in the safe and high-risky categories, while the undecidable patient percentages in the indeterminate category are largely decreased from 61.5% and 58.1% to 16.13%. MISVM improves the overall predictive accuracy from 25.8% to 54.84%.

3.2 Lesion-Level Risky Analysis

As described in Section 2.2, MISVM can also discriminate whether or not the patient has lesions of event risk. Table 3 reports the statistical mean value of safe and high-risky lesions predicted by MISVM. We can see some differences that are in good accordance with clinical observations. For instance, the high-risky lesions are statistically much larger than the safe lesions in VS, AS, length, and HU values. Although the statistics of the lesion distance seem inconsistent with clinical observation, we find that many small non-risky lesions with small distances to the ostia could be a reason of this inconsistency.

Additionally, we found that most high-risky lesions are from LAD. It is consistent to clinical observations: LAD is a major coronary artery that supplies

Table 3. Overall comparison between the risky and safe lesions

	VS	2D AS	3D AS	Length	Width	Distance	MHU	AHU
Risky	136.38	180.29	173.12	13.46	1.08	49.68	507.1	235.45
Safe	29.37	36.43	33.18	5.64	0.81	28.23	351.5	196.63

the most important myocardial territory. Revascularization is more likely performed in LAD than the other coronary arteries [14]. For the MACE samples, we also found about 85% high-risk lesion with neighbors and 15% high-risk lesions without neighbors. Figure 6 (a) and (b) show two representative samples that were correctly predicted as MACE. The lesions in red were classified as risky, while the lesions in green were non-risky. The risky lesions predicted in (a) are relatively larger lesions with high AS and VS. The risky lesions predicted in (b) are spotty calcification. In addition, Figure 6 (c) shows a representative non-event sample that was incorrectly classified, with the red risky lesions. It is interesting to see that these lesions have both the characteristics of large size and spotty calcification, which suggests that CAC scan might not be a complete indicator of MACE. In future work, we plan to include more clinical cues, such as patients' age, gender, history of smoking, hypertension, and diabetics, and biomarkers from blood tests, to develop a more accurate event predictor.

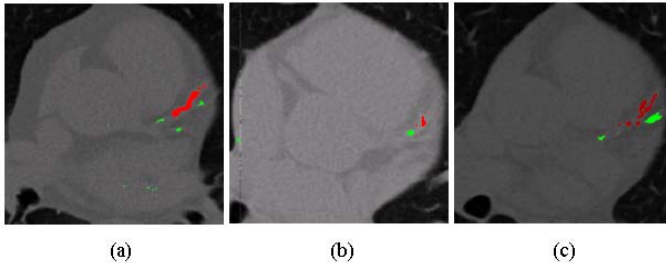


Fig. 6. Representative samples. (a) and (b) are two positive samples that are correctly classified; (c) is a negative sample that is incorrectly classified as positive.

4 Conclusion

In this paper, we proposed a MISVM based lesion-specific CAC quantification framework to improve the event prediction value of CAC scanning. Different from the conventional whole-heart CAC scores, it took account of a more comprehensive lesion-specific CAC information from routine CAC scans, and achieved better prediction sensitivity and specificity. Besides event prediction, it can also be used to characterize culprit or vulnerable calcific lesions, and evaluate their associated event risks. Experimental results demonstrated that it has the potential to significantly improve the clinical predictive value of CAC.

References

1. Detrano, R., et al.: Coronary calcium does not accurately predict near-term future coronary events in high-risk adults. *Circulation* 99(20), 2633–2638 (1999)
2. Agatston, A., Janowitz, W., et al.: Quantification of coronary artery calcium using ultrafast computed tomography. *J. Am. Coll. Cardiol.* 15(4), 827–832 (1990)
3. Callister, T., Cooil, B., Raya, S., et al.: Coronary artery disease: improved reproducibility of calcium scoring with an electron-beam ct volumetric method. *Radiology* 208(3), 807–814 (1998)
4. Greenland, P., LaBree, L., Azen, S., et al.: Coronary artery calcium score combined with framingham score for risk prediction in asymptomatic individuals. *Journal of the American Medical Association* 291, 210–215 (2004)
5. Brown, E., Kronmal, R., et al.: Coronary calcium coverage score: determination, correlates, and predictive accuracy in the multi-ethnic study of atherosclerosis. *Radiology* 247(3), 669–675 (2008)
6. Lakoski, S., et al.: Coronary artery calcium scores and risk for cardiovascular events in women classified as "low risk" based on framingham risk score: The multi-ethnic study of atherosclerosis. *Arch. Intern. Med.* 167(22), 2437–2442 (2007)
7. Greenland, P., et al.: Coronary artery calcium score combined with framingham score for risk prediction in asymptomatic individuals. *JAMA* 291, 210–215 (2004)
8. Qian, Z., Marvasty, I., et al.: Lesion-specific coronary artery calcium quantification better predicts cardiac events. In: *Int'l. Symposium on Biomedical Imaging* (2009)
9. Andrews, R., Tsochantaridis, I., Hofmann, T.: Support vector machines for multiple instance learning. In: *Advances in Neural Information Processing Systems* (2003)
10. Ehara, S., Kobayashi, Y., et al.: Spotty calcification typifies the culprit plaque in patients with acute myocardial infarction: an intravascular ultrasound study. *Circulation* 110(22), 3424–3429 (2004)
11. Pencina, M., D'Agostino, S.R., et al.: Evaluating the added predictive ability of a new marker: From area under the roc curve to reclassification and beyond. *Sta. Med.* 27(2), 157–172 (2008)
12. Fan, R.E., Chen, P.H., Lin, C.J.: Working set selection using the second order information for training svm. *J. Machine Learning Research* 6, 1889–1918 (2005)
13. Hlatky, M., Greenland, P., et al.: Criteria for evaluation of novel markers of cardiovascular risk: A scientific statement from the american heart association. *Circulation* 119(17), 2408–2416 (2009)
14. Virmani, R., Burke, A., et al.: Pathology of the vulnerable plaque. *J. AM Coll. Cardiol.* 47(8), C13–C18 (2006)

Joint Registration and Segmentation of Dynamic Cardiac Perfusion Images Using MRFs

Dwarikanath Mahapatra and Ying Sun

Department of Electrical and Computer Engineering, National University of Singapore, 4 Engineering Drive 3, Singapore 117576
dmahapatra@gmail.com, elesuny@nus.edu.sg

Abstract. In this paper we propose a Markov random field (MRF) based method for joint registration and segmentation of cardiac perfusion images, specifically the left ventricle (LV). MRFs are suitable for discrete labeling problems and the labels are defined as the joint occurrence of displacement vectors (for registration) and segmentation class. The data penalty is a combination of gradient information and mutual dependency of registration and segmentation information. The smoothness cost is a function of the interaction between the defined labels. Thus, the mutual dependency of registration and segmentation is captured in the objective function. Sub-pixel precision in registration and segmentation and a reduction in computation time are achieved by using a multiscale graph cut technique. The LV is first rigidly registered before applying our method. The method was tested on multiple real patient cardiac perfusion datasets having elastic deformations, intensity change, and poor contrast between LV and the myocardium. Compared to MRF based registration and graph cut segmentation, our method shows superior performance by including mutually beneficial registration and segmentation information.

1 Introduction

Dynamic perfusion magnetic resonance (MR) images are characterized by rapid intensity change over a region of interest, low spatial resolution, poor contrast and noise. Therefore, registration or segmentation of the left ventricle (LV) in cardiac perfusion datasets is a challenging task. However, there are certain characteristics of perfusion images which make it appealing for joint registration and segmentation. The LV is characterized by varying levels of intensity over the image acquisition process. While the changing contrast makes registration difficult, it also helps in segmentation. Clear identification of object boundaries leads to greater accuracy in feature extraction and hence improved registration. Motivated by this scenario and the need to implement a computationally efficient method, we propose a Markov random field (MRF) framework for the joint registration and segmentation of the LV in cardiac perfusion datasets.

The first work on joint registration and segmentation [1] used an active contour framework to interleave level set segmentation with a feature based registration method. It successfully segmented and registered portal images to CT

scans. Partial differential equations were used in [2] for joint registration and segmentation while a statistical model was presented in [3]. Wyatt and Noble in [4] combine MRFs and Bayesian estimation for joint registration and segmentation where the use of MRFs is limited to segmentation. There are other methods where registration plays an important role in segmentation and vice-versa. In [5] an image is registered to an atlas or a clearly identified object using level sets. Likewise, including shape information in active contours requires a model of shape variation [6] involving accurate registration of training images. In combining registration and segmentation, the major challenges are to ensure convergence and prevent estimates of registration or segmentation parameters adversely affecting each other. Equally important is to define appropriate energy functions and include relevant information for both processes.

The important contribution of our work is in developing an MRF method for the joint registration and segmentation of cardiac perfusion images. Previous methods are based on active contours which, being iterative, have a high computation time, are likely to be trapped in local minima, and are sensitive to initialization. On the other hand discrete optimization techniques for MRFs, like graph cuts, can find a global or strong local optima in less time. We formulate the joint registration and segmentation problem as one of labeling where each label defines the joint occurrence of displacement field (for registration) and segmentation class. The cost function is a combination of the mutual dependency of registration and segmentation information at every label and a multiresolution graph cut optimization reduces the computation time. The rest of the paper is organized as follows: Section 2 gives details about joint registration and segmentation and our MRF formulation. Section 3 presents our experiments and results and we conclude with Section 4.

2 Theory

2.1 Joint Registration and Segmentation

In registration the objective is to match each pixel in the floating image to the most similar pixel in the reference image and the similarity metric depends on the type of images being used. The displacement field is regularized to give a smooth deformation and the smoothness constraints depend upon the registration framework. For B-spline [7] and other curve based registration methods, curve gradients are used as smoothness constraints. In [8], Shekhovstov et al. used MRFs for non-rigid registration where smoothness depends upon the relative displacement between labels. Since the smoothness formulation is not based on the boundary properties of the object being registered it may result in unexpected deformations of the registered image, especially at object boundaries. This shortcoming is overcome by smoothness criteria based on object features which is integral to joint registration and segmentation.

To achieve joint registration and segmentation between a pair of images the following points have to be kept in mind when formulating the energy function: 1) mutual dependency of registration and segmentation is considered; 2)

registration and segmentation information contribute equally; and 3) estimate of registration parameters does not adversely affect segmentation parameters or vice-versa. The active contour framework in [1] was able to achieve joint registration and segmentation by defining a continuous valued mapping between reference and floating images. A major disadvantage of the active contour framework is the multiple iterations needed for convergence. As with all energy minimization techniques using gradient descent, there is the possibility of the curve being trapped in local minima. The position of the initial curve also influences the final solution. Although this can be overcome by employing a standard segmentation technique to obtain an initial curve, the problems of multiple iterations and getting trapped in local minima still persist. The number of iterations can be greatly reduced by using graph cuts. Graph cuts is based on max flow approach and is very effective in finding the global minima or a strong local minima of discrete MRF energy formulations [9].

2.2 Markov Random Fields

MRFs have been previously used with perfusion images for elastic registration [10] and segmentation [11]. Its energy function takes the following form

$$E(f) = \sum_{s \in P} D_s(f_s) + \sum_{(s,t) \in N} V_{st}(f_s, f_t), \quad (1)$$

where P denotes the set of pixels, f_s denotes the label of pixel $s \in P$ in the floating image and N is the set of neighboring pixel pairs. For joint registration and segmentation f_s gives both the displacement vector and the segmentation class of pixel s , i.e., $f_s = \{x_s^1, x_s^2, L_s\}$ with $x_s = \{x_s^1, x_s^2\}$ denoting displacement along the two axes and L_s denoting the segmentation class ($L_s = 1$ denotes object and $L_s = 0$ denotes background). The labels of the entire set of pixels are denoted by f . $D(f_s) = D_1(x_s) + D_2(L_s, x_s)$, is a unary data penalty function derived from observed data and measures how well label f_s fits pixel s . V_{st} is a pairwise interaction potential that imposes smoothness and measures the cost of assigning labels f_s and f_t to neighboring pixels s and t . The optimization scheme for (1) using graph cuts is discussed later. Next, we discuss each term of the energy function in detail.

Data Penalty Term: D_s assigns a penalty to a pixel s taking on a particular label $f_s = \{x_s^1, x_s^2, L_s\}$. D_s is defined as the sum of two penalty terms. The first term is a function of gradient information and by itself is suitable for registration. The second term includes mutual dependency of registration and segmentation in the penalty term. The following issues are considered in combining the two penalty values: 1) the individual penalties have the same dynamic range, i.e., the difference between their maximum and minimum values should be same; 2) the individual terms are robust for their specific purposes; and 3) the combination of the two terms truly captures the mutual dependency of registration and segmentation. For greater accuracy and robustness a pixel block centered at the pixel is used to calculate the data penalty. We refer to the block centered at pixel s as s_b .

Let $I_f(s_b, i)$ denote the intensity of the i th pixel of block s_b in the floating image I_f and $I_r(s_b, i)$ denote the corresponding pixel intensity in the reference image I_r . The pixel block at $s + x_s$ is denoted as $s_b + x_s$ and the intensity of its i th pixel in I_r is given by $I_r(s_b + x_s, i)$. The corresponding edge orientation angles in the two images are given by $\beta_f(s_b, i)$ and $\beta_r(s_b + x_s, i)$ and the corresponding edge magnitudes are given by $M_f(s_b, i)$ and $M_r(s_b + x_s, i)$. Note that the data penalty is for each pixel but is calculated over a pixel block for robustness. $D_1(x_s)$ incorporates registration information in the objective function and is a function of gradient orientation information. It is given by

$$D_1(x_s) = \frac{1}{2} \left[1 - \frac{\sum_i M_r(s_b + x_s, i) M_f(s_b, i) \cos(\Delta\beta(x_s, s_b, i))}{\sum_i M_r(s_b + x_s, i) M_f(s_b, i)} \right], \quad (2)$$

where $\Delta\beta(x_s, s_b, i) = \beta_r(s_b + x_s, i) - \beta_f(s_b, i)$. Edge information has been successfully used to register contrast enhanced images [12] and is a robust feature in the face of intensity changes. D_1 is a normalized metric that gives values between 0 and 1 with 0 indicating a perfect match. Since D_1 is a penalty, its value is low for greater similarity between the pixel blocks.

The second penalty term, D_2 is a function of the mutual dependency of segmentation class and displacement vectors. The LV and myocardium are identified in I_f by drawing masks around it and I_f is rigidly registered to I_r . The intensities of pixels inside (outside) the mask are used to create Gaussian models of object (background). Let $p_{fo}(s)$ denote the posterior probability of *pixel* $I_f(s)$ belonging to object and $p_{fb}(s)$ denote its probability of belonging to background. The probability of pixel $I_r(s + x_s)$ belonging to object/background is given by $p_{ro}(s + x_s)/p_{rb}(s + x_s)$. Note here that by s we refer to the pixel at s . Similarly, $s + x_s$ refers to the pixel at location $s + x_s$. We do not use i because D_2 is calculated from individual pixel intensity values. Thus, D_2 is given by

$$\begin{aligned} D_2(L_s = 1, x_s) &= 1 - \sqrt{p_{fo}(s) \times p_{ro}(s + x_s)}, \\ D_2(L_s = 0, x_s) &= 1 - \sqrt{p_{fb}(s) \times p_{rb}(s + x_s)} \end{aligned} \quad (3)$$

For registration and segmentation to jointly influence the penalty term, the segmentation information from the reference and floating image are combined as a function of displacement vectors. If $p_{fo}(s)$ and $p_{ro}(s + x_s)$ both have a value greater than 0.5 the pixel is likely to belong to the object and the penalty for $L_s = 1$ is low. The corresponding penalty for $L_s = 0$ is high. If $p_{fb}(s)$ and $p_{rb}(s + x_s)$ have probability values greater than 0.5, indicating a background pixel, the penalty for background class is low and the corresponding penalty for object class is high. When there is ambiguity over the segmentation class, i.e., $p_{fo}(s) \geq 0.5$ and $p_{ro}(s + x_s) < 0.5$ or $p_{fb}(s) \geq 0.5$ and $p_{rb}(s + x_s) < 0.5$ then the square root of the product of the probabilities ensures that both the floating and reference image contribute to the penalty of each label.

Pairwise Interaction Term: This term is used to regularize the solution and combines smoothness constraints due to displacement vector and segmentation class. It is defined as

$$V_{st}(f_s, f_t) = \lambda_{st} \begin{cases} 0.002, & (L_s = L_t) \text{ and } |x_s - x_t| \leq \sqrt{2}, \\ 0.002, & (L_s \neq L_t) \text{ and } |x_s - x_t| \leq 3, \\ \infty, & \text{otherwise.} \end{cases} \quad (4)$$

λ_{st} is a spatially varying weight that depends upon the intensity difference between the neighboring pixels of the floating image and is given by

$$\lambda_{st} = \exp \left\{ - [I_f(s) - I_f(t)]^2 \right\}. \quad (5)$$

When neighboring pixels have the same segmentation class ($L_s = L_t$), they are likely to have similar displacements because pixels on the same object tend to move together. Thus we constrain the maximum relative displacement between the pixels to be $\sqrt{2}$ pixels. If neighboring pixels have different segmentation class ($L_s \neq L_t$), then they can have different displacements since pixels on different objects may move differently. Therefore we allow relative displacement between such pixels to be up to 3 pixels.

Optimization using Graph Cuts: The energy function is optimized using graph cuts [9] which is suitable for discrete MRF labeling problems. Its ability to enforce piecewise coherence makes it especially suitable for vision applications. We represent pixels as nodes V_p in a graph G which also consists of a set of directed edges E that connect nodes. For l labels, l terminal nodes are created. First, ± 6 displacement positions along x and y axis are defined with a step of 1 pixel between two consecutive positions. For every position there are two segmentation classes. The total number of labels are $2 \times (2 \times 6 + 1)^2 = 338$. In the second stage ± 4 displacements along each axis with a step of 0.5 pixel is defined and the total number of labels is 162. Pixel blocks of size 5×5 were used for determining the penalty values.

3 Experiments and Results

Registration Results: Cardiac images were acquired on Siemens Sonata MR scanners following bolus injection of Gd-DTPA contrast agent. The pixel spacing ranges from $(1.5 \times 1.5) - (2.8 \times 2.8)\text{mm}^2$. Contrast agent flows into the right ventricle (RV), then into the LV and finally into the myocardium. The acquired datasets were all in 2D and a total of 10 datasets were used to test our method. Each dataset had 60 frames. In some of the datasets, the images before contrast enhancement did not show the LV and were discarded. The total number of test images were 538 with one reference image from each dataset. The LV and myocardium was identified in each image by drawing masks and then rigidly registered before applying our method. We compared the registration performance of our method (*JRS*) and an MRF method that does not use segmentation information, i.e.,

$$E(f) = \sum_{s \in P} D_1(f_s) + \sum_{(s,t) \in N} V_{st}(f_s, f_t), \quad (6)$$

where $f_s = x_s = \{x_s^1, x_s^2\}$ and V_{st} is given by

$$V_{st}(f_s, f_t) = \begin{cases} 0.002, & |f_s - f_t| \leq \sqrt{2}, \\ \infty, & \text{otherwise.} \end{cases} \quad (7)$$

The reference image (Fig. 1(a)) and floating image (Fig. 1(b)) pair was chosen such that there is intensity change between them as well as noticeable deformations. The deformations are highlighted by the difference image in Fig. 1(c). Dark areas in the difference image are the ones with negative intensity difference and bright areas correspond to positive intensity difference values. Areas in grey have zero intensity difference. The difference image after registration using only MRFs (Fig. 1(d)), does not register all the deformations especially at the boundaries of the RV. The difference image obtained after registration using *JRS* (Fig. 1(e)) shows noticeable improvement in registration performance and inclusion of segmentation information plays a crucial role in it.

The epicardium and endocardium were manually identified in the reference image and all floating images. From the obtained deformation field we register the LV and myocardium in the floating image and calculate the error between the registered contours and those in the reference image. The measured error is the average distance between each point on the registered contour and the nearest point on the reference contour. We also show the outline of the registered contours (in red) on the reference image using *JRS* (Fig. 1(f)) and MRFs (Fig. 1(g)). The contours are better registered using *JRS* as is evident from the distance between the registered epicardium and endocardium. The quantitative error measures before and after registration using *JRS* and MRFs are shown in Table 1. While the average registration error after using MRFs is greater than 1 mm, in case of *JRS* it is well below 1 mm. In fact the maximum registration error for *JRS* hardly goes above 1 mm, with only a few datasets exhibiting that.

The deformed grids obtained from MRFs and *JRS* are shown, respectively, in Figs. 1(h) and (i). The deformation field obtained from MRFs and *JRS* are smooth. Even though segmentation information has been incorporated into the cost function, the smoothness constraint has been appropriately formulated to avoid folding. The registration errors given in Table 1 show the greater accuracy obtained for *JRS*. While MRFs, without any segmentation information, have higher registration error, they do not result in folding of deformation field. The relative displacement threshold in V_{st} ($\sqrt{2}$ in our case) plays an important role in influencing the smoothness of the deformation field. An increase in threshold leads to a disparate displacement field, i.e., neighboring pixels do not have coherent motion. To avoid such issues, the relative displacement between neighboring pixels is constrained to be within $\sqrt{2}$ pixels, which is the maximum distance between neighbors. Such an arrangement is necessary for discrete valued cost functions.

Segmentation Results: The LV and myocardium in each floating image are also separately segmented using graph cuts [11]. Figures 2(a) and (b) show respectively the outline of segmented mask using graph cuts and *JRS* on the same floating image. By segmentation results using *JRS* we refer to the segmentation

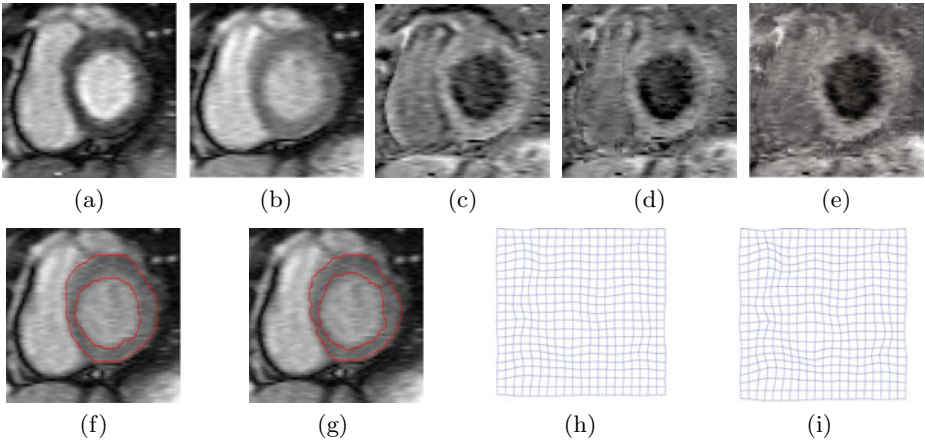


Fig. 1. Registration results for cardiac perfusion images. (a) Reference image; (b) floating image; difference image (c) before registration; (d) after registration using MRFs and (e) after registration using *JRS*. Superimposed outline of registered contours using (f) *JRS* and (g) MRFs. Deformed grids obtained from (h) *JRS* and (i) MRFs.

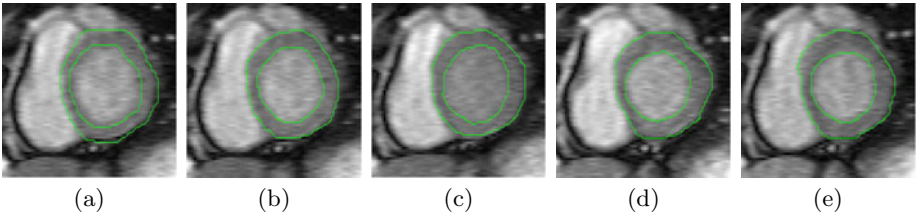


Fig. 2. Segmentation results for cardiac perfusion images. Outline of segmented mask in green using (a) graph cuts; and (b) *JRS*; (c)-(e) show the outline of the deformed mask in three floating images. In each floating image, the deformed mask was obtained by deforming the segmented mask of the reference image according to the corresponding displacement field.

class from the labels of each pixel in the floating image. *JRS* results in greater segmentation accuracy than graph cuts for images with poor contrast between LV and myocardium. The segmented mask from the reference image is deformed using the deformation field of different frames of the sequence. This deformed mask is then superimposed on the corresponding floating image as shown in Figs. 2 (c)-(e). These results show that registration and segmentation mutually benefit each other, i.e., the epicardium and endocardium are not only segmented accurately but also registered to the reference image. Table 1 also shows the different error measures for segmentation. The average Dice Metric (*DM*) and Root Mean Square (*RMS*) values for each dataset are shown in Fig. 3. *JRS* consistently shows higher *DM* and lower *RMS* values than graph cuts. The average *RMS* value for each dataset is close to 1 pixel and the average *DM* is above 90%

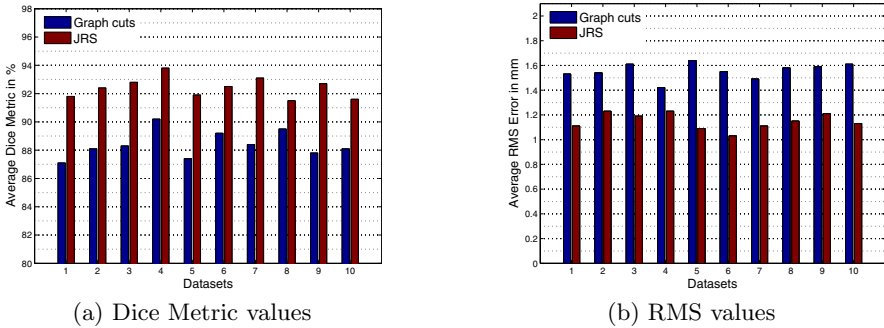


Fig. 3. Quantitative segmentation results for 10 datasets. (a) average DM values; (b) average RMS value. Brown bars show results for JRS and blue bars show results for graph cuts.

Table 1. Summary of registration and segmentation performance on cardiac perfusion datasets. The values indicate average and standard deviations for all datasets.

	Registration Error (mm)			Segmentation Result			
	Before Registration	After Registration		Graph Cuts		JRS	
		MRFs	JRS	RMS (mm)	DM (%)	RMS (mm)	DM (%)
Epicardium	2.2 ± 1.2	1.0 ± 0.2	0.7 ± 0.4	1.45 ± 0.43	88.6 ± 1.6	1.13 ± 0.32	92.1 ± 1.1
Endocardium	2.8 ± 1.0	1.1 ± 0.3	0.6 ± 0.3	1.65 ± 0.31	89.2 ± 0.9	1.11 ± 0.39	92.7 ± 0.8
Overall	2.6 ± 1.1	1.1 ± 0.2	0.5 ± 0.2	1.52 ± 0.34	88.93 ± 1.2	1.11 ± 0.34	92.5 ± 0.9

(values above 80% indicate excellent agreement with manual segmentation). For graph cuts DM values less than 90% are attributed mainly to the poor contrast between LV and myocardium in many images.

4 Conclusion

We have proposed a novel MRF based method for the joint segmentation and non-linear registration of the LV in perfusion cardiac images. Our method is different from previous works using active contours, and MRFs ensure less computation time and high accuracy. The problem was formulated as one of finding the appropriate labels. The labels give the displacement vector and segmentation class for each pixel. The cost function depends on contrast invariant edge information, segmentation class, and mutual dependency of registration and segmentation. The final labels are obtained by minimizing the cost function in a multiresolution graph cut implementation. The coarse to fine graph cut implementation gives sub-pixel accuracy for registration and segmentation. The performance of our method was compared with an MRF based registration method using only gradient information and a graph cut based segmentation method. Quantitative and visual results are shown for the registration of the epicardium and endocardium. For comparing segmentation performance, Dice Metric and

RMS errors between segmented LV and reference manual segmentation are calculated. All sets of results show the superior performance of our method compared to separate registration and segmentation method. In future work we aim to extend our method for 3D datasets and other imaging modalities, and also optimize it for less computation time.

Acknowledgment. The authors thank Siemens Corporate Research, NJ, USA, for providing the datasets, and acknowledge the support by NUS grant R-263-000-470-112.

References

1. Yezzi, A., Zollei, L., Kapur, T.: A variational framework for joint segmentation and registration. In: Proc. MMBIA, pp. 44–51 (2001)
2. An, J.-h., Chen, Y., Huang, F., Wilson, D., Geiser, E.: A variational PDE based level set method for a simultaneous segmentation and non-rigid registration. In: Duncan, J.S., Gerig, G. (eds.) MICCAI 2005. LNCS, vol. 3749, pp. 286–293. Springer, Heidelberg (2005)
3. Flach, B., Sara, R.: Joint non-rigid motion estimation and segmentation. In: Klette, R., Žunić, J. (eds.) IWCIA 2004. LNCS, vol. 3322, pp. 631–638. Springer, Heidelberg (2004)
4. Wyatt, P.P., Noble, J.A.: MAP MRF joint segmentation and registration. In: Dohi, T., Kikinis, R. (eds.) MICCAI 2002, Part I. LNCS, vol. 2488, pp. 580–587. Springer, Heidelberg (2002)
5. Tsai, A., Yezzi, A., Wells, W., Tempany, C., Tucker, D., Fan, A., Grimson, W., Willsky, A.: A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Trans. Med. Imag.* 22(2), 137–154 (2003)
6. Cootes, T., Beeston, C., Edwards, G., Taylor, C.: Unified framework for atlas matching using active appearance models. In: Kuba, A., Sámal, M., Todd-Pokropek, A. (eds.) IPMI 1999. LNCS, vol. 1613, pp. 322–333. Springer, Heidelberg (1999)
7. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans. Med. Imag.* 18, 712–721 (1999)
8. Shekhovtsov, A., Kovtun, I., Hlaváč, V.: Efficient MRF deformation model for non-rigid image matching. *Comput. Vision Image Understand.* 112(1), 91–99 (2008)
9. Boykov, Y., Veksler, O.: Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* 23, 1222–1239 (2001)
10. Mahapatra, D., Sun, Y.: Nonrigid registration of dynamic renal MR images using a saliency based MRF model. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, pp. 771–779. Springer, Heidelberg (2008)
11. Boykov, Y., Funka-Lea, G.: Graph cuts and efficient N-D image segmentation. *Intl. J. Comp. Vis.* 70(2), 109–131 (2006)
12. Sun, Y., Jolly, M.P., Moura, J.M.: Contrast-invariant registration of cardiac and renal MR perfusion images. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) MICCAI 2004. LNCS, vol. 3216, pp. 903–910. Springer, Heidelberg (2004)

3D Radio Frequency Ultrasound Cardiac Segmentation Using a Linear Predictor^{*}

Paul C. Pearlman¹, Hemant D. Tagare^{1,2,3}, Albert J. Sinusas^{3,4},
and James S. Duncan^{1,2,3}

Departments of
¹Electrical Engineering
²Biomedical Engineering
³Diagnostic Radiology, and
⁴Internal Medicine-Cardiology
Yale University, New Haven, CT, USA
paul.pearlman@yale.edu

Abstract. We present an approach for segmenting the left ventricular endocardial boundaries from radio-frequency (RF) ultrasound. The method employs a computationally efficient two-frame linear predictor which exploits the spatio-temporal coherence of the data. By performing segmentation using the RF data we are able to overcome problems due to image inhomogeneities that are often amplified in B-mode segmentation, as well as provide geometric constraints for RF phase-based speckle tracking. We illustrate the advantages of our approach by comparing it to manual tracings of B-mode data and automated B-mode boundary detection using standard (Chan and Vese-based) level sets on echocardiographic images from 28 3D sequences acquired from 6 canine studies, imaged both at baseline and 1 hour post infarction.

1 Introduction

Myocardial strain and strain rate measurements have been shown to be important for understanding cardiac disease. To this end, it has been reasonably well established that speckle tracking from high frame-rate radio-frequency (RF) ultrasound data can be used to generate dense estimates of displacement that can be used to solve for strain [1]. Current speckle tracking algorithms suffer from two drawbacks. First, they are extremely computationally intensive. Second, they perform well only on spatially homogeneous regions of the myocardium, but cannot easily identify such regions.

Segmentation of the left ventricle is particularly challenging in echocardiography due to gross inhomogeneities present in the images that lead to poor contrast between the blood and myocardium, as demonstrated in [2]. Analysis of high frame-rate (30 fps) raw RF images may provide important features that

^{*} This work is supported by grants 5R01HL082640-04 and 5R01HL077810-04.

can be exploited to segment these regions. The primary benefit of using RF as opposed to B-mode images for boundary detection is that, due to the regular structure of the myocardium, the phase of the RF exhibits a temporal coherence not present in the blood pool. This phase information is exploited by looking at complex correlations in state-of-the-art speckle tracking, and its advantages over B-mode are documented in the ultrasound physics literature [1]. This suggests an advantage in using RF data for segmentation. Direct segmentation of the RF also provides a mask that indicates volumes where meaningful cardiac speckle tracking results can be expected and thus a means of simplifying computations.

Few segmentation methods thus far have leveraged RF data for boundary discrimination. Yan et al. [3] and Nillesen et al. [4] use maximum correlation coefficient (MCC) images obtained from RF speckle tracking to segment the left ventricle. These methods suffer from the fact that precomputation of the MCC images is extremely computationally burdensome. Alternatively, Dydenko et al. [5] propose both a spectral autoregressive model and a velocity based model for tissue discrimination. The authors of [5] also establish the variance of the velocity as a meaningful parameter for segmentation.

The key observation in all of the above research is that the temporal correlation in the blood pool and myocardium have different patterns. We also exploit this effect, although we use a model that is computationally much simpler than calculating the MCC image, i.e., linear prediction. Experimental observation shows that this simple model, as opposed to more complex motion models, provides accurate segmentations. This is the main contribution of this paper. We also propose the use of the residues of a linear predictor as a basis for segmentation. These residues are similar to the variance of the velocity adopted in [5]. In addition, as opposed to the single-frame autoregressive model [5], a two-frame linear predictor is more useful because, while the autoregressive fit will vary wildly across the image, the temporal coherence in the signal leads to a more consistent fit that can be used to segment the whole boundary. We adopt a maximum a posteriori (MAP) approach to the segmentation with a probabilistic model that relates subsequent frames of a cardiac image sequence. This estimation is performed numerically by means of a level set active contour. We compare our method to a traditional Chan-Vese intensity based level set segmenter on B-mode images [6]. Additional validation is performed by comparison with manual segmentations.

2 Materials and Methods

2.1 Signal Model

Data Set. The interaction of the base-band RF signal with two specific acoustic media (blood pool and myocardium) is of interest for this work. The blood pool is made up of plasma and other elements that are much smaller than the wavelength of the ultrasound, while the myocardium is made up of primarily muscular fibers interspersed with blood vessels and bile ducts. The fibers make up 90% of the myocardial tissue and are thus responsible for most of the received signal [7]. The

regular structure of the fibers is responsible for the temporal coherence in the data that we exploit with our algorithm. Likewise, the lack of regular structure in the blood pool causes the signal to rapidly decorrelate.

It is of particular interest to us that our data is acquired at a high frame rate so that the motion of the speckle between frames is not significant relative to the boundaries. As a result, a single linear predictor for each medium (blood pool and myocardium) suffices for segmentation. Because of this, our algorithm is computationally efficient.

Preprocessing. The input pulse produced by the ultrasound machine is a real, bandlimited signal. Since the effects of transmission and reflection in ultrasound are linear, the signal we record is also real and bandlimited and can thus be recovered by coherent demodulation [8]. The demodulation results in a vector image consisting of an in-phase and quadrature component and is referred to as the analytic signal. It is this vector image we are interested in segmenting. An example analytic image is shown in Figure 3.

2.2 Linear Predictor

Let I^A be the 3D frame we wish to segment and I^B be the subsequent 3D frame. The domain of I^A is Ω and C is a boundary between two disjoint regions in the image, Ω_C and $\tilde{\Omega}_C$, where $\Omega = \Omega_C \cup \tilde{\Omega}_C$. Then the absolute linear prediction residue of the m^{th} voxel is

$$e_m = \begin{cases} \left| I_m^A - \sum_n \alpha_n \mathcal{R}e \{ I_{m,n}^B \} - i \sum_n \beta_n \mathcal{I}m \{ I_{m,n}^B \} \right|_2, & m \in \Omega_C, \\ \left| I_m^A - \sum_n \tilde{\alpha}_n \mathcal{R}e \{ I_{m,n}^B \} - i \sum_n \tilde{\beta}_n \mathcal{I}m \{ I_{m,n}^B \} \right|_2, & m \in \tilde{\Omega}_C. \end{cases} \quad (1)$$

Note that voxels in Ω_C and $\tilde{\Omega}_C$ have different predictor coefficients; these are given by vectors α , β , $\tilde{\alpha}$, and $\tilde{\beta}$ indexed by n where $I_{m,n}^B$ defines the n^{th} neighbor of voxel I_m^A , and $i = \sqrt{-1}$. A diagram of the predictor is shown in figure 4.

2.3 MAP Estimation

We assume that e_m for $m \in \Omega_C$ and $m' \in \tilde{\Omega}_C$ are distributed normally with probabilities p_1 and p_2 , zero mean, and standard deviations σ_1 and σ_2 , respectively. We seek to segment the image within a level set formulation. Let $\phi_m : \Omega \rightarrow \mathcal{R}$ be the level set function whose zero level set is C , and let H be a smooth Heaviside function. Segmentation is performed by maximizing the following log posterior probability with α , β , $\tilde{\alpha}$, and $\tilde{\beta}$ as defined in equation 1:

$$\begin{aligned} & l(e_m, \phi_m, \alpha, \tilde{\alpha}, \beta, \tilde{\beta}, \sigma_1, \sigma_2) \quad (2) \\ &= \sum_{m \in \Omega_C} \log p_1 (e_m | \alpha, \beta, \sigma_1) + \sum_{m' \in \tilde{\Omega}_C} \log p_2 (e_{m'} | \tilde{\alpha}, \tilde{\beta}, \sigma_2) \\ &= \sum_{m \in \Omega} H(\phi_m) \log p_1 (e_m | \alpha, \beta, \sigma_1) + H(1 - \phi_m) \log p_2 (e_m | \tilde{\alpha}, \tilde{\beta}, \sigma_2) \end{aligned}$$

$$= \sum_{m \in \Omega} H(\phi_m) \left[\log \frac{\sigma_2}{\sigma_1} - \frac{e_m^2}{2\sigma_1^2} + \frac{e_m^2}{2\sigma_2^2} \right] - \left[\log \sqrt{2\pi} \sigma_2 - \frac{e_m^2}{2\sigma_2^2} \right]$$

2.4 Level Set Formulation of the Model

Maximizing the likelihood is equivalent to minimizing its negative, so we define our energy functional as $E(e_m, \phi_m, \alpha, \tilde{\alpha}, \beta, \tilde{\beta}, \sigma_1, \sigma_2) = -l(e_m, \phi_m, \alpha, \tilde{\alpha}, \beta, \tilde{\beta}, \sigma_1, \sigma_2)$. As in [6], we introduce a standard prior on the arc length of the propagating front, $E_{prior}(\phi_m)$. Finally, we introduce a prior on the smoothness of the predictor parameters given by $E_S(\alpha, \tilde{\alpha}, \beta, \tilde{\beta}) = \alpha^T \mathbf{W} \alpha + \tilde{\alpha}^T \mathbf{W} \tilde{\alpha} + \beta^T \mathbf{W} \beta + \tilde{\beta}^T \mathbf{W} \tilde{\beta}$ where \mathbf{W} is a finite difference matrix. The overall energy is thus given by $E(e_m, \phi_m, \alpha, \tilde{\alpha}, \beta, \tilde{\beta}, \sigma_1, \sigma_2) = E_{LP}(e_m, \phi_m, \alpha, \tilde{\alpha}, \beta, \tilde{\beta}, \sigma_1, \sigma_2) + \lambda_{prior} E_{prior}(\phi_m) + \lambda_S E_S(\alpha, \tilde{\alpha}, \beta, \tilde{\beta})$ where λ_S and λ_{prior} are weights on the smoothness priors.

2.5 Optimization

We minimize the energy functional as follows: 1) Initialize ϕ inside the blood pool. 2) Update the other (non-level set) parameters of the model. 3) Update ϕ . We then iterate steps 2 and 3 until a local minima of the energy functional is reached. ϕ is updated by gradient descent, while the other parameters admit closed form solutions.

The maximum likelihood estimates of σ_1 and σ_2 are given by the standard formula for normally distributed random variables. Solving for the predictor parameters requires accumulating the voxels of I^A to form the vectors \mathbf{x}^A and \mathbf{y}^A that contain the values of all of the real and imaginary ordered voxels in Ω_C . $\tilde{\mathbf{x}}^A$ and $\tilde{\mathbf{y}}^A$ are likewise formed from voxels in $\tilde{\Omega}_C$. Similarly, the matrices \mathbf{X}^B , \mathbf{Y}^B , $\tilde{\mathbf{X}}^B$, and $\tilde{\mathbf{Y}}^B$ are generated by defining the rows as the real and imaginary values of the neighborhoods of all ordered voxels in I^B . The estimates for the prediction parameters are thus given by the following matrix equations, referred to as the normal equations:

$$\hat{\alpha} = (\mathbf{X}^{B\top} \mathbf{X}^B + 2\lambda_S \mathbf{W}^\top \mathbf{W})^{-1} \mathbf{X}^{B\top} \mathbf{x}^A \quad (3)$$

$$\hat{\beta} = (\mathbf{Y}^{B\top} \mathbf{Y}^B + 2\lambda_S \mathbf{W}^\top \mathbf{W})^{-1} \mathbf{Y}^{B\top} \mathbf{y}^A \quad (4)$$

$$\hat{\tilde{\alpha}} = (\tilde{\mathbf{X}}^{B\top} \tilde{\mathbf{X}}^B + 2\lambda_S \mathbf{W}^\top \mathbf{W})^{-1} \tilde{\mathbf{X}}^{B\top} \tilde{\mathbf{x}}^A \quad (5)$$

$$\hat{\tilde{\beta}} = (\tilde{\mathbf{Y}}^{B\top} \tilde{\mathbf{Y}}^B + 2\lambda_S \mathbf{W}^\top \mathbf{W})^{-1} \tilde{\mathbf{Y}}^{B\top} \tilde{\mathbf{y}}^A \quad (6)$$

These equations minimize the sum of squared differences between the in-phase and quadrature predictions.

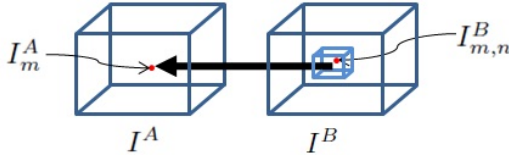


Fig. 1. Linear Predictor for point m

3 Experiments

We acquired 28 sequences of canine RF images from 6 studies using a Phillips iE33 4D Ultrasound imaging system with a sampling rate of 30 frames per second. These images came from both healthy and diseased animals. 15 sequences were acquired at baseline while 13 sequences were acquired one hour post infarction. The infarction was performed by angioplasty occlusion. The proposed algorithm was evaluated at both end-diastole and a frame from ventricular ejection as these time points represent both the minimum and maximum average motion of the left ventricle, respectively. To preserve image acquisition statistics for this analysis and future speckle tracking, the images were left in the Frustum coordinate system, which consists of one scan line dimension and two angular dimensions.

To evaluate the algorithm and compare it with standard techniques, two algorithms were implemented: 1) The Chan-Vese algorithm [6]; 2) Our algorithm based on linear prediction. The Chan-Vese algorithm assumes intensity is homogeneous over each region of the image and only assumes a smoothness prior for the boundary. To properly leverage the temporal continuity in myocardial speckle, the search window for the coefficients of the predictor was chosen such that it included the maximum observed distance between any given speckle in two frames (typically on the order of 30 voxels in the axial dimension and 4 in each angular dimension). This window size is fixed for all experiments. Then, to further reduce complexity, the number of coefficients was decimated so each coefficient represented multiple voxels.

3.1 The Advantage of Segmenting Based on Temporal Coherence

Examples of the performance of the algorithms are shown in figures 2, 4, and 5. The results for the Chan-Vese algorithm are shown in red and our results are shown in blue. In figure 2a, both contours underestimate the boundary on either side of the papillary muscle, but while the proposed algorithm underestimates this boundary to a more significant degree, the Chan-Vese contour leaks out through a mild dropout in the image that does not affect the results of our algorithm. Similar results are shown in figures 2b and 2c for images that contains more significant signal dropout.

Figure 4 shows an example image where, due to poor contrast, the Chan-Vese algorithm fails to adhere to the endocardial border and instead adheres to the inside of the specularly produced by the epicardium, while our proposed

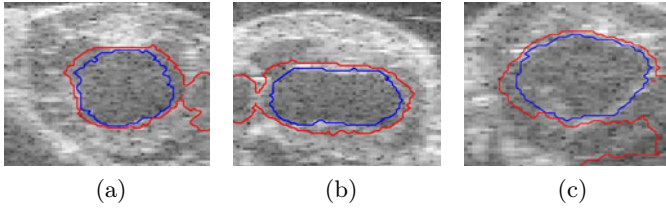
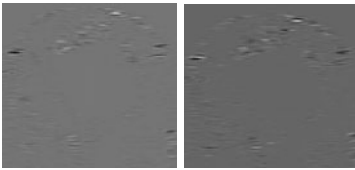


Fig. 2. Chan-Vese contours leak out



(a) In-phase (b) Quadrature

Fig. 3. Slice of analytic image containing myocardium and blood pool

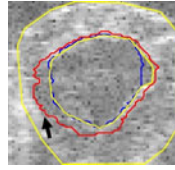


Fig. 4. Chan-Vese contour fails to find boundary

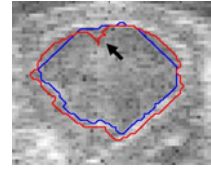


Fig. 5. Chan-Vese contour adheres to acquisition artifact

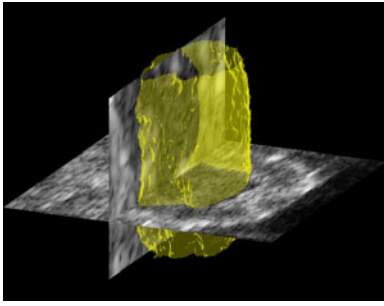
algorithm finds the appropriate boundary (true boundaries in yellow). In the image in figure 5 there is an artifact present in the image that leads to a bright spot in the blood pool. Our proposed algorithm finds the correct boundary in the presence of this artifact, while the Chan-Vese contour adheres to the artifact.

3.2 Results

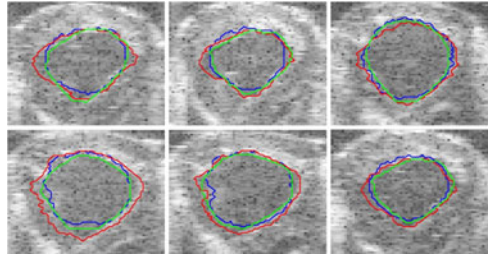
We compared our algorithm and the Chan-Vese algorithm with manually segmented surfaces. Typical segmentations are shown in figures 6a and 6b. Examples showing how the three contours compare are shown in figure 6b for visual evaluation. It was observed by visual inspection that, for most of our data sets, the Chan-Vese model overestimated the size of the blood pool. The error in the proposed model appears to be related to an underestimation of the extent of the blood pool in localized regions.

To quantitatively evaluate the results, we utilized the following three segmentation quality indices: 1) Hausdorff Distance (HD); 2) Mean Absolute Distance (MAD); and 3) the Dice coefficient. It should be noted that 8 Chan-Vese segmentations leaked, failing to find the endocardium, and were thus not included in the quantitative analysis.

Sequences obtained at baseline and those obtained one hour post infarct produced no significant difference in any of the quality indices, so all of the sequences were used to generate the values in Tables 1 and 2.



(a) Endocardial surface produced by Linear Predictor RF Segmenter



(b) Manual (Green), Chan-Vese (Red), Linear Predictor RF Segmenter (Blue)

Fig. 6. Typical segmentations**Table 1.** Sample means and standard deviations of quality indices (n=28) for segmentations obtained on frames at end diastole

Performance Measures	Chan-Vese			Proposed Algorithm		
	Dice	HD	MAD	Dice	HD	MAD
Average	0.643	11.885	1.862	0.725	11.436	1.983
STD	0.048	4.526	0.377	0.061	1.943	0.349

Table 2. Sample means and standard deviations of quality indices (n=28) for segmentations obtained on frames from ventricular ejection

Performance Measures	Chan-Vese			Proposed Algorithm		
	Dice	HD	MAD	Dice	HD	MAD
Average	0.653	9.823	1.812	0.723	9.426	1.872
STD	0.036	4.732	0.497	0.056	2.936	0.363

HD and MAD did not differ greatly between the methods. This result can be explained by the fact that both methods often produce smooth curves that were close to the ground truth with the Chan-Vese contour overestimating and our method slightly underestimating the dimension of the blood pool. As such, the most useful measure in evaluating our method is the Dice coefficient where our algorithm showed a marked improvement over the Chan-Vese method.

Analysis of frames at end diastole and during ventricular ejection produced similar results which demonstrate our proposed algorithm functions both during minimum and maximum motion of the ventricle, which respectively correspond to maximum and minimum temporal coherence in the signal.

4 Discussion and Conclusions

We have presented an active contour RF ultrasound segmentation model that leverages temporal linear prediction within an image sequence. The algorithm

relies solely on features present in the signal rather than *a priori* knowledge of expected shape. By utilizing a feature for segmentation that does not make a piecewise homogeneous assumption for the image, our contour is prevented from leaking out through boundaries that have relatively poor contrast. Since the motion of blood is irregular, performing speckle tracking on the blood pool provides no meaningful information. Our objective function segments based on the same concept, so it is particularly well suited for providing a geometric constraint for RF speckle tracking.

We also showed that our algorithm performs well during both peak and minimum average motion of the ventricle. Thus the algorithm can be extended to segment entire 4D data sets. Nonetheless, solving for the predictor coefficients is not trivial. An extension of this method to larger data sets, 4D segmentation, and joint endocardial/epicardial segmentation will require modifying this stage of the algorithm. This will be accomplished by gradient descent or the Levenberg-Marquardt algorithm.

Other future work includes validation of prediction coefficients and an analysis of their contribution, as opposed to residues, to segmentation. Also, the assumption of a fixed neighborhood for the linear predictor breaks down near the epicardial boundary causing a slight underestimation of the blood pool, which will be addressed through minor changes to our model. Ultimately, this method will be integrated as a preprocessing step into RF speckle tracking algorithms.

Acknowledgments. The authors would like to thank Matt O'Donnell, University of Washington, for helpful discussions and the reviewers for critical feedback.

References

1. Lubinski, M., Emelianov, S., O'Donnell, M.: Speckle tracking methods for ultrasonic elasticity imaging using short time correlation. *IEEE trans. Ultrason Ferroelect. Freq. Control.* 46, 82–96 (1999)
2. Qian, X., Tagare, H.D.: Overcoming dropout while segmenting cardiac ultrasound images. In: *ISBI*, pp. 105–108 (2006)
3. Yan, P., Jia, C.X., Sinusas, A., Thiele, K., O'Donnell, M., Duncan, J.S.: LV Segmentation through the analysis of radio frequency ultrasonic images. *Inf. Process Med. Imaging* 20, 233–244 (2007)
4. Nillesen, M., Lopata, R.G.P., Huisman, H.J., Thijssen, J.M., Kapusta, L., Korte, C.L.: 3D cardiac segmentation using temporal correlation of radio frequency ultrasound data. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009. LNCS*, vol. 5762, pp. 927–934. Springer, Heidelberg (2009)
5. Dydenko, I., Friboulet, D., Gorce, J.M., D'Hooge, J., Bijmens, B., Magnin, I.E.: Towards ultrasound cardiac image segmentation based on the radiofrequency signal. *Med. Image Anal.* 7(3), 353–367 (2003)
6. Chan, T.F., Vese, L.A.: Active Contours Without Edges. *IEEE Transactions on Image Processing* 10, 266–277 (2001)
7. Shung, K., Thieme, G.: *Ultrasonic scattering in biological tissues*. CRC Press, Boca Raton (1993)
8. Langton, C.: Hilbert transform, analytic signal, and the complex envelope. *Signal Processing and Simulation Newsletter* (1999)

Automated Interventricular Septum Thickness Measurement from B-Mode Echocardiograms

Navneeth Subramanian¹, Dirk Padfield², Sheshadri Thiruvenkadam¹,
Anand Narasimhamurthy¹, and Sigmund Frigstad³

¹ GE Global Research, Whitefield Road, Bangalore, India
`navneeth.s@ge.com`

² GE Global Research, One Research Circle, Niskayuna, NY

³ GE Vingmed Ultrasound, Trondheim, Norway

Abstract. In this work, we address the problem of automated measurement of the interventricular septum thickness, one of the key parameters in cardiology, from B-mode echocardiograms. The problem is challenging due to high levels of noise, multi modal intensity, weak contrast due to near field haze, and non rigid motion of the septum across frames. We introduce a complete system for automated measurement of septum thickness from B-mode echocardiograms incorporating three main components: a 1D curve evolution algorithm using region statistics for segmenting the septum, a motion clustering method to locate the mitral valve, and a robust method to calculate the septum width from these inputs in accordance with medical standards. Our method effectively handles the challenges of such measurements and runs in near real time. Results on 57 patient recordings showed excellent agreement of the automated measurements with expert manual measurements.

1 Introduction

Interventricular septum thickness (or diameter) (IVSd) is one of the key parameters in cardiology since it is, along with LV size, one of the main indicators of cardiac hypertrophy. It can be used as a screening parameter for septal hypertrophy and has also shown a correlation to 24h ambulatory blood pressure. Unfortunately, the manual measurement of the septum thickness on echocardiograms suffers from large inter- and intra-observer variability based on the experience of the user. Further variability results from different acquisitions since it is not easy to pick the right plane to scan, and this challenge is compounded by misalignment between available acoustic windows for imaging and the patient's heart. These issues can lead to undesirable variations in the diagnosis, resulting in missed detection of subjects and unnecessary further testing of healthy patients. Automating this measurement can avoid this variability and can save physicians time and effort.

Close observation reveals that the IVSd measurement is a complex cognitive task performed by the cardiologist. An end-to-end IVSd measurement in accordance with the American Society of Echocardiography (ASE) guidelines [1]

involves segmentation of the septum, identification of the tip of the mitral valve, and measurement of the thickness orthogonal to the centerline of the septum. The aforementioned procedure must be repeated for all frames of a multi-frame recording, and the septum thickness is reported on the frame corresponding to end-diastole.

Our contribution is a solution for end-to-end automation of the complex septum measurement workflow on PLAX B-mode echocardiograms. Our method incorporates three main components: a 1D curve evolution algorithm for segmenting the septum on each frame, a motion-clustering based method to localize the mitral valve, and a robust method to calculate the septum width from these inputs in accordance with medical standards. We know of no previous attempt at end-to-end septum measurement automation. For this reason, we compare against previously developed approaches for septum segmentation and identification of the mitral valve, individually.

Any approach to segment the septum from ultrasound images faces several challenges. Firstly, the intensity in and around the septum region is multi modal, hence purely intensity and histogram based schemes are challenged. For example, the approach proposed in [2] uses multi-level thresholding for measurement on a single instantaneous frame. This method, however, has an inherent difficulty on tough clinical cases where near-field haze is encountered. Secondly, the speckle noise inherent in ultrasound images makes edge-based methods unreliable. These methods are sensitive to initialization and are plagued by the near-field haze which leads to low contrast at the upper part of the septum. For any approach to segment the septum across several frames reliably, robustness to noise and initialization is needed. Region based active contour [AC] approaches have typically offered robustness to curve placement and noise, but methods using global statistics [3] are not ideal for segmenting objects from multi-modal data, as in our case. Recent works [4,5] have addressed this issue by locally estimating statistics around each pixel, although at a higher computational cost. There are also efforts that introduce prior shape information into segmentation schemes based on [AC] [6,7], and these methods have been reasonably successful in handling low contrast and boundary gaps, but, to be effective, they need to be carefully trained to acquire a rich description of shape statistics. However, the shape of the septum varies vastly across patients and with view angles, and undergoes large non rigid motion across frames. These factors make it difficult to build a shape atlas of the septum to use within a segmentation approach.

We propose a local region based variational energy to solve the segmentation problem. One main difference as compared to usual 2D [AC] methods is in our representation of the septum. As a result of insights gained from clinical data, we represent the septum using two 1D profiles which correspond to the top and bottom septum boundaries respectively. This representation allows capture of regional statistics in and around the septum, and model interactions between the boundaries (e.g. width) of the septum, and is computationally easier. With the above terms, our segmentation energy was able to effectively deal with the data challenges and viable for real time implementation.

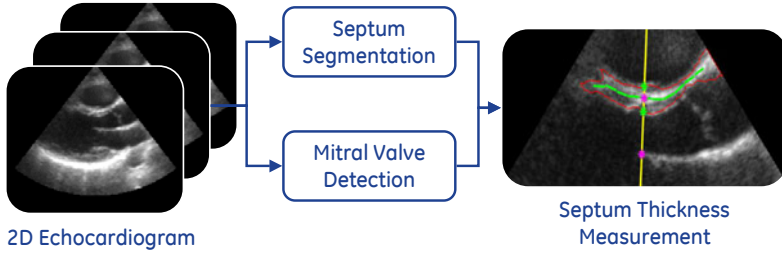


Fig. 1. Fully automatic IVSd measurement: The right figure shows the segmented septum in red, the detected mitral valve tip as a purple star, the fit of piecewise lines to the septum medial axis as green lines, the computation of the orthogonal line in yellow, and calculation of the points used to define septum thickness as green stars

Next, a critical step for defining the measurement line of the septum thickness is the identification of the mitral valve leaflet. Martin *et al.* [8] propose a semi-automatic method where a rough segmentation is first obtained and is then refined using active contours with dynamic program minimization. However, as outlined by the authors, the approach is sensitive to the manual initialization of the contours at the muscle-leaflet junction, precluding its application in automated measurement. We propose a simple method, motivated by the observation that valve leaflets are the fastest moving cardiac structures. The novelty of our method is that it is completely unsupervised without the need for prior training or initialization as compared to previous approaches.

Finally, given the segmented septum and the location of the mitral valve tip, the thickness measurement is given by the length of the septum region along the line that is orthogonal to the center line of the septum region and passing through the mitral valve tip. Our end to end method was validated on a large variety of patient recordings, many with challenging image quality, and provided reproducible measurement that follow the guidelines of the ASE. The automated measurements closely correspond with manual measurements generated by cardiologists from two separate institutions.

2 Methods

Our approach divides the problem into three steps that can be solved independently: septum segmentation, mitral valve tip detection, and septum thickness measurement as shown in Figure 1.

2.1 Septum Segmentation

As mentioned in the introduction, rather than utilize 2D contour evolution approaches, we formulate our search space over pairs of smooth 1D profiles. This representation enables easy access to regional statistics in and around the septum, and model interactions between the top and bottom septum boundaries,

in our case, to broadly constrain the width of the segmented septum. The width term is crucial for effective segmentation under weak contrast at the top part of the septum due to near field haze. Also from a computational standpoint, our 1D-profile based formulation works over a smaller set of minimization arguments than regular [AC] representations. Thus the resulting segmentation is very fast making it feasible for real time use.

Given an ultrasound image $I : \Omega \rightarrow R$, $\Omega = [a, b] \times [c, d]$, the goal is to search for two smooth 1D functions $f, g : [a, b] \rightarrow [c, d]$, whose profiles represent the top and bottom parts of the septum. We assume that close to the septum boundary, the image intensity is piecewise constant. Hence we seek an f and g that separate local neighborhoods into three regions (see Figure 2), each of which we model with Gaussian statistics. We denote the septum region, the region between the 1D profiles of f and g , as R^s . We denote the neighborhood region above the septum as R^a , i.e. between profiles of g and $g + \Delta$, where Δ is some pre-defined interval. Similarly, the neighborhood below the septum is R^b , between profiles of $f - \Delta$ and f . Rather than represent the entire septum region by one distribution, we divide Ω into K disjoint rectangles, $\Omega = \bigcup_{i=1}^K B_i$, $B_i = (a_i, b_i) \times (c, d)$, which yields multiple, more accurate distributions. In this way, the image intensity is expected to be Gaussian in the regions $R^s \cap B_i$, $R^a \cap B_i$, and $R^b \cap B_i$. Under this construction, $\mu^a = [\mu_1^a, \mu_2^a, \dots, \mu_K^a]$, $\sigma^a = [\sigma_1^a, \sigma_2^a, \dots, \sigma_K^a]$ gives the set of distributions in each of the regions $R^a \cap B_i$, similarly μ^b, σ^b gives the distribution in each $R^b \cap B_i$, and μ^s, σ^s gives the distribution in each $R^s \cap B_i$. To find the septum region, we minimize the following energy over the space of smooth 1D functions $f, g : [a, b] \rightarrow [c, d]$, with distributions $\mu^a, \sigma^a, \mu^s, \sigma^s, \mu^b, \sigma^b$:

$$\begin{aligned}
 E(f, g, \mu^a, \sigma^a, \mu^s, \sigma^s, \mu^b, \sigma^b) = & \sum_{i=1}^K \left[\int_{a_i}^{b_i} \int_f^g \left(\frac{(I - \mu_i^s)^2}{(\sigma_i^s)^2} + \ln(\sigma_i^s) \right) \right. \\
 & + \int_{a_i}^{b_i} \int_{f-\Delta}^f \left(\frac{(I - \mu_i^b)^2}{(\sigma_i^b)^2} + \ln(\sigma_i^b) \right) + \int_{a_i}^{b_i} \int_g^{g+\Delta} \left(\frac{(I - \mu_i^a)^2}{(\sigma_i^a)^2} + \ln(\sigma_i^a) \right) \left. \right] \\
 & + \lambda_{smooth} \int_a^b \left(\sqrt{1 + (f')^2} + \sqrt{1 + (g')^2} \right) + \lambda_{width} \int_a^b (f + w - g)^2 \quad (1)
 \end{aligned}$$

The data term drives f, g to take values such that the intensity distribution in each of $R^s \cap B_i$, $R^a \cap B_i$, and $R^b \cap B_i$ are Gaussian, given by $\mu^a, \sigma^a, \mu^s, \sigma^s, \mu^b, \sigma^b$. The smoothness terms for f and g are governed by parameter λ_{smooth} . The last term constrains the width of the region R^s to be close to the expected septum width (w) and is balanced by λ_{width} . To minimize E , we use the Euler Lagrange equations of Equation 1 to iteratively solve for f, g , and $\mu^a, \sigma^a, \mu^s, \sigma^s, \mu^b, \sigma^b$, using an explicit finite difference scheme and given an initial guess f_0, g_0 . Figure 2 illustrates our segmentation evolution on a sample image.

2.2 Mitral Valve Detection

Locating the mitral valve leaflet is critical to septum measurement since clinical guidelines require measurement to be done at the level of the mitral valve in

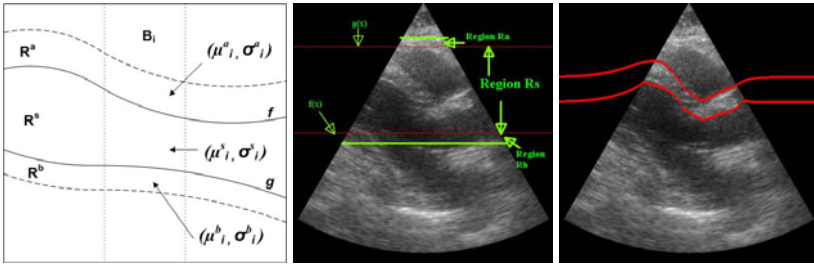


Fig. 2. f, g represent the lower and upper boundaries of the septum, respectively. The initialization of the functions (middle), and final segmentation of the septum (right). Regional statistics in the septum (R^s), above the septum (R^a), and below the septum (R^b) are used to drive evolution of the functions f, g .

the end-diastole frame. This task is challenging since the mitral valve is a fast moving and thin structure, but we introduce a method to locate the valve leaflet tip that leverages this characteristic. The main advantages of our method is that it is fast and is completely unsupervised without the need for prior training or initialization. It consists of computing a motion map, identifying candidate locations corresponding to high motion, and computing a representative location corresponding to the anterior tip of the mitral valve. The motion map is computed via frame differencing considering a window of several frames. It is necessary to consider more than one set of frame differences because the mitral valve leaflet is not clearly visible in some frames so that only one frame difference may not yield many locations corresponding to the valve leaflet. Next, pixels corresponding to high motion are identified as candidate locations, and these are grouped into clusters using a K -means clustering algorithm. The cluster center is then representative of the valve location. In our case, we are interested only in the mitral valve, and the cluster center corresponds with the mitral valve tip. The algorithm steps are illustrated in Figure 3.

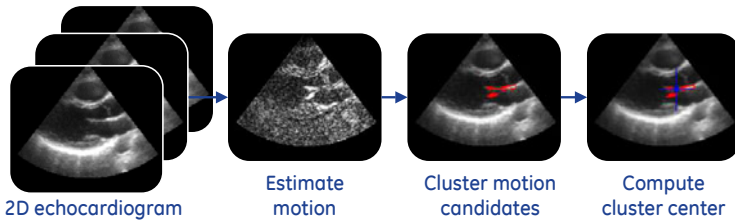


Fig. 3. Block diagram of the method for localizing the mitral valve leaflet

2.3 Septum Thickness Measurement

The manual measurement protocol of the ASE specifies that the measurement line should be drawn orthogonal to the center line of the septum region and passing through the mitral valve tip. We embed this requirement directly into our septum thickness algorithm, which is demonstrated in Figure 4. To represent the center line, our method first calculates the medial axis of the septum region. To make the algorithm computationally efficient, we then divide the medial axis into K regions (for example $K = 10$ works well) and fit line segments to each of these regions using linear regression. These K line segments are shown as green lines in Figure 4. Given the K line segments, we find the line through the mitral valve tip p_m with the smallest distance to each line segment $\overline{p_s p_e}$ (where p_s and p_e are the start and end point of the line, respectively). Each of these lines is orthogonal to its corresponding fit line segment by definition, and they intersect at point p_i that is found as $p_i = p_s + t * (p_e - p_s)$ where $t = -\frac{(p_s - p_m) \cdot (p_s - p_e)}{\|p_e - p_s\|^2}$ is the distance to p_i from p_s . Given the intersection point p_i and the mitral valve tip location p_m , the Euclidean distance between these two points is calculated. This method is repeated for all of the line segments fit to the septum region, and the line with the shortest distance is chosen for the septum thickness calculation. The septum thickness is found as the distance between the two points where this line intersects the septum segmentation mask.

3 Results

We evaluated our framework on two databases of B-mode parasternal long axis (PLAX) recordings, representing a total of 57 different patients. Each recording has several cardiac cycles (≈ 3 cycles/recording). The datasets were acquired on GE Vivid series ultrasound scanners at clinical sites in two countries. Two cardiologists independently performed manual measurements of IVSd on all datasets. Both cardiologists used the ECG to identify the end-diastole frame, on which the septum thickness was measured using calipers in B-mode. One of the cardiologists also scored the images as good, moderate, or poor based on the image quality and difficulty in imaging the patient.

The performance of the automated measurement algorithm was evaluated by comparing its measurement value on the selected frame with the cardiologists' measurements as shown in Figure 4. We report a mean error of 0.88 ± 0.96 mm (min=0.0 & max=5.0 mm) with respect to cardiologist 1, and an error of 1.17 ± 0.92 mm (min=0.0 & max=3.6 mm) with respect to cardiologist 2. The inter-observer weighted kappa statistic was 0.58. The weighted kappa statistic for our algorithm (0.56 vs. cardiologist 1, 0.49 vs. cardiologist 2) compares favourably with the inter-observer value. Figure 5 shows results on example good, moderate and poor quality recordings as defined by the cardiologists. Note the robustness of the algorithm to near-field haze and poor contrast.

In a clinical application, one uses the septum measurement to classify patients as normal, mildly abnormal, or hypertrophic. Standard recommendations on septum thickness for each of these categories are separated by approximately 2mm.

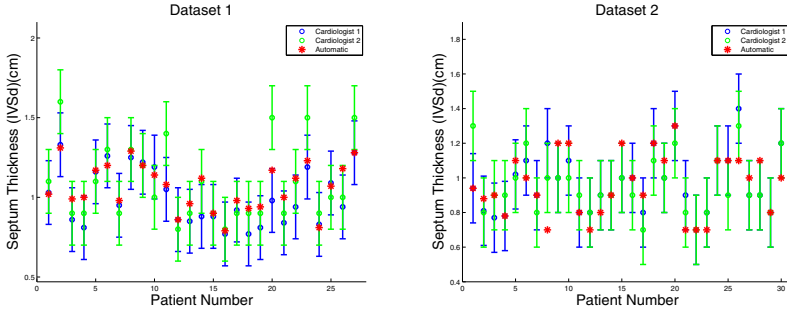


Fig. 4. Comparison of automatic versus manual measurements for two datasets representing a total of 57 patient recordings. The plot on the left shows results for 27 cases, and that on the right shows 30 cases. The blue and green circles are the manual measurements from two different expert cardiologists, and each circle shows the ± 2 mm range considered acceptable by the ASE standards. The red stars show the measurements from the automated method, which, in most cases, are within the range of the manual measurements.

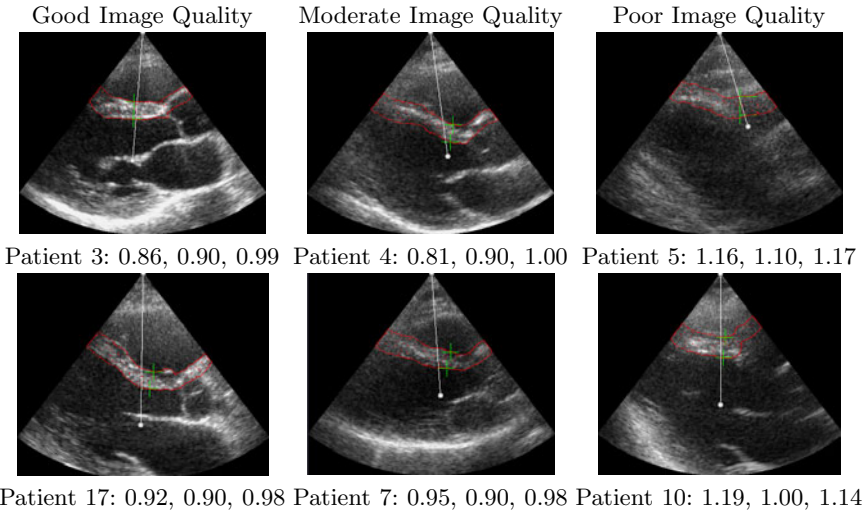


Fig. 5. Visual comparison of example measurement results from Dataset 1. The three columns represent good, moderate, and poor image acquisitions, respectively, as classified by expert cardiologists. The text under each figure lists the patient number (see the left plot of Figure 4) followed by the thickness measurement by cardiologist 1, cardiologist 2, and the automatic system, respectively. The red outline defines the automatic septum segmentation, and the green “+” signs indicate the locations of the septum width measurement. The white line is given as a reference to show the line connecting the tip of the probe with the detected mitral valve tip location.

Hence, defining the success criterion as the automated approach being within 2mm of the cardiologists measurement is clinically acceptable. Using this criterion, the success ratio of the automatic septum measurement solution (within ± 2 mm of the expert measurement) was found to be 89% (51/57 cases) with respect to cardiologist 1 and 82 % (47/57 cases) with respect to cardiologist 2. Our method runs in near real-time, with an average computation time of 0.1 sec/frame on a 2.6GHZ PC with 2GB RAM.

4 Conclusions

We have developed an automatic approach for septum thickness measurement consisting of septum segmentation using 1D curve evolutions based on regional statistics, mitral valve tip detection, and measurement line computation. The framework and its steps were developed in accordance with the clinical standards of the ASE, and they run in near real time. Our framework was validated on 57 B-mode PLAX recordings and was within the ± 2 mm guidelines specified by the ASE in an average of 85% of the cases. This automated measurement solution compares favorably with manual measurements by expert cardiologists and is attractive for clinical application.

Acknowledgements. The authors would like to thank Feng Lin, Kajoli Krishnan and Kai Thomenius for valuable insight and discussions.

References

1. Lang, R., et al.: Recommendations for chamber quantification: a report from the American Society of Echocardiography Guidelines and Standards Committee. *Journal of the American Society of Echocardiography* 18(12), 1440–1463 (2005)
2. Moladoust, H., et al.: Estimation of septal wall thickness by processing sequential echocardiographic images. *Iranian Cardiovascular Research Journal* 3(1) (2009)
3. Chan, T.F., Vese, L.A.: Active contours without edges. *IEEE Trans. Image Processing* 10, 266–277 (2001)
4. Guo, W., Chen, Y.: Using non-parametric kernel to segment and smooth images simultaneously. In: *ICIP*, pp. 217–220 (2006)
5. Lankton, S., Tannenbaum, A.: Localizing region-based active contours. *IEEE Trans. Image Processing* 17, 2029–2039 (2008)
6. Leventon, M., Grimson, W.L., Faugeras, O.: Statistical shape influence in geodesic active contours. In: *CVPR*, vol. 1, pp. 316–323 (2000)
7. Chen, Y., Thiruvankadam, S., et al.: On the incorporation of shape priors into geometric active contours. In: *IEEE VLISM*, pp. 145–152 (2001)
8. Martin, S., Daanen, V., et al.: Tracking of the mitral valve leaflet in echocardiography images. In: *IEEE Intl. Symp. on Biomedical Imaging, ISBI* (2006)

Probabilistic-Driven Oriented Speckle Reducing Anisotropic Diffusion with Application to Cardiac Ultrasonic Images

G. Vegas-Sanchez-Ferrero^{1,*}, S. Aja-Fernandez¹, M. Martin-Fernandez¹,
A.F. Frangi³, and C. Palencia²

¹ Laboratorio de Procesado de Imagen, Universidad de Valladolid, Spain

² Dep. Matematica Aplicada, Universidad de Valladolid, Spain

³ CISTIB, Universidad Pompeu Fabra, Spain

Abstract. A novel anisotropic diffusion filter is proposed in this work with application to cardiac ultrasonic images. It includes probabilistic models which describe the probability density function (PDF) of tissues and adapts the diffusion tensor to the image iteratively. For this purpose, a preliminary study is performed in order to select the probability models that best fit the statistical behavior of each tissue class in cardiac ultrasonic images. Then, the parameters of the diffusion tensor are defined taking into account the statistical properties of the image at each voxel. When the structure tensor of the probability of belonging to each tissue is included in the diffusion tensor definition, a better boundaries estimates can be obtained instead of calculating directly the boundaries from the image. This is the main contribution of this work. Additionally, the proposed method follows the statistical properties of the image in each iteration. This is considered as a second contribution since state-of-the-art methods suppose that noise or statistical properties of the image do not change during the filter process.

Keywords: Anisotropic diffusion, probability models, diffusion tensor, structure tensor, cardiac ultrasonic images.

1 Introduction

Ultrasound is a non-invasive imaging modality and a low-cost way to help diagnosing, which is widespread for many medical applications. However, ultrasonic (US) images are characterized by the presence of speckle, which is a granular pattern that degrades resolution and adds spatial noise to the image [1]. Therefore, many speckle removal filters have been proposed in the literature [2,3,4].

Speckle in ultrasound images can be seen as a random process whose statistical features depend on the tissue class. The existence of a deterministic component

* Thanks to Gerencia Regional de Salud, Consejeria de Sanidad y Junta de Castilla y Leon (GRS 292/A/08, SAN126/VA32/09, SAN126/VA33/09, VA027/A07) and Ministerio de Ciencia e Innovacion (MTM2007-63257) for funding.

due to specular reflections of the echo-pulse depends on the number of obstacles of the tissue (scatters) into the resolution cell and their size in comparison with the wavelength of ultrasound signal. Thus, many of the speckle removal filters are based on local statistics.

Among those filters, we focus on anisotropic diffusion based filters since the proposed filter can be considered as a probabilistic extension of them. One of the most popular approaches is Speckle Reducing Anisotropic Diffusion (SRAD) [2], which extends the well known Perona Malik's anisotropic diffusion avoiding to threshold the norm of the gradient. Instead of that threshold, an estimate of the coefficient of variation of noise is used. Results and stability depends on a good estimation of the local statistics as it was demonstrated in [3].

In the case of Oriented Speckle Reducing Anisotropic Diffusion (OSRAD) [4], no logarithmic compression is supposed and the speckle statistics are assumed to be modeled by a Rician distribution. In the case of high SNR, speckle statistics are approximated by a Gaussian distribution. This is a reasonable assumption when resolution cells have a large number of scatters and there exists a deterministic component, which is the case of the myocardial tissue for the raw signal obtained by the transducer. However, besides logarithmic compression, an interpolation process is performed to reconstruct the image and, therefore, the probabilistic model is affected. Additionally, the diffusion schemes are iterative, so the Rician distribution supposition does not hold in each iteration. So, because of logarithmic compression and interpolation, the statistics of speckle in reconstructed images do not follow the original distributions of the raw signal. Hence, many recent works study different probability models in order to obtain a suitable model to describe the statistical behavior of the speckle [1,5,6].

In this work, we propose a novel anisotropic diffusion filter with application to cardiac ultrasonic images, which includes the probability models that describe PDFs of different tissue classes and adapt iteratively to the filtered image. Probability models for each tissue class were selected from a training database using real-life cases for cardiac ultrasonic images. A mixture PDF model is adopted for fitting probabilistic models of tissue classes to the histogram of the whole image. With this methodology, a Bayesian scheme can be applied to estimate the probability of belonging to each of the tissue classes for every image voxel.

When the structure tensor of the probability of belonging to each tissue is combined in the definition of the diffusion tensor, better boundaries estimates can be obtained instead of directly calculating them from the image. The structure tensor of the probability of belonging to each tissue allows defining a diffusion tensor, which takes into consideration the boundaries between different tissue classes for filtering purposes. This is the main contribution of this work and, to the best of our knowledge, no similar approach has been considered in the literature for this purpose.

The paper is organized as follows. In section 2, the stastical study is explained and the distributions are selected. Section 3 describes the filtering method. Section 4 presents some experiments and results on synthetic and real images. In section 5, we conclude analyzing the results.

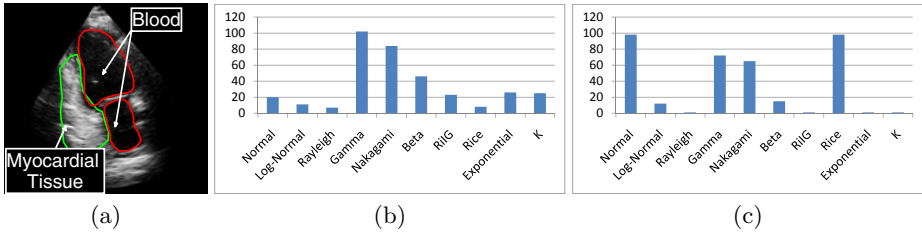


Fig. 1. (a) Example of the regions for blood and myocardial tissue. (b) Blood. (c) Myocardial Tissue. For blood tissue class, the Gamma distribution performs best. For myocardial tissue, the best performance is achieved for Rice or normal distributions.

2 Probabilistic Estimation of Tissue Classes in US Images

In this section, the goodness-of-fit of several probabilistic distributions is presented. The study is extended to a representative set of distributions that were proposed in the literature: Gamma [15], Log-Normal [5], Rayleigh [15,6], Normal [5], Nakagami [16], Beta, Rician Inverse Gaussian (RiIG) [6], Rice [7], Exponential and K [1]. A set of 120 two-dimensional images of size 1024×768 and 8 bits per pixel were obtained from a clinical machine Philips iE33 ultrasonographic system¹ with the software PMS5.1 Ultrasound iE33 4.0.1.357 scanned from human subjects.

The methodology used to carry out the study is the same used in [15] where a χ^2 goodness-to-fit test is performed. We used a significance value $\alpha = 0.05$ for the test. In order to avoid spatial correlation, the image is subsampled by a factor of 6. Two different tissue classes were considered: blood and myocardium.

In Fig. 1, an example of the analyzed tissues is presented as well as the results of passing the goodness-of-fit test for blood and myocardial tissue. In Fig. 1b we can see the better performance of the gamma distribution in the case of blood, this result holds with that one obtained in [5]. In the case of the myocardial tissue (Fig. 1c), Rice distribution and Normal distributions are the ones with the best fit. This result is coherent to the interpretation of quasi periodic scatterers, which gives rise to a Rician model of speckle [7] and confirms the hypothesis of the OSRAD filter [4]. Due to the simplicity of estimating Normal parameters for estimating mixtures of PDFs, we chose the Normal distribution as a good candidate for this tissue class.

3 Probabilistic Directional SRAD

When the histogram of the ultrasonic image is considered as the contribution of different kinds of tissues which follow different PDFs, the mixture of PDFs arises as a natural way to fit the histogram. This way, when parametric distributions are considered, the parameters of each independent variable have to be calculated.

¹ Philips Healthcare, Andover, MA, USA.

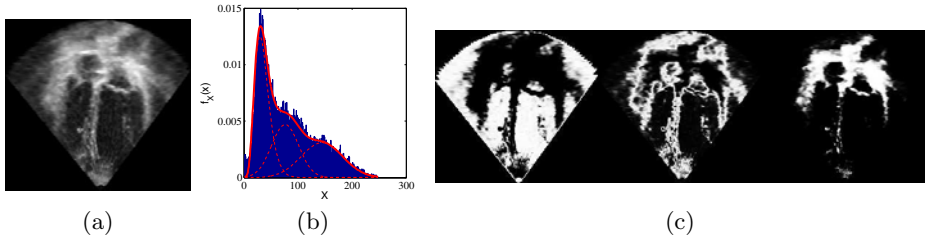


Fig. 2. (a) Clinical ultrasound image. (b) Finite Mixture Model. (c) Tissue class probability map for each voxel of the image. The good fit of the mixture model of the image (a) becomes clear as well as the structure of the ownership to each class where the boundaries of the left and right ventricles are clearly defined.

The most widespread method for this purpose is the Expectation-Maximization (EM) algorithm, which allows estimating the parameters of the mixture. We decided to adopt the unsupervised learning of mixture models method [8], since it is capable of selecting the number of mixture components from a maximum (k_{max}) and it simultaneously integrates model selection and estimation.

Let X be a random variable (and x a sample) that follows a finite mixture PDF of K components where $K \in [1, k_{max}]$. Its PDF can be written as follows:

$$f_X(x) = \sum_{k=1}^K a_k f_{X_k}(x|\Theta_k) \tag{1}$$

where $a_k > 0$ and $\sum_{k=1}^K a_k = 1$ are the probabilities of belonging to each class of the mixture, and Θ_k are the parameters of the class \mathcal{C}_k , which follow a Gaussian or Gamma distribution. The probability of belonging to each class for each voxel (i.e. each sample x) is the following:

$$P(x \in \mathcal{C}_k | X = x, \Theta_k) = \frac{P(X = x | x \in \mathcal{C}_k, \Theta_k)P(x \in \mathcal{C}_k | \Theta_k)}{\sum_{i=1}^K P(X = x | x \in \mathcal{C}_i, \Theta_i)P(x \in \mathcal{C}_i | \Theta_i)} \tag{2}$$

This expression allows us to calculate a set of images that indicate the probability of belonging to each class for every voxel and can be seen as the partial volume contribution of each class. Fig 2 shows an ultrasound image, its finite mixture PDF, and the probability ownership to each class. In Fig 2b a good fit of the mixture model of the image in Fig. 2a becomes clear as well as the structure the ownership to each class (Fig. 2c).

The main advantage of this set of images is that a structure tensor can be derived from each probability image, which is able to extract the directions of minimal probability change even when low gradients are presented. This way, an anisotropic filter can be defined that takes the advantage of these directions in the probability map for each tissue class. We will refer to it as POSRAD (Probabilistic Oriented Speckle Reducing Anisotropic Diffusion). This filter makes use

of the extension of the scalar component of the diffusion equation to a diffusion matrix which takes the advantage of the local tissue orientation as it was proposed in [4,9]. However, in our case we calculate the local orientation of the probability density function for each tissue class and combine them in the following way.

As in [9], we calculate the structure tensor of the probability density function for each tissue class as:

$$T_k = G_\sigma * (\nabla_\sigma P(x \in \mathcal{C}_k | X = x, \Theta_k) \cdot \nabla_\sigma P(x \in \mathcal{C}_k | X = x, \Theta_k)^T) \quad (3)$$

where G_σ is a Gaussian kernel of standard deviation σ , and $\nabla_\sigma P(x \in \mathcal{C}_k | X = x, \Theta_k)$ is the gradient of the probability density function for each tissue class filtered with a Gaussian kernel of standard deviation σ . Finally, let $\lambda_1^k \geq \lambda_2^k \geq \lambda_3^k$ be the eigenvalues and $(\mathbf{v}_1^k, \mathbf{v}_2^k, \mathbf{v}_3^k)$ their respective eigenvectors. The local orientation of the maximal variation of probability of the class \mathcal{C}_k is given by \mathbf{v}_1^k , and the local orientation of the minimal variation is given by \mathbf{v}_3^k .

Let consider the following diffusion equation:

$$\begin{cases} u(x, 0) = u_0 \\ \frac{\partial u}{\partial t} = \text{div}(D \nabla u) \end{cases} \quad (4)$$

where the matrix D is expressed in a diagonal form with the eigenvectors $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ and eigenvalues $\lambda_1, \lambda_2, \lambda_3$. Since we have K structure tensors (each tissue class with probability density function), we choose the eigenbase of the structure tensor with maximal λ_1^k : $\hat{k} = \arg \max_k (\lambda_1^k)$. This base gives the orientation of the maximal variation of probability among all the classes. The interpretation of this choice is that we choose as boundary the one with the maximal gradient of the probability density function over all tissue classes. This way, the most probable boundary is preserved in the filtering process. In the basis of $T_{\hat{k}}$, namely $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, the diffusion matrix D is defined as:

$$D = E \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} E^T \quad \text{where} \quad \begin{cases} \lambda_1 = 1 - \|\nabla_{\mathbf{e}_1, \sigma} P(x \in \mathcal{C}_{\hat{k}} | X = x, \Theta_{\hat{k}})\|_2 \\ \lambda_2 = 1 - \|\nabla_{\mathbf{e}_2, \sigma} P(x \in \mathcal{C}_{\hat{k}} | X = x, \Theta_{\hat{k}})\|_2 \\ \lambda_3 = 1 \end{cases} \quad (5)$$

where $\|\cdot\|_2$ is the 2-norm, $\nabla_{\mathbf{e}_i, \sigma}$ is the directional derivative in the direction \mathbf{e}_i filtered with a Gaussian kernel with a standard deviation σ , and E is the matrix whose columns are the eigenvectors $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. Note that this definition performs a diffusion filtering in the direction of the minimal variation of probability (\mathbf{e}_3) while preserves the maximal variation of probability since $\|\nabla_{\mathbf{e}_1, \sigma} P(x \in \mathcal{C}_{\hat{k}} | X = x, \Theta_{\hat{k}})\|_2$ will have a value closed to 1. Note that the discrete approximations of $\|\nabla_{\mathbf{e}_1, \sigma} P(x \in \mathcal{C}_{\hat{k}} | X = x, \Theta_{\hat{k}})\|_2$ and $\|\nabla_{\mathbf{e}_2, \sigma} P(x \in \mathcal{C}_{\hat{k}} | X = x, \Theta_{\hat{k}})\|_2$ are bounded in $[0, 1]$, thus $\lambda_1, \lambda_2 \in [0, 1]$.

The main advantage of this definition is that, in homogeneous areas D becomes isotropic and in the presence of boundaries it becomes more anisotropic and with a main orientation aligned along the boundaries thus preserving the main contours. Additionally, the evolution of the probabilistic models can be followed through the iterative process since the finite mixture model is re-estimated

in each iteration. This is an important advantage of the presented methodology over other state-of-the-art filters, which assume that the statistics of the image/noise do not change during the filtering process. Since no assumptions can be made about the changes of the statistics of the image, a finite mixture of Gaussians is adopted during the filtering process except for the first iteration.

4 Experiments and Results

In this section we present some experiments with synthetic and real ultrasound images in order to compare our technique to other state-of-the-art algorithms. First we consider the 2D synthetic image presented in [4] which is publicly available[2] and allows us to compare the proposed method to other algorithms. This image is presented in Fig 3.a which has 6 regions of constant intensity. The image is corrupted with a Gaussian multiplicative noise with $\sigma_n = 0.5$ (Fig. 3.b). In Fig. 3.c the filtered image with the proposed method is presented. The parameters of the filter were: 200 iterations, time step 0.5, $\sigma = 1$ and $k_{\max} = 6$.

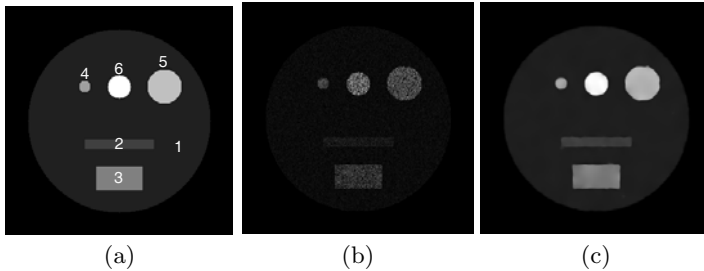


Fig. 3. (a) Original synthetic image of [4]. (b) Noisy image with multiplicative noise with $\sigma = 0.5$. (c) Filtered image with the proposed filter.

Table 1. Results on filtering synthetic 2D image with $\sigma_n = 0.5$, presented as Mean \pm Standard Deviation of the intensity in each region (R1 - R6) of the Fig. 3.a

Method	R1	R2	R3	R4	R5	R6	M_f	FOM
Initial	10	20	40	50	60	80		1
Noisy image	10.03 \pm 5.00	2.26 \pm 9.57	39.39 \pm 20.23	54.86 \pm 22.09	58.72 \pm 29.55	81.12 \pm 40.59		0.193
Median	10.00 \pm 0.65	18.58 \pm 1.96	37.42 \pm 4.77	38.42 \pm 7.53	57.40 \pm 4.89	75.80 \pm 9.77	0.01	0.434
Lee	10.09 \pm 1.42	19.34 \pm 1.93	38.55 \pm 5.64	48.66 \pm 7.27	56.18 \pm 7.20	77.82 \pm 11.96	0.02	0.224
Kuan	10.11 \pm 1.42	19.33 \pm 1.90	38.53 \pm 5.47	48.64 \pm 7.26	57.13 \pm 6.93	77.71 \pm 11.52	0.02	0.222
P&M	9.97 \pm 2.39	19.95 \pm 9.10	39.38 \pm 20.17	54.90 \pm 22.02	58.72 \pm 29.53	81.12 \pm 40.58	0.00	0.323
homotopic P&M	9.55 \pm 1.53	18.56 \pm 6.00	37.80 \pm 13.56	51.46 \pm 11.37	56.31 \pm 19.87	77.85 \pm 26.69	0.00	0.533
Cattie et al.	10.02 \pm 0.89	20.02 \pm 1.22	39.39 \pm 10.36	54.86 \pm 19.65	58.72 \pm 17.97	81.12 \pm 34.55	0.00	0.626
Flux	10.01 \pm 0.61	20.13 \pm 0.93	39.58 \pm 4.08	55.12 \pm 4.93	58.90 \pm 5.49	81.36 \pm 8.82	0.33	0.693
Rudin et al.	10.01 \pm 0.87	15.11 \pm 0.41	37.49 \pm 2.08	47.16 \pm 3.15	56.97 \pm 2.62	79.10 \pm 6.64	0.04	0.852
Rudin et al. attach	9.85 \pm 0.89	15.04 \pm 0.42	37.44 \pm 2.10	47.14 \pm 3.17	56.93 \pm 2.63	79.07 \pm 6.67	0.03	0.846
SRAD, expl., N_4	10.12 \pm 0.93	19.95 \pm 1.60	39.66 \pm 4.63	53.82 \pm 5.07	58.89 \pm 7.07	81.45 \pm 10.44	0.17	0.361
SRAD, expl., N_8	10.11 \pm 0.98	20.00 \pm 1.55	39.64 \pm 5.03	54.15 \pm 4.36	58.96 \pm 7.25	81.48 \pm 10.53	0.16	0.347
SRAD, impl., N_4	10.47 \pm 0.59	20.66 \pm 0.90	41.36 \pm 2.74	57.51 \pm 2.96	60.89 \pm 2.06	84.89 \pm 4.33	0.24	0.375
SRAD, impl., N_8	10.42 \pm 0.39	20.42 \pm 0.40	41.22 \pm 2.54	59.14 \pm 0.81	60.69 \pm 1.02	84.13 \pm 1.37	0.43	0.615
SRAD, impl., N_9	10.12 \pm 0.93	19.94 \pm 1.64	39.64 \pm 4.58	53.67 \pm 4.89	58.86 \pm 6.95	81.37 \pm 10.27	0.18	0.387
DPAD, expl., N_9	10.11 \pm 0.97	20.00 \pm 1.55	39.63 \pm 4.93	54.03 \pm 4.15	58.94 \pm 7.08	81.43 \pm 10.22	0.17	0.341
DPAD, impl., N_4	10.48 \pm 0.38	20.65 \pm 0.66	41.54 \pm 2.19	57.43 \pm 3.06	60.61 \pm 2.13	85.20 \pm 3.73	0.32	0.524
DPAD, impl., N_9	10.44 \pm 0.35	20.43 \pm 0.54	41.09 \pm 2.00	58.35 \pm 1.11	60.89 \pm 0.86	85.00 \pm 2.02	0.45	0.731
OSRAD	10.28 \pm 0.51	20.09 \pm 0.82	40.38 \pm 2.38	54.95 \pm 1.69	59.64 \pm 3.73	83.52 \pm 4.26	0.47	0.768
POSRAD	10.09 \pm 0.68	19.91 \pm 0.85	39.49 \pm 1.53	51.90 \pm 3.14	58.11 \pm 2.11	79.57 \pm 3.95	0.55	0.865

² <http://serdis.dis.ulpgc.es/~krissian/HomePage/Demos/OSRAD/OSRAD.html>

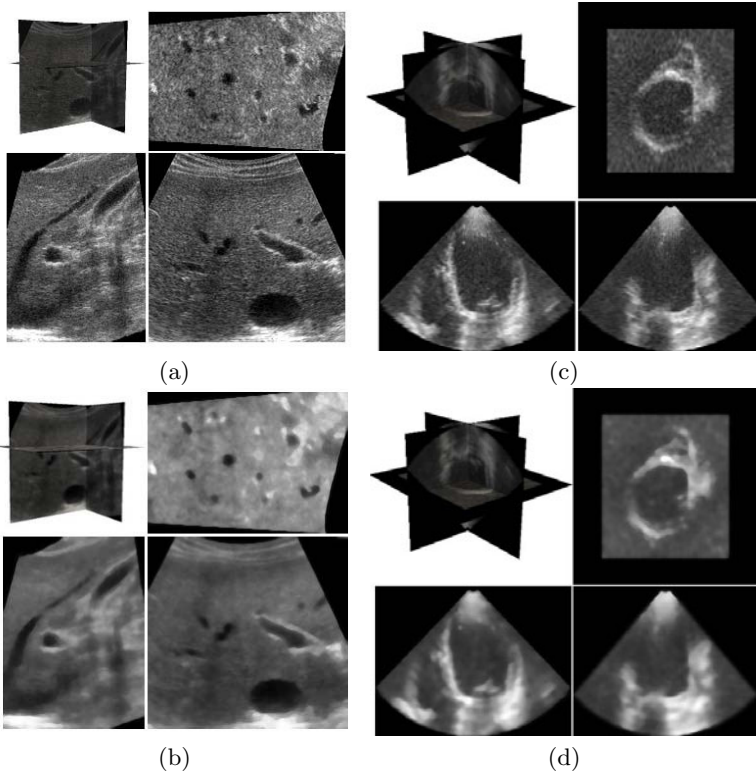


Fig. 4. (a) Noisy 3D ultrasound of a liver. (b) Filtered 3D ultrasound of a liver. (c) Noisy 3D ultrasound of a Heart. (d) Filtered 3D ultrasound of a Heart. The proposed method achieve a good edge preservation with an efficient noise removal in both cases.

No suppositions were made concerning the probability distributions of the tissue classes, so a Gaussian mixture model was used.

Since these parameters are the same used by OSRAD, a direct comparison with the results in [4] can be established. We use the same quality measures for validation: M_f and figure of merit (FOM)³, see [4] for more details. Both values are comprised between 0 and 1. The highest the value, the better the filter. Results presented in Table I show that the best FOM and M_f is achieved by our method. These results show that a better performance can be achieved when variations of the probability for each class is taken into consideration since the diffusion tensor becomes more anisotropic near the most probable edges while homogeneous regions are filtered in an isotropic way.

We now consider a 3D ultrasound image of a liver generated using the Stradwin v3.8 software (Cambridge University, Cambridge, UK), which has been used

³ The quality measure M_f was recalculated since it measures the quality of the filter compared to the others filters.

commonly in the literature [4]. These data were acquired by a freehand system and are available on Cambridge University⁴. The image is a volume with $201 \times 193 \times 142$ voxels with an isotropic resolution of $0.5 \times 0.5 \times 0.5 \text{ mm}^3$. Both, the original and the filtered image are presented in Fig 4 (a-b). Visual inspection shows that a good edge preservation is obtained with an efficient noise removal. The parameters for this experiment were: 200 iterations, time step 0.5, $\sigma = 1$ and $k_{\max} = 6$ with a Gaussian mixture model.

As a final experiment, we consider a real 3D ultrasonic image of the heart obtained with the same clinical machine of section 2. In this experiment we use the finite mixture model of Gaussian and Gamma distributions, 100 iterations, time step 0.5, $\sigma = 1$ and $k_{\max} = 4$. Results are presented in Fig 4 (c-d) were a good edge preservation is observed. An efficient noise removal is achieved in the left ventricle when compared to the noisy image.

5 Conclusion

In this work we have presented a new diffusion filter that takes into account the statistical properties of the image. This methodology offers an advantage over the follow-up of statistical properties of the image in each iteration while preserving and enhancing the structures. Experiments on synthetic and real images show that the proposed method preserves edges of the structures better than others and performs a good restoration.

References

1. Nillesen, M.M., Lopata, R.G., Gerrits, I.H., Kapusta, L., Thijssen, J.M., de Korte, C.L.: Modeling envelope statistics of blood and myocardium for segmentation of echocardiographic images. *Ultrasound Med. Biol.* 34(4), 674–680 (2008)
2. Yu, Y., Acton, S.: Speckle reducing anisotropic diffusion. *IEEE Trans. Img. Proc.* 11(11), 1260–1270 (2002)
3. Aja-Fernandez, S., Alberola-Lopez, C.: On the estimation of the coefficient of variation for anisotropic diffusion speckle filtering. *IEEE Trans. Img. Proc.* 15(9), 2694–2701 (2006)
4. Krissian, K., Westin, C.F., Kikinis, R., Vosburgh, K.: Oriented speckle reducing anisotropic diffusion. *IEEE Trans. Img. Proc.* 16(5), 1412–1424 (2007)
5. Tao, Z., Tagare, H.D., Beaty, J.D.: Evaluation of four probability distribution models for speckle in clinical cardiac ultrasound images. *IEEE Trans. Med. Imag.* 25(11), 1483–1491 (2006)
6. Eltoft, T.: Modeling the amplitude statistics of ultrasonic images. *IEEE Trans. Med. Imag.* 25(2), 229–240 (2006)
7. Wagner, R., Smith, S., Sandrik, J., Lopez, H.: Statistics of speckle in ultrasound B-scans. *IEEE Trans. Sonics Ultrason.* 30(3), 156–163 (1983)
8. Figueiredo, M., Jain, A.: Unsupervised learning of finite mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 381–396 (2002)
9. Krissian, K., Aja-Fernández, S.: Noise-driven anisotropic diffusion filtering of MRI. *IEEE Trans. Img. Proc.* 18(10), 2265–2274 (2009)

⁴ <http://mi.eng.cam.ac.uk/~rwp/stradwin/>

Patient Specific Models for Planning and Guidance of Minimally Invasive Aortic Valve Implantation

I. Waechter¹, R. Kneser¹, G. Korosoglou², J. Peters¹,
N.H. Bakker³, R.v.d. Boomen³, and J. Weese¹

¹ Philips Research Aachen, Germany

² Universitaet Heidelberg, Kardiologie, Germany

³ Interventional X-Ray, Philips Healthcare, Best, The Netherlands

Abstract. Recently, new techniques for minimally invasive aortic valve implantation have been developed generating a need for planning tools that assess valve anatomy and guidance tools that support implantation under x-ray guidance. Extracting the aortic valve anatomy from CT images is essential for such tools and we present a model-based method for that purpose. In addition, we present a new method for the detection of the coronary ostia that exploits the model-based segmentation and show, how a number of clinical measurements such as diameters and the distances between aortic valve plane and coronary ostia can be derived that are important for procedure planning. Validation results are based on accurate reference annotations of 20 CT images from different patients and leave-one-out tests. They show that model adaptation can be done with a mean surface-to-surface error of 0.5mm. For coronary ostia detection a success rate of 97.5% is achieved. Depending on the measured quantity, the segmentation translates into a root-mean-square error between 0.4 – 1.2mm when comparing clinical measurements derived from automatic segmentation and from reference annotations.

1 Introduction

Valvular heart diseases are among the most prominent causes of heart failure and premature cardiac death. Aortic valve stenosis is a very common valvular disease and aortic valve replacement in open surgery has been conducted on these patients for decades. This is, however, a very invasive and expensive treatment. In addition, it is considered too high risk or contraindicated for many patients [1]. In the last decade, techniques for minimally invasive aortic valve implantation have been developed [1,2] that offer a new treatment option. The artificial valve is mounted on a stent which is delivered through a catheter, either transfemoral, transsubclavian, or transapical, under X-ray guidance. Accurate assessment of the valve anatomy is essential as, e.g. occlusion of the coronary ostia by the valve leaflets or the rim of the stent is life threatening. Furthermore, methods to guide implantation are important as the aortic valve anatomy is not clearly visible when using X-ray imaging.

In this paper, we present a method to extract the aortic valve anatomy from CT images. The method is based on a model-based segmentation approach previously developed to segment the heart from CT images [3]. In addition, a new method for the detection of anatomical landmarks is presented that exploits the model-based segmentation. Within the reference model, the likely regions of the coronary ostia have been encoded to support their automatic detection. The aortic valve model refers to a heart phase with closed aortic valve as overpacing is used to stop the heart during implantation. In addition, it covers not only the aortic valve and bulbus, but also the aorta and the left ventricle, which are of relevance when using the transfemoral or transapical approach, respectively. The model differs in that respect from the work of Ionasec et al. [4] who have presented methods to generate a dynamic model of the aortic valve only.

Furthermore, we show how a number of measurements that are important for stent selection and procedure planning (see [5] for relevant measurements) can be derived from the model and information encoded therein. Examples are the distance of the coronary ostia to the valve plane and diameter measurements along the outflow tract, the aortic bulbus, and the ascending aorta. Extraction of the measurements is done automatically, which distinguishes the work from Gesat et al. [6] who presented a planning system for transapical valve implantation relying on user selected anatomical landmarks.

The paper is structured as follows: In section 2, we present our approach for the segmentation of the aortic valve and the new anatomical landmark detection method that we use to localize the coronary ostia. In addition, we describe how we determine the planes for measurements and how we compute the desired measurements. Section 3 includes validation results for the segmentation accuracy, landmark detection, and measurement accuracy.

2 Method

2.1 Model-Based segmentation

Previously, a method for segmenting the heart chambers and the great vessels [3,7] was presented. The method detects the heart with an adapted Generalized Hough Transform and positions a generic heart model in the image. Mesh adaptation is done by iterating boundary detection and mesh deformation.

The pose of the heart chambers is improved by minimizing the external energy E_{ext} with respect to the parameters of a global similarity transformation during mesh deformation. After 20 iterations, the point of gravity of the most distal descending aorta segment is computed, and the descending aorta is detected in the corresponding axial image slice. Another 20 iterations are performed to refine adaptation of the heart chambers and the most distal descending aorta segment, where a linear transformation is assigned to each model part and the external energy E_{ext} is minimized with respect to the parameters of the multi-linear transformation. Finally, 30 iterations of deformable adaptation are performed. The tubular segments building the great vessels are successively activated during these iterations. Mesh deformation is done by adding shape constraints to

the energy $E = E_{ext} + \alpha E_{int}$. In particular, the internal energy penalizes deviations from the generic heart model undergoing a multi-linear transformation to account for pose and shape variability.

For acceleration, mesh adaptation is largely performed with a reduced mesh resolution. In addition, already well-adapted mesh parts are frozen during deformable adaptation. Robustness is achieved by assigning optimal boundary detection functions with Simulated Search [8].

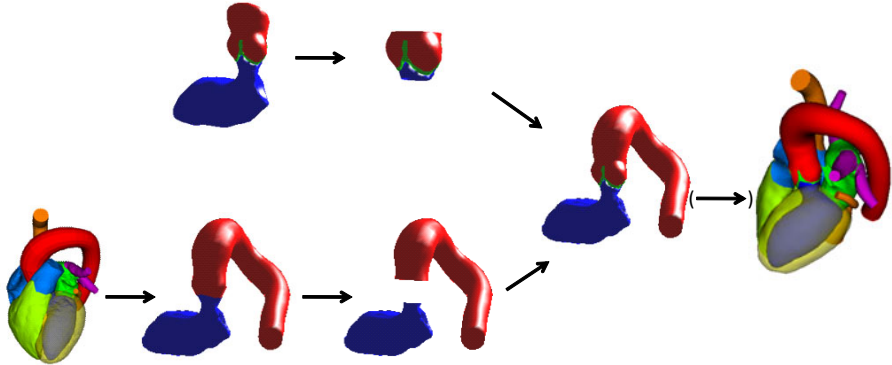


Fig. 1. Merging two models: The upper model contains the detailed aortic valve and aortic bulbus. These structures are cut and inserted into the lower, whole heart model.

Within the existing heart model, the aortic valve and aortic bulbus have only roughly been modelled. Therefore, a more detailed model of the aortic valve, aortic bulbus, and left ventricle was created. For that purpose, these structures were manually annotated (and reviewed by a clinician) in 20 CT datasets resulting in 20 reference meshes. These were used to build a mean mesh model and to train boundary detection functions by simulated search [8]. The aortic valve model was subsequently integrated into the heart model (see Fig. 1). For that purpose, both models were adapted to the same patient to bring them in the same coordinate system. Both meshes were cut with identical cut planes, one below the aortic valve and one above the aortic bulbus. Finally, the aortic valve model was attached to the left ventricle and the aorta of the heart model. The other parts of the heart model can be added if required by the application. All mesh-related information, like anatomical affiliations and boundary detectors, were transferred from one of the models to the merged model.

2.2 Ostia Finder

The shape, diameter, appearance, and position of the coronary ostia vary substantially between patients (see Fig. 2). This makes it difficult to detect them: including the ostia directly in the model is difficult as their position varies too much; a landmark detection method is also difficult as similar structures in the dataset may lead to many false detection results. We propose, therefore, a new approach. The basic idea is to use information from model-based segmentation

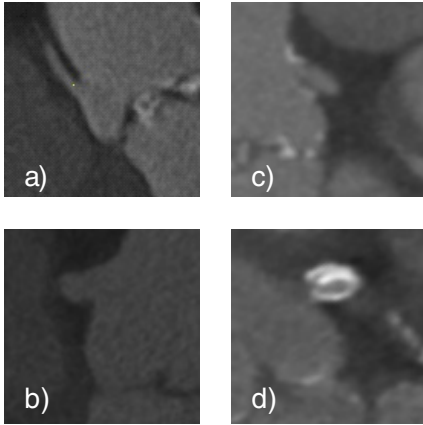


Fig. 2. Examples of Ostia: a) thin and pointing straight up b) thick and curving downwards c) pointing straight down, very near to aortic bulbus d) poor visibility of first part then calcified

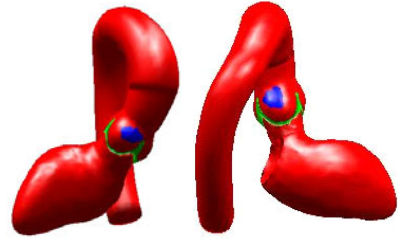


Fig. 3. Ostia candidate patches

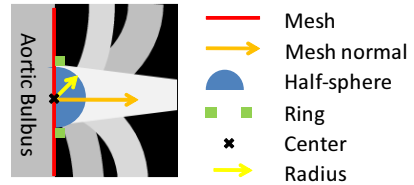


Fig. 4. Ostia search pattern with potential shapes of ostia

and to restrict landmark detection for the coronary ostia to a well-defined area on the surface of the aortic bulbus.

To encode the area for coronary ostia detection in the model, the mesh vertex closest to a coronary ostium is determined and the triangles linked to the vertex are marked. This is done for all manually determined coronary ostia positions in the training images. Fig. 3 shows the resulting candidate patches of both coronary ostia. They have a diameter between 1cm and 1.5cm.

Coronary ostia detection itself presumes that model-based segmentation has been performed. The candidate patch is determined for the actual image and pattern matching is performed on the surface of the patch. The pattern consist of a half-sphere $S(c, r)$ and a ring $R(c, r)$ (see Fig. 4), where r is the radius of the half-sphere and the inner radius of the ring and c is the centre of both. The orientation of the half-sphere and the ring is given by the mesh normal at the centre c . Assuming that the correct ostium position is characterized by a high image intensity inside the ostium that is surrounded by a ring with low image intensity, the likelihood for a coronary ostium is computed according to

$$L(c, r) = \frac{\sum_{x \in S(c, r)} \min\{I(x), I_{AB}\}}{\sum_{x \in S(c, r)} 1} - \alpha \cdot \frac{\sum_{x \in R(c, r)} \min\{I(x), I_{AB}\}}{\sum_{x \in R(c, r)} 1}, \quad (1)$$

where $I(x)$ is the image intensity at voxel x , I_{AB} denotes the average image intensity in the aortic bulbus and α is a weighting factor. The threshold I_{AB} is used to reduce the influence of calcifications and it is computed using the model-based segmentation result. The coronary ostium position is found with a complete search on the candidate patch while varying r between 2mm and 4mm. The weighting factor α was chosen to be 0.5.

2.3 Planes for Measurements

For computing the measurements for procedure planning and providing supporting functionality such as automatic reformatting, a number of planes must be defined. These planes are derived from the adapted shape model and information encoded therein.

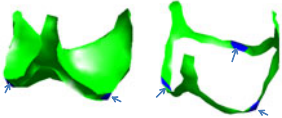


Fig. 5. Aortic valve plane landmarks, encoded on mesh

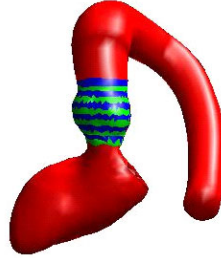


Fig. 6. Rings encoded on bulbus

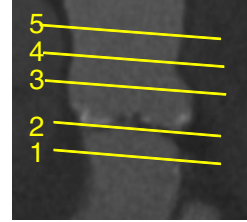


Fig. 7. Planes for diameter measurements

The aortic valve plane is defined as the plane which touches all three leaflets of the valve from beneath. To facilitate its computation, three landmarks are placed on the mesh on the basal ring of the aortic annulus (see Fig. 5). After segmentation by model adaptation, the aortic valve plane is computed from the position of the three landmarks.

To define the planes for diameter measurements of the left ventricular outflow tract (1), aortic valve annulus (2), the middle of the aortic bulbus (3), the sinutubular junction (4) and the ascending aorta (5), several rings are encoded on the mesh (see Fig. 6, 7). After model adaptation, planes are derived from respective rings by regression analysis.

2.4 Measurements

Diameters along aortic bulbus. The diameter information is important for selecting the proper stent [5]. The approach proposed for the measurements differs from common clinical practice, where they are currently derived from manually adjusted reformats or coronal slices of the CT image. Manual slice selection in combination with the non-circular cross-sections bears, however, a significant risk of over- or underestimating the respective quantity. The proposed 3D-based approach aims at delivering more consistent measurements.

To characterize the diameter, the mesh is cut by the respective plane. Then, an ellipse [9] or an inner and outer circle is fitted to the vertices of the resulting contour. The ventricular outflow tract and the aorta are normally elliptical, so here an ellipse is fitted using direct least square fitting (see Fig. 8, left). The aortic annulus and bulbus, however, have the shape of a rounded triangle. Here, the radius of the inner and outer circle are estimated from the vertices of the resulting contour (see Fig. 8, right).

Distance between ostia and aortic valve plane. The distance between ostia and aortic valve plane is important to evaluate the risk of ostia occlusion. The valve leaflets are pressed against the wall of the aortic bulbus by the implanted aortic valve and it is essential to verify that they cannot occlude the coronary ostia [5]. The difference of the distance of the ostia to the valve plane and the outer radius of the valve annulus can be used to assess the risk. Additionally, the distance is important for the decision, whether a specific stent can be implanted, where the upper part of the stent could occlude the coronary ostia [5].

Given the location of the ostia and the aortic valve plane, computation of the distances between the ostia and the aortic valve plane is straightforward (see Fig 9). The right image of Fig 9 also shows why it is difficult to determine the valve plane manually in 2D slices: As the valve is not a planar structure it touches the plane only at three points. If a slice like the one in the image is chosen to select the plane manually, the result is not correct.

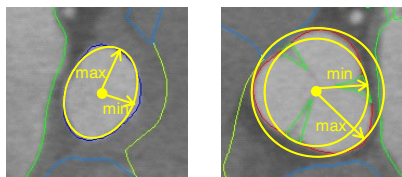


Fig. 8. Left: Fitted ellipse with short and long axis, right: fitted inner and outer circle

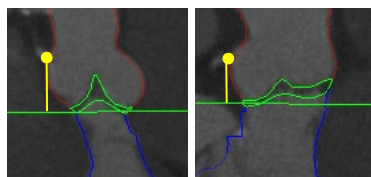


Fig. 9. Distance ostia to aortic valve plane

3 Evaluation Setup and Results

For the evaluation, a leaving-one-out study was conducted. 20 CT datasets (9 with calcified valves) with reference segmentations were available for this study. Always 19 of the 20 dataset were used to train the model. This includes computation of the mean shape, boundary detectors and the ostia search regions. The model was then used to segment the remaining dataset and the measurements were conducted as described above. Then, the mean surface-to-surface error [3] of the automatic segmentation and the reference segmentation were computed and the measurements from the automatic segmentation were compared with the measurements derived from the reference segmentation.

The segmentation was successful in all 20 datasets. The mean surface-to-surface error for left ventricular outflow tract, aortic valve and aortic bulbus was $0.5 \pm 0.4\text{mm}$. Example images of the segmentation are shown in Fig. 10. The localization of the ostia was successful for 39 out of 40 ostia (20 right, 20 left ostia). For the failure case, the landmark was detected in the coronary artery but not in the centre of the ostium. For the remaining 39 ostia, the left coronary ostium was detected with root mean squared (RMS) error of $1.2 \pm 0.6\text{mm}$ and the right coronary ostium with an RMS error of $1.0 \pm 0.8\text{mm}$. The valve plane was always detected successfully. The centre of the valve annulus was detected

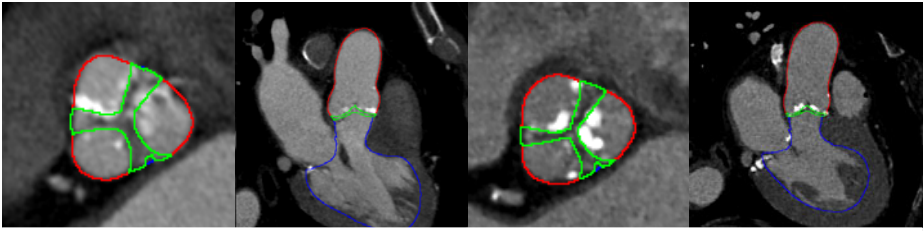


Fig. 10. Examples of aortic valve segmentation

with a RMS error of $0.8 \pm 0.3\text{mm}$ and the normal of the plane had an RMS error of $1.8 \pm 1.0^\circ$.

For the different proposed measurements, the range of values (mean, stdev, min, max) was computed on the reference meshes and the RMS errors between the measurements derived from the automatically segmented data and the measurements derived from the reference meshes were computed. An overview of the measurements is given in Table [1](#).

Table 1. Evaluation of measurements (all values in mm)

Measurement	mean	stdev	min	max	RMS Error
Distance left ostium - aortic valve plane	16.6	3.3	10.4	22.7	0.9
Distance right ostium - aortic valve plane	17.5	2.8	13.2	22.0	0.6
Short axis diameter, LV outflow tract	19.7	1.9	15.5	23.9	1.0
Long axis diameter, LV outflow tract	28.2	2.3	23.5	33.7	1.2
Inner diameter, aortic valve annulus	22.6	2.2	18.3	27.4	0.8
Outer diameter, aortic valve annulus	28.0	2.1	24.4	32.2	1.0
Inner diameter, mid bulbus	29.6	3.0	22.8	35.2	0.6
Outer diameter, mid bulbus	35.8	3.9	28.3	43.3	0.7
Short axis diameter, sinutubular junction	26.4	2.5	21.7	31.0	0.7
Long axis diameter, sinutubular junction	28.2	3.0	22.1	32.9	1.0
Short axis diameter, ascending aorta	27.1	2.2	22.3	31.7	0.5
Long axis diameter, ascending aorta	28.5	2.1	23.8	32.9	0.4

4 Conclusions

We have presented a model-based approach for the extraction of the aortic valve geometry from CT images and showed on 20 CT data sets that model adaptation can be done with a surface-to-surface error of 0.5mm. Furthermore, we presented a new approach for the detection of the coronary ostia that applies a simple pattern search method in regions encoded on the model. Accurate model adaptation in combination with the locality introduced in this way avoids misleading detection results and enables coronary ostia detection with a success rate of 97.5%. We believe that this approach can be generalized and also be used for the detection of other anatomical landmarks.

Furthermore, the resulting aortic valve model and information encoded therein was used to derive a number of measurements for planning minimally invasive aortic valve procedures. Compared to the measurements derived from the reference annotations, automatic segmentation resulted in RMS errors of the measurements between 0.4 and 1.2mm. Respective errors are comparable or smaller than the difference reported, for instance, for diameter measurements of the aortic valve annulus between CT and transthoracic echo ($1.22 \pm 1.3\text{mm}$) or CT and transesophageal echo ($1.52 \pm 1.1\text{mm}$) [10]. As inaccuracies in sizing the aortic valve annulus can result in various complications such as paravalvular leaks, valve embolization or migration, centric valvular leaks, premature leaflet deterioration or post-procedure heart block, we think that our approach has not only the potential to improve consistency of clinical measurements, but may also contribute to an improved clinical outcome.

In the future, the model may also be used to support interventional guidance. Similar as proposed for electrophysiological procedures, the aortic valve model may be overlaid onto X-ray images for that purpose. Another approach to support guidance would be the combination with methods to localize and track the aortic valve prosthesis as proposed by Karar et al. [11].

References

1. Cribier, et al.: Early experience with percutaneous transcatheter implantation of heart valve prosthesis for the treatment of end-stage inoperable patients with calcific aortic stenosis. *J. Am. Coll. Cardiol.* 43(4), 698–703 (2004)
2. Walther, et al.: Transapical minimally invasive aortic valve implantation: multicenter experience. *Circulation* 116(11 Suppl.), 1240–1245 (2007)
3. Ecabert, et al.: Automatic model-based segmentation of the heart in CT images. *IEEE TMI* 27(9), 1189–1201 (2008)
4. Ionasec, et al.: Dynamic model-driven quantitative and visual evaluation of the aortic valve from 4D CT. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part I. LNCS*, vol. 5241, pp. 686–694. Springer, Heidelberg (2008)
5. Tops, et al.: Noninvasive evaluation of the aortic root with multislice computed tomography implications for transcatheter aortic valve replacement. *JACC Cardiovascular Imaging* 1(3), 321–330 (2008)
6. Gessat, et al.: A planning system for transapical aortic valve implantation. *SPIE Medical Imaging* 7261, 72611E.1–72611E.12 (2009)
7. Peters, et al.: Segmentation of the heart and major vascular structures in cardiovascular CT images. *SPIE Medical Imaging* 6914, 691417.1–691417.12 (2008)
8. Peters, et al.: Optimizing boundary detection via simulated search with applications to multi-modal heart segmentation. *Medical Image Analysis* 14(1), 70–84 (2010)
9. Fitzgibbon, et al.: Direct least square fitting of ellipses. *IEEE Trans. PAMI* 21(5), 476–480 (1999)
10. Messika-Zeitoun, et al.: Multimodal assessment of the aortic annulus diameter: implications for transcatheter aortic valve implantation. *J. Am. Coll. Cardiol.* 55(3), 186–194 (2010)
11. Karar, et al.: Localization and tracking of aortic valve prosthesis in 2D fluoroscopic image sequences. *SPIE Medical Imaging* 7261, 72611Q.1–72611Q.8 (2009)

A Framework for Using Diffusion Weighted Imaging to Improve Cortical Parcellation*

Matthew J. Clarkson^{1,2}, Ian B. Malone², Marc Modat¹, Kelvin K. Leung^{2,1}, Natalie Ryan², Daniel C. Alexander¹, Nick C. Fox², and Sébastien Ourselin^{1,2}

¹ Centre for Medical Image Computing, University College London, WC1E 6BT, UK

² Dementia Research Centre, UCL Institute of Neurology, London, WC1N 3BG, UK

Abstract. Cortical parcellation refers to anatomical labelling of every point in the cortex. An accurate parcellation is useful in many analysis techniques including the study of regional changes in cortical thickness or volume in ageing and neurodegeneration. Parcellation is also key to anatomic apportioning of functional imaging changes. We present preliminary work on a novel algorithm that takes an entire cortical parcellation and iteratively updates it to better match connectivity information derived from diffusion weighted imaging. We demonstrate the algorithm on a cohort of 17 healthy controls. Initial results show the algorithm recovering artificially induced mis-registrations of the parcellation and also converging to a group-wise average. This work introduces a framework to investigate the relationship between structure and function, with no a-priori knowledge of specific regions of interest.

1 Introduction

Cortical parcellation is a labelling of every point in the cortex [5]. This labelling delineates regions throughout the cortex, enabling subsequent analysis on a per-region basis, and is an essential tool for neuroimaging research. An accurate cortical parcellation is of use when analysing properties such as thickness [4] or white matter connectivity [7]. As the parcellation is a key behind-the-scenes component, it is essential for understanding the underlying anatomy and both the cross-sectional and longitudinal changes that occur in neurodegenerative diseases.

Using magnetic resonance imaging (MRI), previous parcellation methods have looked at labelling parts of the cortex, such as prominent sulci, to sub-divide the volume of each hemisphere [10], or using watershed techniques to label all sulci [8]. Subsequently, Fishl et. al. [5] developed a comprehensive method to label

* This work was undertaken at UCL/UCLH which received a proportion of funding from the Department of Health's NIHR Biomedical Research Centres funding scheme. The Dementia Research Centre is an Alzheimer's Research Trust coordinating centre. KKL is supported by TSB grant M1638A. MC is supported by TSB grant M1638A and UCLH/UCL CBRC grant 168. NCF is funded by the Medical Research Council (UK).

all voxels in the cortex. This method, based on a probabilistic atlas derived from manually labelled datasets is incorporated into the widely used FreeSurfer software package.

However, it is known that the topographical boundaries in a cortical parcellation may not be visible in structural MRI scans and only have a limited correspondence to underlying cytoarchitecture [1]. Recently it has been shown that connectivity profiles derived from diffusion data can elucidate functionally distinct regions in the cortex [6,2]. These methods infer a parcellation directly from diffusion weighted imaging (DWI) data by clustering together voxels with similar connectivity. Thus far, these voxel based methods have been applied to regions of the brain such as the supplementary motor cortex (SMA) and pre-SMA [6] and also to define 9 subregions of the cingulate cortex [2]. Most recently, surface based methods have been developed to register brains based on regional connectivity [9] and also analyse which areas have sufficient connectivity to drive a parcellation [3]. We describe work that combines DWI and volumetric MRI data to modify an existing parcellation of the whole cortex, with the hypothesis that a parcellation based on functional connectivity provides a parcellation that is more functionally and cytoarchitecturally correct when compared with conventional topology driven parcellation. Importantly we first asked whether these approaches give concordant or different results, without pre-judging which is better since this judgement may be dependent upon the use to which the parcellation is put.

2 Methods

We first describe several pre-processing steps, before describing the proposed algorithm. For these experiments, we use the cortical parcellation, and spherical surface model produced by FreeSurfer [4] [5] and probabilistic tractography implemented in MRtrix [2] [11], however, the proposed algorithm could be applied to any triangulated surface model representing the cortex with pointwise connectivity established. Furthermore, we pre-process our T1-weighted (T1w) images using the FreeSurfer pre-processing pipeline. The FreeSurfer pre-processing pipeline includes steps that linearly resample to 1mm iso-tropic voxels, performs N3 intensity inhomogeneity correction, registers to Talairach space and then performs another white matter specific intensity normalisation before continuing on to perform a cortical thickness and volume based processing stream. These FreeSurfer pre-processing steps are not discussed as part of this paper.

2.1 Combining Structural and Connectivity Information

For each subject's T1w scan, we run the FreeSurfer cross sectional pipeline, and extract the triangulated white matter surface, the inflated spherical surface and the 36 anatomical labels, for both the left and right side of the brain.

¹ <http://surfer.nmr.mgh.harvard.edu/fswiki>

² <http://www.brain.org.au/software/>

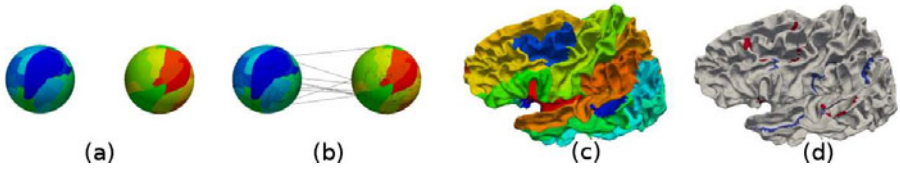


Fig. 1. Overview of method. (a) Spherical representations of the left and right cortical surfaces are extracted using FreeSurfer. (b) Connectivity information from MRtrix is added (only the first 500 tracts are shown). (c) The parcellation boundaries are optimised using the bi-spherical model, and then mapped back to the white matter surface. (d) The difference can be visualized, and the updated parcellation used in further cortical analysis.

FreeSurfer guarantees the ordering of points listed in the white matter surface matches the ordering of points in the corresponding inflated spherical surface, for each side, thereby enabling simple mapping between the two. The left and right spherical mesh models were combined into a single file³, with each sphere offset by $\pm(100, 100, 100)$ mm to aid visualization, and the left hemisphere labelled with anatomical regions 0-35, and the right hemisphere labelled with the same anatomical regions, but with labels 36-71, illustrated in Fig. 1(a).

For each subject's DWI scans, Fibre Orientation Distributions (FODs) were computed by Constrained Spherical Deconvolution with a maximum harmonic degree of eight^[11], using the MRtrix package. Probabilistic tractography was performed by seeding 100,000 tracts within a cortical mask using MRtrix with default parameters. The tracts were then filtered to only store the endpoints of tracts that started and finished within the cortical mask (i.e. those that connect grey matter). Typically this resulted in circa. 70,000 tracts per subject.

The diffusion data was registered (9 degrees-of-freedom) to the T1w scans, and the tracts and FreeSurfer white matter surfaces transformed into T1w image space. For each end of each tract, we find the closest point in either the left or right white matter surface, and then the corresponding point in the left or right spherical models. Each tract is thus attached to the spherical model, and is illustrated in Fig. 1(b). In practice, the number of triangles in the white matter mesh is far greater than the number of tracts, leading to a sparsely populated model. Thus we subsample each spherical model, and re-attach the tracts to the nearest spherical surface point. The output of the pre-processing is a single bi-spherical model containing two labelled spheres representing the left and right cortical hemispheres and a large number of lines representing the connectivity, as determined from the tractography.

2.2 Cortical Parcellation Optimisation Algorithm

The algorithm we propose is a group-wise optimisation, designed to update each subject's cortical parcellation, to achieve consistency between subjects in the

³ <http://www.vtk.org>

within subject connectivity. We desire a mean connectivity profile to remove bias caused by any one individual. Implicitly, we are assuming that the group of subjects should have similar connectivity. This may not be the case under various disease conditions, so to investigate the framework, we concentrate on healthy controls.

Consider a dataset of P points in \mathbb{R}^3 , where with each point $\mathbf{p}_{i=1\dots P}$ we associate a single label $l_{j=1\dots L}$. Consider also that the dataset contains T lines, denoted by $t_{k=1\dots T}$ where $t_k^{i,i'}$ denotes the fact that the k -th line connects point i with i' and $l(t_k^i)$ and $l(t_k^{i'})$ denote the label values associated with line k at points i and i' respectively. For a given configuration of labels, we can compute a symmetric 2D histogram H , where each bin $H(l_j, l_{j'})$ contains the number of label l_j connected to label $l_{j'}$ via the set of T lines. Thus:

$$H(l_j, l_{j'}) = \frac{1}{T} \sum_{k=1}^T \begin{cases} 1 & \text{if } l(t_k^i) = l_j \text{ and } l(t_k^{i'}) = l_{j'} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Thus, for a given bi-spherical model (see Fig. 1(b)), we can compute a 2D histogram of label connectivity. Furthermore for N 2D histograms of label connectivity, we define the mean connectivity histogram as

$$\hat{H}(l_j, l_{j'}) = \frac{1}{N} \sum_{n=1}^N H_n(l_j, l_{j'}) \quad (2)$$

Finally for two 2D histograms H_n and $H_{n'}$, we define a sum of squared difference histogram similarity measure as

$$S(H_n, H_{n'}) = \sum_{j=1}^L \sum_{j'=1}^L (H_n(l_j, l_{j'}) - H_{n'}(l_j, l_{j'}))^2 \quad (3)$$

Algorithm 1 outlines a single subject optimisation that matches a single subject's bi-spherical model to a label connectivity histogram. This could be used in a stand alone sense if the label connectivity histogram was known a-priori from another process.

The overall optimisation proceeds in a group-wise fashion, and is shown in Algorithm 2. For a group of N bi-spherical models, an average connectivity histogram is calculated, and then for each subject, the parcellation is optimised using Algorithm 1 to drive towards the average. Within this combined framework, there are many alternatives that could be suggested, some of which are outlined in section 3. Fig. 1(c) and (d) show the labels mapped to the cortical surface, and after an update shows the difference image, illustrating updated region boundaries.

2.3 Experiments and Results

17 healthy control subjects were selected, and each scanned on a 3T Siemens TrioTim scanner, with a 32 channel head coil, using an MPRAGE sequence

Algorithm 1. Individual subject optimisation

1. Until a maximum number of iterations
 2. Evaluate $S(H_n, \hat{H})$
 3. For each point \mathbf{p}_i find neighbours with different labels
 4. If the number of differing labels > 0
 5. For each differing label calculate $\nabla S(H_n, \hat{H})$ using eqn (3) by swapping label
 6. Store the label l_i that minimises eqn (3) in a list
 7. Update labels using list created in step 6
 8. Evaluate $S(H_n, \hat{H})$. If this is larger than step 2, stop
 9. Goto 1.
-

Algorithm 2. Group-wise optimisation

1. Until a maximum number of iterations
 2. For each subject $n = 1 \dots N$
compute histogram $H_n(l_j, l_{j'})$ using eqn (1)
 3. Compute the mean connectivity histogram using eqn (2)
 4. For each subject $n = 1 \dots N$
optimise the parcellation values using Algorithm 1
 5. Goto 1.
-

(TE=2.9ms, inversion interval TR=2200ms, TI=900ms). T1-weighted volumetric images were obtained with a 28.2-cm field of view and 256×256 acquisition matrix to provide 208 continuous 1.1mm thick slices. DWI images were obtained (TE=91ms, TR=6800ms, 64 gradient directions, $b=1000 \text{ smm}^{-2}$, and $8 \times B_0$ images) with a 24-cm field of view and 96×96 acquisition matrix, providing 55 continuous 2.5mm slices. FreeSurfer's cortical processing stream was run on each subject's T1 image, and MRtrix's tractography performed using a single B_0 image and 64 gradient directions.

For each of the 17 subjects, the histogram of label connectivity was generated. The labels were then mis-registered by selecting every other region (i.e. regions 0, 2, 4...70) in turn, and updating the neighbours of each boundary point to that region label, similar to dilating a boundary in an intensity image. We did this 1, 2...5 times, resulting in 5 datasets of 17 mis-labelled models. We then used Algorithm 1 to optimise the labels, on a per-subject basis, to match them to their original label connectivity histogram. We visually inspected the results. For all the levels of mis-registration, and for all subjects, the algorithm moved the boundaries of the regions. It could be seen that in some cases the apparent mis-registration increased. Even in this case, there is no gold standard, as the diffusion data may suggest a different parcellation to the initial topographical parcellation.

Fig. 2 shows a visualisation of the results for a single subject. Fig. 2(a) shows the bi-spherical model showing left and right hemispheres, with parcellation labels attached. Fig. 2(b) shows a difference image, the thick blue and red lines

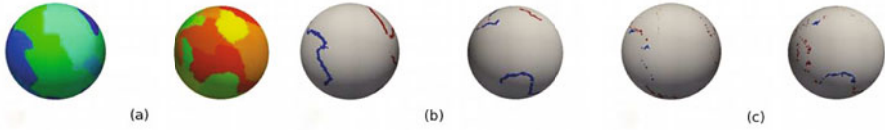


Fig. 2. Results of Algorithm 1 (a) Spherical representations of the left and right cortical surfaces. (b) A difference image showing the change in region label before running Algorithm 1 (c) A difference image showing the change in region label after running Algorithm 1

shows boundaries that have been moved. Fig. 2(c) shows the boundaries after the optimisation algorithm has run. It can be seen that Algorithm 1 has reduced the size of the thick blue and red lines, indicating that the algorithm converged back towards the original label connectivity histogram. Fig. 2(c) also shows that some noise has been introduced, most likely caused by lack of regularisation at the boundary.

Then, for all 17 subjects the full group-wise optimisation Algorithm 2 was run. The aim of the experiment was to see how well Algorithm 2 converged, and to see the effects of the diffusion imaging. Again we have no gold standard, as we are hypothesising that the parcellation should change to something that's only visible on diffusion images, which have not been manually segmented. We ran the main loop, Algorithm 2 for 5 iterations, by which point the internal loop running Algorithm 1 had converged. Fig. 3 illustrates the movement of the parcellation boundary on a single subject. Once the groupwise registration had converged, the difference in the boundary can be seen.

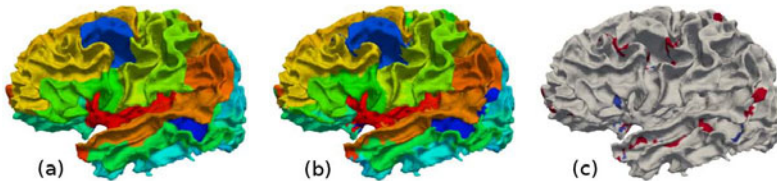


Fig. 3. Results of Algorithm 2 (a) Initial (FreeSurfer) atlas based parcellation. (b) After groupwise optimisation using Algorithm 2, the boundaries have moved. (c) A difference image showing the difference between (a) and (b).

3 Discussion and Conclusion

We have demonstrated an algorithm to update a cortical parcellation using diffusion imaging. This work shows promising and interesting results. Importantly it shows that with a simple optimisation scheme, tract based connectivity does indeed alter an atlas based topographical parcellation, implying that a topography

based parcellation is not entirely concordant at a functional level between individuals. This is not at all surprising given the well described differences between structure, function and cytoarchitecture correlates [6,2] and may be particularly relevant in functionally driven applications. While we are aware that we have only performed visual inspection, it is already interesting to see that there may be some regions where the boundary only moves a little, indicating a good concordance between structure and connectivity, and regions where the boundary moves a lot. Further work will investigate the reasons for this.

There are many opportunities for improvement of this algorithm, both on a purely technical basis, and on a more philosophical one. We started this work, inspired by the Spherical Demons paper [12]. We aimed to compute a force at each vertex along the boundary. The algorithm attaches connectivity lines directly to points to avoid repeatedly searching for the closest line. However, it is difficult to match the resolution of the tractography to the resolution of the spherical mesh. This results in a sparse number (approx 20-30%) of boundary points having connection data. For these results, we down-sampled the spherical mesh, using larger triangles, and attached the connections to the closest available point. This increased the percentage of points having connections, but effectively reduced the resolution. We need to study the effect of different tract seeding strategies, the number of tracts, and the masks used.

Our optimisation was simple, and could certainly be improved. We have not yet implemented any form of regularisation which leads to small islands appearing in the parcellation. This could be remedied by maintaining an explicit boundary model, enforcing boundaries are not broken. Alternatively, rather than optimise on a point by point basis, a regular grid of control points could be used. Furthermore, the connectivity histogram is complex, and could be smoothed [9] or a more global optimisation scheme adopted.

The question of how closely structure and function should match, on an individual and group-wise basis, is an interesting one. It has been shown that there can be significant differences between subjects [1]. It may be better to simply deal with each subject individually. The voxel based methods described in [6,2] provide a way to search for clusters, purely within the tractography information with no a priori knowledge of the existence of any parcellation. A combination of an existing parcellation from an atlas, and also a clustering technique on the tracts rather than a clustering on the connectivity histogram may provide a stronger indicator of distinct regions.

For applications where the scale of interest is large, such as thickness measurements over large regions of the cortex, then deriving fine grained boundary detail may provide little statistical difference. Techniques such as this and [6,2] may provide greater insight in areas where from a clinical perspective there is a need to sub-divide a given region. Another interesting area of study would be in longitudinal analysis. Over time, changing regional boundaries may confound atlas based parcellation algorithms. Using diffusion imaging and a model of how a disease process might affect tracts over time, may provide a way for connectivity information to iteratively update the parcellation boundaries.

In conclusion, we have demonstrated a prototype framework to iteratively update an atlas based parcellation using diffusion imaging. Initial results prove promising, and potentially open up many more areas of research.

References

1. Amunts, K., Malikovic, A., Mohlberg, H., Schormann, T., Zilles, K.: Brodmann's areas 17 and 18 brought into stereotaxic space-where and how variable? *Neuroimage* 11(1), 66–84 (2000)
2. Beckmann, M., Johansen-Berg, H., Rushworth, M.F.S.: Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *J. Neurosci.* 29(4), 1175–1190 (2009)
3. Cathier, P., Mangin, J.F.: Registration of cortical connectivity matrices. In: Conference on Computer Vision and Pattern Recognition Workshop, CVPRW 2006, p. 66 (17-22, 2006)
4. Fischl, B., Dale, A.M.: Measuring the thickness of the human cerebral cortex from magnetic resonance images. *PNAS* 97(20), 11050–11055 (2000)
5. Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M.: Automatically parcellating the human cerebral cortex. *Cereb Cortex* 14(1), 11–22 (2004)
6. Johansen-Berg, H., Behrens, T.E.J., Robson, M.D., Drobnjak, I., Rushworth, M.F.S., Brady, J.M., Smith, S.M., Higham, D.J., Matthews, P.M.: Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex.. *Proc. Natl. Acad. Sci. U. S. A.* 101(36), 13335–13340 (2004)
7. Park, H.J., Kim, J.J., Lee, S.K., Seok, J.H., Chun, J., Kim, D.I., Lee, J.D.: Corpus callosal connection mapping using cortical gray matter parcellation and DT-MRI. *Hum. Brain Mapp.* 29(5), 503–516 (2008)
8. Rettmann, M.E., Han, X., Xu, C., Prince, J.L.: Automated sulcal segmentation using watersheds on the cortical surface. *Neuroimage* 15(2), 329–344 (2002)
9. Roca, P., Rivière, D., Guevara, P., Poupon, C., Mangin, J.F.: Tractography-based parcellation of the cortex using a spatially-informed dimension reduction of the connectivity matrix. *Med. Image Comput. Comput. Assist Interv.* 12(Pt. 1), 935–942 (2009)
10. Thompson, P.M., Schwartz, C., Toga, A.W.: High-resolution random mesh algorithms for creating a probabilistic 3D surface atlas of the human brain. *Neuroimage* 3(1), 19–34 (1996)
11. Tournier, J.D., Calamante, F., Connelly, A.: Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *Neuroimage* 35(4), 1459–1472 (2007)
12. Yeo, B.T.T., Sabuncu, M.R., Vercauteren, T., Ayache, N., Fischl, B., Golland, P.: Spherical demons: fast diffeomorphic landmark-free surface registration. *IEEE Trans. Med. Imaging* 29(3), 650–668 (2010)

Tract-Based Probability Densities of Diffusivity Measures in DT-MRI

Çağatay Demiralp and David H. Laidlaw

Brown University, USA

Abstract. We evaluate probability density functions of diffusivity measures in DTI fiber tracts as biomarkers. For this, we estimate univariate and bivariate densities, such as joint probability densities of the tract arc length and FA, MD, RD, and AD, in the transcallosal fibers in the brain. We demonstrate the utility of estimated densities in hypothesis testing of differences between a group of patients with VCI and a control group. We also use the estimated densities in classifying individual subjects in these two groups. Results show that these estimates and derived quantities, such as entropy, can detect group differences with high statistical power as well as help obtain low classification errors.

1 Introduction

Tract-based approaches to analysis of diffusion brain data sets are attracting interest because they can geometrically localize quantities while reducing sensitivity of associated statistics to computational errors and noise (e.g., [1,2,3]). In this context, we propose using probability densities in characterizing diffusion along diffusion-tensor MRI (DTI) fiber tracts generated from diffusion-weighted imaging data sets. We experiment with well known diffusivity measures such as fractional anisotropy (FA), mean diffusivity (MD), radial diffusivity (RD), and axial diffusivity (AD) [4]. We estimate univariate and bivariate diffusion densities of these four measures nonparametrically, using histogram and kernel density estimation methods. We use the estimated densities for hypothesis testing of differences between a group of patients with vascular cognitive impairment (VCI) and a control group. Results indicate that probability density estimates can detect group differences with high statistical significance.

Contributions. Our main contribution is to introduce tract-based probability density functions, including the joint density of tract arc length and scalar diffusivity measures, as potential metrics for characterizing diffusivity in DT-MRI brain data sets. Our approach is a simple and effective addition to existing methods. We also show that probability density entropy itself is a useful biomarker.

2 Related Work

Using voxel or integral curve (as in tractography) representations of fiber tracts, previous research has proposed tract-based analysis of diffusion-derived measures in order to increase the specificity and robustness of related statistics [1,2,3]. A typical approach in these earlier studies is to find a representative skeleton (or curve) first and then project the diffusivity measures of individual subjects on this representative. For example, tract-based spatial statistics (TBSS) has been proposed to overcome some of the disadvantages

of approaches using voxel-based morphometry (VBM) [1]. TBSS reduces the sensitivity of statistics to alignment problems by projecting subject data onto a voxel skeleton of the white matter as characterized by FA. Using tractography makes the incorporation of fiber tract arc length into the analysis easier. One of the joint densities evaluated here is the probability of a diffusivity value given an arc-length distance. Parametrization of tracts based on arc-length distance has already been shown to be useful [5], and tract arc length by itself has been proposed as a metric [6]. As in TBSS, previous studies using tractography have also been able to detect differences between groups of subjects by analyzing diffusivity measures projected onto an arc-length parametrized representative tract [2,3].

Recently, probability density functions (PDFs) have been used as shape descriptors in cortical folding analysis and in quantifying FA change with respect to the thickness of the surface sheet covering the corticospinal tract [7,8].

While earlier tract-based analysis methods essentially collapse bundles into a single representative subset (skeleton or curve) and run the subsequent analysis through this subset, our approach considers each bundle directly and separately, modeling the full spectrum of individual bundle's data as a sample from a probability distribution characterizing the corresponding subject. Also, our use of PDFs is simple, general (it can be applied to any tracts and combine any number of variables), and truly spatial, since we orient tracts.

3 Data Collection

The underlying premise of using tract-based probability densities in clinical research is that they have different values in patients with and without known white-matter injury. In order to verify this, we generated tractograms of a group of patients with vascular cognitive impairment (VCI) and a group of healthy control subjects and compared the group differences using probability densities. VCI is a general term for vascular cognitive deficits caused by injuries to the white matter from cerebrovascular diseases. VCI typically manifests itself with problems in speech, language, and the ability to organize thoughts and track actions. Its effects on memory, however, are considered mild in comparison to Alzheimer's disease. Often, deficits in cognitive domains caused by VCI are similar to those common in individuals of advanced age. Below we give only a summary of our data collection process; details can be found in [6].

Subjects. Our subjects were 19 patients with VCI and 20 healthy individuals who served as a control group. The subjects were age-matched. Diffusion-weighted MRI (DWI) sequences for each subject's brain were acquired on a 1.5T Siemens Symphony scanner with the following parameters in 12 bipolar diffusion encoding gradient directions: thickness = 1.7 mm, FOV = 21.7 cm \times 21.7 cm, TR = 7200 ms, TE = 156 ms, b = 1000, and NEX = 3.

Fiber-tract Generation. From acquired sequences, we calculate tensors and then derive the three principal eigenvalues and eigenvectors for each image voxel after interleaving the three sets of DWIs. We then generate fiber-tract models of the whole brain for each subject by integrating (by second-order Runge-Kutta integration) the major

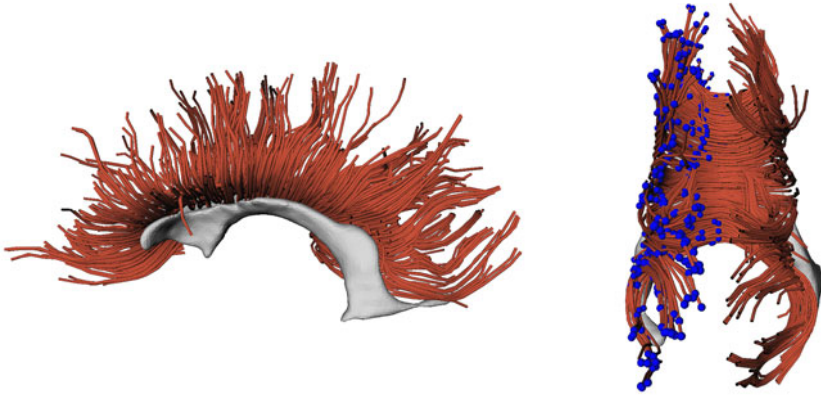


Fig. 1. Transcallosal fiber tracts obtained from a healthy subject. The gray surface representing the ventricles is displayed for anatomical correspondence. Left: Sagittal view. Right: Tilted coronal view, where spheres (blue in color) indicate the “beginning” of each tract.

eigenvector field of the diffusion tensor field bidirectionally at seed points (see Figure 1). Transcallosal fibers, defined as all trajectories passing through the corpus callosum, for each participant were selected manually by a rater using a custom interactive visualization program. Selected models were checked by two experts for anatomical correctness. Since tract of interest (TOI) selection was based on each participant’s own anatomy, registration was not necessary.

Orienting Tracts. Since we compute the joint probabilities of arc length and the diffusivity measures, we need to assign start- and end-point designations consistently within and across fiber tracts and across subjects. We achieve this by simply computing the start-to-end vector for each curve and iteratively reorienting the curves until all the vectors are in the same half-space; we repeat this interactively between data sets.

4 Methods

Density Estimation. Since we do not assume a particular distribution of diffusivity measures along tracts, we take a nonparametric approach to estimation. We use two widely studied estimators, histogram and kernel estimators [9,10]. Let $X_1^N = \{X_1, \dots, X_N\}$ be independent and identically distributed (i.i.d.) random variables drawn from a density $f(x)$. The histogram estimator of f is obtained simply by dividing the range of the samples into bins of width h and counting the number of samples falling into each bin.

Kernel density estimation (KDE) addresses some of the obvious disadvantages of histograms, including non-differentiability and reference-point dependency. The kernel estimator of f using X_1^N is given by $\hat{f}(x) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x-X_i}{h}\right)$, where $K(x)$ is some kernel satisfying $\int_{-\infty}^{\infty} K(x)dx = 1$ and h is the kernel *bandwidth*, analogous to the histogram bin width. We omit here multivariate versions of the equations; they can be found in any good multivariate statistical analysis book (e.g., [9,10]).

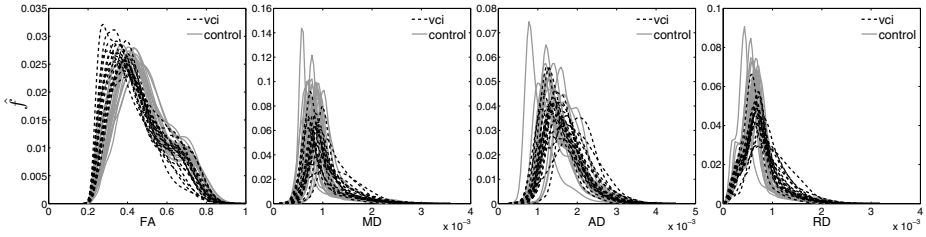


Fig. 2. Univariate density estimates of FA, MD, AD, and RD measures for control subjects and subjects with VCI. They suggest an increase in variances of MD, AD, RD densities with VCI. Also, the shifts in the densities of the two groups are consistent with previous reports: while FA decreases, MD, RD and, to a lesser extent, AD increase with atrophy in the brain.

Both bin width and kernel size can affect the results significantly. Therefore we use a data-driven rule originating in the L_2 theory of histograms, minimizing the mean integrated square error (MISE) [10,11]. We set the histogram bin width $h = 3.5\hat{\sigma}N^{-\frac{1}{d+2}}$, where d is the dimensionality of the data (e.g., $d = 2$ for a bivariate distribution) and N is the number of samples. We select $\hat{\sigma} = \min\{s, IQR/1.349\}$ as discussed in [12], where s is the standard deviation of samples and IQR is the inter-quantile range. We use a Gaussian kernel with both univariate and bivariate estimators. As for the kernel bandwidth h , we apply the “normal scale” rule introduced in [9], which is also asymptotically optimal.

Non-parametric Multivariate Hypothesis Testing. Both histogram and kernel estimates (kernel estimators in their discretized forms) essentially provide a vector of values identifying the underlying distribution. These “density vectors” can be used for testing differences between groups, which requires a multivariate test statistic. We use the permutation Hotelling’s T^2 test [13]. Hotelling’s T^2 statistic is a multivariate extension of Student’s t statistic and permutation is a common method for running non-parametric tests using scores obtained with parametric statistics.

Entropy. One of the usual quantities computed over probability densities is entropy (or Shannon entropy) [14]. While entropy has different interpretations in different domains, it can be viewed here as a measure of randomness in data (i.e. uniformity of its distribution). We compute the entropy of the density f using the estimated density as $H(\hat{f}) = -\sum_x x\hat{f}(x)$ and evaluate it in group comparison of our subjects.

Classification. Classification is now a standard approach to exploring medical data sets. We use a support vector machine (SVM) classifier to assess the usefulness of quantities derived from tract-based diffusivity probability densities. SVM is a maximum-margin classifier minimizing the classification error while maximizing the geometric margin between the classes [15]. It is particularly suitable for two-class problems.

5 Results

We estimate univariate densities of FA, MD, AD, and RD as well as joint densities of these measures with fiber tract arc length, where arc length is normalized to 1, in each

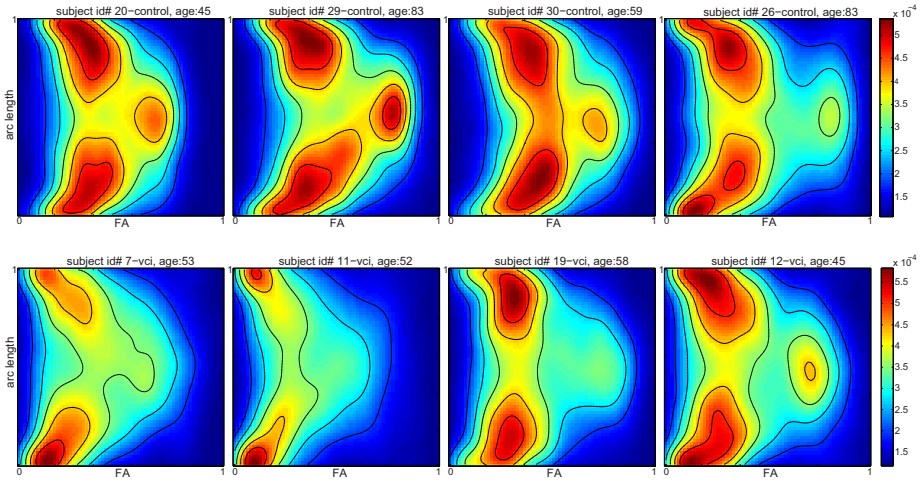


Fig. 3. Kernel density estimates of FA and arc length joint PDF for 8 subjects—4 controls (first row) and 4 patients (second row). The third local maximum (on the mid-right) starts disappearing with VCI. Also, observe, in the last column, the similarity between a healthy subject of advanced age and a young patient.

subject’s transcallosal fiber tract. Overall, results show that the tract-based probability densities help detect differences between groups and classify individual subjects with high accuracy. Figure 2 shows kernel estimates for univariate FA, MD, AD, and RD probability densities. Even in this simple univariate form, estimated PDFs are rich in information. They suggest an increase in variances of MD, AD, RD densities with VCI. Also the shifts in the estimated densities of the two groups support previous reports that, while FA decreases, MD, RD and, to a lesser extent, AD increase with atrophy in the brain. Figure 3 shows kernel estimates for the joint probability of FA with the arc length. These 2D density results are visually informative and can help generate hypotheses. For example, it is clear from the figures that the third local maximum (on the mid-right) starts disappearing with VCI and age. Similarly, the decrease in MD with VCI is clear in MD and the arc length joint density shown in Figure 4.

5.1 Group Comparisons

Using Entropies. We expect neurological diseases to affect the entropy of probability distributions of diffusivity-related measures. We test this hypothesis for all the measures using their univariate and bivariate densities. The results show that MD, RD, and AD entropies increase significantly while FA decreases (see Figure 5 and Table 1a). This outcome is in line with previous findings that AD is more sensitive to atrophy than FA [16].

Density Vectors. Using vectorized “raw” density values, we compare the group of subjects with VCI with our control group, employing the multivariate permutation

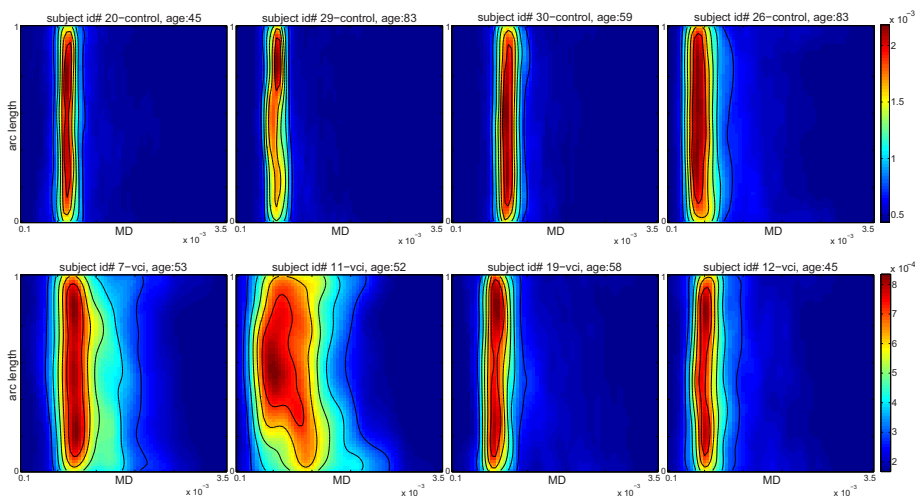


Fig. 4. Kernel density estimates of MD and arc length joint PDF for 8 subjects—4 controls (first row) and 4 patients (second row). The joint PDFs tend to spread out (i.e., increase in variance) with VCI. As seen in the last column, aging can have an effect similar to VCI.

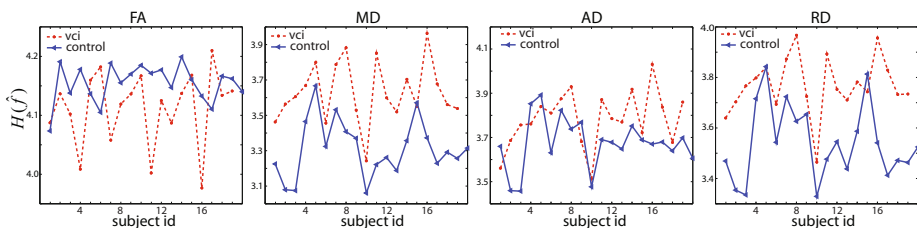


Fig. 5. Entropies for univariate probability density estimates of FA, MD, AD, and RD for all subjects. MD, AD, and RD density entropies significantly increase with VCI, but the FA entropy decreases, though not significantly (see also Table 1b).

Table 1. (a) p -values for the group comparison (VCI vs. control) using entropies of the estimated probability densities. Except for FA, the entropy measure was able to detect the group difference for all densities. (b) p -values for the group comparison (VCI vs. control) using probability density vectors. Histogram estimates of univariate densities were able to detect the group differences.

	(a)		(b)	
	univariate	bivariate	univariate	
	histogram	kernel	histogram	kernel
FA	0.36797	0.01239	0.0297	0.8379
MD	0.00000	0.00000	0.00000	0.00000
AD	0.00106	0.00878	0.0042	0.00016
RD	0.00000	0.00000	0.00000	0.00000

	univariate
FA	0.00078
MD	0.00000
AD	0.00005
RD	0.00002

Table 2. Classification results for our 39 subjects, obtained using an SVM classifier with 10-fold cross-validation. Data points used for classification are univariate histogram density estimates.

	density	sensitivity	specificity	error
FA	83.3%	81%	17.9%	
MD	90%	94.7%	8%	
AD	77.2%	88.2%	17.9%	
RD	89.5%	90%	10.2%	

Hotelling’s test [13,17]. We reduce the dimensionality of the density vectors using principal component analysis (PCA) before running the test, while preserving 99% of the variance. The results show that the density vectors of univariate estimates are able to detect group differences (see Table 16).

5.2 Classifying Normal and Pathology

We ran an SVM with a polynomial kernel on 39 data sets using 10-fold cross-validation. Using density estimates of MD yields the best classification. Table 2 summarizes the results.

6 Discussion and Conclusions

We have proposed a simple and practical addition to existing tract-based multi-subject diffusion data analysis techniques, which are essentially based on collapsing bundles into a representative subset and sampling diffusion measures along this subset. Our approach considers all the tracts in a given bundle and treats the associated diffusivity measures as samples from a PDF characterizing the subject. We estimate this PDF using the histogram and kernel estimators. While using the histogram density estimator is fast and simple, the kernel estimator provides smoother results with useful differential properties.

Given a reasonable choice of variables, estimated PDFs are likely to be descriptive and rich in information, and therefore well-suited for analysis of individual and group differences in diffusion data sets. In the case examined here, both PDFs themselves and their entropy measures appear to be good indicators of group differences, although additional work is required to understand how this generalizes to other cases. Similarly, our initial results on the classification of subjects based on densities suggest the potential utility of PDFs in detecting individual differences. Note that using natural distance measures on PDFs such as the Bhattacharyya distance and Kullback-Leibler (KL) divergence could further improve the analysis discussed above. Finally, although we have not dealt with alignment issues, the probabilistic approach discussed here can also increase the robustness of the analysis to registration errors in fully automated settings.

Acknowledgments. This work has been supported by NIH grant 1R01EB00415501A1.

References

1. Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., Behrens, T.E.: Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage* 31(4), 1487–1505 (2006)
2. O'Donnell, L.J., Westin, C.F., Golby, A.J.: Tract-based morphometry. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part II*. LNCS, vol. 4792, pp. 161–168. Springer, Heidelberg (2007)
3. Goodlett, C.B., Fletcher, P.T., Gilmore, J.H., Gerig, G.: Group statistics of DTI fiber bundles using spatial functions of tensor measures. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part I*. LNCS, vol. 5241, pp. 1068–1075. Springer, Heidelberg (2008)
4. Song, S.K., Sun, S.W., Ramsbottom, M.J., Chang, C., Russell, J., Cross, A.H.: Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. *NeuroImage* 17(3), 1429–1436 (2002)
5. Corouge, I., Fletcher, P.T., Joshi, S., Gouttard, S., Gerig, G.: Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis. *Medical Image Analysis* 10(5), 786–798 (2006)
6. Correia, S., Lee, S.Y., Voorn, T., Tate, D.F., Paul, R.H., Zhang, S., Salloway, S.P., Malloy, P.F., Laidlaw, D.H.: Quantitative tractography metrics of white matter integrity in diffusion-tensor MRI. *NeuroImage* 42(2), 568–581 (2008)
7. Awate, S.P., Yushkevich, P.A., Song, Z., Licht, D.J., Gee, J.C.: Multivariate high-dimensional cortical folding analysis, combining complexity and shape, in neonates with congenital heart disease. In: Prince, J.L., Pham, D.L., Myers, K.J. (eds.) *Information Processing in Medical Imaging*. LNCS, vol. 5636, pp. 552–563. Springer, Heidelberg (2009)
8. Zhang, H., Awate, S.P., Das, S.R., Woo, J.H., Melhem, E.R., Gee, J.C., Yushkevich, P.A.: A tract-specific framework for white matter morphometry combining macroscopic and microscopic tract features. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009*. LNCS, vol. 5762, pp. 141–149. Springer, Heidelberg (2009)
9. Silverman, B.W.: *Density Estimation for Statistics and Data Analysis*. In: *Monographs on Statistics and Applied Probability*. Chapman and Hall, London (1986)
10. Scott, D.W.: *Multivariate Density Estimation: Theory, Practice, and Visualization*. Wiley Interscience, Hoboken (1992)
11. Freedman, D., Diaconis, P.: On the histogram as a density estimator: L_2 theory. *Probability Theory and Related Fields* 57(4), 453–476 (1981)
12. Wand, M.P.: Data-based choice of histogram bin width. *The American Statistician* 51(1), 59–64 (1997)
13. Good, P.I.: *Permutation, Parametric, and Bootstrap Tests of Hypotheses*. Springer Series in Statistics. Springer, New York (2004)
14. Cover, T.M., Thomas, J.A.: *Elements of Information Theory*. Wiley Interscience, New York (1991)
15. Bishop, C.M.: *Pattern Recognition and Machine Learning*. Springer, Heidelberg (2006)
16. Metwalli, N., Benatar, M., Usher, S., Bhagavatheeshwaran, G., Hu, X., Carew, J.: Sensitivity of axial and radial diffusivity to the neuropathology of amyotrophic lateral sclerosis. *NeuroImage* 47 (Suppl. 1), S128–S128 (2009)
17. Efron, B., Tibshirani, R.J.: *An Introduction to the Bootstrap*. Chapman & Hall, New York (1993)

Inference of a HARDI Fiber Bundle Atlas Using a Two-Level Clustering Strategy

Pamela Guevara^{1,2*}, Cyril Poupon^{1,2}, Denis Rivière^{1,2}, Yann Cointepas^{1,2},
Linda Marrakchi^{1,2}, Maxime Descoteaux⁴, Pierre Fillard^{1,2,3},
Bertrand Thirion^{1,2,3}, and Jean-François Mangin^{1,2}

¹ Neurospin, CEA, Gif-sur-Yvette, France

² Institut Fédératif de Recherche 49, Gif-sur-Yvette, France

³ Parietal Team, INRIA Saclay-Ile-de-France, Saclay, France

⁴ Université de Sherbrooke, Sherbrooke, Canada

Abstract. This paper presents a method inferring a model of the brain white matter organisation from HARDI tractography results computed for a group of subjects. This model is made up of a set of generic fiber bundles that can be detected in most of the population. Our approach is based on a two-level clustering strategy. The first level is a multiresolution intra-subject clustering of the million tracts that are computed for each brain. This analysis reduces the complexity of the data to a few thousands fiber bundles for each subject. The second level is an inter-subject clustering over fiber bundle centroids from all the subjects using a pairwise distance computed after spatial normalization. The resulting model includes the large bundles of anatomical literature and about 20 U-fiber bundles in each hemisphere.

1 Introduction

An interesting way to compare the different DW-MRI acquisition schemes, diffusion models and tractography algorithms proposed in the literature lies in the exploitation of the large sets of generated tracts to infer atlases of the fiber bundles. Hence, this paper presents a method taking as input the sets of diffusion-based tracts of a population of subjects and producing as output a list of generic fiber bundles that can be detected in most of the population. The usual strategies proposed for the reconstruction of fiber bundles follow two complementary ideas. The first approach is based on regions of interest (ROI) used to select or exclude tracts [1]. The second strategy is based on tract clustering using pairwise similarity measures [2]. This last approach is potentially less intensive in terms of user interaction and can also embed predefined knowledge represented by a bundle template [3]. Furthermore, when applying the clustering after spatial normalization with a set of tracts stemming from several subjects, this strategy can help to discover new reproducible bundles. The fiber clustering approach has

* This project was supported by grants from Région Ile-de-France (France), French Embassy in Chile, CONICYT Chile and University of Concepción (Chile)

been successfully used to map the well-known fiber bundles of deep white matter (DWM). Until now, short fibers of superficial white matter (SWM) have been barely considered. The cartography of fiber bundles of SWM is a complex and unachieved task for the human brain. In a recent paper, Oishi et al. performed a group analysis to study SWM using a voxel-based approach relying on linear brain normalization [4]. They could identify only four U-fiber bundles because of the blurring occurring with such a normalization. In this paper we show that the fiber clustering approach can overcome this weakness.

We propose a two-level strategy chaining intra and inter-subject fiber clustering. The first level can be viewed as a compression procedure reducing a huge set of fibers to a few thousand bundles. This step is developed following a multi-resolution paradigm. A key point is the use of a voxel-based parcellation of white matter, allowing the analysis of any number of fibers. This parcellation produces small fiber subsets that can be split further using additional clustering performed in the space of fiber extremities. The second level is an inter-subject clustering of the resulting fiber bundles. This group analysis relies on a pairwise distance between bundles computed after affine spatial normalization. A simulation is performed to prove that affine normalization is sufficient to create consistent clusters in the bundle space.

2 Material and Method

2.1 Diffusion and Tractography Datasets

Analysis was performed for twelve subjects of the NMR public database. This database provides high quality T1-weighted images and diffusion data acquired with a GE Healthcare Signa 1.5 Tesla Excite scanner. The diffusion data presents a high angular resolution (HARDI) based on 200 directions and a b-value of 3000 s/mm^2 (voxel size of $1.875 \times 1.875 \times 2 \text{ mm}$). DW-weighted data were acquired using a twice refocusing spin echo technique compensating Eddy currents to the first order. Geometrical distortions linked to susceptibility artifacts were corrected using a phase map acquisition. T1 and DW-weighted data were automatically realigned using a rigid 3D transform.

The diffusion Orientation Distribution Function was reconstructed in each voxel using a spherical deconvolution of fiber Orientation Distribution Function. It is a spherical deconvolution transform (SDT) reconstructed from q-ball imaging with a constrained regularization [5], using a maximum spherical harmonic (SH) order 8 and a Laplace-Beltrami regularization factor $\lambda = 0.006$.

Tracts were reconstructed using a deterministic tractography algorithm [6] provided by BrainVISA public software (<http://brainvisa.info>). Tractography was initiated from two seeds in each voxel of the mask (voxel size of $0.94 \times 0.94 \times 1.2 \text{ mm}$), in both retrograde and anterograde directions, according to the maximal direction of the underlying ODF. Tracking parameters included a maximum curvature angle of 30° and a minimum and maximum fiber length of 20 mm and 200 mm, respectively, leading to a set of about 1.5 millions tracts per subject.

2.2 The Two-Level Clustering Method

Intra-subject clustering. The intra-subject clustering follows a multiresolution strategy including five steps (see Fig. 1A):

Step 1: Hierarchical decomposition: The complete tract set is segmented into four parts, called *fiber subsets*: right hemisphere, left hemisphere, inter-hemispheric and cerebellum tracts. The following steps are applied separately to each subset.

Step 2: Fiber segmentation based on length: The subset is split into different *fiber groups*, containing fibers of similar length. While looking insignificant, this second step is of key importance because it partially overcomes a big limitation of the voxel-based approach: the difficulty to separate fiber bundles with different shapes overlapping for a large part of their voxel support.

Step 3: Voxel-based clustering: Each fiber group obtained in the preceding step is divided using a connectivity-based parcellation of white matter voxels. A T2 fiber mask is calculated containing voxels crossed by tracts. The mask is randomly parcellated using a geodesic k-means leading to about 12,000 parcels per fiber group. Parcels are clustered using an average-link hierarchical clustering (HC) based on a parcel connectivity measure defined as the number of tracts passing through the pair of parcels. The tree resulting from the hierarchical clustering is analyzed in order to get an adaptive partition where each cluster contains ideally only one putative fiber bundle. The tree analysis discards small isolated clusters and split the large clusters until reaching sizes compatible with the largest actual bundles. Finally, each cluster mask is used to extract corresponding diffusion-based *fiber clusters*.

Step 4: Extremity-based fiber clusters subdivision: Fiber clusters are divided into several thin and regular bundles, called *fascicles*, based on the fiber extremities. For each cluster, the list of fiber extremities is converted into a voxel-based density image. This image is segmented by a 3D watershed into maxima-based regions. These regions are used to divide the extracted fiber clusters into several fascicles, each one composed of the fibers whose extremities pass through two particular regions.

Step 5: Fiber fascicle centroids clustering: This step considers all the fascicles from all the fiber length groups of a subset. It consists in a second clustering, aiming to agglomerate fiber fascicles that were over-segmented in the fiber clusters subdivision step (*Step 4*) or in the length-based segmentation step (*Step 2*). For this, a centroid tract, localized in the center of each fascicle, is computed as a representative of the fascicle. It is determined as the tract minimizing a distance to the rest of the fascicle fibers. The distance measure employed is a symmetrized version of the mean closest point distance [72,3]. Centroids from all the fascicles of a set are clustered using an average-link hierarchical clustering over a pairwise distance between centroids. The distance is defined as the maximum distance between corresponding points, which is more stringent than the mean closest point distance. This step aims only at regrouping fascicles that have very similar shapes and positions.

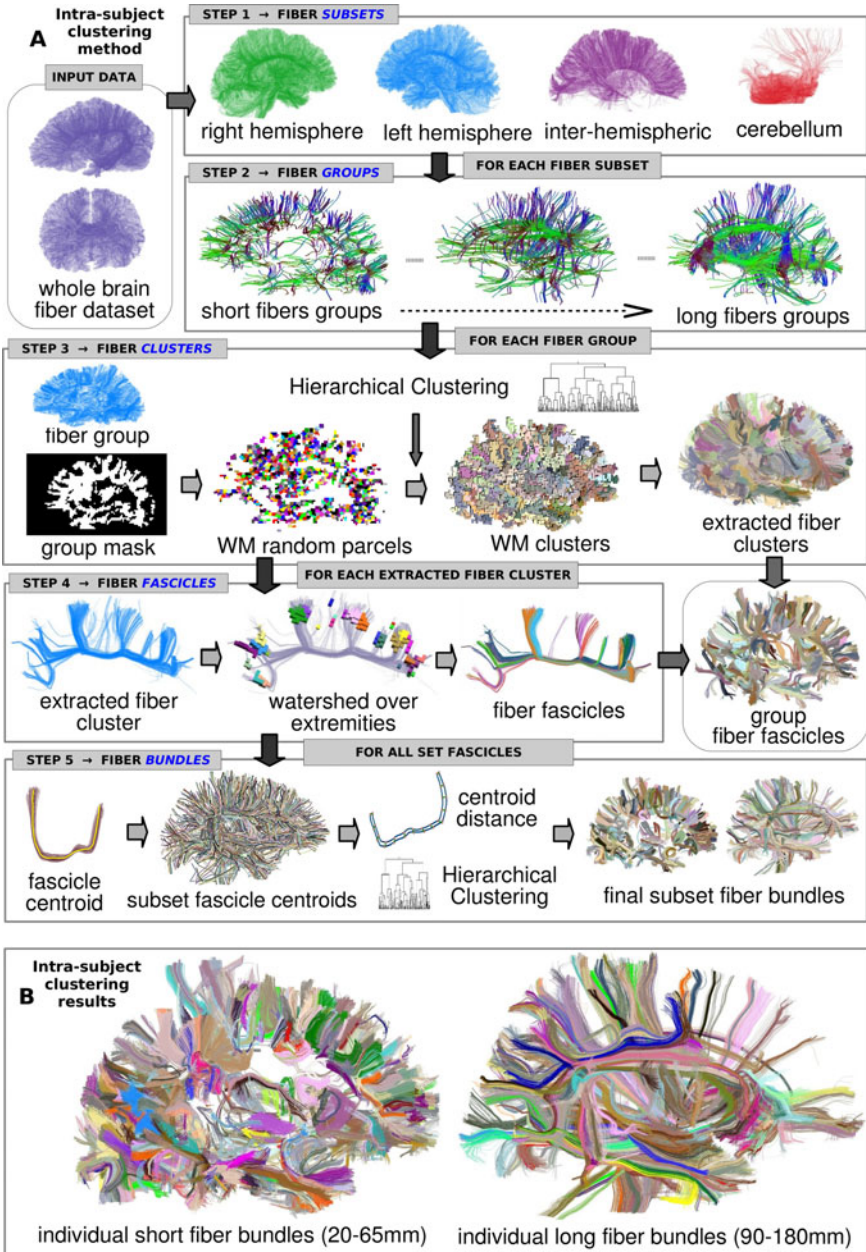


Fig. 1. A: General scheme of the intra-subject clustering: Step 1: Hierarchical decomposition, **Step 2:** Fiber segmentation based on length, **Step 3:** Voxel-based clustering, **Step 4:** Extremity-based fiber clusters subdivision, **Step 5:** Fiber fascicles centroids clustering. **B: Example results for the intra-subject clustering.** Short and long bundles obtained for the right hemisphere fiber set of one subject.

Inter-subject clustering. The second clustering level aims at matching the putative bundles produced by the previous level across the population of subjects. This step is very similar to the clustering performed in section 2.2 (*Step 5*) but this time the calculation considers the bundles obtained from all the subjects in the same fiber subset. A centroid is first calculated for each bundle using the mean of the two mean closest point distances [7][2][3]. Once all the centroids are computed, they are transformed to the Talairach space (TS) using an affine transformation estimated from the T1-weighted image. Then, a bundle centroid affinity graph is computed using the maximum distance between corresponding points, normalized by the bundle length. This restraining measure puts a focus on matching bundles with similar shapes and positions in TS. The affinity graph is used to compute an average-link HC. The resulting tree is analyzed in order to extract only very tight clusters, where the distance between all the fibers within a cluster is inferior to a maximum distance (tM_{dcp}). The resulting clusters are discarded if they do not contain more than half of the subjects. A final procedure aims at relaxing the constraints in order to recover some instances of the generic bundles that were missed during the stringent clustering analysis. The goal is to be less demanding on the match between centroids, which is specially important for the subjects that present a deficient normalization in Talairach space. For each non attributed centroid, we compute the distance to each of the centroids of the tight clusters. When the distance to the nearest neighbor is below a threshold, the non attributed centroid is added to the final generic bundle representation. Most of the added centroids belong to long fiber bundles.

Inter-subject clustering validation: In order to study the behavior of the inter-subject clustering over a population of subjects aligned with affine registration, we created a simulated dataset of fiber bundles centroids. First, one subject of the NMR database was selected to generate a set of 200 simulated bundle centroids. These bundle centroids were fibers selected from the right hemisphere of this subject with a minimum pairwise distance across the set. The distance used was the maximum distance between corresponding points. The minimum distance was set to 12 mm (see Fig. 2A₁). The obtained bundle centroids set was transformed to the space of each one of the eleven remaining subjects of the database, using a non-rigid transform, calculated between T1 images using MedINRIA (<http://www-sop.inria.fr/asclepios/software/MedINRIA>). Hence, we obtained a set of ground truth clusters, each one containing a centroid in each subject (see Fig. 2A₂). In addition, 500 fibers from each subject were selected to simulate noise. These fibers were pairwise separated by a minimum distance equal to 11 mm (see. Fig. 2A₄). For each subject, we got a fiber dataset of 700 fibers (200 centroids and 500 added noise fibers), leading to a total number of 8400 fibers for the twelve subjects. We applied the inter-subject clustering to the fibers dataset, with the maximum distance within clusters (tM_{dcp}) varying from 5 to 25 mm. Resulting clusters were analyzed and compared with the ground truth. First, only clusters containing centroids from a minimum of seven different subjects were selected. Then, a cluster was counted as *recovered* only if all its centroids belonged to the same simulated cluster, otherwise, it was counted

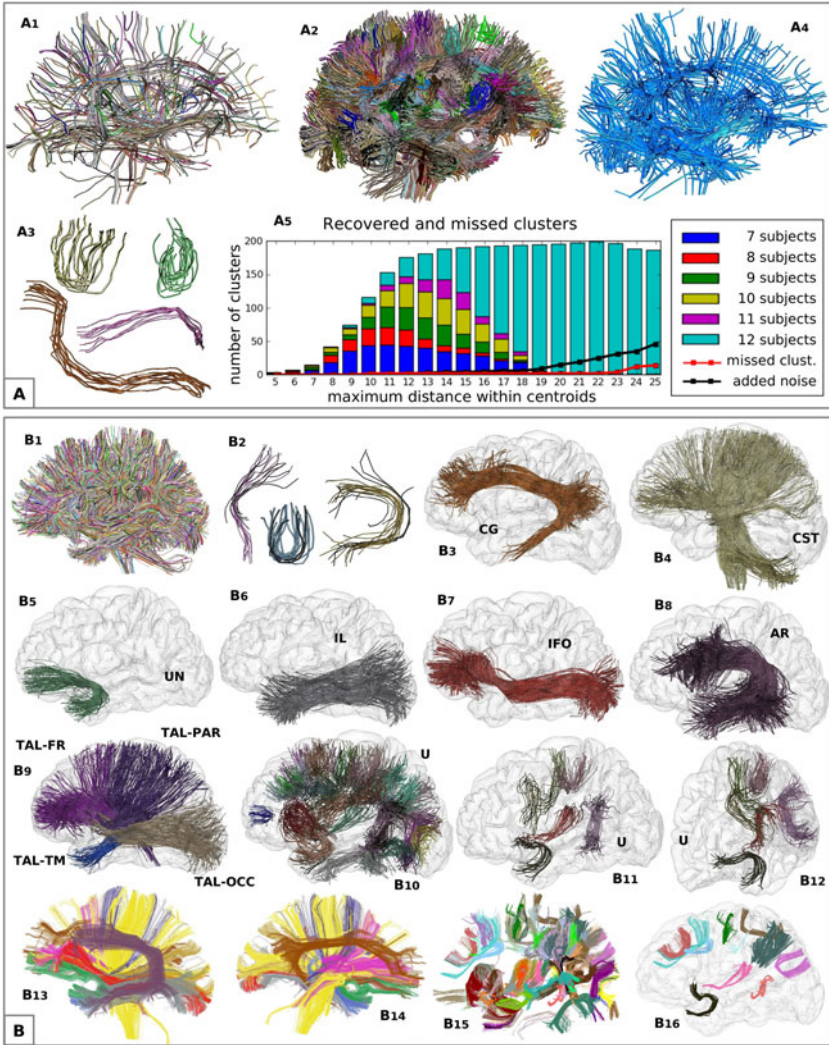


Fig. 2. A: Inter-subject clustering validation: **A1:** the original 200 fibers selected as simulated centroids. **A2:** simulated bundles for the 12 subjects in talairach space. **A3:** a selection of bundles from A2. **A4:** noise fibers set (500 fibers) of one subject. **A5:** inter-subject clustering simulation results presenting recovered clusters (color bars), missed clusters (red line) and clusters with added noise (black line). **B: Inter-subject clustering results for twelve subjects:** **B1:** centroids of one subject. **B2:** examples of tight clusters and the added centroids by the nearest neighbor procedure (in black). **B3-B9:** long generic bundles centroids (manually labelled): cingulum (CG), corticospinal tract (CST), uncinate (UN), inf. longitudinal (IL), inf. fronto-occipital (IFO), arcuate (AR), and thalamic radiations (TAL-(FR,PAR,TM,OC)). **B10-B12:** selections of generic short bundles centroids. **B13-B16:** generic bundles for one subject. B15 shows all the U-fibers while B16 contains a selection of U-fibers.

as a *missed* cluster. Fig. 2A₅ presents the simulation results as a function of the distance tM_{dcp} . From the analysis, we note, as expected, that the number of recovered clusters (color bars) increases with tM_{dcp} , as the number of subjects in the clusters (indicated by different colors, from 7 to 12). For distances tM_{dcp} superior to 11 mm, a large number of clusters was recovered, but for distances between 11 and 15 mm, most of the clusters miss some centroids. This behavior is accepted by the method, which adds a cluster to the model as soon as it includes more than half of the subjects. The red line indicates the number of missed clusters, which is very low. These are most of the time fused with other clusters. The black line shows the number of recovered clusters that contain also added noise fibers. Finally, a large number of clusters made up of only noise fibers was found but discarded by the method because none of these clusters had fibers from more than six different subjects.

3 Results

Individual fiber bundles were extracted for twelve subjects. The results obtained for one subject (right hemisphere) are presented in Fig. 1B. Inter-subject clustering of the left hemisphere is presented in Fig. 2B. Ten DWM generic fiber bundles were identified in all the subjects. Some of them are split into several generic fascicles: Inferior fronto-occipital fasciculus, Cortico-spinal tract, Arquate fasciculus, Uncinate fasciculus, Inferior Longitudinal fasciculus, Cingulum, and Frontal, Temporal, Parietal and Occipital Thalamic Radiations. Twenty generic U-Fiber bundles (SWM) occurring in at least seven subjects were detected in the left hemisphere model.

4 Discussion and Conclusion

As for any fiber tracts analysis method, our results depend strongly on the quality of the tractography results. Our method can not detect bundles that are not tracked in individuals. Also, spurious bundles can not be differentiated from real bundles if they are reproducible across subjects. Besides, anomalous final bundles can be found due to errors in the propagation mask. Since this mask defines where fibers are tracked, bundles can be erroneously cutted or fused. Nevertheless, independently of the tracking results, our method is a powerfull tool to extract the main bundles that constitute the dataset.

Our method is able to analyze huge fiber datasets and infer a model of the generic bundles present in a population. The first level, composed by an intra-subject clustering, can be seen as a compression of information and a filtering, where bundles representing the individual whole white matter structure are identified. The second level, an inter-subject clustering, deals with a reasonable number of bundle centroids from a population of subjects and is capable to extract generic bundles present in most of the subjects. Long known bundles were identified, but the result of major significance is the capability to identify generic short association bundles. Hence, our approach will scale up easily to the 1 mm

spatial resolution that can now be achieved with highly parallel imaging or very high fields. This spatial resolution is bound to highlight a myriad of U-fiber bundles and better delineate other bigger bundles crossing.

The general idea of hierarchical decomposition underlying our method will be pushed further with the use of a segmentation of the deep grey matter structures to improve the clustering of their connections with the cortex. The corresponding tracts will be filtered out and clustered independently.

We have shown that the affine registration to standard space is sufficient to align reasonably the deep tracts across all the subjects. Each U-fiber bundle inferred in this paper did require a reasonable alignment of the bundles of only half of the subjects, which happens in the most stable brain regions. However, increasing the number of generic U-fiber bundles, will require an improvement of the spatial normalisation used to compare bundles across subject. Further work will lead us to improve iteratively the spatial normalization using the inferred bundles as constraints in order to better align other bundles. Moreover, we will compare our results with the strategy computing bundles after computing an average atlas of diffusion data based on high quality diffeomorphic spatial normalization. Deciding which strategy is the best is one of the goals of our research.

References

1. Catani, M., Thiebaut de Schotten, M.: A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex* 44(8), 1105–1132 (2008)
2. O’Donnell, L.J., Kubicki, M., Shenton, M.E., Dreusicke, M.H., Grimson, W.E.L., Westin, C.F.: A method for clustering white matter fiber tracts. *AJNR Am. J. Neuroradiol.* 27(5), 1032–1036 (2006)
3. O’Donnell, L., Westin, C.F.: Automatic tractography segmentation using a high-dimensional white matter atlas. *IEEE Transactions on Medical Imaging* 26(11), 1562–1575 (2007)
4. Oishi, K., Zilles, K., Amunts, K., Faria, A., Jiang, H., Li, X., Akhter, K., Hua, K., Woods, R., Toga, A.W., Pike, G.B., Rosa-Neto, P., Evans, A., Zhang, J., Huang, H., Miller, M.I., van Zijl, P.C.M., Mazziotta, J., Mori, S.: Human brain white matter atlas: Identification and assignment of common anatomical structures in superficial white matter. *Neuroimage* (July 2008)
5. Tournier, J.D., Calamante, F., Connelly, A.: Robust determination of the fibre orientation distribution in diffusion mri: non-negativity constrained super-resolved spherical deconvolution. *Neuroimage* 35(4), 1459–1472 (2007)
6. Perrin, M., Poupon, C., Cointepas, Y., Rieul, B., Golestani, N., Pallier, C., Rivière, D., Constantinesco, A., Le Bihan, D., Mangin, J.F.: Fiber tracking in q-ball fields using regularized particle trajectories. *Inf. Process Med. Imaging* 19, 52–63 (2005)
7. Corouge, I., Gouttard, S., Gerig, G.: Towards a shape model of white matter fiber bundles using diffusion tensor MRI. In: *ISBI 2004* (2004)

DTI Based Diagnostic Prediction of a Disease via Pattern Classification

Madhura Ingalhalikar¹, Stathis Kanterakis¹, Ruben Gur²,
Timothy P.L. Roberts^{3,*}, and Ragini Verma¹

¹ Section of Biomedical Image Analysis,
University of Pennsylvania, Philadelphia, PA, USA

{Madhura.Ingalhalikar,Stathis.Kanterakis,Ragini.Verma}@uphs.upenn.edu

² Brain Behavior Laboratory, University of Pennsylvania, Philadelphia, PA, USA
gur@bbi.med.upenn.edu

³ Center for Autism Research, Children's Hospital of Philadelphia,
Philadelphia, PA, USA
robertstim@email.chop.edu

Abstract. The paper presents a method of creating abnormality classifiers learned from Diffusion Tensor Imaging (DTI) data of a population of patients and controls. The score produced by the classifier can be used to aid in diagnosis as it quantifies the degree of pathology. Using anatomically meaningful features computed from the DTI data we train a non-linear support vector machine (SVM) pattern classifier. The method begins with high dimensional elastic registration of DT images followed by a feature extraction step that involves creating a feature by concatenating average anisotropy and diffusivity values in anatomically meaningful regions. Feature selection is performed via a mutual information based technique followed by sequential elimination of the features. A non-linear SVM classifier is then constructed by training on the selected features. The classifier assigns each test subject with a probabilistic abnormality score that indicates the extent of pathology. In this study, abnormality classifiers were created for two populations; one consisting of schizophrenia patients (SCZ) and the other with individuals with autism spectrum disorder (ASD). A clear distinction between the SCZ patients and controls was achieved with 90.62% accuracy while for individuals with ASD, 89.58% classification accuracy was obtained. The abnormality scores clearly separate the groups and the high classification accuracy indicates the prospect of using the scores as a diagnostic and prognostic marker.

1 Introduction

High dimensional pattern classification methods like support vector machine (SVM) can be used to capture multivariate relationships among various anatomical regions for more effectively characterizing group differences as well as quantifying the degree of pathological abnormality associated with each individual.

* The authors would like to acknowledge support from the NIH grants: MH079938, DC008871 and MH060722.

In this paper, we propose a method for region-based pattern classification that creates abnormality classifiers using information from DTI data in anatomically meaningful regions.

Classification not only elucidates regions that are sufficiently affected by pathology so as to be able to differentiate the group from that of healthy controls, but also assigns each individual with a probabilistic score indicating the degree of abnormality. Such a score can be used as a diagnostic tool to predict the extent of disease or group difference and assess progression or treatment effects.

Pattern classification has been previously implemented in neuroimaging studies, mainly in structural imaging [1,2,3,4,5] and a few using DTI information [6,7]. The process can be divided into 4 steps: feature extraction, feature selection and classifier training and testing or cross-validation. Pattern classification methods differ on the basis of *features extracted*. Typical methods for feature extraction use direct features like structural volumes and shapes [2,5], principal component analysis (PCA) [6], automated segmentation [4], wavelet decomposition [1], and Bayes error estimation [7]. High dimensionality of the feature space and limited number of samples pose a significant challenge in classification [2]. To solve this problem, *feature selection* is performed in order to produce a small number of effective features for efficient classification and to increase the generalizability of the classifier. Feature selection methods that are commonly used for structural neuro-imaging data involve filtering methods (e.g. ranking based on Pearson's correlation coefficient) and/or wrapper techniques (e.g. recursive feature elimination (RFE)). Classifiers are then trained on the selected features either using linear classifiers like linear discriminant analysis (LDA) or non-linear classifiers like k-nearest neighbors (kNN) [7] and support vector machines (SVM) [1,2,4,5]. Prior to their application to a new test subject, it is important that the classifier be cross-validated, for which jack-knifing (leave-one-out) is a popular technique. When applied to any subject individual, these cross-validated classifiers produce an abnormality score that can be combined with clinical scores to aid in diagnosis and prognosis.

While most classification work used structural images few studies did attempt DTI-based classification. Caan et al. [6] used fractional anisotropy (FA) and linear anisotropy images as their features, followed by dimensionality reduction using PCA and trained the data using linear discriminant analysis (LDA). Wang et al. [7] used a k-NN classifier trained on full brain volume FA and geometry maps. The former used a linear classifier incapable of capturing complex non-linear relationships, while the latter used a non-linear classification albeit on the full brain.

We describe a classification methodology based on DTI features of anisotropy and diffusivity computed from anatomically meaningful regions of interest (ROI). The regions are based on an atlas and do not require a priori knowledge of affected region. An efficient two step feature ranking and selection technique is implemented. A non-linear SVM classifier is then trained and cross-validated using the leave-one-out paradigm. The classifiers are then applied to 2 different

datasets: one includes patients with SCZ and another with ASD, where each individual is assigned a classification score that quantifies degree of pathology relative to the population. Despite the difference in size of the dataset and the heterogeneity of the pathology, we obtained classifiers with high cross-validation accuracy. This establishes the applicability of the classifier scores for aiding diagnosis and prognosis. The classifier also produces a ranking of features, anatomical regions in our case, which contribute towards the separation of groups, thereby making the classification score of abnormality clinically meaningful.

2 Methods

Our method of creating ROI-based classifiers using DTI-based information involves 4 steps: a. Feature extraction b. Feature selection c. Classifier training and d. Cross-validation.

2.1 Feature Extraction

Our method begins with spatial normalization of all the tensor images to a standard atlas created by Wakana et al. [8], with 176 anatomical regions labeled. A high dimensional elastic registration introduced that involves matching tensor attributes, is used and all scans are registered to the template [9]. In the earlier studies that involved classification [6,7], the whole brain was used as a feature. This is challenging for training classifiers as the sample size is small while the data dimensionality is very high. As a result of the inter-individual structural variability, using fewer features computed from small anatomical regions of interest may lower the dimensionality while maintaining the spatial context, thereby producing more robust features [4]. Therefore we use the ROI's from Wakana et al.'s [8], diffusion tensor template, as shown in figure 1(a). The FA and mean diffusivity (TR) are computed from the spatially normalized tensor images for each subject and then averaged over each ROI. Thus, each subject is associated with a feature vector (eq. 1) that involves 'n' ROI's. Eq. 1 describes the feature vector where each component is the average value of the scalar in that ROI.

$$f_s = (ROI_1^{FA}, ROI_2^{FA} \dots ROI_n^{FA}, ROI_1^{Tr}, ROI_2^{Tr} \dots ROI_n^{Tr}) \quad (1)$$

2.2 Feature Selection

Identifying the most characteristic features is critical for minimizing the classification error. This can be achieved by ranking and selection of the relevant features and eliminating all the redundant features. We use a two step feature ranking and selection as suggested in most of the pattern classification theories, where the the two steps involve a filter and a wrapper yielding best results [10]. We filter the features using minimum redundancy and maximum relevance (mRMR) ranking introduced by Peng et al. [11]. This method involves computation of the mutual information between 2 variables x and y which is defined by their probability density functions $p(x)$, $p(y)$ and $p(x, y)$.

$$I(x; y) = \int \int p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy \quad (2)$$

Based on the mutual information values, Peng et al. have defined 2 criteria to find the near optimal set of features: 1. Maximum relevance and 2. Minimum redundancy. In maximum relevance, the selected feature x_i is required individually to have the largest mutual information with the class label c which reflects the highest dependency on the target class. If we assume the optimal subset of features is S that consists of m features, the maximum relevance of the whole dataset is obtained by $\max(D(S, c))$ where D is given by equation 3.

$$D = \frac{1}{|S|^2} \sum I(x_i; c) \quad (3)$$

It is likely that features with maximum relevance are highly redundant i.e. the features highly depend upon each other. The minimum redundancy criteria therefore is given by $\min(R(S))$ where R is obtained from equation 4.

$$R = \frac{1}{|S|^2} \sum I(x_i; x_j) \quad (4)$$

Optimizing D and R values simultaneously, we get the ideal set of features that can be implemented in classification of the data. Details of the mRMR method can be found in [11]. For selecting the feature subset after ranking, the cross validation error from the classifier is computed sequentially for the ranked features and plotted against the number of features. From this error plot, the area with consistently lowest error is marked and the smallest number of features from that area are chosen as the candidate feature set [11].

Once a feature subset is chosen from the mutual information based filter, a recursive backward elimination (RFE) wrapper is applied over this subset. Unlike the filter, the wrapper explicitly minimizes the classification error by recursively excluding one feature with the constraint that the resultant feature set has lower leave one out (LOO) error [12].

2.3 Training Classifiers Using SVM

A non-linear support vector machine (SVM) is amongst the most powerful pattern classification algorithms, as it can obtain maximal generalization when predicting the classification of previously unseen data compared to other nonlinear classifiers [13]. By using a kernel function, it maps the original features into higher dimensional space where it computes a hyperplane that maximizes the distance from the hyperplane to the examples in each class. Having found such a hyperplane, the SVM can then predict the classification of an unlabeled example by mapping it into the feature space and checking on which side of the separating plane the example lies.

The input to the classifier consists of feature matrix and a vector of class labels consisting of 1 and -1 values defining the two classes (the controls are labeled

as 1 and the patients as -1, in our case). The classifier performs a non-linear mapping $\phi : R^d \rightarrow H$ by using a kernel function, which maps the feature space R^d to a higher dimensional space H . We use the Gaussian radial basis function as a kernel function.

Based on the distance from the hyperplane, the classifier computes a probabilistic score between 1 and -1 for each test subject. When the probabilistic score is ≥ 0 , the subject is classified as class 1 (controls), otherwise as class 2 (patients). The probabilistic classifier score, therefore, represents the level of abnormality in the subject.

2.4 Cross-Validation

We compute the probabilistic abnormality score for each subject by implementing the leave one out (LOO) cross validation method. In the LOO validation, one sample is chosen for testing, while other samples are trained using the methods described in section 2.3. The classifier is evaluated based on the classification result of the test subject. By repeatedly leaving each subject out as a test subject, obtaining its abnormality score, and averaging over all the left-out subjects we obtain the average classification rate. Validation is also performed by plotting the receiver operating characteristic (ROC) curves. The ROC is a plot of sensitivity vs. (1-specificity) of the classifier when the discrimination threshold is varied.

3 Results

Two datasets were used for carrying out the classification:

- The SCZ dataset consisted of 37 female controls and 27 female SCZ patients. The DWI images were acquired on Siemens 3T scanner with $b=1000$ and 64 gradient directions.
- The ASD dataset consisted of 25 children with ASD and 23 typically developing kids. The DWI images were acquired on Siemens 3T scanner with $b=1000$ and 30 gradient directions.

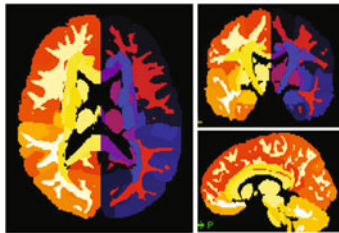


Fig. 1. (a) ROI's introduced by Wakana et al. [8] showing 176 structures

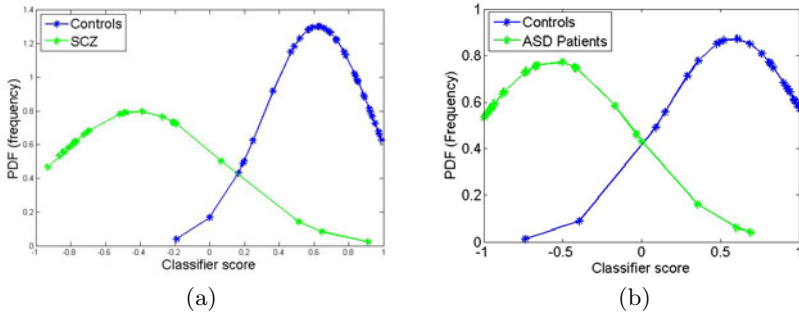


Fig. 2. Plot showing LOO classification (abnormality) scores vs. frequency of occurrence of the score value (PDF) where green indicates patients and blue indicates controls. Each subject's score is represented with a star. (a) SCZ data where the stars overlap as a result of a large number of samples. The controls and patients are very well separated with 90.62% accuracy. (b) ASD dataset shows 89.58% classification even with less number of samples.

The SCZ dataset has a large sample size in comparison with the ASD dataset. Moreover, the SCZ data consists of a cohesive population as only females are considered, while the ASD data consists of a heterogeneous population. We chose such different datasets to test the versatility of our classification method.

As mentioned in section 2, the scans were spatially normalized to the DTI template. Subsequently, the mean FA and MD measures for each ROI were derived. Feature selection was performed using mutual information values and the RFE wrapper as described in section 2.2. After employing the filter, 130 top features for SCZ dataset were chosen while for ASD 150 top features were picked. The selection of number of features was based on the error plot as described in section 2.2. From this feature subset, the RFE wrapper chose 25 top features for SCZ and 17 top features for ASD by minimizing the classification error. For the non-linear classifier, a suitable kernel size was determined based on the ROC curves, after testing different σ values ranging from 0.001 to 1. It was found that the kernel size ranging from 0.05 - 0.15 was optimal. The C value, which is the trade off parameter between training error and SVM margin, was tested for 1,5,7,10 and 50, and it was observed that values between 5 and 50 were a better choice for our classification.

Using the LOO method, we computed the classifier score for each subject. Figure 2 shows the classification results. A normal probability density (PDF) which represents the likelihood of each LOO score is plotted against the LOO score. The average LOO classification for SCZ data was 90.62% (6 subjects misclassified) while for ASD data it was 89.58% (5 subjects misclassified). The ROC curves are shown in figure 3. Both the datasets shows a steep curve with a large area under the curve (AUC) of 0.905 for SCZ and 0.913 for ASD (fig. 3(a and b)). We also applied our method to test asymptomatic relatives of SCZ patients. The abnormality scores of the relatives were closer to the patients than the healthy controls, suggesting similar brain abnormality patterns.

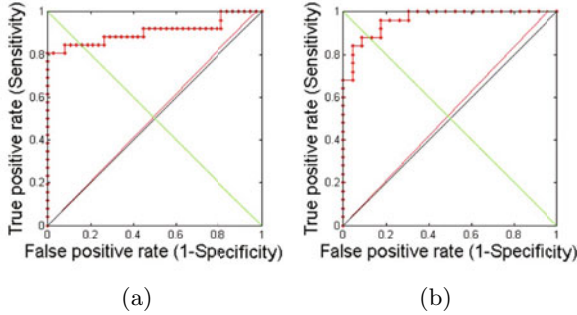


Fig. 3. ROC curves for classification between controls and patients. The area under the curve (AUC) defines the level of performance. (a) SCZ data with an AUC of 0.905 (b) ASD with an AUC of 0.913.

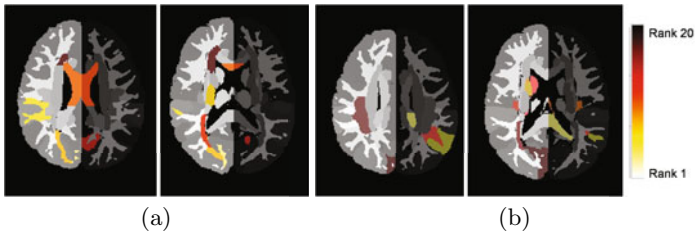


Fig. 4. Top ranked ROI's mapped on template image. These ROI's contributed largely towards classification based two step feature selection.(a) SCZ (b) ASD data.

Besides computing the abnormality score, it is important to know which areas of the brain significantly contribute towards the classification. The selected features are the most discriminating features and contribute most towards group differences between controls and patients. Fig. 4 shows the top 25 ROI's for SCZ (fig. 4(a)) and 17 ROI's for ASD (fig. 4(b)) that were chosen by the wrapper are overlaid on the template image and color coded according to their rank.

4 Conclusion

We have presented a classification methodology based on anatomically meaningful DTI features for identification of white matter abnormality in diseases like SCZ and ASD. Use of ROI based features reduces dimensionality and simplifies the clinical interpretation of pathology induced changes. Superior experimental results over different populations indicate that our method can successfully classify patients and controls, despite the sample size and heterogeneity of the population. Based on the abnormality score of the test subject, this method can quantify the degree of pathology and the score can be potentially used as a clinical biomarker.

Thus, our method can be applied in various studies involving disease progression, treatment effects etc. The application to family members underlines the applicability of the method to prognosis. Future work includes extending the classification method for controlling for age and treatment effects. Moreover, we can add other features like radial and axial diffusivity measures or utilize the whole tensor in log-euclidean form as our feature for further investigation.

References

1. Lao, Z., Shen, D., Xue, Z., Karacali, B., Resnick, S.M., Davatzikos, C.: Morphological classification of brains via high-dimensional shape transformations and machine learning methods. *Neuroimage* 21(1), 46–57 (2004)
2. Golland, P., Grimson, W.E.L., Shenton, M.E., Kikinis, R.: Detection and analysis of statistical differences in anatomical shape. *Med. Image Anal.* 9(1), 69–86 (2005)
3. Yushkevich, P., Joshi, S., Pizer, S.M., Csernansky, J.G., Wang, L.E.: Feature selection for shape-based classification of biological objects. *Inf. Process Med. Imaging* 18, 114–125 (2003)
4. Fan, Y., Shen, D., Gur, R.C., Gur, R.E., Davatzikos, C.: Compare: classification of morphological patterns using adaptive regional elements. *IEEE Trans Med. Imaging* 26(1), 93–105 (2007)
5. Pohl, K.M., Sabuncu, M.R.: A unified framework for mr based disease classification. *Inf. Process Med. Imaging* 21, 300–313 (2009)
6. Caan, M.W.A., Vermeer, K.A., van Vliet, L.J., Majoie, C.B.L.M., Peters, B.D., den Heeten, G.J., Vos, F.M.: Shaving diffusion tensor images in discriminant analysis: a study into schizophrenia. *Med. Image Anal.* 10(6), 841–849 (2006)
7. Wang, P., Verma, R.: On classifying disease-induced patterns in the brain using diffusion tensor images. *Med. Image Comput. Comput. Assist Interv.* 11(Pt 1), 908–916 (2008)
8. Wakana, S., Jiang, H., Nagae-Poetscher, L.M., van Zijl, P.C.M., Mori, S.: Fiber Tract-based Atlas of Human White Matter Anatomy. *Radiology* 230(1), 77–87 (2004)
9. Ingalhalikar, M., Yang, J., Davatzikos, C., Verma, R.: Dti-droid: Diffusion tensor imaging-deformable registration using orientation and intensity descriptors. *International Journal of Imaging Systems and Technology* 20(2), 99–107 (2010)
10. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. *J. Mach. Learn. Res.* 3, 1157–1182 (2003)
11. Peng, H., Long, F., Ding, C.: Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal. Intell.* 27(8), 1226–1238 (2005)
12. Rakotomamonjy, A.: Variable selection using svm based criteria. *J. Mach. Learn. Res.* 3, 1357–1370 (2003)
13. Vapnik, V.N.: An overview of statistical learning theory. *IEEE Trans Neural Netw.* 10(5), 988–999 (1999)

Probabilistic Anatomical Connectivity Using Completion Fields

Parya MomayyezSiahkal and Kaleem Siddiqi

Centre for Intelligent Machines, School of Computer Science, McGill University
{pmamay,siddiqi}@cim.mcgill.ca

Abstract. Diffusion magnetic resonance imaging has led to active research in the analysis of anatomical connectivity in the brain. Many approaches have been proposed to model the diffusion signal and to obtain estimates of fibre tracts. Despite these advances, the question of defining probabilistic connectivity indices which utilize the relevant information in the diffusion MRI signal to indicate connectivity strength, remains largely open. To address this problem we introduce a novel numerical implementation of a stochastic completion field algorithm, which models the diffusion of water molecules in a medium while incorporating the local diffusion MRI data. We show that the approach yields a valid probabilistic estimate of connectivity strength between two seed regions, with experimental results on the MICCAI 2009 Fibre Cup phantom [\[1\]](#).

Keywords: Diffusion-MRI, Tractography, Connectivity, 3D stochastic completion field.

1 Introduction

During the past decade advances in diffusion magnetic resonance imaging have led to the development of a variety of tractography algorithms which reconstruct connectivity patterns between distinct cortical and subcortical areas in the human brain. The early deterministic tractography methods, based on streamline tracking and diffusion tensor imaging [\[2\]](#), verified the inherent potential of diffusion-MRI to capture anatomical connectivity non-invasively. However, these approaches suffered from the limitation that they did not explicitly consider the effects of partial volume averaging, the underlying noise in the data and other imaging artifacts. More recent probabilistic tractography algorithms [\[3,4,5\]](#) have attempted to address these concerns by providing an approximation of the underlying noise and incorporating it into the fibre propagation framework. Whereas some of these latter schemes employ Monte-Carlo simulations to perform tractography for many iterations [\[3\]](#), others use front propagation approaches to evolve a surface from a seed region and then find the minimum cost trajectories between each voxel and the seed region by back propagating through a “time of arrival” map [\[4\]](#). In these approaches a measure of connectivity is either defined based on the number of times a voxel has been passed through [\[3,6\]](#) or by considering the lowest confidence value of all tract segments along a reconstructed

streamline [4]. In a related but distinct approach, Fletcher *et al.* propose a front evolution scheme to compute a connectivity measure between two regions of interest without explicitly obtaining fibre tracts first [7].

A different class of algorithms tackles the problem of fibre tractography and white matter connectivity by solving a partial differential equation (PDE) and finding the minimal path given the PDE [8,9]. O'Donnell *et al.* use diffusion tensor estimation and solve for the steady state concentration of the diffusion equation governed by the Fick's law in an anisotropic medium. A Riemannian metric is then used to compute geodesic distances between two points, which in turn is interpreted as their degree of connectivity. Batchelor *et al.* improve this framework by adding a convection term which incorporates a measure of the anisotropy in the diffusion MRI data at each voxel. The main advantage of these algorithms is their potential to incorporate all the relevant information contained in orientation distribution functions (ODFs), or any other diffusion model, into the tractography framework.

In work that is somewhat distinct from the above methods Momayyez *et al.* have modeled the Brownian motion of water molecules by a probabilistic 3D random walk, leading to an application of stochastic completion fields to fibre tractography [10]. While this framework is also PDE-based, the given PDE describes the evolution of the probability distribution of particles following the random walk, as indicated by the Fokker-Planck equation. The present article extends their results by developing a new, unconditionally stable numerical implementation, which allows for the local diffusion data to be incorporated directly into the stochastic completion field framework. The extension allows for probabilistic estimates of fibre tracts between two seed regions with the property that connectivity strength emerges as an inherent characteristic. We validate the new approach by comparing it with a state-of-the-art probabilistic tractography algorithm on the MICCAI 2009 Fiber Cup Phantom [1].

2 3D Stochastic Completion Fields

Using an extension of a model introduced by Williams and Jacobs [11], a *3D stochastic completion field* represents the probability that a particle undergoing Brownian motion passes through a given location (x, y, z) with orientation (θ, ϕ) in 3D, while completing a path between two seed regions. In a 3D random walk, particles tend to travel in straight lines with slight changes in orientation controlled by two successive deviations in the osculating and binormal planes, which are respectively proportional to the diffusion parameters σ_ϕ^2 and σ_θ^2 . Additionally, a fraction of particles $(1 - e^{-\frac{1}{\zeta}})$ decays per unit time step. While σ_ϕ and σ_θ maintain the prior assumption of smoothness, the latter decay rate ζ gives more support to shorter paths. Furthermore, the inherent Markov assumption of the 3D random walk leads to a notion of a *stochastic source field* and a *stochastic sink field*. The former represents the probability of passing through any state (x, y, z, θ, ϕ) for a particle which begins in a source state $(x_p, y_p, z_p, \theta_p, \phi_p)$, while the latter represents the probability of reaching a final

sink state $(x_q, y_q, z_q, \theta_q, \phi_q)$ from any initial state. The *stochastic completion field* is computed by scalar multiplication of the source and sink fields.

While Monte-Carlo simulation provides a direct way to compute the stochastic completion field, a more efficient and powerful alternative is to discretize the Fokker-Planck equation which expresses the evolution in time of the probability density function of the particles analytically [10]:

$$\frac{\partial P}{\partial t} = -\sin \theta \cos \phi \frac{\partial P}{\partial x} - \sin \theta \sin \phi \frac{\partial P}{\partial y} - \cos \theta \frac{\partial P}{\partial z} + \frac{\sigma_\phi^2}{2} \frac{\partial^2 P}{\partial \phi^2} + \frac{\sigma_\theta^2}{2} \frac{\partial^2 P}{\partial \theta^2} - \frac{1}{\zeta} P. \quad (1)$$

Here P is the probability that a particle passes through state (x, y, z, θ, ϕ) in the course of its random walk from a source to a sink region.

Whereas the results in [10] demonstrate a type of proof of concept, they are based on a simple first-order finite difference scheme to solve for the final steady-state distribution of P , and as such are quite limited. In particular, such a numerical scheme suffers from instability for complicated fibre pathways with highly curving or twisting tracts. Moreover, the estimation of the completion field in [10] is performed in a manner that is independent of the local diffusion-MRI data, with the final connectivity pattern being obtained by post-multiplication in a Bayesian framework. The Bayesian approach further requires empirical adjustments of parameters which is a significant challenge without prior knowledge of the fibre pathways being estimated.

3 An Unconditionally Stable Numerical Scheme

The Fokker-Planck equation given by Eq. 1 is a multi-dimensional PDE which can be solved by implementing an operator splitting strategy, where each time unit is divided into six subunits, each of which formulates separate numerical scheme for each variable. When performing operator splitting, it is usually sufficient to have stable numerical estimation for each operator in order to make the overall scheme stable. For the x , y and z dimensions which are included in the advection part of Eq. 1, we propose to use the Lax-Wendroff scheme which is second-order accurate in time and space. Since in Eq. 1 all the advection coefficients $(-\sin \theta \cos \phi, -\sin \theta \sin \phi, -\cos \theta)$ are always smaller than or equal to one, instability is not an issue. The main challenge in terms of numerical estimation is for the σ (diffusion) terms. These terms can get large for fibre tracts with highly curved or twisted segments. To handle this we propose the use of the more complex implicit and unconditionally stable Crank-Nicholson numerical scheme. The numerical estimation is again second order in time and space but because of its implicit character, it involves matrix computation. However the method eliminates the need for thresholds on maximum curvature or torsion values and makes the local incorporation of the diffusion MRI data into the framework possible. The overall numerical scheme is summarized by the following set of equations, where t denotes the time parameter:

$$P_{x,y,z,\theta,\phi}^{t+\frac{1}{6}} = P_{x,y,z,\theta,\phi}^t - \sin \theta \cos \phi \frac{\Delta t}{2\Delta x} (P_{x+\Delta x,y,z,\theta,\phi}^t - P_{x-\Delta x,y,z,\theta,\phi}^t) \\ + \frac{(\sin \theta \cos \phi)^2 \Delta t^2}{2\Delta x^2} (P_{x+\Delta x,y,z,\theta,\phi}^t - 2P_{x,y,z,\theta,\phi}^t + P_{x-\Delta x,y,z,\theta,\phi}^t)$$

$$\begin{aligned}
 P_{x,y,z,\theta,\phi}^{t+\frac{1}{3}} &= P_{x,y,z,\theta,\phi}^{t+\frac{1}{6}} - \sin\theta \sin\phi \frac{\Delta t}{2\Delta y} \left(P_{x,y+\Delta y,z,\theta,\phi}^{t+\frac{1}{6}} - P_{x,y-\Delta y,z,\theta,\phi}^{t+\frac{1}{6}} \right) \\
 &\quad + \frac{(\sin\theta \sin\phi)^2 \Delta t^2}{2\Delta y^2} \left(P_{x,y+\Delta y,z,\theta,\phi}^{t+\frac{1}{6}} - 2P_{x,y,z,\theta,\phi}^{t+\frac{1}{6}} + P_{x,y-\Delta y,z,\theta,\phi}^{t+\frac{1}{6}} \right) \\
 P_{x,y,z,\theta,\phi}^{t+\frac{1}{2}} &= P_{x,y,z,\theta,\phi}^{t+\frac{1}{3}} - \cos\theta \frac{\Delta t}{2\Delta z} \left(P_{x,y,z+\Delta z,\theta,\phi}^{t+\frac{1}{3}} - P_{x,y,z-\Delta z,\theta,\phi}^{t+\frac{1}{3}} \right) \\
 &\quad + \frac{(\cos\theta)^2 \Delta t^2}{2\Delta z^2} \left(P_{x,y,z+\Delta z,\theta,\phi}^{t+\frac{1}{3}} - 2P_{x,y,z,\theta,\phi}^{t+\frac{1}{3}} + P_{x,y,z-\Delta z,\theta,\phi}^{t+\frac{1}{3}} \right) \\
 P_{x,y,z,\theta,\phi}^{t+\frac{2}{3}} &- \frac{\sigma_\phi^2 \Delta t}{4(\Delta\phi)^2} \left(P_{x,y,z,\theta,\phi+\Delta\phi}^{t+\frac{2}{3}} - 2P_{x,y,z,\theta,\phi}^{t+\frac{2}{3}} + P_{x,y,z,\theta,\phi-\Delta\phi}^{t+\frac{2}{3}} \right) = \\
 &\quad P_{x,y,z,\theta,\phi}^{t+\frac{1}{2}} + \frac{\sigma_\phi^2 \Delta t}{4(\Delta\phi)^2} \left(P_{x,y,z,\theta,\phi+\Delta\phi}^{t+\frac{1}{2}} - 2P_{x,y,z,\theta,\phi}^{t+\frac{1}{2}} + P_{x,y,z,\theta,\phi-\Delta\phi}^{t+\frac{1}{2}} \right) \\
 P_{x,y,z,\theta,\phi}^{t+\frac{5}{6}} &- \frac{\sigma_\theta^2 \Delta t}{4(\Delta\theta)^2} \left(P_{x,y,z,\theta+\Delta\theta,\phi}^{t+\frac{5}{6}} - 2P_{x,y,z,\theta,\phi}^{t+\frac{5}{6}} + P_{x,y,z,\theta-\Delta\theta,\phi}^{t+\frac{5}{6}} \right) = \\
 &\quad P_{x,y,z,\theta,\phi}^{t+\frac{2}{3}} + \frac{\sigma_\theta^2 \Delta t}{4(\Delta\theta)^2} \left(P_{x,y,z,\theta+\Delta\theta,\phi}^{t+\frac{2}{3}} - 2P_{x,y,z,\theta,\phi}^{t+\frac{2}{3}} + P_{x,y,z,\theta-\Delta\theta,\phi}^{t+\frac{2}{3}} \right) \\
 P_{x,y,z,\theta,\phi}^{t+1} &= e^{-\frac{1}{\zeta}} P_{x,y,z,\theta,\phi}^{t+\frac{5}{6}}. \tag{2}
 \end{aligned}$$

It is clear from Eq. 2 that the computation at each voxel is local and hence allows for parallelization and reduction in time complexity of the algorithm. In addition, for every voxel distinct diffusion σ_θ , σ_ϕ and decay ζ coefficients can be specified using properties of the local diffusion-MRI data. Thus, the applicability of the stochastic completion field model to fibre tractography is significantly extended. We re-emphasize that although the new numerical scheme is computationally more complex than that of [10] (the run time increases by approximately 40%), the method in [10] has limited practical utility for curved fibre pathways.

4 Local Incorporation of Diffusion-MRI Data

To tune the diffusion parameters σ_ϕ , σ_θ and the decay parameter ζ locally, we use information from the fibre orientation distribution (FOD) function at each voxel, calculated using the approach of Anderson [12]. The basic idea is to decrease diffusion (and correspondingly increase the likelihood of trajectories) in directions that are aligned with fiber orientations supported by the data. Specifically, σ_ϕ and σ_θ are determined by finding the closest FOD maximum to each (ϕ, θ) direction and computing the two angular differences between a given direction and the chosen maximum. The local corresponding σ_ϕ and σ_θ values are then set equal to these angular differences. The decay rate ζ is also chosen based on the angular difference between a given direction (ϕ, θ) and the chosen FOD maximum. When the angular difference is greater than $\pi/4$, the decay coefficient is made very small to suppress trajectories in directions that are far from the FOD maxima. When the (ϕ, θ) direction closely matches the FOD maximum, a fixed decay coefficient is assigned to it. Overall this strategy allows parameters

to be set automatically from the FODs and it supports smoothly varying fibre tracts while avoiding an excessive penalty for curved fibre pathways.¹

To reconstruct the connectivity pattern, two independent probability distributions are evolved based on Eq. 1. In the first, the particles are initially placed in one seed region to generate the source field. In the second the particles are constrained to end in the other seed region to generate the sink field. The final stochastic completion field, which is a probabilistic representation of the connectivity pattern, is given by the multiplication of these two fields.

5 Connectivity Strength

Defining an index of anatomical connectivity is itself a topic of great interest in the neuroscience community. Most of the connectivity measures proposed thus far in the literature are based on tractography algorithms. Two of the most commonly used connectivity measures reported in the literature consists of: (1) *voxel counting*: counting the number of times an ROI has been reached from a seed region when employing a Monte-Carlo simulation [3][5][6] or, (2) finding the *weakest link*: assigning a weight to each tract segment along a reconstructed fibre pathway based on some weight function or on the previous Monte-Carlo approach and considering the lowest confidence value of all tract segments along the streamline [4]. A limitation of these methods is that they are heuristic in nature. They do not incorporate relevant information along the reconstructed fibre tract into the connectivity measure in a principled fashion. They also suffer from reduced connectivity values with distance from the seed region.

The stochastic completion field model offers a potential solution. The algorithm's output is the steady-state probability distribution of all the states when joining the source and sink regions. In other words, the probability of reaching each state has been calculated during the evolution of the probability distributions from the source and the sink regions while integrating all the computed probabilities along the connectivity pattern. Consequently, the probability values computed for the source and sink ROIs during the computation actually represent the measure of connectivity between them. Moreover, choosing the decay rate based on the approximate length of the fibre tract eliminates the length dependency issue. To obtain the connectivity strength between the source and sink ROIs, it is therefore sufficient to compute the average probability values of all the voxels included in these regions.

¹ Ultimately, in order to make the results obtained for different fibre tracts comparable, the decay rate has to be chosen relative to the length of each fibre. To accomplish this the algorithm is run once using a decay rate proportional to the Euclidean distance between the source and sink regions. Following this run, the obtained connectivity pattern is used to approximate the length of the underlying fibre tracts. This length is then used to set the fixed decay rate for a second run of the algorithm.

6 Experiments

6.1 Diffusion-MRI Phantom Data

The proposed algorithm was tested on the MICCAI 2009 Fibre cup phantom [1], for which ground truth fibre tracts are available. The phantom is made of hydrophobic acrylic fibres whose diameter is of the same scale as that of myelinated axons [2]. Two diffusion-MRI datasets have been provided with different spatial resolutions of $3 \times 3 \times 3 \text{mm}^3$ and $6 \times 6 \times 6 \text{mm}^3$ of which we have selected the higher resolution images with an associated b -value of 1500 s/mm^2 .

6.2 Results

The results of the algorithm for six different source and sink region pairs are shown in Fig. 1, which depicts the steady-state probability distribution for each pair. Based on the ground truth, it is known that the first four pairs are associated with a true fibre tract while the latter two are not. For the purpose of comparison, we have implemented a probabilistic fiber tractography algorithm similar to that of [5], where a residual bootstrapping scheme is first used to estimate the uncertainty of a voxels FOD using a single HARDI acquisition. The estimated uncertainties are then used in a Monte-Carlo simulation of a probabilistic streamline tractography algorithm. Confidence values are also assigned to the reconstructed streamlines using a weakest link approach. The probabilistic streamline tractography was run for 500 iterations, with a seed frequency of 10 per voxel for all ROIs and a threshold of 75° for the maximum turning angle from one voxel to the next. The latter threshold was optimized empirically.

Table 1 summarizes the connectivity indices obtained from our algorithm and from probabilistic tractography. The connectivity index for each scenario is computed by taking the average of the corresponding values for all the voxels included in the sink and source ROIs.

Table 1. Connectivity indices obtained for six different seed region pairs

Algorithm	(a)	(b)	(c)	(d)	(e)	(f)
Stochastic Completion Field	0.17152	0.22445	0.04716	0.24342	0.00551	0.07559
Probabilistic Tractography	0.00008	0.24051	0.00819	0	0	0

There are several observations that can be made from Table 1. First, the connectivity indices obtained from probabilistic tractography span a wide range, which can be due to the heuristic nature of the algorithm and its dependency on length. Moreover, while the algorithm assigns a connectivity index of zero to cases (e) and (f), which are not associated with a true fibre pathway, it is also not capable of finding the true connectivity for case (d), which is a straight fibre tract and should be an easy pathway to reconstruct. The connectivity

² For details, visit <http://www.lnao.fr/spip.php?article107>

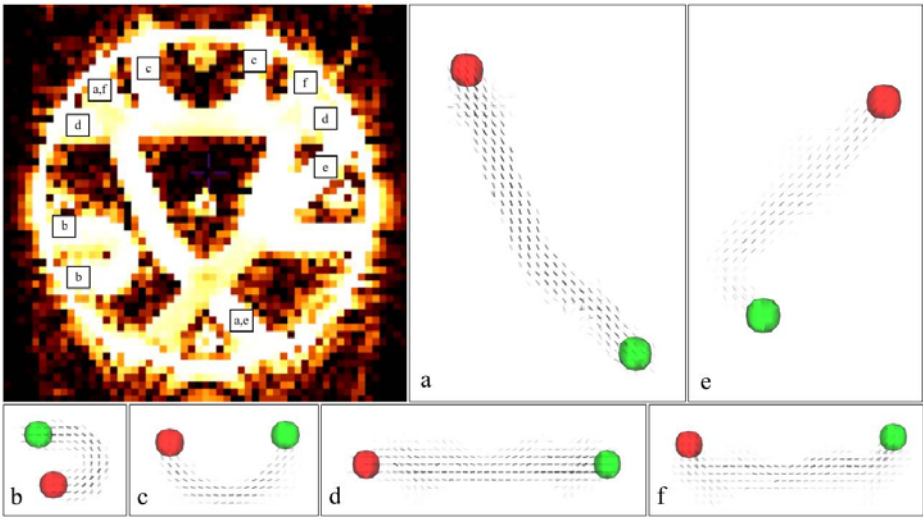


Fig. 1. Results of the local completion field algorithm. The probability of a state is inversely proportion to the transparency of the associated orientation (The reader is encouraged to zoom-in on the electronic version to get a better sense of the 3D characteristics). Top-Left: Mean diffusivity image and the seed regions used.

measures obtained by the completion field algorithm, on the other hand, are comparatively in the same range for cases (a), (b) and (d), which represent true connections. The value obtained for case (e) is very small, showing the low confidence of the algorithm in finding a connection. For cases (c) and (f), however, the algorithm fails to assign a higher connectivity to tract (c). This is due to the fact that the true fibre tracts form a bottle neck with another pathway along the streamline. As a result, at a local scale the algorithm favors the smoother connections leading to a lower connectivity value for case (d) and a higher value for case (f). This failure can be essentially solved by acquiring higher resolution data. In the meantime, one of our future goals is to modify the algorithm to incorporate constraints such as curvature consistency to be able to handle such ambiguous situations.

7 Conclusion

We have addressed the problem of anatomical connectivity by extending the stochastic completion field algorithm introduced in [10]. A new, unconditionally stable numerical implementation has been developed which allows for the local incorporation of the diffusion data into the framework. The extension provides a basis for theoretically derived measures of connectivity based on probabilistic estimates of fibre tracts. The results obtained on the phantom demonstrate the inherent ability of the algorithm to compute consistent connectivity measures and to reconstruct the most probable fibre pathways between two seed regions.

This approach can find highly curving fibre tracts without the need to impose thresholds on maximum curvature or torsion. The algorithm has low computational complexity, which is independent of the size of the seed regions, and the automatic tuning of its parameters is a further advantage.

References

1. Poupon, C., Rieul, B., Kezele, I., Perrin, M., Poupon, F., Mangin, J.: New diffusion phantoms dedicated to the study and validation of high-angular-resolution diffusion imaging (HARDI) models. *J. Magnetic Resonance in Medicine* 60, 1276–1283 (2008)
2. Mori, S., Crain, B.J., Chacko, V.P., Zijl, P.C.M.: Three dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of Neurology* 45, 265–269 (1999)
3. Behrens, T.E., Berg, H.J., Jbabdi, S., Rushworth, M.F., Woolrich, M.W.: Probabilistic diffusion tractography with multiple fibre orientations: what can we gain? *NeuroImage* 34, 144–155 (2007)
4. Parker, G.J.M., Wheeler-Kingshott, C.A.M., Barker, G.J.: Estimating distributed anatomical connectivity using fast marching methods and diffusion tensor imaging. *IEEE Transaction on Medical Imaging* 21, 505–512 (2002)
5. Berman, J.I., Chung, S., Mukherjee, P., Hess, C.P., Han, E.T., Henry, R.G.: Probabilistic streamline q-ball tractography using the residual bootstrap. *NeuroImage* 39, 215–222 (2008)
6. Yo, T.S., Anwender, A., Descoteaux, M., Fillard, P., Poupon, C., Knösche, T.R.: Quantifying brain connectivity: A comparative tractography study. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009*. LNCS, vol. 5761, pp. 886–893. Springer, Heidelberg (2009)
7. Fletcher, P.T., Tao, R., Jeong, W.K., Whitaker, R.T.: A volumetric approach to quantifying region-to-region white matter connectivity in diffusion tensor MRI. In: Karssemeijer, N., Lelieveldt, B. (eds.) *IPMI 2007*. LNCS, vol. 4584, pp. 346–358. Springer, Heidelberg (2007)
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: Dohi, T., Kikinis, R. (eds.) *MICCAI 2002*. LNCS, vol. 2488, pp. 459–466. Springer, Heidelberg (2002)
9. Batchelor, P.G., Hill, D.L.G., Atkinson, D., Calamante, F., Connelly, A.: Fibre-tracking by solving the diffusion-convection equation. In: *Proceedings of Int. Soc. Mag. Reson. Med., Honolulu*, vol. 10 (2002)
10. Momayyez, P., Siddiqi, K.: 3D stochastic completion fields for fiber tractography. In: *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA 2009)*, Miami Beach, pp. 178–185 (2009)
11. Williams, L., Jacobs, D.: Stochastic completion fields: A neural model of illusory contour shape and salience. *Neural Computation* 9, 837–858 (1997)
12. Anderson, A.W.: Measurement of fiber orientation distributions using high angular resolution diffusion imaging. *J. Magnetic Resonance in Medicine* 54, 1194–1206 (2005)

Reconstruction of Scattered Data in Fetal Diffusion MRI

Estanislao Oubel¹, Meriam Koob², Colin Studholme³, Jean-Louis Dietemann²,
and François Rousseau¹

¹ LSIT, UMR 7005, CNRS-Université de Strasbourg, France

² LINC, UMR 7191, CNRS-Université de Strasbourg, France

³ UCSF, Radiology Department, San Francisco, CA 94143

Abstract. In this paper we present a method for reconstructing D-MRI data on regular grids from sparse data without assuming specific diffusion models. This is particularly important when studying the fetal brain *in utero*, since registration methods applied for movement and distortion correction produce scattered data in spatial and angular (gradient) domains. We propose the use of a groupwise registration method, and a dual spatio-angular interpolation by using radial basis functions (RBF). Experiments performed on adult data showed a high accuracy of the method when estimating diffusion images in unavailable directions. The application to fetal data showed an improvement in the quality of the sequences according to criteria based on fractional anisotropy (FA) maps, and differences in the tractography results.

1 Introduction

Diffusion Magnetic Resonance Imaging (D-MRI) is an imaging modality increasingly used for studying the normal and pathological development of the fetal brain. Since the limited sensitivity of fetal ultrasound to detect and depict the maturational processes of the developing white matter, D-MRI can be considered as one of the most promising methods for studying *in vivo* the human white matter development.

Eddy current-induced image distortions and patient motion during prolonged acquisitions cause image misalignment in D-MRI sequences, invalidating the assumption of a consistent relationship between image space and anatomy in the image processing. Such distortions are generally assumed to be affine, and usually corrected by coregistration with the T_2 -weighted image (T_2^{epi}) by using affine transformation models and mutual information (MI) [1]. When imaging a fetus, there exists the additional problem of movement, and the acquired slices are no longer a regular sample of the volume [2][3]. This may be compensated by relaxing the 3D transformation model to a set of transformations applied to each slice independently. When these transformations are applied to the original sequence, scattered data are generated in the spatial and angular (gradient) domains, requiring an interpolation in both fields. Recently, Jiang *et al.* [3] have presented

an interpolation method assuming that the local diffusion properties can be represented by a rank-2 tensor model. However, this model cannot describe voxels containing multiple fibers with different orientations, a condition referred to as intravoxel orientational heterogeneity (IVOH) [4].

In this paper, we present a reconstruction method independent of the diffusion model than can be used with more complex diffusion models like Gaussian mixture models or higher-order diffusion tensors. Distortion corrections are performed by applying groupwise registration with an affine slice-by-slice transformation model. A dual RBF-based interpolation in the spatial and angular domains was then used for reconstruction on a regularly sampled grid.

2 Method

2.1 Origin of Sparse Data

Typically, a D-MRI sequence consists of a set of $N + 1$ regularly sampled images $\mathcal{S} = \{S_0, S_1, \dots, S_N\}$ where S_0 is the image obtained without diffusion weighting, and $S_{i=1:N}$ are diffusion-weighted (DW) images obtained with diffusion-sensitizing gradients G_i of direction U_i and strength b . Ideally, S_0 and S_i are related by the *Stejskal-Tanner* equation, but distortions caused by eddy currents and fetal motion invalidate this relationship. Image registration techniques can be applied to restore the lost spatial correspondence between S_0 and S_i , but a reconstruction from sparse data is then required.

Let us express the original sequence as $\mathcal{S} = \{(X, \Theta, S(X, \Theta))\}_{X \in \Omega_s, \Theta \in \Omega_a}$ where Ω_s is a regular spatial grid, and Ω_a is a regular angular grid on the unit sphere. After correction, \mathcal{S} becomes $\mathcal{S}' = \{(X', \Theta', S(X, \Theta))\}_{X \in \Omega_s, \Theta \in \Omega_a}$, where $X' = A_i^z(X)$ are the transformed spatial points, and $\Theta' = R_i^z(\Theta)$ are the coordinates of U_i corrected with the rotational component R_i^z of A_i^z . In these equations, the superscript “ z ” has been added to indicate the slice-to-slice nature of the transformations. Differently of (X, Θ) , the transformed coordinates (X', Θ') do not belong to regular grids, and \mathcal{S}' results sparse.

2.2 Registration Approach

Distortion correction methods relying on the registration of S_i to S_0 may fail when applied to fetal D-MRI, probably a consequence of the large differences between both images [3]. In this paper, we propose a method that takes advantage of the similarity between DW images to first ensure their joint alignment. The mean of the registered images is characterized by a higher SNR than images S_i , and provides a better depiction of the anatomical structure of the brain. These properties allow an accurate registration to S_0 , necessary to map all the sequence in its space of coordinates. This method is illustrated and explained in Figure 1. In essence, the method is similar to the one propose by Guimond *et al.* [5]. The main difference is the separation of the registration into (1) a groupwise registration of the DW images, and (2) the registration of the resulting mean image with the anatomical image.

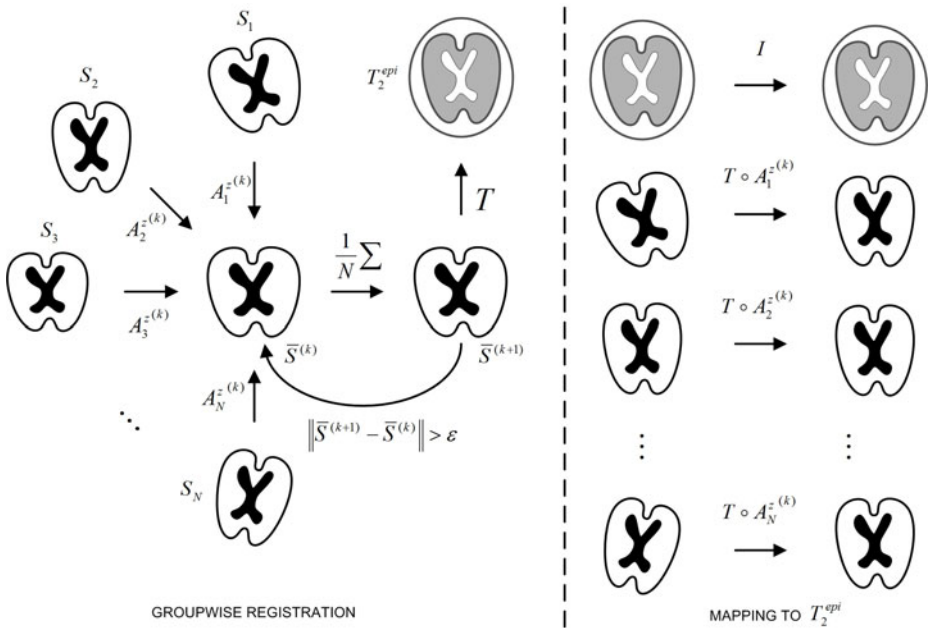


Fig. 1. Groupwise registration method. Initially all images S_i are registered to an arbitrary DW image S^* taken as reference. Then, all images are resampled and averaged to obtain a first average $\bar{S}^{(1)}$, which becomes the new reference. At iteration k , the images S_i are registered to the mean $\bar{S}^{(k)}$ to obtain the transforms $A_i^z(k)$. After convergence, $\bar{S}^{(k)}$ is registered to T_2^{epi} , and the composition of transformations $T \circ A_i^z(k)$ is applied to map all images S_i into the coordinate system of T_2^{epi} .

2.3 Sparse Interpolation

To interpolate values on a regular grid from scattered data, we have used radial basis functions (RBF). RBFs have already been applied for interpolation on spherical geodesic grids in the context of numerical weather prediction, outperforming linear interpolation strategies [6]. The idea behind RBF interpolation is that every point has an influence on a neighborhood according to some functional $\phi(r)$, where r is the distance from the point. Then, the value of the function at the point P is given by a linear combination of the ϕ 's centered in the points P_i :

$$y(P) = \sum_{i=0}^{N-1} w_i \phi(\|P - P_i\|) \tag{1}$$

where the weights w_i are calculated by solving a linear system of equations for the function to agree with the observations at points P_i .

In this paper, we have used a Gaussian function as RBF since this function tends to zero for high r , and the influence of points P_i distant from P can be neglected. This allows considering only points in a neighborhood $\mathcal{N}(P)$ of P

for interpolation, which reduces the computational complexity of the method. $\mathcal{N}(P)$ was formed by points P_i falling inside the *support region* of the Gaussian function, defined in the context of this paper as the interval $[-s_\phi, +s_\phi]$ so that $\phi(s_\phi) = 0.01 \times \phi(0)$.

In our case, each point contains spatial and angular coordinates which must be considered separately because of the difference in scale between both types of coordinates. This situation is different from the problem dealt in [6] where only an interpolation in the sphere is required. To take into account these differences, we propose to modify Equation 1 by replacing the single RBF with the product of an spatial (ϕ) and an angular (ψ) RBF:

$$y((X, \Theta)) = \sum_{i=0}^{N-1} w_i \phi(|X - X_i|) \psi(|\Theta - \Theta_i|) \quad (2)$$

where $X = (x, y, z)$ are the spatial coordinates, and $\Theta = (\phi, \theta)$ the spherical coordinates of the sampling vector U_i . Differently from Equation 1, Equation 2 allows a dual interpolation in two different unrelated spaces. In Equation 2, $|X - X_i|$ represents the Euclidean distance, whereas $|\Theta - \Theta_i|$ is the geodesic distance over the unit sphere.

3 Materials and Experiments

3.1 Image Data

Fetal MRI was performed on a 1.5 T Siemens Avanto MRI Scanner (SIEMENS, Erlangen, Germany) at the Hautepierre Hospital (Strasbourg, France) using an 6-channel phased array coil combined to the spine array positioned around the mother abdomen. An axial spin echo single-shot echo-planar sequence was acquired along 30 non-collinear diffusion gradient encoding directions with a b value of $700s/mm^2$. The following pulse sequence parameters were used: TR=6800 ms; TE=99 ms; FOV=250 × 250 mm^2 ; matrix = 128 × 128; 41 contiguous axial slices of 3.5 mm thickness covering the whole fetal brain; no gap; number of excitations = 2. The total imaging time was 7mi 10s. Pregnant women were briefed before the exam and signed informed consent. To reduce motion artifacts, fetal sedation was obtained with 1 mg of flunitrazepam given orally to the mother 30 mi before the exam. The study was approved by the local ethics committee.

For validation purposes, three D-MRI sequences of the brain were acquired for an adult healthy subject in the following conditions: (i) static in supine position (\mathcal{S}^{ref} , the reference), and (ii) static with the head rotated with respect to the reference (\mathcal{S}^{rot}).

3.2 Slice to Volume Registration Accuracy

Initially we want to explore the ability to recover slice to volume alignment for typical but known motion, on typical anatomical structures. Fetal data are not suitable for assessing accuracy since motion artifacts are always present to some

degree. Therefore only the adult dataset \mathcal{S}^{ref} was used to this aim. Central slices of the volume were modified with a specific transformation and then registered *independently* to the whole volume. The displacements were chosen from a uniform distribution with a varying range of $[-8, +8]mm$ for translations in each direction, and between $[-10, +10]^\circ$ for each rotation. These ranges of variation represent movements much larger than those observed in real fetal data.

The accuracy was assessed by computing a registration error measured on a set of 4 points P_i within every slices as follows: $RMS = (\frac{1}{N} \sum_{i=1}^N TRE_i)^{\frac{1}{2}}$, where TRE is the target registration error defined as $TRE = \|P_i - \hat{T}^{-1}(T^*(P_i))\|^2$. T^* denotes the known applied motion transformation, and \hat{T} is the estimated geometric transformation. P_i are the corners of the intersection between the bounding box containing the brain, and each slice. The error previously defined provides thus an upper bound of the registration error for the region of interest.

3.3 Evaluation of RBF Interpolation

Leave-one-out test. A leave-one-out test by using the adult data \mathcal{S}^{ref} was performed for evaluating the capability of recovering non-acquired DW images from the available measurements. This test estimates the image \mathcal{S}_j^{ref} at a point (X_j, Θ_j) from $\mathcal{S}^{ref} \setminus \mathcal{S}_j^{ref}$ and $\mathcal{N}(X_j) \setminus X_j$, which in terms of RBF interpolation can be expressed as:

$$\hat{\mathcal{S}}^{ref}(X_j, \Theta_j) = \sum_{i=0, i \neq j}^{N-1} w_i \phi(|X_j - X_i|) \psi(|\Theta_j - \Theta_i|) \quad (3)$$

We have then computed the RMS error between $\hat{\mathcal{S}}^{ref}(X_j, \Theta_j)$ and $\mathcal{S}^{ref}(X_j, \Theta_j)$ for a set of random points distributed uniformly over the brain.

DW image estimation from a rotated sequence. In this experiment, DW images of \mathcal{S}^{ref} were estimated from \mathcal{S}^{rot} after performing a registration to put \mathcal{S}^{rot} into the frame of reference of \mathcal{S}^{ref} . Differently from the leave-one-out experiment, here all DW images of \mathcal{S}^{rot} at all spatial positions were employed for estimation.

3.4 Diffusion Descriptors

To evaluate the performance of the registration method, we have also considered two criteria for quality assessment of the resulting FA image.

Mean FA. The average value of FA over the cerebrospinal fluid (CSF), \overline{FA}_{cfs} is expected to be close to zero because of the isotropic diffusion properties of the CSF. Registration errors may induce an increase of this measure, since voxels belonging to the CSF in some DW images may be matched with voxels belonging to the gray or white matter in others.

Entropy. Nielsen *et al.* [7] have compared polynomial and affine distortion correction, and observed a reduction in erroneous regions of FA maps along with a more spiky FA distribution in favor of the polynomial registration. Netsch and van Muiswinkel [8] have reported an increase in the sharpness of FA maps after distortion correction. These observations correspond to lower H_{FA} values.

4 Results

The slice-to-volume test was applied to S_0^{ref} and S_i^{ref} . RMS values higher than 0.2 mm were considered as registration failures, and discarded for analysis. Under this criterion, successful registrations were obtained in 95% of the applied transformations. The obtained RMS errors were $0.162 \pm 0.004\text{ mm}$ for S_0^{ref} , and $0.105 \pm 0.011\text{ mm}$ for S_i^{ref} . Figure 2 (a) shows this error distribution.

The RMS error between the original sequence S^{ref} and its leave-one-out estimation \hat{S}^{ref} was lower than 0.001 for $s_\phi \in [2.0, 4.5]\text{ mm}$, and $s_\psi \in [0.6, 0.9]\text{ rad}$. The minimum values were obtained for $s_\phi = 3.5\text{ mm}$ (the inter-plane resolution), suggesting that the use of only *in-plane* neighbors is not sufficient for interpolation purposes, and adjacent slices must be included. Higher values of s_ϕ make the interpolated value dependent of remote regions which could not present the same diffusion properties, and the error starts increasing. No dependence of s_ψ was found in this experiment.

Figure 2 (b) shows the results when estimating S^{ref} from S^{rot} . In this case the error does depends on s_ψ , being lower for higher s_ψ values. After a given value (0.72 in our case) the error starts increasing again, because of the influence of distant gradient directions.

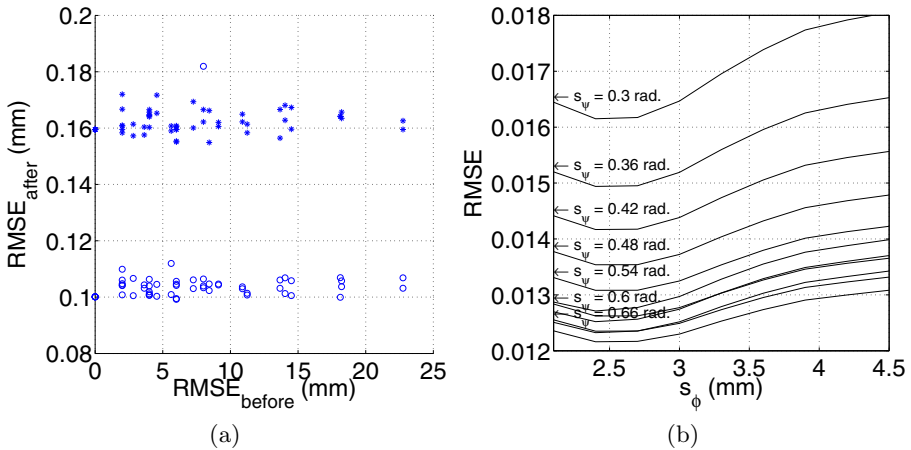


Fig. 2. (a) Slice to volume registration accuracy. This Figure shows the RMS error before and after registration for S_i (o) and T_2^{epi} (+). Each point corresponds to a specific applied transformation. (b) Estimation of S^{ref} from S^{rot} . The figure shows the normalized RMS error for different support regions of the RBF functions ϕ and ψ .

Table 1. Diffusion descriptors based on FA maps. H_{FA} = Entropy of FA, \overline{FA}_{csf} = mean FA in the CSF. For each column, the best value is shown in bold.

Method	Fetus #1		Fetus #2		Fetus #3		Fetus #4	
	H_{FA}	\overline{FA}_{csf}	H_{FA}	\overline{FA}_{csf}	H_{FA}	\overline{FA}_{csf}	H_{FA}	\overline{FA}_{csf}
Original	2.04	0.14	2.17	0.26	2.15	0.23	2.31	0.24
Manufacturer	2.11	0.12	2.09	0.21	2.20	0.20	2.33	0.19
Our approach	1.93	0.08	1.78	0.13	1.72	0.16	2.20	0.10

Table 1 compares the values of \overline{FA}_{cfs} and \overline{FA}_{cfs} for the original sequences, the values provided by the scanner’s manufacturer, and after applying the reconstruction method presented in this paper.

To perform the tractography, an expert radiologist traced regions containing the following bundles on T_2^{se} images: (i) corpus callosum, (ii) pyramidal tract, and (iii) Middle cerebellar peduncle. These regions were used for seeding the tractography after propagation to the T_2^{epi} image by using affine registration, and to check the presence of specific bundles. Tensor was estimated by using a standard least squares method, and the tractography was performed by applying a streamline method. In both cases we have used the algorithms implemented in Slicer¹. Figure 3 shows an example of the obtained results.

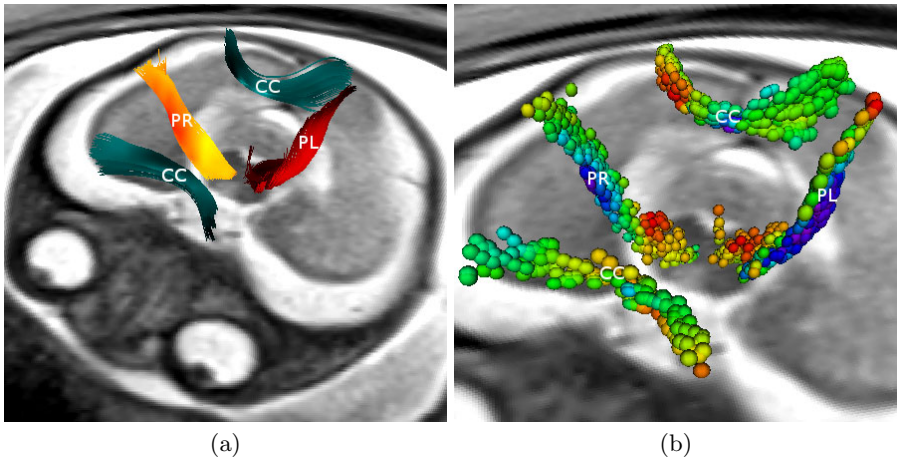


Fig. 3. Tractographies performed on a sequence corresponding to a fetus of 28 weeks of gestational age showing the corpus callosum (CC), the right pyramidal tract (PR), and the left pyramidal tract (PL). (a) Fibers. (b) Tensors.

¹ <http://www.slicer.org>

5 Discussion and Conclusions

In this paper, we have presented a method for reconstructing fetal D-MRI sequences from sparse data. A groupwise registration method based on slice-by-slice affine transformations was applied to compensate motion and eddy-current distortions, and a dual spatio-angular interpolation based on RBFs was used to estimate signal values on regular sampling grids. As the proposed method does not assume any diffusion model, the generated data can be used to study diffusion patterns even in IVOH. Experiments with adult data showed a high accuracy for the slice-to-volume registration, and for the estimation of DW images along unavailable gradient directions. In fetuses, the method improved the quality of the sequences as evidenced by the lower values of \overline{FA}_{cfs} and H_{FA} with respect to the original ones. The tractography provided different results for the original and reconstructed sequences, but they must be quantified and compared with objective criteria in order to assess their clinical significance.

Acknowledgements

The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013 Grant Agreement no. 207667). This work is also funded by NIH Grant R01 NS055064 and a CNRS grant for collaboration between LSIT and BICG.

References

1. Rohde, G.K., Barnett, A.S., Basser, P.J., Marenco, S., Pierpaoli, C.: Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI. *Magn. Reson. Med.* 51(1), 103–114 (2004)
2. Rousseau, F., Glenn, O.A., Iordanova, B., Rodriguez-Carranza, C., Vigneron, D.B., Barkovich, J.A., Studholme, C.: Registration-based approach for reconstruction of high-resolution in utero fetal MR brain images. *Acad. Radiol.* 13(9), 1072–1081 (2006)
3. Jiang, S., Xue, H., Counsell, S., Anjari, M., Allsop, J., Rutherford, M., Rueckert, D., Hajnal, J.V.: Diffusion tensor imaging (DTI) of the brain in moving subjects: application to in-utero fetal and ex-utero studies. *Magn. Reson. Med.* 62(3), 645–655 (2009)
4. Tuch, D.S., Reese, T.G., Wiegell, M.R., Makris, N., Belliveau, J.W., Wedeen, V.J.: High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magn. Reson. Med.* 48(4), 577–582 (2002)
5. Guimond, A., Meunier, J., Thirion, J.-P.: Average brain models: a convergence study. *Comput. Vis. Image Understand* 77, 192–210 (2000)
6. Carfora, M.F.: Interpolation on spherical geodesic grids: a comparative study. *J. Comput. Appl. Math.* 210(1-2), 99–105 (2007)
7. Nielsen, J.F., Ghugre, N.R., Panigrahy, A.: Affine and polynomial mutual information coregistration for artifact elimination in diffusion tensor imaging of newborns. *Magn. Reson. Imaging* 22(9), 1319–1323 (2004)
8. Netsch, T., van Muiswinkel, A.: Quantitative evaluation of image-based distortion correction in diffusion tensor imaging. *IEEE Trans. Med. Imaging* 23(7), 789–798 (2004)

Symmetric Positive-Definite Cartesian Tensor Orientation Distribution Functions (CT-ODF)

Yonas T. Weldeselassie¹, Angelos Barmpoutis², and M. Stella Atkins¹

¹ School of Computing Science, Simon Fraser University

² Dept. of Computer & Information Science and Engineering, University of Florida*

Abstract. A novel method for estimating a field of orientation distribution functions (ODF) from a given set of DW-MR images is presented. We model the ODF by Cartesian tensor basis using a parametrization that explicitly enforces the positive definite property to the computed ODF. The computed Cartesian tensors, dubbed Cartesian Tensor-ODF (CT-ODF), are symmetric positive definite tensors whose coefficients can be efficiently estimated by solving a linear system with non-negative constraints. Furthermore, we show how to use our method for converting higher-order diffusion tensors to CT-ODFs, which is an essential task since the maxima of higher-order tensors do not correspond to the underlying fiber orientations. We quantitatively evaluate our method using simulated DW-MR images as well as a real brain dataset from a post-mortem porcine brain. The results conclusively demonstrate the superiority of the proposed technique over several existing multi-fiber reconstruction methods.

1 Introduction

Diffusion tensor imaging (DT-MRI) is a non-invasive imaging technique that measures the self-diffusion of water molecules in the body, thus capturing the microstructure of the underlying tissues. Second order symmetric positive definite (SPD) tensors have commonly been used to model the diffusivity profile at each voxel with the assumption of a single coherent fiber tract per voxel. Under this assumption diffusivity was defined as $d(g) = g^T D g$ where g is the diffusion weighting magnetic gradient vector and D is the 2^{nd} order tensor to be estimated from a set of DW-MR images. This model, despite its simplicity and robustness, has been shown to be incorrect in regions containing intra-voxel orientational heterogeneity such as crossing and merging of fiber bundles [1,2].

Several methods have been proposed to overcome the single fiber orientation limitation of second order tensors. Tuch et al. [1] proposed the use of diffusion imaging

* Implementation code of our method is available through Matlab Central. This work was partially funded by the Canadian Natural Sciences and Engineering Research Council (NSERC). We thank Dr. Tim Dyrby, Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital, Hvidovre, Denmark, for providing the porcine dataset.

with diffusion weighting gradients applied along many directions distributed almost isotropically on the surface of a unit sphere; a method known as high angular resolution diffusion imaging (HARDI). In contrast to rank 2 tensors, this method does not assume any a priori knowledge about the diffusivity profile. A number of approaches have been proposed to compute the ensemble-average diffusion propagator $P(\mathbf{r}, t)$ of HARDI data. These methods include q-ball imaging (QBI) [3], diffusion spectrum imaging (DSI) [4], and diffusion orientation transform (DOT) [5]. These methods, collectively known as q-space imaging techniques, identify multiple fibers components by calculating the probability distribution function (PDF) of the diffusion process in each voxel, based on the Fourier transform relationship between the PDF of diffusion displacement and the diffusion weighted signal attenuation in q-space. DSI performs a discrete Fourier transform to obtain $P(\mathbf{r}, t)$, which requires a time intensive Cartesian sampling in q-space and hence is impractical for routine clinical use. QBI method takes measurements on a q-space ball and approximates the radial integral of the displacement probability distribution function by the spherical Funk-Radon transform. One problem with QBI is that the estimated diffusion orientation distribution function (ODF) is modulated by a zeroth-order Bessel function that induces spectral broadening of the diffusion peaks. DOT computes ODF at a fixed radius by expressing the Fourier transform in spherical coordinates and evaluating the radial part of the integral analytically assuming signals decay can be described by either a mono or a multi-exponential model, where the latter requires data acquisition over multiple concentric spheres, a time consuming proposition.

Another approach for multi-fiber reconstruction is to describe the apparent diffusion coefficient (ADC) by higher order diffusion tensors (e.g. 4^{th} and 6^{th}) that generalize the 2^{nd} order tensors and have the ability to approximate multi-lobed functions [6]. Several methods have been proposed for estimating 4^{th} order tensors with positive definite constraints [7,8,9] as well as for processing higher order tensor fields [10]. This approach is attractive not only because the rich set of processing and analysis algorithms developed for second order tensor fields can be extended for higher order tensors but also unlike spherical harmonics basis, the local maxima of higher order tensors can be easily computed [11,12] due to their simple polynomial form. Unfortunately, the use of higher order diffusion tensors has been confined to the estimation of tensor ADC profiles, although it is now known that the local maxima of ADC profiles estimated using higher order tensors generally can not be used to directly represent the orientations for the intravoxel crossing fibers [2,13,14].

In this paper, we propose the use of higher order SPD Cartesian tensors to model ODF profiles and present a novel method for estimating tensor field of ODF profiles from a given set of Diffusion-Weighted MR images. In our technique the ODF is modeled by Cartesian tensor basis using a parametrization that explicitly enforces the positive definite property to the computed distribution functions. The computed Cartesian tensor ODFs (CT-ODFs) are SPD tensors whose coefficients can be efficiently estimated by solving a linear system with non-negative constraints.

We quantitatively evaluate our method and demonstrate the superiority of the proposed technique over several existing multi-fiber reconstruction methods.

There are two main contributions in this paper: 1) We present a novel method for positive-definite CT-ODF estimation from DW-MRI. To the best of our knowledge there is no existing ODF model in literature that imposes explicitly the positivity to the estimated ODF, which is naturally a positive-valued spherical function. 2) We present a use full application of our method for converting higher-order diffusion tensor ADC profiles to CT-ODFs. We should emphasize that this is an essential task since the maxima of higher-order tensors do not correspond to the underlying fiber orientations. On the other hand, our method computes Cartesian Tensor ODFs whose maxima can be computed analytically and correspond to the true axonal orientations.

2 Method

2.1 Symmetric Positive-Definite Cartesian Tensors of Even Orders

Any spherical function $f(\mathbf{g})$ can be approximated by higher order Cartesian tensors:

$$f(\mathbf{g}) = \sum_{i=1}^3 \sum_{j=2}^3 \cdots \sum_{l=1}^3 g_i g_j \cdots g_l C_{i,j,\dots,l} \quad (1)$$

where g_i is the i -th component of the 3-dimensional unit vector \mathbf{g} , and $C_{i,j,\dots,l}$ are the coefficients of the l -th order tensor.

When approximating certain spherical functions in DW-MRI, we are interested in tensors of even orders with full symmetry, due to the antipodal symmetric nature of the DW-MR signal acquisition. In this case of symmetry, those tensor coefficients which correspond to the same monomial $g_1^a g_2^b g_3^c$ are equal to each other (e.g. $C_{2,2,2,1} = C_{2,2,1,2} = C_{2,1,2,2} = C_{1,2,2,2}$, since they all correspond to the monomial $g_1 g_2^3$).

Furthermore, if the approximated function $f(\mathbf{g})$ is a positive-valued function, the Cartesian tensor should be positive-definite, i.e. $f(\mathbf{g}) > 0 \forall \mathbf{g} \in S_2$. Therefore Eq. 1 needs to be re-parametrized such that this positivity property is adhered to. In this work, we use the higher-order positive-definite tensor parametrization that has been recently proposed in [9]. According to this parametrization, any non-negative spherical function can be written as a positive-definite L^{th} order homogeneous polynomial in 3 variables, which is expressed as a sum of squares of $(L/2)^{\text{th}}$ order homogeneous polynomials $p(g_1, g_2, g_3; \mathbf{c})$, where \mathbf{c} is a vector that contains the polynomial coefficients.

$$f(\mathbf{g}) = \sum_{j=1}^M \lambda_j p(g_1, g_2, g_3; \mathbf{c}_j)^2 \quad (2)$$

The parameters λ_j in Eq. 2 are non-negative weights. This parametrization approximates the space of L^{th} order SPD tensors and the approximation accuracy depends on how well the set of vectors \mathbf{c}_j sample the space of unit vectors \mathbf{c} . It has been shown that by constructing a large enough set of well sampled vectors \mathbf{c}_j , we can achieve any desired level of accuracy [9].

2.2 Positive-Definite Cartesian Tensor ODF (CT-ODF) Profiles

The Diffusion-Weighted MR signal for a given magnetic gradient orientation \mathbf{g} and gradient weighting b , can be modeled using the standard multi-fiber reconstruction framework as follows

$$S(\mathbf{g}, b) = \int_{S_2} w(\mathbf{v})B(\mathbf{v}, \mathbf{g}, b)d\mathbf{v} \quad (3)$$

where the integration is over all unit vectors \mathbf{v} , $B(\mathbf{v}, \mathbf{g}, b)$ is a basis function, and $w(\mathbf{v})$ is a non-negative spherical function that can be seen as a mixing/weighting function. There have been several proposed models for the basis function $B()$ such as a Rigaut-type function [15], von Mises-Fisher distribution [16] and others. In all of these models, the integral in Eq. 3 cannot be computed analytically, thus one needs to approximate the space of unit vectors \mathbf{v} by a discrete set of vectors $\mathbf{v}_1, \dots, \mathbf{v}_K$. In this case Eq. 3 is correctly discretized by $S(\mathbf{g}, b) = \sum_{k=1}^K w_k B(\mathbf{v}_k, \mathbf{g}, b)$ iff there are at most K underlying neural fibers that are oriented necessarily along the vectors \mathbf{v}_k . Another problem with the aforementioned discretization is that the function $w()$ is not anymore continuous over the sphere (it equals to w_k for \mathbf{v}_k and it is zero everywhere else).

The main idea in this paper is to avoid the above unnatural discretization of the space of orientations, by using a blending function $w()$, which can be appropriately decomposed so that: 1) it is positive-definite, and 2) is continuous over the sphere. In this work, we model such blending function as a L^{th} order SPD tensor (say 4^{th}) by plugging Eq. 2 into Eq. 3 as follows

$$S(\mathbf{g}, b) = \int_{S_2} \sum_{j=1}^M \lambda_j p(v_1, v_2, v_3; \mathbf{c}_j)^2 B(\mathbf{v}, \mathbf{g}, b)d\mathbf{v} \quad (4)$$

where v_1, v_2, v_3 are the three components of the unit vector \mathbf{v} .

Given a data set of DW-MRI signal attenuations S_i/S_0 associated with magnetic gradient orientations g_i and diffusion weighting b-value b , the coefficients of a L^{th} order positive-definite CT-ODF can be estimated by minimizing the following energy function with respect to the unknown polynomial-weighting coefficients λ_j

$$E = \sum_{i=1}^N \left(S_i/S_0 - \sum_{j=1}^M \lambda_j \int_{S_2} p(v_1, v_2, v_3; \mathbf{c}_j)^2 B(\mathbf{v}, \mathbf{g}_i, b)d\mathbf{v} \right)^2 \quad (5)$$

In order for the basis function $B()$ to reflect the signal attenuation of a single and highly oriented fiber response, we require the basis function to be a Gaussian that represents the diffusion process which is highly restricted perpendicular to the orientation \mathbf{v} . This is given by

$$B(\mathbf{v}, \mathbf{g}, b) = \lim_{\delta \rightarrow +\infty} e^{-\delta(\mathbf{v}^T \mathbf{g})^2} \quad (6)$$

Here we should emphasize that the model in Eq. 6 agrees with the properties of the DW-MR signal response, i.e. it takes maximum and minimum values for

diffusion sensitizing gradient orientations \mathbf{g} that are perpendicular and parallel to the underlying fiber orientation \mathbf{v} respectively. Moreover, δ is such that it captures information about b and mean diffusivity (D) and can be adjusted by altering either b or D . So this ‘symmetry’ can be simplified by using only δ in Eq. 6.

In order to compute the CT-ODF, we need to solve Eq. 5 for λ_j 's. This problem can be rewritten into an equivalent linear system problem $\mathbf{B}\mathbf{x} = \mathbf{y}$ where \mathbf{x} is an M -dimensional vector of the unknown λ_j , \mathbf{y} is an N -dimensional vector containing the given signal attenuations S/S_i and \mathbf{B} is a matrix of size $N \times M$ with the elements $\mathbf{B}_{i,j} = \int_{S_2} p(v_1, v_2, v_3; \mathbf{c}_j)^2 B(\mathbf{v}, \mathbf{g}_i, b) d\mathbf{v}$.

This linear system is solved for the non-negative \mathbf{x} using the efficient non-negative least squares algorithm and runs in 12ms/voxel. We can then easily compute the CT-ODF coefficients by multiplying the solution vector with a matrix \mathbf{C} , (i.e. $\mathbf{C}\mathbf{x}$), where the matrix \mathbf{C} is of size $\frac{(2+L)!}{2(L)!} \times M$ that contains monomials formed by the vectors \mathbf{c}_j . Note that L is the order of the CT-ODF and $\frac{(2+L)!}{2(L)!}$ is the number of the unique coefficients in an L^{th} -order Cartesian tensor. In the case of 4^{th} -order CT-ODFs, the multiplication $\mathbf{C}\mathbf{x}$ gives the 15 unique coefficients of a positive-definite tensor.

We applied our proposed method for estimating 4^{th} -order CT-ODFs ($L = 4$), using a set of $M = 321$ polynomial coefficients \mathbf{c}_j and $\delta = 200$. Regarding the parameter δ , we performed several experiments using different values $\delta > 100$ and we obtained similar fiber orientations density profiles, which shows that our method is not sensitive to the selection of the value of δ .

2.3 Computing CT-ODF from Higher-Order Diffusion Tensor

Now, we present an application of our proposed framework for computing the coefficients of a CT-ODF from a given higher-order diffusion tensor and diffusion weighting b -value b , which is an essential task since the maxima of higher-order tensors do not correspond to the underlying fiber orientations. Given a higher-order diffusion tensor, the coefficients of the corresponding CT-ODF are computed by using the technique we presented in the previous section as follows

$$\mathbf{C}\mathbf{B}^{-1}exp(-b\mathbf{G}\mathbf{t}) \tag{7}$$

where the matrices \mathbf{C} and \mathbf{B} are as defined in the previous section, \mathbf{G} is of size $N \times \frac{(2+L)!}{2(L)!}$ and contains only monomials constructed from N unit vectors \mathbf{g}_i uniformly distributed on the unit sphere, and \mathbf{t} is a vector of size $\frac{(2+L)!}{2(L)!}$ that contains the unique coefficients of the given higher-order diffusion tensor. For example, in the case of 4^{th} -order tensors, the 15 unique coefficients are given in the vector \mathbf{t} , and \mathbf{G} is of size $N \times 15$.

Note that in Eq. 7 the exponential function $exp()$ acts in an element-by-element fashion. Furthermore, the matrix inversion in Eq. 7 should be performed using non-negative least squares, as it has been shown in the previous section.

3 Experimental Results

In this section, we present experimental results of the proposed method applied to simulated as well as real DW-MRI data from a post-mortem porcine brain.

3.1 Synthetic Dataset

The proposed method was tested on a synthetic dataset by comparing the actual fiber orientations with the maxima of estimated CT-ODFs. In order to compare our results with spherical deconvolution techniques, included is also the results obtained using MOW [15], QBI [3], DOT [5] and MOVMF [16] methods by computing the maxima of either the PDF or ODF profiles of the corresponding methods.

The data was generated by simulating the MR signal from two crossing fibers whose orientations are $(\cos 20^\circ, \sin 20^\circ, 0)$ and $(\cos 100^\circ, \sin 100^\circ, 0)$ using the realistic diffusion MR simulation model in [17] with b -value = $1500\text{s}/\text{mm}^2$ and 81 gradient directions. Six distinct Rician noise levels were added to the simulated data and for each noise level the experiments were repeated 100 times.

Figure 1 shows a plot of the means and standard deviations of deviation angles between the actual fiber orientations and the maxima of estimated CT-ODFs. For the particular noise level with std. dev. = 0.08 the deviation angles for all the methods are reported in the adjacent table. Also notice that in this experiment the deviation angle of the computed orientations is compared to its closest actual fiber orientation because the crossing fibers are weighted equally in generating the MR signals. The results demonstrate the superiority of the proposed method over QBI, DOT, MOVMF and MOW methods.

3.2 Real Dataset

Here we present CT-ODFs computed from high-quality DWI on post-mortem pig brain, which resemble the human brain in neuroanatomical complexity and

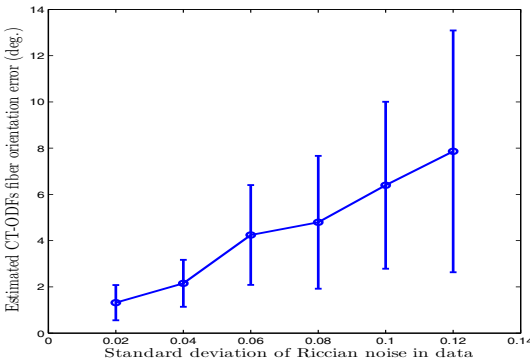


Table of errors (deg.)
noise st. dev. = 0.08

Method	Mean	St. dev.
QBI	9.125	± 4.545
DOT	6.645	± 3.720
MOVMF	5.624	± 3.514
MOW	5.010	± 2.955
CT-ODF	4.7926	± 2.8726

Fig. 1. Deviation angle between actual fiber orientations and maxima of estimated CT-ODFs using a simulated 2-fiber crossing data with orientations $(\cos 20^\circ, \sin 20^\circ, 0)$ and $(\cos 100^\circ, \sin 100^\circ, 0)$ at different levels of Rician noise

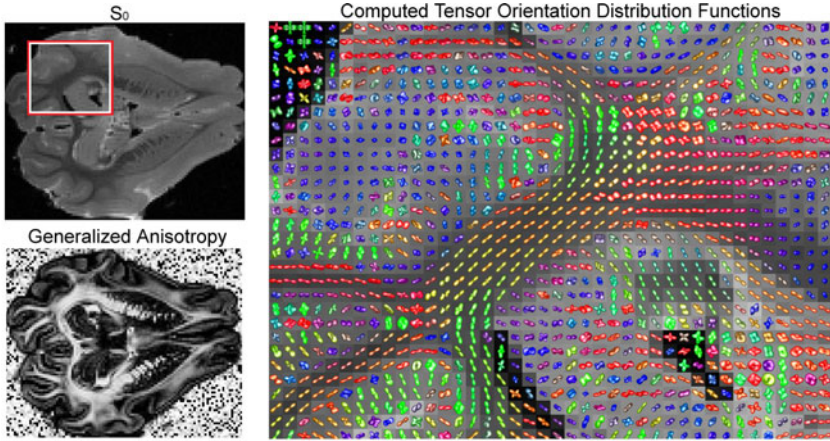


Fig. 2. CT-ODF field estimated by the proposed technique using data from a post-mortem porcine brain

where perfusion fixation was used to ensure that tissue characteristics were comparable to in vivo conditions [18]. Images are acquired using a pulsed gradient spin echo pulse sequence with echo time of $60ms$, 128×128 matrix with 10 slices, and voxel size of $0.5 \times 0.5 \times 0.5mm^3$. Three image were collected without diffusion weighting ($b \sim 0s/mm^2$) and 61 DWI with gradient strength $61mT/m$, gradient duration $23ms$, and gradient separation $30ms$. Each of these image sets used different diffusion gradients with approximate b values of $3146s/mm^2$. Fig. 2 shows CT-ODFs computed using our method along with generalized anisotropy and S_0 images. As can be verified in the generalized anisotropy image; the branching, bending and crossing of white matter tracts are correctly depicted by the computed CT-ODFs.

4 Conclusions

We presented a novel technique to estimate ODFs modeled as SPD high order tensors from DW-MR images. The performance of the proposed method is compared against several existing ODF measures on a synthetic dataset with different noise levels and outperformed the other methods. We also demonstrated the use of our method on a real DT-MR image obtained from a post-mortem porcine brain. Our results clearly demonstrate that crossing and merging of fibers are correctly depicted with CT-ODFs. Since higher order tensors have been used for segmentation, registration and computations of anisotropy measures for DT images, it remains to be seen if these tasks can be performed with the proposed CT-ODF fields as well.

References

1. Tuch, D., Weisskoff, R., Belliveau, J., Wedeen, V.J.: High angular resolution diffusion imaging of the human brain. In: ISMRM, p. 321 (1999)
2. Alexander, D., Barker, G., Arridge, S.: Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. *MRM* 48(2), 331–340 (2002)
3. Tuch, D.: Q-ball imaging. *MRM* 52(6), 1358–1372 (2004)
4. Wedeen, V., Hagmann, P., Tseng, W., Reese, T., Weisskoff, R.: Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *MRM* 54(6), 1377–1386 (2005)
5. Özarslan, E., Shepherd, T., Vemuri, B., Blackband, S., Mareci, T.: Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). *NeuroImage* 31(3), 1086–1103 (2006)
6. Ozarslan, E., Mareci, T.: Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *MRM* 50(5), 955–965 (2003)
7. Barmpoutis, A., Hwang, M., Howland, D., Forder, J., Vemuri, B.: Regularized positive-definite fourth order tensor field estimation from DW-MRI. *Neuroimage* 45(1S1), 153–162 (2009)
8. Ghosh, A., Deriche, R., Moakher, M.: Ternary quartic approach for positive 4th order diffusion tensors revisited. In: ISBI, pp. 618–621 (2009)
9. Barmpoutis, A., Vemuri, B.C.: A unified framework for estimating diffusion tensors of any order with symmetric positive-definite constraints. In: ISBI, pp. 1385–1388 (2010)
10. Yassine, I., McGraw, T.: 4th order diffusion tensor interpolation with divergence and curl constrained Bézier patches. In: ISBI, pp. 634–637 (2009)
11. Bloy, L., Verma, R.: On computing the underlying fiber directions from the diffusion orientation distribution function. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, p. 8. Springer, Heidelberg (2008)
12. Schultz, T., Seidel, H.: Estimating crossing fibers: A tensor decomposition approach. *TVCG* 14(6), 1635–1642 (2008)
13. Zhan, W., Stein, E., Yang, Y.: Mapping the orientation of intravoxel crossing fibers based on the phase information of diffusion circular spectrum. *NeuroImage* 23(4), 1358–1369 (2004)
14. Von dem Hagen, E., Henkelman, R.: Orientational diffusion reflects fiber structure within a voxel. *MRM* 48(3), 454–459 (2002)
15. Jian, B., Vemuri, B., Özarslan, E., Carney, P., Mareci, T.: A novel tensor distribution model for the diffusion-weighted MR signal. *NeuroImage* 37(1), 164–176 (2007)
16. Kumar, R., Barmpoutis, A., Vemuri, B., Carney, P., Mareci, T.: Multi-fiber reconstruction from DW-MRI using a continuous mixture of von Mises-Fisher distributions. In: CVPR Workshop, pp. 1–8 (2008)
17. Söderman, O., Jönsson, B.: Restricted diffusion in cylindrical geometry. *JMR, Series A* 117(1), 94–97 (1995)
18. Dyrby, T., Baaré, W., Alexander, D., Jelsing, J., Garde, E., Søgaard, L.: An ex vivo imaging pipeline for producing high-quality and high-resolution diffusion-weighted imaging datasets. In: HBM (2010)

Model-Free and Analytical EAP Reconstruction via Spherical Polar Fourier Diffusion MRI

Jian Cheng^{1,2}, Aurobrata Ghosh¹, Tianzi Jiang², and Rachid Deriche¹

¹ Athena Project Team, INRIA Sophia Antipolis – Méditerranée, France

² Center for Computational Medicine, LIAMA, Institute of Automation,
Chinese Academy of Sciences, China
Jian.Cheng@sophia.inria.fr

Abstract. How to estimate the diffusion Ensemble Average Propagator (EAP) from the DWI signals in \mathbf{q} -space is an open problem in diffusion MRI field. Many methods were proposed to estimate the Orientation Distribution Function (ODF) that is used to describe the fiber direction. However, ODF is just one of the features of the EAP. Compared with ODF, EAP has the full information about the diffusion process which reflects the complex tissue micro-structure. Diffusion Orientation Transform (DOT) and Diffusion Spectrum Imaging (DSI) are two important methods to estimate the EAP from the signal. However, DOT is based on mono-exponential assumption and DSI needs a lot of samplings and very large b values. In this paper, we propose Spherical Polar Fourier Imaging (SPFI), a novel model-free fast robust analytical EAP reconstruction method, which almost does not need any assumption of data and does not need too many samplings. SPFI naturally combines the DWI signals with different b -values. It is an analytical linear transformation from the \mathbf{q} -space signal to the EAP profile represented by Spherical Harmonics (SH). We validated the proposed methods in synthetic data, phantom data and real data. It works well in all experiments, especially for the data with low SNR, low anisotropy, and non-exponential decay.

1 Introduction

Diffusion MRI is a non-invasive technique to explore the complex white matter by probing the diffusion process of water molecules. EAP has the full information about the diffusion process. Estimating the EAP is at the heart of dMRI. When the narrow pulse condition is met, EAP is related with the signal attenuation by a Fourier transform.

$$P(\mathbf{R}) = \int E(\mathbf{q})e^{-2\pi i\mathbf{q}\cdot\mathbf{R}}d\mathbf{q}, \quad (1)$$

where \mathbf{R} is the displacement vector in \mathbf{R} -space, and \mathbf{q} is the reciprocal vector in \mathbf{q} -space. There are many articles about Orientation Distribution Function (ODF) in High Angular Resolution Diffusion Imaging (HARDI) [1][2]. But ODF is just one of the features of EAP and it has no radial information.

Historically, DTI was proposed by assuming $P(\mathbf{R})$ as a Gaussian distribution [3]. It is actually a model-based method to estimate $P(\mathbf{R})$, which cannot describe the non-Gaussian diffusion. Diffusion Spectrum Imaging (DSI) [4] is a well known model-free

method to estimate EAP. The main shortcoming of DSI is that it uses a numerical Fourier Transform and needs very large range of b values and a lot of samplings.

Diffusion Orientation Transform (DOT) [6] is a fast analytical method based on mono-exponential decay assumption on $E(\mathbf{q})$. It was proposed to relax the Gaussian assumption in DTI to the assumption of mono-exponential decay. This assumption lets us have the full information about $E(\mathbf{q})$ in the whole 3D \mathbf{q} -space from the $E(q_0\mathbf{u})$ only in a single shell. Then EAP profile at a given radius R_0 could be calculated analytically. However the estimated EAP is the true PDF convolved by the Fourier transform of the function $E(q, \mathbf{u})^{q^2/q_0^2} E(q, \mathbf{u})^{-1}$ [6], where $\mathbf{q} = q\mathbf{u}$, $q = \|\mathbf{q}\|$. It was shown surprisingly that the estimated $\tilde{P}(R\mathbf{r})$ in some synthetic experiments is sharper than the real $P(R\mathbf{r})$. But since this effect comes from the intrinsic modeling error, it is still not clear whether DOT can work well in the complex real data with non-exponential decay, low anisotropy and low SNR. The author in [6] has extended mono-exponential model to multi-exponential model so that it can reduce the modeling error and work for the data from multiple shells. However, it is impractical because a nonlinear fitting is needed for every direction [6], suffering from limited samples, local minima, computational complexity, and an analytic solution exists only when three b values satisfy an arithmetic process. Diffusion Propagator Imaging (DPI) [7] is another analytical estimation which assumes the $E(\mathbf{q})$ can be represented as the form of the solution of 3D Laplacian equation. It seems to work well just with small number of samplings. However, that assumption is unrealistic for $E(\mathbf{q})$, since $E(0)$ does not exist based on that assumption.

$$E(\mathbf{q}) = \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l a_{n,l,m} R_n(\|\mathbf{q}\|) Y_l^m(\mathbf{u}) \quad B_{n,l,m}(\mathbf{q}) = R_n(\|\mathbf{q}\|) Y_l^m(\mathbf{u}) \quad (2)$$

$$R_n(\|\mathbf{q}\|) = \kappa_n(\zeta) \exp\left(-\frac{\|\mathbf{q}\|^2}{2\zeta}\right) L_n^{1/2}\left(\frac{\|\mathbf{q}\|^2}{\zeta}\right) \quad \kappa_n(\zeta) = \left[\frac{2}{\zeta^{3/2}} \frac{n!}{\Gamma(n+3/2)}\right]^{1/2} \quad (3)$$

Compared with Laplacian equation modeling in [7], Spherical Polar Fourier Expression (SPFE) of $E(\mathbf{q})$ seems better. It was proposed to sparsely represent $E(\mathbf{q})$ [8]. See formulae (2), (3), where $Y_l^m(\mathbf{u})$ is the l order m degree Spherical Harmonic (SH) basis and $R_n(q)$ is the Gaussian-Laguerre polynomial basis. Since $B_{n,l,m}(\mathbf{q})$ is the orthonormal basis in R^3 , any $E(\mathbf{q})$ could be represented by a linear combination of $\{B_{n,l,m}\}$. While Laplacian equation modeling can not. In [8], the authors proposed two methods to estimate the coefficients $\{a_{n,l,m}\}$, a least square fit and a nonlinear robust estimation which considers the Rician noise. After $\{a_{n,l,m}\}$ are estimated, a inner product of $a_{l,n,m}$ and a kernel $b_{n,l,m}$ was used to calculate some features of $P(\mathbf{R})$, e.g. the ODF in [1], EAP profile [8]. The problem of [8] is that the kernel $b_{n,l,m}$ needs to be calculated numerically from FFT for every direction or calculated for one direction then rotated by Wigner rotation matrix for other directions. For EAP profile it can bring some numerical error since the kernel has some delta functions inside.

In this paper, based on the SPFE in [8], we propose Spherical Polar Fourier Imaging (SPFI), a novel technique for model-free analytical reconstruction of the EAP profile from the signals in different Q-shells. It is a linear transformation from the coefficients $\{a_{n,l,m}\}$ of the signal $E(\mathbf{q})$ to the coefficients $\{c_{l,m}\}$ of EAP profile $P(R_0)$ represented by SH for a given R_0 . First we deduce the transform for EAP profile. Then, we perform the

method in some non-exponential synthetic data, a challenging phantom data and a real monkey data with several b values.

2 Analytical EAP Profile Estimation Based on SPF

Our method is close in spirit to the methods in DOT [6] and DPI [7]. Adding a strong assumption (in DOT) or choosing a good representation of $E(\mathbf{q})$ (in DPI and SPFI) will dramatically simplify the Fourier transform in (1).

2.1 Estimation of EAP Profile

SPFE is a kind of orthonormal basis representation and it was shown in [8] that it can sparsely represent the diffusion signal. See formula (2). After we estimate the coefficients of the signal via the methods in [8], we proved that there is a linear, analytical solution to get the coefficients of the EAP profile at a given radius R_0 under SH representation. Firstly consider the plane wave equation in spherical coordinates in (4)

$$e^{\pm 2\pi i \mathbf{q} \cdot \mathbf{R}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l (\pm i)^l j_l(2\pi q R) Y_l^m(\mathbf{u}) Y_l^m(\mathbf{r}), \quad (4)$$

where $j_l(x)$ is the l -th order spherical Bessel function. $j_l(x) = \sqrt{\frac{\pi}{2x}} J_{l+0.5}(x)$, $J_{l+0.5}(x)$ is the Bessel function of the first kind. Then put the formula (2) into formula (1).

$$\begin{aligned} P(R_0 \mathbf{r}) &= 4\pi \int \left\{ \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l a_{n,l,m} R_n(q) Y_l^m(\mathbf{u}) \right\} \left\{ \sum_{l'=0}^{\infty} \sum_{m'=-l'}^{l'} (-i)^{l'} j_{l'}(2\pi q R_0) Y_{l'}^{m'}(\mathbf{u}) Y_{l'}^{m'}(\mathbf{r}) \right\} d\mathbf{q} \\ &= 4\pi \sum_{l=0}^L \sum_{m=-l}^l (-1)^{l/2} \sum_{n=0}^N \underbrace{\left\{ \int_0^{\infty} j_l(2\pi q R_0) R_n(q) q^2 dq \right\}}_{I_{l,n}(R_0)} a_{n,l,m} Y_l^m(\mathbf{r}) \end{aligned} \quad (5)$$

where we use the orthonormal property of SHs, i.e. $\int_{S^2} Y_l^m(\mathbf{u}) Y_{l'}^{m'}(\mathbf{u}) d\mathbf{u} = \delta_{ll'} \delta_{mm'}$, and define $I_{l,n}(R_0) = \int_0^{\infty} j_l(2\pi q R_0) R_n(q) q^2 dq$. It should be noted that in formula (2), if we fix the SH as the spherical basis and use another radial basis $R'_n(q)$, the equation (5) also holds. We choose Gaussian-Laguerre radial basis because it could sparsely represent $E(\mathbf{q})$ [8]. Considering in (3) $L_n^{1/2}(x) = \sum_{i=0}^n l_n^i x^i$, $l_n^i = (-1)^i \binom{n+0.5}{n-i} \frac{1}{i!}$, we have

$$I_{l,n}(R_0) = \frac{\kappa_n(\zeta) \zeta^{1.25}}{2 \sqrt{R_0}} \sum_{i=0}^n l_n^i \int_0^{\infty} x^{2i+1.5} J_{l+0.5}(2\pi R_0 \sqrt{\zeta} x) \exp(-0.5x^2) dx \quad (6)$$

Based on the property of Bessel function [9], we have

$$\int_0^{\infty} x^\mu \exp(-\alpha x^2) J_\nu(\beta x) dx = \frac{\beta^\nu \Gamma(0.5\nu + 0.5\mu + 0.5)}{2^{\nu+1} \alpha^{0.5(\mu+\nu+1)} \Gamma(\nu+1)} {}_1F_1\left(\frac{\mu+\nu+1}{2}; \nu+1; -\frac{\beta^2}{4\alpha}\right) \quad (7)$$

where ${}_1F_1$ is the confluent hypergeometric function of the first kind. ${}_1F_1(a; b; x) = \sum_{k=0}^{\infty} \frac{(a)_k x^k}{(b)_k k!}$, $(a)_k = (a(a+1)\dots(a+k-1))$, with $(a)_0 = 1$. Here in (6) $\alpha = 0.5$, $\beta = 2\pi R_0 \sqrt{\zeta}$, $\nu = l + 0.5$, $\mu = 2i + 1.5$, then we get

$$I_{l,m}(R_0) = \frac{\kappa_n(\zeta) \zeta^{0.5l+1.5} \pi^{l+0.5} R_0^l}{\Gamma(l+1.5)} \sum_{i=0}^n l_n^i 2^{0.5l+i-0.5} \Gamma(0.5l+i+1.5) 1F1\left(\frac{2i+l+3}{2}; l+\frac{3}{2}; -2\pi^2 R_0^2 \zeta\right)$$

Put it into (5), we have

$$P(R_0 \mathbf{r}) = \sum_{l=0}^L \sum_{m=-l}^l \left\{ 4(-1)^{l/2} \frac{\zeta^{0.5l+1.5} \pi^{l+1.5} R_0^l}{\Gamma(l+1.5)} \sum_{n=0}^N f_{n,l,m}(\zeta, R_0) a_{n,l,m} \right\} Y_l^m(\mathbf{u}) \quad (8)$$

$$c_{l,m} = 4(-1)^{l/2} \frac{\zeta^{0.5l+1.5} \pi^{l+1.5} R_0^l}{\Gamma(l+1.5)} \sum_{n=0}^N f_{n,l,m}(\zeta, R_0) a_{n,l,m} \quad (9)$$

$$f_{n,l,m}(\zeta, R_0) = \kappa_n(\zeta) \sum_{i=0}^n (-1)^i \binom{n+0.5}{n-i} \frac{1}{i!} 2^{0.5l+i-0.5} \Gamma(0.5l+i+1.5) 1F1\left(\frac{2i+l+3}{2}; l+\frac{3}{2}; -2\pi^2 R_0^2 \zeta\right) \quad (10)$$

Now we get the linear transform from $\{a_{n,l,m}\}$ to $\{c_{l,m}\}$, which could be implemented as an matrix multiplication. Moreover please note the important difference between SPFI and DOT. Here our transformation is independent with the data, since $f_{n,l,m}(\zeta, R_0)$ is just dependent on ζ and R_0 . Once we give a R_0 and the basis, we have the transform. While in DOT, the transform is dependent on the ADC, which could brings some numerical errors. See appendix A in [6] for error analysis. Similarly with the appendix in [6], here the confluent hypergeometric function $1F1$ could also be calculated from an analytical solution. Actually in practice a numerical truncated approximation of $1F1$ is also acceptable, since in SPFI we just need the values of $1F1$ only at the fixed value $-2\pi^2 R_0^2 \zeta$ and the transform matrix just needs to be calculated only once. Another important similarity with DOT is that if we just choose $N = 0$ in radial part, our transform will be the DOT, which could be seen from the formulae (9), (10). That is true since the order 0 of the radial basis follows mono-exponential decay. However, in SPFI we should use $N \geq 1$ to describe anisotropic decay, since the order 0 in (2) is just an isotropic part.

2.2 Zero Displacement Probability

The $P_0 = P(0)$ is the probability of water molecules that minimally diffuse within the diffusion time Δ [5]. The P_0 map could be used in tissue segmentation and some other applications [5]. In SFPI, we can easily estimate P_0 from (2), or by setting $R_0 = 0$ in (6), (5). These two ways are equivalent. Since $\int_0^\infty \exp(-st) t^\alpha L_n^\alpha(t) dt = \frac{\Gamma(\alpha+n+1)(s-1)^n}{n! s^{\alpha+n+1}}$ [9] and $\int_{S^2} Y_l^m(\mathbf{u}) d\mathbf{u} = \sqrt{4\pi} \delta_{l0} \delta_{m0}$ we have

$$P_0 = \iiint E(\mathbf{q}) d\mathbf{q} = \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l a_{n,l,m} \left\{ \int_0^\infty R_n(q) q^2 \right\} \left\{ \int_{S^2} Y_l^m(\mathbf{u}) d\mathbf{u} \right\} \quad (11)$$

$$= \sqrt{8\pi} \sum_{n=0}^N (-1)^n \kappa_n(\zeta) \zeta^{1.5} \frac{\Gamma(n+1.5)}{n!} a_{n,0,0} \quad (12)$$

$$= \boxed{4 \sqrt{\pi} \zeta^{\frac{3}{2}} \sum_{n=0}^N (-1)^n \sqrt{\frac{\Gamma(n+1.5)}{n!}} a_{n,0,0}} \quad (13)$$

2.3 Implementation of Methods

The Implementation includes two steps. The first step is to estimate coefficients $\{a_{n,l,m}\}$ from the signals $\{E(\mathbf{q}_i)\}$. The second step is the linear analytical transform proposed above from $\{a_{n,l,m}\}$ to $\{c_{l,m}\}$, which is actually independent of the first step. The whole estimation error is just from the first step, since the second step is analytical and compact. [8] suggested two methods to estimate $\{a_{n,l,m}\}$, a linear least square (LS) fitting with regularization in the radial and spherical parts, and a non-linear PDE based optimization process, which considers the Rician noise. Here we choose the LS method, known to be faster, in the first step. We suggest that the Rician correction could be performed directly on the DWI data as a pre-processing step [10,11], although in our experiments to perform an appropriate comparison of methods we did not perform any Rician correction.

For LS estimation, denote signal vector by $E = [E(\mathbf{q}_i)]_{S \times 1}$, the basis matrix by $M = [R_n(\mathbf{q})Y_l^m(\mathbf{u})]_{S \times 0.5(L+1)(L+2)(N+1)}$, and the spherical and radial regularization diagonal matrices respectively by $L = [l(l+1)]$ and $N = [n(n+1)]$. Then the coefficient vector $A = (M^T M + \lambda_l L^T L + \lambda_n N^T N)M^T E$, where λ_l and λ_n are the regularization terms for spherical and radial parts. For the second step, the linear transformation in (9), (10) could be also implemented as a matrix multiplication, i.e. $C = FA$. Thus we can combine this two steps or perform separately. To combine two steps, $C = \{F(M^T M + \lambda_l L^T L + \lambda_n N^T N)M^T\}E$. The matrix of the whole process, $F(M^T M + \lambda_l L^T L + \lambda_n N^T N)M^T$, is independent with E and needs to be calculated only once for the whole data set. It makes our method extremely fast. Another option is to separate these two steps. and store $a_{n,l,m}$ once it is estimated in the first step. Then estimating the coefficients for EAP profile at a given R_0 could be performed directly on the stored $\{a_{n,l,m}\}$. That means we just need to calculate $\{a_{n,l,m}\}$ once and could re-calculate different EAP profiles in different radii very fast. The main computation complex is in the estimation of $a_{n,l,m}$. But it is still very fast if least square fitting is used.

There are two important points to consider in the implementation. The first one is about $E(0)$. If a data set has several b values, b_1, b_2, \dots, b_N , we actually use $N+1$ b values, considering $b_0 = 0$ and $E(0) = 1$ for any $\mathbf{u} \in S^2$, which makes our estimation more reasonable and accurate. Otherwise, there is no warranty for the estimated signal $\tilde{E}(0) = 1$. Another advantage is that for the single shell HARDI data, considering $b = 0$ can let us have 2 shells, which will largely improve the results. The second one is how to determine the parameter ζ in basis. The authors in [8] proposed an experience strategy for it, which is dependent on the radial truncation order N . However, we think the parameter should be just dependent on the signal, not on the basis order. Considering $E(\mathbf{q}) = \exp(-4\pi^2\tau q^2 D)$, $b = 4\pi^2\tau q^2$, and a typical diffusion coefficient of $D = 0.7 \times 10^{-3} \text{mm}^2/\text{s}$, a typical b -value $b = 3000 \text{s}/\text{mm}^2$, we set $\zeta = \frac{1}{8\pi^2\tau \times 0.7 \times 10^{-3}}$. If $4\pi^2\tau = 1$, then ζ is about 700. In our experiments we always set $\zeta = 700$.

3 Results on Synthetic, Phantom and Real Data

Synthetic data. Gaussian mixture model $S(\mathbf{q}_s) = \sum_{i=1}^M p_i G_i(\mathbf{q}_s)$, $G_i(\mathbf{q}_s) = \exp(-q_s^2 \mathbf{u}_s^T \mathbf{D}_i \mathbf{u}_s)$ has been used widely to generate synthetic data [2,13]. However, it could bias the results in favor of those methods assuming a model based on Gaussian

Table 1. For each configuration in each column, the left part and the right part show, respectively for Gaussian mixture model and non-Gaussian mixture model, the percentage of correct number of detected maximum of the estimated EAP profile and the mean of angular error. The first four rows recorded the performance of DOT on single shell data [6] with 81 gradient directions on the hemisphere. The last row is the results of our methods using 4 shells.

b value	(1, [1.1,0.5,0.5]e-3, 10, —)	(2, [1.3,0.4,0.4]e-3, 10, 90)	(2, [1.7,0.3,0.3]e-3, 35, 60)	(2,[1.7,0.3,0.3]e-3, 20, 65)
500	(28.0%,13.5°) (55.7%,12.6°)	(35.0%,27.3°) (49.9%,25.4°)	(0.4%,26.8°) (0.3%,26.8°)	(3.4%,28.5°) (3.6%,29.4°)
1000	(82.1%,8.9°) (72.5%,10.7°)	(61.4%,18.9°) (57.4%,20.6°)	(0.4%,26.8°) (1.8%,26.8°)	(9.2%,28.8°) (6.9%,28.2°)
2000	(80.1%,8.3°) (49.0%,11.1°)	(67.5%,13.0°) (56.7%,18.4°)	(20.7%,26.14°) (15.8%,25.7°)	(46.7%,9.4°) (36.2%,11.2°)
3000	(23.1%,10.9°) (9.5%,13.4°)	(44.6%,16.1°) (32.7%,21.5°)	(58.9%,7.8°) (47.3%,7.8°)	(87.5%,5.0°) (76.1%,6.4°)
4 shells	(99.3%,6.7°) (89%,8.9°)	(96.1%,9.1°) (83.5%,12.3°)	(81.8%,4.8°) (62.1%,6.5°)	(95.2%,4.0°) (82.8%,5.5°)

mixture or mono/multi-exponential decay. Here we choose both Gaussian mixture and non-Gaussian mixture to validate our methods. We set $S(\mathbf{q}) = \sum_{i=1}^M p_i f_i(\mathbf{q})$, $f_i(\mathbf{q}) = G(\mathbf{q})$ for a Gaussian mixture model and $f(\mathbf{q}) = 0.5G(\mathbf{q}) + 0.5T(\mathbf{q})$, $T(\mathbf{q}) = \exp(-\sqrt{2q^2 \mathbf{u}^T \mathbf{D} \mathbf{u}})$ for a non-Gaussian mixture model. It could be proved that the ODF of $T(\mathbf{q})$ are the same as the ODFs of $G(\mathbf{q})$, although they have different EAPs [9]. The EAP of $T(\mathbf{q})$ is $P(R\mathbf{r}) = \frac{16\pi}{\sqrt{|\mathbf{D}|(4+4\pi^2 R^2 \mathbf{r}^T \mathbf{D}^{-1} \mathbf{r})^2}}$. Thus we have the analytical ground truth of EAP. We use the same way in [2] to add Rician noise with $SNR = 1/\sigma$. SNR is defined as the ratio of maximal signal intensity of $S(0) = 1$ to the standard deviation σ of complex Gaussian noise. We reconstructed EAP profile $P(R_0\mathbf{r})$ at $R_0 = 15\mu\text{m}$ from our method and DOT in different configurations of signal generators with different fiber numbers (1 or 2), eigenvalues of \mathbf{D} , SNR, angle between 2 fibers. See Table 1, where (2, [1.7,0.3,0.3]e-3, 35, 60) means two fibers ($M = 2$), eigenvalues are [1.7,0.3,0.3]e-3, angle is 60° . and so on. For each configuration, data in 4 shells ($b=500,1000,2000,3000s/mm^2$) were generated for 1000 trials. For DOT, 4 order SH with $\lambda = 0.006$ was chosen for single shell data. For SPFI, we use all data in 4 shells and chose $N = 1$ for $SNR = 10$, and $N = 2$ for others and $L = 4$, $\lambda_l = 1e - 8$, $\lambda_n = 1e - 8$, $\zeta = 700$ for all experiments. We recorded the percentage of correct number of detected maximum of estimated EAP profile and the mean of angular error. The experiments showed that SPFI works better in the configurations with low anisotropy, much noise and non-exponential decay. Please note that we did not compare our method with DOT in multi-exponential model [6], because it is impractical as discussed in the introduction part.

Phantom data. We applied SPFI to a public phantom data with 3 shells with b-values of 650,1500,2000 s/mm^2 respectively. This data has been used in the fiber cup contest in MICCAI 2009 to evaluate tracking methods [12]. The anisotropy of this data is very low, which makes it hard to detect the fibers. We believe that it is complex enough to evaluate different reconstruction methods and tracking methods. We compare our reconstruction method using 3 shells with DOT using one shell ($b=2000$) using Laplacian regularization term $\lambda = 0.006$ [2]. For SPFI, we choose $L = 4$, $\lambda_l = 5e - 8$ in the spherical part and $N = 1, \lambda_n = 1e - 9$ in the radial part [8]. Two crossing areas with EAP profiles in $R_0 = 15\mu\text{m}, 17\mu\text{m}$ were chosen for visualization using min-max normalization [1]. To perform a fair comparison, we also tune the Laplacian regularization term lambda to 0.002 and 0.02 for region B. The results were shown in Fig. 1. It shows that SPFI could work well in the data with low anisotropy and non-exponential decay, which agrees with the results for synthetic data. that the method using 3 shells is

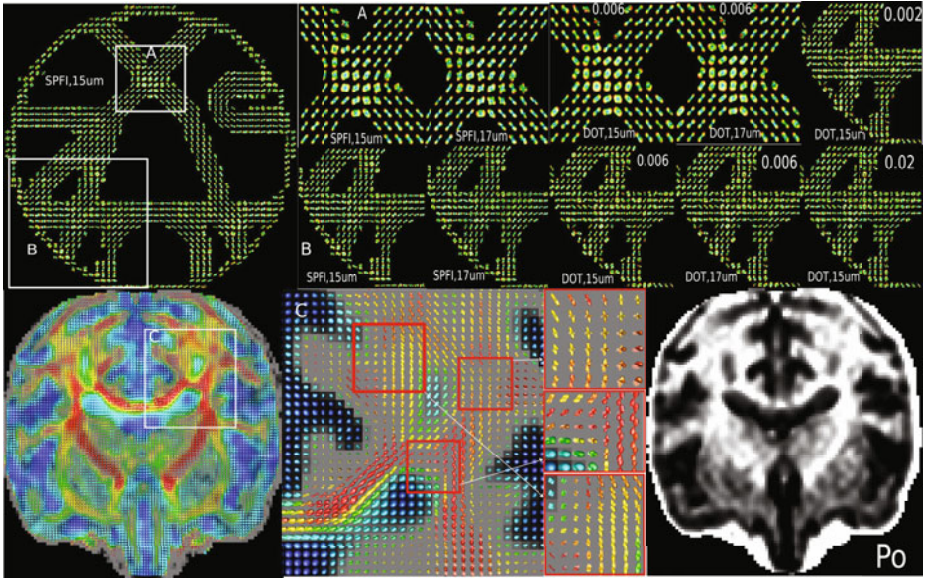


Fig. 1. First row: phantom data, from left to right: whole view of $P(R_0\mathbf{r})$, $R_0 = 15\mu\text{m}$ and $P(R_0\mathbf{r})$, $R_0 = 15\mu\text{m}, 17\mu\text{m}$ in region A and B of phantom data, calculated from SPFI and DOT. Region B was shown for DOT with $\lambda = 0.002, 0.006, 0.02$. Second row: real data result from SPFI, from left to right: whole view of $P(R_0\mathbf{r})$, $R_0 = 15\mu\text{m}$ and EAP in region C and P_o of the real data

better. The bad performance of DOT is probably because of the modeling error of the mono-exponential assumption.

Real data. We tested our method using real monkey data with 3 shells ($b=500, 1500, 3000$), 30 directions at each b value, TE/TR/matrix= $120\text{ms}/6000\text{ms}/128 \times 128$. We set $L = 4, N = 2, \lambda_l = 5e - 9, \lambda_n = 1e - 9$ and show the results of $P(R_0\mathbf{r})$ at $R_0 = 15\mu\text{m}$. The glyphs were colored by GFA calculated from EAP profile [1]. Please note that we did not do any normalization here, e.g. min-max normalization [1]. That is because of two reasons. 1) the EAP profiles in white matter seem sharp enough and the profiles in CSF and gray matter are almost isotropic. 2) we will lost the radial information if we do some normalization. Please note that the radial information in EAP is also important compared with its peaks. It might be used to infer the axonal diameter and it is sensitive to white-matter anomalies [14]. From the results, we can see that the CSF has the largest probability (glyph size) compared with white matter and gray matter, just like the visualization of tensors in DTI. Tensors cannot illustrate and recover crossing fibers, while EAP profiles can. We also show in Fig. 1 the P_o calculated from SPFI.

4 Conclusion

We proposed Spherical Polar Fourier Imaging (SPFI), a novel model-free fast robust analytical EAP profile reconstruction method based on Spherical Polar Fourier expression (SPFE) of the signal in \mathbf{q} -space. It provides a linear analytical closed form to estimate

the EAP profile from the signal under SPFE. It is a linear transformation that is independent of data. This transformation matrix is just calculated only once for a whole data set, which makes SPFI extremely fast. SPFI can avoid the error from unrealistic assumptions and can naturally combine the signals with different b-values. The experimental results from synthetic data, phantom data and real data show that SPFI can perform better than DOT, especially for the data with low anisotropy, low SNR and non-exponential decay.

Acknowledgment. This work was partly supported by the Natural Science Foundation of China (30730035), the National Key Basic Research and Development Program of China (2007CB512305), the National High Technology Research and Development Program of China (2009AA02Z302), the External Cooperation Program of the Chinese Academy of Sciences (GJHZ200826), the French ANR “Neurological and Psychiatric diseases” NucleiPark and the France-Parkinson Association.

References

1. Tuch, D.S.: Q-ball imaging. *Magnetic Resonance in Medicine* 52, 1358–1372 (2004)
2. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Regularized, fast and robust analytical q-ball imaging. *Magnetic Resonance in Medicine* 58, 497–510 (2007)
3. Basser, P.J., Mattiello, J., LeBihan, D.: MR diffusion tensor spectroscopy and imaging. *Biophysical Journal* 66, 259–267 (1994)
4. Wedeen, V.J., Hagmann, P., Tseng, W.Y.I., Reese, T.G., Weisskoff, R.M.: Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance In Medicine* 54, 1377–1386 (2005)
5. Wu, Y.C., Alexander, A.L.: Hybrid diffusion imaging. *NeuroImage* 36, 617–629 (2007)
6. 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). *NeuroImage* 31, 1086–1103 (2006)
7. Descoteaux, M., Deriche, R., Bihan, D.L., Mangin, J.F., Poupon, C.: Diffusion propagator imaging: Using Laplace’s equation and multiple shell acquisitions to reconstruct the diffusion propagator. In: Prince, J.L., Pham, D.L., Myers, K.J. (eds.) *Information Processing in Medical Imaging*. LNCS, vol. 5636, pp. 1–13. Springer, Heidelberg (2009)
8. Assemlal, H.E., Tschumperlé, D., Brun, L.: Efficient and robust computation of PDF features from diffusion MR signal. *Medical Image Analysis* 13, 715–729 (2009)
9. Gradshteyn, I., Ryzhik, I.: *Table of Integrals, Series, and Products*. Elsevier, Amsterdam (2007)
10. Tristán-Vega, A., Aja-Fernández, S.: DWI filtering using joint information for DTI and HARDI. *Medical Image Analysis* 14, 205–218 (2010)
11. Descoteaux, M., Wiest-Daessle, N., Prima, S., Barillot, C., Deriche, R.: Impact of Rician adapted non-local means filtering on HARDI. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part II*. LNCS, vol. 5242, pp. 122–130. Springer, Heidelberg (2008)
12. Poupon, C., Rieul, B., Kezele, I., Perrin, M., Poupon, F., Mangin, J.: New diffusion phantoms dedicated to the study and validation of high-angular-resolution diffusion imaging (HARDI) models. *Magnetic Resonance In Medicine* 60(6), 1276–1283 (2008)
13. Aganj, I., Lenglet, C., Sapiro, G., Yacoub, E., Ugurbil, K., Harel, N.: Reconstruction of the orientation distribution function in single and multiple shell q-ball imaging within constant solid angle. *Magnetic Resonance in Medicine* (2009)
14. Cohen, Y., Assaf, Y.: High b-value q-space analyzed diffusion-weighted MRS and MRI in neuronal tissues: A technical review. *NMR in biomedicine* 15, 516–542 (2002)

Group-Wise Diffeomorphic Diffusion Tensor Image Registration

Xiujuan Geng, Hong Gu, Wanyong Shin,
Thomas J. Ross, and Yihong Yang

National Institute on Drug Abuse, NIH
gengx@nida.nih.gov

Abstract. We propose an unbiased group-wise diffeomorphic registration technique to normalize a group of diffusion tensor (DT) images. Our method uses an implicit reference group-wise registration framework to avoid bias caused by reference selection. Log-Euclidean metrics on diffusion tensors are used for the tensor interpolation and computation of the similarity cost functions. The overall energy function is constructed by a diffeomorphic demons approach. The tensor reorientation is performed and implicitly optimized during the registration procedure. The performance of the proposed method is compared with reference-based diffusion tensor imaging (DTI) registration methods. The registered DTI images have smaller shape differences in terms of reduced variance of the fractional anisotropy maps and more consistent tensor orientations. We demonstrate that fiber tract atlas construction can benefit from the group-wise registration by producing fiber bundles with higher overlaps.

1 Introduction

Diffusion tensor magnetic resonance imaging (DTI) is an emerging imaging modality to non-invasively measure water diffusion in biological tissues [1]. It plays an important role in studying brain white matter microstructure and anatomical connectivity. DTI registration is necessary to estimate correspondences among different diffusion tensor (DT) images, study white matter alterations in developing and disease populations, and build a white matter atlas. DTI registration techniques have been first proposed in [2]. In contrast to scalar images, DT images are multidimensional at each voxel, and more information such as shape and orientation of the tensors can be used in registration. Recently, many DTI registration methods have been presented using whole tensors [3,4], features extracted from tensors [5,6], multi-channels [7], and multi-contrasts [8].

In building white matter and fiber tract atlases, group-wise registration is crucial to map each image to an unbiased common space [9]. In [10], the fractional anisotropy (FA) extracted from DTI is used and registered to a common space following the work in [11]. Zhang *et al.* [12] proposed an atlas construction method using the information encoded in tensors, especially the orientation information, which may enable more accurate alignment of fiber tracts. This method estimates the unbiased atlas iteratively. The input DT images are first

averaged, and each image is registered to it. All deformed images are averaged again to get the second atlas; the procedure is repeated until the atlas converges [13]. Bampoutis *et al.* [14] recently developed a group-wise registration of 4th-order tensor images and showed the improvement compared to 2nd-order tensor registration using synthetic fiber crossing data.

In this paper, we propose a group-wise DTI registration technique that extends the implicit-reference group-wise (IRG) registration framework proposed by Geng *et al.* [15]. The DTI IRG registration simultaneously estimates diffeomorphic transformations and the deformed DT images converge to the group average. The unbiased atlas can be obtained by averaging the deformed images. Tensor reorientation is applied using “Finite Strain” strategy [2] during the estimation to preserve the geometric features of the tensor fields [9]. The Log-Euclidean metric is used to define the similarity cost and for tensor interpolation. Compared to the work in [12], our method does not compute the average input images as the reference and update the atlas by repeating the registration multiple times. We jointly estimate transformations that deform each pair of DT images to have similar shapes. Another major difference is that the registration algorithm in [12] estimates transformations using an enhanced piecewise affine framework, our registration adapts large deformation diffeomorphic demons approach [4].

The FA map variance and the directional consistency after the IRG DTI registration were compared to reference-based registration. Results show smaller registration errors using the group-wise registration. To demonstrate the advantage in building a fiber tract atlas, we constructed the uncinate fasciculus (UF) atlas of a group of healthy normal subjects using group-wise and reference DTI registration methods. The proposed group-wise registration produces higher bundle overlap and therefore a sharper UF atlas.

2 Implicit-Reference Group-Wise DTI Registration

Tensor distance computation, interpolation and reorientation are involved in DTI registration. The Affine-invariant metric, a Riemannian metric, is the natural metric for diffusion tensors that are positive semi-definite. Here we use Log-Euclidean metric since it is computationally efficient and provides similar performance compared to Riemannian metric [16]. The distance between two tensors is defined as:

$$D(T_i, T_j) = (\text{Trace}(\log T_i - \log T_j)^2)^{\frac{1}{2}}. \quad (1)$$

The log of a rotated tensor can be calculated as $\log(R^T T R) = R^T \log(T) R$. The interpolation of log tensors can be performed using the Euclidean metric after taking the log transformation of the tensors. These properties make the registration convenient by applying a log transform of the input DTI data before registration, using the Euclidean metric to estimate the transformations and performing exponential transform to convert log tensors back to tensors.

2.1 Algorithm Overview

The scalar-based IRG registration with a small deformation elastic model can be stated as an optimization problem of finding a set of N transformations that deform N images to an implicit reference by minimizing the following function:

$$C = \sum_{i=1}^N \sum_{j=1}^{i-1} C_{Sim}(I_i(\Phi_{iS}), I_j(\Phi_{jS})) + \sum_{i=1}^N C_{Reg}(\Phi_i).$$

Φ_{iS} is the transformation to be estimated to map image I_i to the implicit reference I_S . C_{Sim} is the similarity cost between deformed I_i and I_j . C_{Reg} is the regularization cost and can be defined as an elastic differential operator [15].

To map DT images with large shape differences, a similar framework as diffeomorphic demons was used to estimate transformations in diffeomorphisms. Unlike the work in [4], the exact finite strain differential is not implemented here, and the diffeomorphic transformations are ensured by constraining velocity fields small enough at each step. The optimization scheme is to separate the energy function to two parts by introducing another transformation variable Ψ . The minimization of the similarity and regularization costs can be separated. A quadratic regularization can be performed efficiently with a convolution kernel.

The group-wise diffeomorphic DTI registration is formulated as estimating a set of transformations that minimize the cost function:

$$C = \sum_{i=1}^N \sum_{j=1}^{i-1} C_{Sim}(T_i \circ \Phi_{iS}^t, T_j \circ \Phi_{jS}^t) + \sum_{i=1}^N D(\Phi_{iS}^t, \Psi_{iS}^t)^2 + \sum_{i=1}^N C_{Reg}(\Psi_{iS}^t, v_{iS}^t), \quad (2)$$

where v_{iS}^t is the velocity field from T_i to the implicit reference at time t , $\Phi_{iS}^t = \int_{\tau=0}^t v_{iS}^\tau d\tau$, $D(\Phi_{iS}^t, \Psi_{iS}^t)$ is the distance between Φ and Ψ to constrain Ψ close to Φ , and the last term is the regularization term to ensure smooth velocity and transformation fields, which is defined as an isotropic differentiable operator $\nabla(\cdot)$ in this work. The similarity cost can be defined as the squared Euclidean distance after log transform of the tensors. Let T_i denote the log transformed tensor image for simplicity. The cost function in Eq.(2) can be rewritten as:

$$\begin{aligned} & \sigma \sum_{i=1}^N \sum_{j=1}^{i-1} \int_{\Omega} \|T_i \circ \Phi_{iS}^t(x) - T_j \circ \Phi_{jS}^t(x)\|^2 dx + \rho \sum_{i=1}^N \int_{\Omega} \|\Phi_{iS}^t(x) - \Psi_{iS}^t(x)\|^2 dx \\ & + \sum_{i=1}^N (\lambda_{\Psi} \int_{\Omega} \|\nabla(\Psi_{iS}^t(x))\|^2 dx + \lambda_v \int_{\Omega} \|\nabla(v_{iS}^t(x))\|^2 dx), \end{aligned} \quad (3)$$

where Ω represents the tensor image space, and σ , ρ , λ_{Ψ} and λ_v are weighting parameters for each cost term. We set σ to $\frac{1}{N-1} \sum_{j=1, \neq i}^N \frac{1}{\|T_i \circ \Phi_{iS}^t(x) - T_j \circ \Phi_{jS}^t(x)\|^2}$ for each Φ_{iS} , varied ρ to keep the largest update field less than 1 at each iteration, and set λ_{Ψ} and λ_v to be 0.5 and 0.5.

The deformed tensor image $T_i \circ \Phi_{iS}^t$ can be computed as:

$$R_{iS} T_i(\Phi_{iS}^t) R_{iS}^T, \quad (4)$$

where R_{iS} is the reorientation matrix extracted from the Jacobian of the inverse Φ_{iS} in Eulerian space, or from the Jacobian and taking its transpose afterwards [2]:

$$R = ((J(\Phi) \cdot J(\Phi)^T)^{-\frac{1}{2}} J(\Phi))^T. \tag{5}$$

2.2 Optimization of the Cost Function

After linear approximation of the similarity cost function, the first two terms in Eq.(3) become

$$\begin{aligned} \sigma & \parallel \sum_{j=1, \neq i}^N (T_i \circ \Psi_{iS}^t - T_j \circ \Psi_{jS}^t) + \nabla T_i \circ \Psi_{iS}^t \cdot v_{iS}^t \parallel^2 + \rho \int_{\Omega} \|\Phi_{iS}^t(x) - \Psi_{iS}^t(x)\|^2 dx \\ & = \parallel \left[\sum_{j=1, \neq i}^N \sqrt{\sigma} (T_i \circ \Psi_{iS}^t - T_j \circ \Psi_{jS}^t) \right]_0 + \left[\sqrt{\sigma} \nabla T_i \circ \Psi_{iS}^t \right]_{\sqrt{\rho} I} v_{iS}^t \parallel^2, \end{aligned} \tag{6}$$

where I denotes the identity transformation field. The update of v_{iS}^t can be calculated by setting the above equation to zero and solving for v_{iS}^t :

$$v_{iS}^t = \frac{-\sum_{j=1, \neq i}^N (\nabla T_i \circ \Psi_{iS}^t)' (T_i \circ \Psi_{iS}^t - T_j \circ \Psi_{jS}^t)}{\parallel \nabla T_i \circ \Psi_{iS}^t \parallel^2 + \frac{\rho}{\sigma} I}. \tag{7}$$

The optimization of the Eq.(3) except the similarity term can be done by convolving a Gaussian kernel with v_{iS} and Ψ_{iS} alternatively [17]. Smoothing v_{iS} is more fluid like registration, and smoothing Ψ_{iS} is more diffusion like registration. The iterative process is done for each image simultaneously. The cost function is similar to a recent 4D registration study with multichannel diffeomorphic Demons algorithm [18].

The implementation of the algorithm can be summarized as follows: (1) take the log transform of the input tensor images; (2) initialize Φ_{iS} and Ψ_{iS} to be identity fields; (3) reorient T_i using Eqs.(4, 5); (4) let $\Phi_{iS}^n = \Psi_{iS} \circ (I + v_{iS}^n)$ and estimate the velocity field v_{iS}^{n+1} according to Eq.(7); (5) regularize v by taking the Gaussian kernel of it: $v_{iS}^{n+1} = \lambda_v K * v_{iS}^{n+1}$; (6) let $\Psi_{iS}^{n+1} = \lambda_{\Psi} K * \Psi_{iS}^n \circ (I + v_{iS}^{n+1})$, and $\Phi_{iS}^{n+1} = \Psi_{iS}^{n+1}$; (7) repeat steps 3–6 until convergence, $C_{Sim} < \epsilon$, or $n > N_{max}$; (8) apply the exponential transform to convert the log tensors back to tensors and apply the estimated Φ_{iS} to them to generate deformed DT images.

3 Experiments and Results

We demonstrate the performance of the proposed method using 10 DT images acquired from a 3T Siemens MRI scanner. An EPI-based spin echo sequence was used to acquire diffusion-weighted images (DWIs). The whole brain was covered with 35 axial slices and a 128×128 in-plane matrix, resulting in a resolution of $1.72 \times 1.72 \times 4.0mm^3$. Besides one non-diffusion weighted reference image, 12 directions were used to apply the diffusion-sensitive gradients with a b-factor of

1000s/mm². For EPI, TR/TE = 5000/87ms, BW = 1700Hz/Pixel, and NEX=4. The DT images were reconstructed from DWIs. B0 maps were used to affine align each map to the average of the original maps, and the transformation matrices were applied to the DT images to obtain affine aligned images.

To compare the IRG DTI registration with reference-based methods, each DT image was selected as the reference, and all other images were registered to it using the same diffeomorphic DTI registration framework with the same parameter settings. Therefore, 10 groups of reference-based registrations were performed and compared with the IRG registration. The standard deviation (STD) of FA maps and the directional consistency (DC) of tensor images after registration were computed to evaluate different methods. The FA and DC were computed from the deformed tensor images. The DC is defined as the absolute value of the dot product of the major eigenvectors of two tensor images, ranging from 0 (two directions are perpendicular) to 1 (directions are consistent).

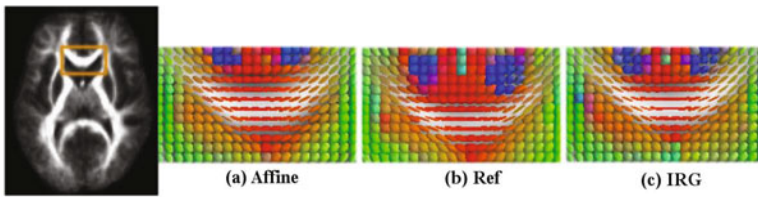


Fig. 1. Average tensor images in the genu of the corpus callosum after (a) affine alignment, (b) reference-based, and (c) group-wise DTI registration

Fig. 1 shows a subregion of the averaged tensor images in the genu of the corpus callosum before diffeomorphic registration (after affine), after a typical reference-based and the proposed IRG DTI registration. The average (using Log-Euclidean metric) of the deformed tensor images has pointier shapes after diffeomorphic registration which may be due to the more consistent directions of the deformed tensors. Compared to the reference-based method, the IRG method produced an average shape closer to the average of the affine aligned images.

Fig. 2 plots the STD of FA maps and the DC maps after affine, a typical reference-base and IRG DTI registration. The maps are shown on a mask generated from thresholding the average of all deformed FA maps after all registrations at 0.3. The same mask was used to compute the RMS of the STD of FA maps and the average of DC maps for 10 reference-based registrations and the group-wise registration. The RMS of each registration method is listed in Tab. 1. The IRG method produced the least FA variance compared to all reference-based methods. For each registration, nine DC maps were generated and the average of each DC map is plotted in Fig. 3. To compare the registration with image I_i as the reference, the DC map was computed between the deformed I_i , and all other deformed images after IRG registration. The IRG registered tensors with higher directional consistency compared to all reference-based registrations with inconsistent performances varying according to the reference selection.

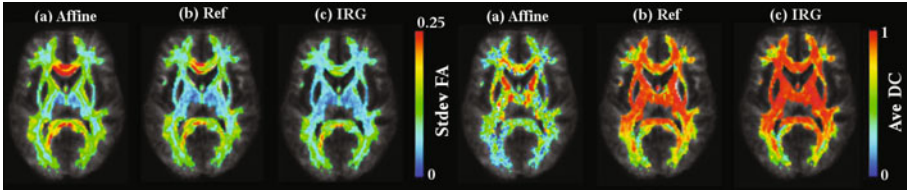


Fig. 2. Standard deviation of FA and DC maps after (a) affine alignment, (b) reference-based, and (c) group-wise DTI registration

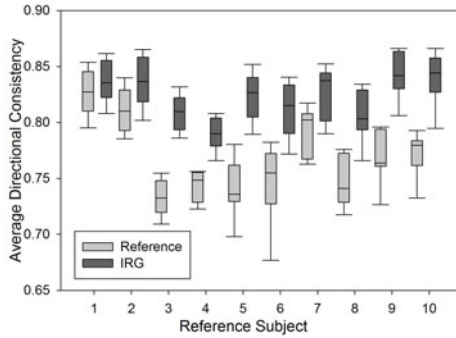


Fig. 3. Average directional consistency values after 10 reference-based registrations and the group-wise registration

The construction of an UF atlas was used to demonstrate the advantage of the group-wise DTI registration. Whole brain white matter fiber tracts were generated using the software “Trackvis” [19] from the registered DTI data. An anatomical atlas provided in AFNI [20] was used to provide the ROIs of right amygdala and medial prefrontal cortex (mPFC). The anatomical atlas was affine aligned to the average of the deformed FA maps. The ROIs were diluted 2mm along the boundary to overlap with the white matter tracts. The UF bundles were extracted for each subject by selecting fibers passing both right amygdala and mPFC. All points on the fiber tracks were projected to the image grid. The number of points in each voxel was counted and the values in each voxel were normalized to 0 to 1, where 0 means no fiber passing the voxel and 1 indicates this voxel has the highest fiber density. We defined the fiber bundle to be the regions

Table 1. RMS of the standard deviation of the deformed FA maps after reference and group-wise registration

method	ref-1	ref-2	ref-3	ref-4	ref-5	ref-6	ref-7	ref-8	ref-9	ref-10	IRG
RMS	.0933	.0955	.111	.108	.114	.115	.103	.108	.111	.108	.0889

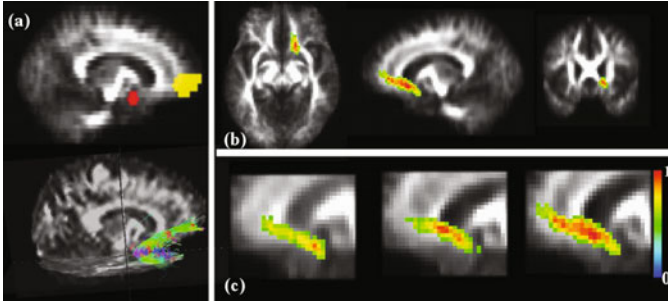


Fig. 4. Uncinate fasciculus atlas generated using group-wise and reference-based DTI methods. (a) Amygdala and mPFC and the UF fibers from a single subject. (b) The UF atlas built from group-wise registration. (c) The sagittal view of the UF atlas built by (from left to right): affine alignment, reference-based registration, and group-wise registration. The color represent the relative overlap of fiber bundles.

with the fiber density greater than 0.5. The fiber bundle atlas was computed by averaging each binarized fiber bundle. To evaluate the fiber bundle atlas, the relative overlap at each voxel x was computed: $RO(x) = n(x)/N$, where $n(x)$ is the number of subjects that have fibers at x , and N is the total number of subjects. Fig. 4 shows the UF atlas after affine, a typical reference-based and IRG registration. The average overlaps of the UF bundles were 0.531, 0.616 and 0.632 after affine, reference-based and IRG methods, respectively.

4 Discussion

Overall, our results show that the IRG DTI registration provides better performance in terms of smaller FA variation and higher tensor directional consistency compared to reference-based methods under the same diffeomorphic registration framework. Using the UF as an example, the group-wise registration provides higher overlap of the deformed UF bundles, and therefore improves the construction of the fiber tract atlas.

Even though the unbiased group-wise DTI registration produces less error than the pair-wise reference-based method, it is still difficult to match fiber tracts across subjects due to the large variations and sometimes different topologies of the same structure. Combining fiber bundle information in the registration procedure may help to produce better fiber bundle alignment [21]. Since 2nd order tensors have limitations in representing intra-voxel fiber crossings, tensor-based registrations may not map the complex white matter structures accurately. Higher order tensor registration [14] and registration of orientation distribution functions [22] have the potential to overcome these difficulties.

Acknowledgment. This work was sponsored by the National Institute on Drug Abuse, Intramural Research Program, National Institutes of Health.

References

1. Basser, P.J., Pierpaoli, C.: Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor mri. *Journal of Magnetic Resonance Series B*, 111, 209–219 (1996)
2. Alexander, D., Pierpaoli, C., Basser, P., Gee, J.: Spatial transformations of diffusion tensor magnetic resonance images. *IEEE Trans. Med. Imaging* 20(11), 1131–1139 (2001)
3. Chiang, M.C., Leow, A.D., Klunder, A.D., Dutton, R.A., Barysheva, M., Rose, S.E., McMahon, K.L., de Zubicaray, G.I., Toga, A.W., Thompson, P.M.: Fluid registration of diffusion tensor images using information theory. *IEEE Trans. Med. Imaging* 27(4), 442–456 (2008)
4. Yeo, B.T., Vercauteren, T., Fillard, P., Peyrat, J.M., Pennec, X., Golland, P., Ayache, N., Clatz, O.: Dt-refind: Diffusion tensor registration with exact finite-strain differential. *IEEE Trans. Med. Imaging* 99 (2009)
5. Yap, P.T., Wu, G., Zhu, H., Lin, W., Shen, D.: Timer: Tensor image morphing for elastic registration. *NeuroImage* 47, 549–563 (2009)
6. Li, H., Xue, Z., Guo, L., Wong, S.T.: Simultaneous consideration of spatial deformation and tensor orientation in diffusion tensor image registration using local fast marching patterns. In: Prince, J.L., Pham, D.L., Myers, K.J. (eds.) *Information Processing in Medical Imaging*. LNCS, vol. 5636, pp. 63–75. Springer, Heidelberg (2009)
7. Park, H.J., Kubicki, M., Shenton, M.E., Guimond, A., McCarley, R.W., Maierd, S.E., Kikinis, R., Jolesz, F.A., Westin, C.F.: Spatial normalization of diffusion tensor mri using multiple channels. *NeuroImage* 20(4), 1995–2009 (2003)
8. Ceritoglu, C., Oishi, K., Li, X., Chou, M.C., Younes, L., Albert, M., Lyketsos, C., van Zijl, P.C., Miller, M.I., Mori, S.: Multi-contrast large deformation diffeomorphic metric mapping for diffusion tensor imaging. *NeuroImage* 47, 618–627 (2009)
9. Peyrat, J.M., Sermesant, M., Pennec, X., Delingette, H., Xu, C., McVeigh, E.R., Ayache, N.: A computational framework for the statistical analysis of cardiac diffusion tensors: Application to a small database of canine hearts. *IEEE Trans. on Med. Imaging* 26(11), 1500–1514 (2007)
10. Goodlett, C., Davis, B., Jean, R., Gilmore, J., Gerig, G.: Improved correspondence for dti population studies via unbiased atlas building. In: *MICCAI*, pp. 260–267 (2007)
11. Joshi, S., Davis, B., Jomier, M., Gerig, G.: Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage* 23, S151–S160 (2004)
12. Zhang, H., Yushkevich, P.A., Rueckert, D., Gee, J.C.: Unbiased white matter atlas construction using diffusion tensor images. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part II*. LNCS, vol. 4792, pp. 211–218. Springer, Heidelberg (2007)
13. Guimond, A., Meunier, J., Thirion, J.: Average brain models: A convergence study. *Computer Vision and Image Understanding* 77(2), 192–210 (2000)
14. Barmpoutis, A., Vemuri, B.C.: Groupwise registration and atlas construction of 4th-order tensor fields using the r+ riemannian metric. In: *MICCAI*, pp. 640–647 (2007)
15. Geng, X., Christensen, G.E., Gu, H., Ross, T.J., Yang, Y.: Implicit reference-based group-wise image registration and its application to structural and functional mri. *NeuroImage* 47(4), 1341–1351 (2009)

16. Kindlmann, G., Estepar, R.S.J., Niethammer, M., Haker, S., Westin, C.F.: Geodesic-loxodromes for diffusion tensor interpolation and difference measurement. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part I. LNCS, vol. 4791, pp. 1–9. Springer, Heidelberg (2007)
17. Cachier, P., Ayache, N.: Isotropic energies, filters and splines for vector field regularization. *Journal of Mathematical Imaging and Vision* 20(3), 251–265 (2004)
18. Peyrat, J.M., Delingette, H., Sermesant, M., Xu, C., Ayache, N.: Registration of 4d cardiac ct sequences under trajectory constraints with multichannel diffeomorphic demons. *IEEE Trans. on Med. Imaging* (2010)
19. Wang, R., Benner, T., Sorensen, A., Wedeen, V.: Diffusion toolkit: A software package for diffusion imaging data processing and tractography. In: Proceedings of the 15th Annual Meeting of ISMRM (2007)
20. Cox, R.W.: AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research* 29, 162–173 (1996)
21. Ziyang, U., Sabuncu, M.R., O'Donnell, L.J., Westin, C.F.: Nonlinear registration of diffusion mr images based on fiber bundles, pp. 351–358 (2007)
22. Geng, X., Ross, T.J., Zhan, W., Gu, H., Chao, Y.P., Lin, C.P., Christensen, G.E., Schuff, N., Yang, Y.: Diffusion mri registration using orientation. In: Information Processing in Medical Imaging, pp. 626–637 (2009)

Fast and Accurate Reconstruction of HARDI Data Using Compressed Sensing

Oleg Michailovich¹ and Yogesh Rathi²

¹ Department of ECE, University of Waterloo

² Brigham and Women's Hospital, Harvard Medical School

Abstract. A spectrum of brain-related disorders are nowadays known to manifest themselves in degradation of the integrity and connectivity of neural tracts in the white matter of the brain. Such damage tends to affect the pattern of water diffusion in the white matter – the information which can be quantified by diffusion MRI (dMRI). Unfortunately, practical implementation of dMRI still poses a number of challenges which hamper its wide-spread integration into regular clinical practice. Chief among these is the problem of long scanning times. In particular, in the case of High Angular Resolution Diffusion Imaging (HARDI), the scanning times are known to increase linearly with the number of diffusion-encoding gradients. In this research, we use the theory of compressive sampling (aka compressed sensing) to substantially reduce the number of diffusion gradients without compromising the informational content of HARDI signals. The experimental part of our study compares the proposed method with a number of alternative approaches, and shows that the former results in more accurate estimation of HARDI data in terms of the mean squared error.

1 Introduction

The human brain consists of about 10^{11} nerve cells which can be further subdivided into about 1000 different cell types, a complexity that far exceeds that of other organs of the body. An additional complexity is evident in the way in which the component cells of the brain interconnect and function. In contrast to other types of the cells, each neuron communicates with many target cells by means of its peculiar protoplasmatic protrusion, called an axon. Moreover, axons with similar destinations tend to form bundles, known as neural fiber tracts. Taken together, these fibers play a pivotal role in the determination of brain connectivity, which is presently far from being completely understood. Consequently, by reconstructing the pattern of connectivity of the neural tracts in both healthy and diseased subjects, it is possible to obtain an abundance of valuable diagnostic information that could be used for early diagnostics of brain-related disorders, for assessing the damage caused to the brain by stroke, tumors or injuries, as well as for planning and monitoring of neurosurgeries [1], [2]. In this regard, diffusion MRI (dMRI) appears to be an unparalleled tool of diagnosis imaging of the brain, as it is the only imaging modality which can “sense” the local orientation of neural fibers tracts.

Diffusion tensor imaging (DTI) is the most widespread type of dMRI, which describes the local diffusion of water molecules using a unimodal Gaussian model [3], [4], [5]. Unfortunately, the performance of DTI is known to deteriorate considerably at the spatial locations where neural tracts cross, touch upon each other, or diverge [6], [7], [8], [9]. To overcome this intrinsic limitation of DTI, the High Angular Resolution Diffusion Imaging (HARDI) technique has been proposed [10], [8], [11], [12] which is capable of capturing *multi-modal* diffusion patterns through sampling the corresponding diffusion signals over a spherical shell in k -space.

The superiority of HARDI over DTI suggests the possibility of further improvement of the diagnostic value of dMRI [13]. Unfortunately, this superiority comes at a significant price: HARDI requires using a substantially larger number of diffusion-encoding gradients N , with a typical N being in the range between 60 and 100 (as compared to $N \in [25, 30]$ in the case of DTI). As the total scanning time increases linearly with N , HARDI-based analysis is currently deemed to be “too slow” to be used in clinical applications involving children or patients with dementia.

The above deficiency of HARDI can be overcome using the theory of *compressive sampling* (aka compressed sensing) (CS) [14], [15], [16], which provides a framework to recover HARDI signals from a much smaller number of measurements as compared to the standard range of N . Formalizing such a reconstruction approach constitutes the main contribution of this work. It should be noted that there already exists a body of works in which the theory of CS has been used for reconstruction of grayscale (i.e. T1, T2, PD) MR images from their sub-critical samples in k -space [17], [18], [19]. The proposed method, on the other hand, is different in that it applies the tools of CS in the diffusion domain (rather than in the spatial domain), as is detailed in the sections that follow.

2 Compressed Sensing

A continuum of all possible orientations in \mathbb{R}^3 can be identified with the points of the unit sphere $\mathbb{S}^2 := \{\mathbf{u} \in \mathbb{R}^3 \mid \|\mathbf{u}\|_2 = 1\}$. Let $\mathbf{p} = (x, y, z)$ be the spatial coordinate of an arbitrary voxel within a volume of interest. Then, the diffusion signal $S(\mathbf{u})$ at \mathbf{p} can be formally described as a real-valued function defined over \mathbb{S}^2 . In practical settings, the signal $S(\mathbf{u})$ can only be measured along N discrete orientations $\{\mathbf{u}_k\}_{k=1}^N$. The most fundamental question in this regard has always been: what is the minimum number of diffusion directions N required to unambiguously represent the signal S in terms of its discrete values S_k ?

A particularly important answer to the above question is offered by the theory of CS. To summarize some fundamental results of the theory, we consider S to be a member of a Hilbert space, endowed with a standard inner product. Since diffusion measurements are linear, the discrete measurements $S_k := S(\mathbf{u}_k)$ can be expressed in the form of inner products $S_k = \langle S, \varphi_{j_k} \rangle$, with $\{\varphi_{j_k}\}_{k=1}^N$ being a subset of a sampling basis $\{\varphi_i\}_{i \in \mathcal{I}}$. Moreover, let $\{\psi_j\}_{j \in \mathcal{J}}$ be another basis in the signal space, such that any S can be expressed according to

$$S(\mathbf{u}) = \sum_{j \in \mathcal{J}} c_j \psi_j(\mathbf{u}), \quad \forall \mathbf{u} \in \mathbb{S}^2, \quad (1)$$

where \mathcal{J} denotes the set of indices over which the basis functions ψ_j are counted. Note that, in a more general setup, the set $\{\psi_j\}_{j \in \mathcal{J}}$ may be overcomplete, in which case it becomes a frame. However, independently of whether it is a basis or a frame, in what follows, the set $\{\psi_j\}_{j \in \mathcal{J}}$ is assumed to be finite, with the total number of its elements being equal to M .

Using the above definitions, we make the following two assumptions:

- S is assumed to be *sparsely representable* by $\{\psi_j\}_{j \in \mathcal{J}}$, which implies that the number K of non-zero coefficients c_j in (1) is significantly less than M .
- The bases $\{\varphi_i\}_{i \in \mathcal{I}}$ and $\{\psi_j\}_{j \in \mathcal{J}}$ are assumed to be *incoherent*, implying that the value of $\mu = \sup_{i,j} |\langle \varphi_i(\mathbf{u}), \psi_j(\mathbf{u}) \rangle|$ is relatively small.

On Conditions (a) and (b), the theory of CS proves that an accurate approximation of S is possible from only $\mathcal{O}(\mu^2 \log(M)K)$ of its measurements, as opposed to $\mathcal{O}(M)$ in the case of standard sampling theory. Moreover, this approximation is computable through solution of a convex optimization problem [14], [15], [16].

The above considerations suggest that the applicability of CS to HARDI depends on the availability of a basis/frame $\{\psi_j\}_{j \in \mathcal{J}}$ for which the assumptions (a) and (b) above would be valid. Such a basis has been recently introduced in [20], where it is called a basis of *spherical ridgelets* – the functions which have been specifically designed to provide sparse representation of HARDI signals. Moreover, as the energy of spherical ridgelets is distributed alongside the great circles of \mathbb{S}^2 , the ridgelets appear to be very incoherent with respect to the Dirac sampling basis. All this makes spherical ridgelets an ideal candidate to be used for CS-based recovery of HARDI signals, as it is demonstrated next.

3 Spherical Ridgelets

Spherical ridgelets are constructed using the fundamental principles of wavelet theory. Let $\kappa_\rho(x) = \exp\{-\rho x(x+1)\}$, where $\rho > 0$ and $x \geq 0$. Also, let $\kappa_j(x) = \kappa_\rho(2^{-j}x)$ be a scaled version of κ_ρ , where $j \in \mathbb{N}$. Then, the Gaussian-Weierstrass scaling function $K_{j,\mathbf{v}} : \mathbb{S}^2 \rightarrow \mathbb{R}$ at resolution $j \in \mathbb{N}$ and orientation $\mathbf{v} \in \mathbb{S}^2$ is given by

$$K_{j,\mathbf{v}}(\mathbf{u}) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \kappa_j(n) P_n(\mathbf{u} \cdot \mathbf{v}), \quad \forall \mathbf{u} \in \mathbb{S}^2 \quad (2)$$

where P_n denotes the Legendre polynomial of order n and the dot stands for the Euclidean inner product in \mathbb{R}^3 . Consequently, following [20], the semi-discrete frame \mathbb{F} of spherical ridgelets can be defined as

$$\mathbb{F} := \{\Psi_{j,\mathbf{v}} \mid \mathbf{v} \in \mathbb{S}^2, j \in \mathbb{N} \cup \{-1\}\}, \quad (3)$$

where

$$\Psi_{j,\mathbf{v}} = \frac{1}{2\pi} \begin{cases} \mathcal{R}\{K_{0,\mathbf{v}}\}, & \text{if } j = -1, \\ \mathcal{R}\{K_{j+1,\mathbf{v}} - K_{j,\mathbf{v}}\}, & \text{if } j \in \mathbb{N} \end{cases} \quad (4)$$

with \mathcal{R} standing for the Funk-Radon transform (FRT) [21]. Note that, for an arbitrary $j \in \mathbb{N} \cup \{-1\}$ and $\mathbf{v} \in \mathbb{S}^2$, the FRT of $K_{j,\mathbf{v}}(\mathbf{u})$ can be computed as

$$\mathcal{R}\{K_{j,\mathbf{v}}\}(\mathbf{u}) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \lambda_n \kappa_j(n) P_n(\mathbf{u} \cdot \mathbf{v}), \quad \forall \mathbf{u} \in \mathbb{S}^2 \tag{5}$$

where

$$\lambda_n = \begin{cases} 2\pi(-1)^{n/2} \frac{1 \cdot 3 \cdots (n-1)}{2 \cdot 4 \cdots n}, & \text{if } n \text{ is even} \\ 0, & \text{if } n \text{ is odd.} \end{cases} \tag{6}$$

The set \mathbb{F} in (3) is infinite-dimensional, and hence is not suitable for practical computations. To define a discrete counterpart of \mathbb{F} , one has first to restrict the values of the resolution index j to a finite set $\{-1, 0, 1, \dots, J\}$, where J defines the highest level of “detectable” signal details. Additionally, the set of all possible orientations $\mathbf{v} \in \mathbb{S}^2$ of spherical ridgelets needs to be discretized as well. To find a proper discretization scheme, we first note that the construction in (4) is dyadic, which suggests that the bandwidth of spherical ridgelets increases proportionally to 2^j . Owing to the fact that the dimension of the space of spherical harmonics of degree n is equal to $(n+1)^2$, it is reasonable to define the number of discrete orientations to be a function of j . Specifically, let m_0 be the smallest spherical order resulting in $\kappa_0(m_0) \leq \epsilon$ for some predefined $0 < \epsilon \ll 1$ (e.g. $\epsilon = 10^{-6}$). Then, one can define the number of ridgelet orientations at resolution j to be equal to $M_j = (2^{j+1}m_0 + 1)^2$. Subsequently, for each M_j , let $\{\mathbf{v}_{i,j}\}_{i=1}^{M_j}$ be a set of M_j distinct points on \mathbb{S}^2 . A finite-dimensional set of spherical ridgelets can then be defined as

$$\mathbb{F}_d = \bigcup_{j=-1}^J \{\Psi_{j,\mathbf{v}_{i,j}}\}_{i=1}^{M_j}, \tag{7}$$

where the subscript d stands for “discrete”. It should be emphasized that the set \mathbb{F}_d is finite, as it consists of a total of $M = \sum_{j=-1}^J (2^{j+1}m_0 + 1)^2$ spherical ridgelets.

Given a measurement set of N diffusion-encoding orientations $U_N := \{\mathbf{u}_k\}_{k=1}^N$, one can use (4) and (5) to compute the values of the spherical ridgelets in \mathbb{F}_d over U_N . The resulting vales can be then stored into an $N \times M$ matrix \mathcal{A} . Subsequently, if $c \in \mathbb{R}^M$ is defined to be a (column) vector of ridgelet coefficients and $y := [S(\mathbf{u}_1), S(\mathbf{u}_2), \dots, S(\mathbf{u}_N)]^T$, then the measurement model can be formally expressed as $\mathcal{A}c = y + e$, where e is an error vector that accounts for both measurement and model errors.

Assuming that $\|e\|_2 \leq \eta$, an optimal estimate c should satisfy $\|\mathcal{A}c - y\|_2 \leq \eta$. Unfortunately, since $N < M$, the above condition by itself falls short to define an optimal c . However, if the coefficients c are known to be sparse w.r.t. \mathbb{F}_d , then the theory of CS states that a useful estimate of the representation coefficients can be found by solving

$$c^* = \arg \min_c \|c\|_1 \text{ subject to } \|\mathcal{A}c - y\|_2 \leq \eta. \tag{8}$$

Note that (8) is a convex optimization problem, whose solution can be found using standard tools of optimization theory. In the present work, the problem has been solved using the ℓ_1 -magic software freely available at <http://www.acm.caltech.edu/l1magic/>.

For the sake of comparison, we also estimate the representation coefficients c by computing the minimum ℓ_2 -norm solution satisfying $\|\mathcal{A}c - y\|_2 \leq \eta$. It is important to point out that such solutions are typically *not* sparse, in which case there are no *a priori* guarantees of optimality of resulting estimates. It will be shown however that, in the case of spherical ridgelet analysis, the ℓ_2 -solutions appear to be quite informative. Note that the ℓ_2 -estimates are attractive for they admit a close form expression which is

$$c^* = \mathcal{A}^T (\mathcal{A} \mathcal{A}^T)^{-1} y. \quad (9)$$

4 Results

To test the performance of the proposed method, HARDI data acquired from an dMRI phantom [22] were used. Specifically, the data were acquired using 3 mm isotropic acquisition over a square field-of-view of size 19.2 cm, $b = 2000$ s/mm², and 128 diffusion orientations, uniformly distributed over the sphere. (For more details on the data acquisition setup see www.lnao.fr/spip.php?article107.) The acquired HARDI signals were low-pass filtered using the regularized estimation approach of [23] to suppress the effect of measurement noise. The signals thus obtained were regarded as the *original* HARDI signals against which their CS-based estimates were compared. In this work, normalized mean squared error (NMSE) has been used to assess the performance of different reconstruction methods.

To simulate CS acquisition, the full set of 128 diffusion directions was reduced to subsets of $N = 16, 18, 20, \dots, 32$ spherical points. These subsets were chosen in such a way that their points covered \mathbb{S}^2 in a quasi-uniform manner (similarly to the original coverage by 128 points). For each value of N , its associated sensing matrix \mathcal{A} was constructed with $J = 1$, $\rho = 0.5$ and $m_0 = 4$. Subsequently,

Table 1. NMSE ($\pm\sigma$) obtained by different reconstruction methods

N	Spherical Ridgelets (SR)		Spherical Harmonics (SH)	
	ℓ_1 -minimization	ℓ_2 -minimization	ℓ_1 -minimization	ℓ_2 -minimization
16	0.0148 \pm 0.0045	0.0176 \pm 0.0047	0.0883 \pm 0.0329	0.5145 \pm 0.0211
18	0.0116 \pm 0.0042	0.0147 \pm 0.0039	0.0646 \pm 0.0304	0.4011 \pm 0.0257
20	0.0107 \pm 0.0032	0.0127 \pm 0.0034	0.0440 \pm 0.0183	0.2774 \pm 0.0319
22	0.0085 \pm 0.0031	0.0120 \pm 0.0031	0.0326 \pm 0.0138	0.1976 \pm 0.0257
24	0.0083 \pm 0.0030	0.0118 \pm 0.0028	0.0257 \pm 0.0097	0.1411 \pm 0.0252
26	0.0077 \pm 0.0023	0.0115 \pm 0.0026	0.0195 \pm 0.0071	0.1130 \pm 0.0246
28	0.0069 \pm 0.0018	0.0113 \pm 0.0024	0.0158 \pm 0.0045	0.0790 \pm 0.0230
30	0.0067 \pm 0.0016	0.0112 \pm 0.0020	0.0146 \pm 0.0037	0.0651 \pm 0.0145
32	0.0066 \pm 0.0015	0.0109 \pm 0.0019	0.0121 \pm 0.0033	0.0566 \pm 0.0140

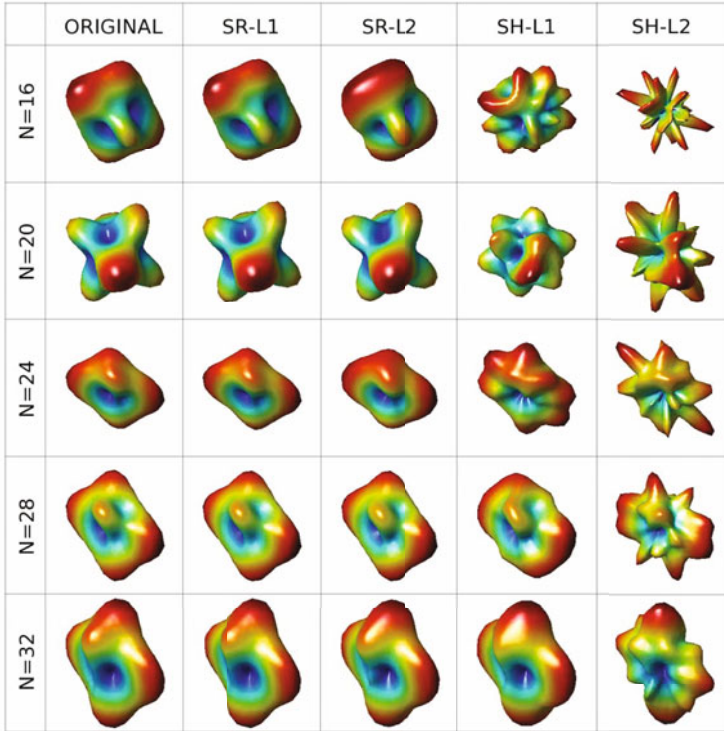


Fig. 1. HARDI signals recovered by different reconstruction algorithms

the optimal ridgelet coefficients were computed using two different approaches, namely by solving the minimum ℓ_1 -norm problem (8) and using (9). (These two estimates will be referred to below as SR-L1 and SR-L2, respectively.) Note that, in the case of SR-L1, the value of $\eta = 0.12$ in (8) was found by trial and error to result in the most accurate reconstruction in terms of NMSE. Finally, the estimated ridgelet coefficients were used to evaluate their associated HARDI signals at the points of the original data set according to (1), followed by the computation of NMSE.

As an alternative to spherical ridgelets, the applicability of spherical harmonics (SHs) to the problem of CS-based reconstruction of HARDI signals has been investigated as well [12], [24]. To this end, the orthogonal set of SHs of orders up to $n = 8$ inclusive was used to define the sensing matrix \mathcal{A} . Note that, since the diffusion signals are symmetric and real-valued, only the even-ordered SHs have been employed. Analogously to the case of spherical ridgelets, the SH-based reconstruction was carried out twice, viz. using both (8) and (9). The respective estimates will be referred to below as SH-L1 and SH-L2, correspondingly.

Table 1 summarizes the values of NMSE (\pm one standard deviation) obtained in the above experiments. Note that each value in the table has been computed by averaging the results of 500 independent estimations. One can clearly see that the

ridgelets-based ℓ_1 -reconstruction (SR-L1) performs incomparably better than SH-L1. Moreover, the SH-based approach is also outperformed by SR-L2, which performs surprisingly well in the case of spherical ridgelet analysis (despite the inability of ℓ_2 -norm minimization to take advantage of the sparse nature of the ridgelet coefficients). The results of Table 1 are also supported by Fig. 1 which depicts a subset of the original HARDI signals together with their estimates computed by the compared algorithms for different values of N .

5 Discussion and Conclusions

Central to the theory of compressed sensing is the notion of sparse representation. In particular, if a signal of interest is known to have a sparse representation in the domain of a linear transform, then it can be accurately reconstructed from a much smaller number of its discrete measurements, as it is required by the classical sampling theory. In the case of HARDI data analysis, the most frequently used representation basis is that of spherical harmonics (SHs). Unfortunately, SHs do not provide sparse representation of HARDI signals, which makes them inappropriate for the use in CS-based estimation. The basis of spherical ridgelets, on the other hand, can represent HARDI signals using a relatively small number of representation coefficients. As demonstrated by Table 1, the above fact allows one to reconstruct HARDI signals from as few as 20 of their measurements with the reconstruction error being about 1%. This result suggests that the CS-based acquisition of HARDI signals has a potential of substantially shortening the scanning times in dMRI, which is known to be a major bottleneck in clinical applications of this important tool of diagnostic imaging.

References

1. Mori, S.: Introduction to diffusion tensor imaging. Elsevier, Amsterdam (2007)
2. Johansen-Berg, H., Behrens, T.E.J.: Diffusion MRI: From quantitative measurements to in-vivo neuroanatomy, 1st edn. Academic Press, London (2009)
3. Basser, P.J., Mattiello, J., LeBihan, D.: Estimation of the effective self-diffusion tensor from the NMR spin echo. *J. Magn. Reson. Imag.* B 103(3), 247–254 (1994)
4. Bihan, D.L., Mangin, J.F., Poupon, C., Clark, C., Pappata, S., Molko, N., Chabriat, H.: Diffusion tensor imaging: Concepts and applications. *Journal of Magnetic Resonance Imaging* 13, 534–546 (2001)
5. 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, 1077–1088 (2003)
6. Frank, L.: Anisotropy in high angular resolution diffusion-tensor MRI. *Magnetic Resonance in Medicine* 45, 935–939 (2001)
7. Alexander, D.C., Barker, G.J., Arridge, S.R.: Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. *Journal of Magnetic Resonance Imaging* 48(2), 331–340 (2002)

8. Tuch, D.S., Reese, T.G., Wiegell, M.R., Makris, N., Belliveau, J.W., Wedeen, V.J.: High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine* 48, 577–582 (2002)
9. Alexander, D.C.: Multiple-fiber reconstruction algorithms for diffusion MRI. *Annals of the New York Academy of Science* 1064, 113–133 (2005)
10. Frank, L.R.: Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magnetic Resonance in Medicine* 47, 1083–1099 (2002)
11. Anderson, A.W.: Measurement of fiber orientation distributions using high angular resolution diffusion imaging. *Magnetic Resonance in Medicine* 54, 1194–1206 (2005)
12. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Apparent diffusion coefficients from high angular resolution diffusion images: Estimation and applications. *Magnetic Resonance in Medicine* 56(2), 395–410 (2006)
13. Behrens, T., Johansen-Berg, H., Jbabdi, S., Rushworth, M., Woolrich, M.: Probabilistic diffusion tractography with multiple fiber orientations: What can we gain? *NeuroImage* 34, 144–155 (2007)
14. Candes, E., Romberg, J.: Quantitative robust uncertainty principles and optimally sparse decompositions. *Foundations of Computational Mathematics* 6(2), 227–254 (2006)
15. Candes, E., Romberg, J., Tao, T.: Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory* 52(2), 489–509 (2006)
16. Donoho, D.: Compressed sensing. *IEEE Transactions on Information Theory* 52(4), 1289–1306 (2006)
17. Lustig, M., Donoho, D., Pauly, J.M.: Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine* 58(6), 1182–1195 (2007)
18. Jung, H., Ye, J.C., Kim, E.Y.: Improved k-t BLASK and k-t SENSE using FOCUSS. *Physics in Medicine and Biology* 52, 3201–3226 (2007)
19. Jung, H., Sung, K., Nayak, K.S., Kim, E.Y., Ye, J.C.: k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI. *Magnetic Resonance in Medicine* 61, 103–116 (2009)
20. Michailovich, O., Rathi, Y.: On approximation of orientation distributions by means of spherical ridgelets. *IEEE Transactions on Image Processing* 19(2), 461–477 (2010)
21. Groemer, H.: Geometric applications of Fourier series and spherical harmonics. Cambridge University Press, Cambridge (1996)
22. Poupon, C., Rieul, B., Kezele, I., Perrin, M., Poupon, F., Mangin, J.F.: New diffusion phantoms dedicated to the study and validation of high-angular-resolution diffusion imaging (HARDI) models. *Magn. Reson. Med.* 60(6), 1276–1283 (2008)
23. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Regularized, fast, and robust analytical Q-ball imaging. *Magnetic Resonance in Medicine* 58, 497–510 (2007)
24. Hess, C.P., Mukherjee, P., Han, E.T., Xu, D., Vigneron, D.R.: Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. *Magnetic Resonance in Medicine* 56, 104–117 (2006)

General and Efficient Super-Resolution Method for Multi-slice MRI

D.H.J. Poot^{1,2}, V. Van Meir², and J. Sijbers²

¹ BGR, Erasmus Medical Center, Rotterdam

² Visionlab, University of Antwerp, Antwerp

Abstract. In this paper, a method is developed that reconstructs a high resolution image from an arbitrary set of multi-slice 3D MR images with a high in-plane resolution and a low through-plane resolution. Such images are often recorded to increase the efficiency of the acquisition. With a model of the acquisition of MR images, which is improved compared to previous super-resolution methods for MR images, a large system with linear equations is obtained. With the conjugated gradient method and this linear system, a high resolution image is reconstructed from MR images of an object. Also, a new and efficient method to apply an affine transformation to multi-dimensional images is presented. This method is used to efficiently reconstruction the high resolution image from multi-slice MR images with arbitrary orientations of the slices.

Keywords: Super-resolution, multi-slice imaging, reconstruction, conjugated gradients, affine transformation, multi-dimensional imaging.

1 Introduction

This paper describes how a high resolution isotropic 3D image can be reconstructed from a series of multi-slice 3D MR images, in which the slice thickness is substantially larger than the in plane resolution. Since this method improves the resolution in all directions up to the in-plane resolution, it is a super-resolution method. The motivation of this work is that multi slice images are often acquired with anisotropic voxels. The slice thickness can be substantially higher than the in-plane resolution in order to increase the Signal to Noise Ratio (SNR). As will be shown, several of these multi-slice MR images, recorded with different slice orientations, can be combined into a single, high resolution 3D image with isotropic voxels. The advantage of multi slice images, compared to full 3D acquisitions is that it is possible to interleave the acquisition of slices. That is, while waiting for the (T_1) relaxation of the magnetization of a slice, (a part of) the k-space of various other slices can be excited and recorded. In general, when the repetition time (TR) is limited by the T_1 decay, it is possible to record multi slice images significantly faster than full 3D images with the same resolution [1].

Previously, several attempts have been made to improve the resolution of MR images. The methods of Peled et. al. [2] and Carmi et al. [3] try to improve the in-plane resolution. The validity of such methods was questioned by Scheffler [4],

since each in-plane shifted image acquires the same points in k-space, only introducing a linear phase shift in the k-space samples. This means that each of the shifted images does contain the same information, except for measurement noise, with no possibility to improve the resolution. A different method, presented by Greenspan et al. [5], improves the resolution in the slice direction, i.e. the direction in which the different slices of a multi-slice MR image are recorded. Several multi-slice MR images with different positions in the slice direction are combined. This method is not limited to the original resolution as the acquisition in the slice direction is not band limited. However, only MR images with identical orientation can be combined. This limitation is partially removed by the method of Shilling et al. [6], in which multi-slice MR images rotated around a common frequency encoding axis are combined. It allows the reconstruction of high resolution slices with iterative projection reconstruction algorithms. In their method they state the projection as a linear system and solve the high resolution image from the set of linear equations with iterative solvers that are also used in Computed Tomography (CT) reconstructions.

In our work, the method of Shilling et al. [6] is extended to allow for any orientation of the slices of the multi-slice MR images, i.e. the images do not need to be rotated around a common frequency encoding axis. Furthermore, the projection via matrix multiplication is reformulated as affine transformation, followed by a filter operation, which reduces the number of computations substantially. Finally, the method presented in this paper uses the Conjugated Gradient method to solve the large linear system in a low number of iterations.

2 Methods

2.1 Introduction

The acquisition of multiple MR images with the same contrast can be seen as acquiring multiple samples from the same object. For each slice of a multi-slice MR image, the MR acquisition records a part of the k-space of the object. With the discrete Fourier transform, a projection of the object intensities in each slice is reconstructed on a discrete grid. Since MR acquisitions record a limited part of the k-space, the intensity at a grid point does not depend exclusively on the intensity of the object at the location of the grid point, but a low-pass filter is applied to the excited slice of the object before sampling the intensities at the grid nodes. An alternative interpretation is that the intensity value of each voxel of the MR image is obtained by multiplying a properly shifted version of a (3D) sampling function with the object.

2.2 Model of the MRI Acquisition

Let \mathbf{o} ($n_o \times 1$) be a vector containing the intensities of the continuous object o at the 3D grid points \mathbf{x}_m , $m \in \{1, \dots, n_o\}$, where \mathbf{x}_m is a coordinate in object space. Let \mathbf{S}_j ($n_{S_j} \times 1$) be the samples of the $j^{\text{th}} \in \{1, \dots, N\}$ MR image S_j at the grid nodes \mathbf{y}_l , $l \in \{1, \dots, n_{S_j}\}$, where \mathbf{y}_l is a coordinate in the space of

the j^{th} MR image, and let the measurement noise be described by \mathbf{e}_j ($n_{S_j} \times 1$). When the coordinates \mathbf{x} and \mathbf{y} are linked by a coordinate transform T_j , the acquisition of the MR images can be modeled with:

$$\mathbf{S} = \mathbf{X}\mathbf{o} + \mathbf{e}, \tag{1}$$

where

$$\mathbf{S} = [\mathbf{S}_1^T \dots \mathbf{S}_N^T]^T, \quad \mathbf{X} = [\mathbf{X}_1^T \dots \mathbf{X}_N^T]^T, \quad \mathbf{e} = [\mathbf{e}_1^T \dots \mathbf{e}_N^T]^T, \tag{2}$$

and

$$X_j(l, m) = w(T_j(\mathbf{x}_m) - \mathbf{y}_l), \tag{3}$$

where w is the sampling function. This sampling function is assumed to be separable in MR- image coordinates, $w(\mathbf{y}) = \prod_{i=1}^3 w_i(y_i)$, which can, without loss of generality, be numbered 1,2, and 3 for the read, phase, and slice encoding directions, respectively. The sampling function w is (implicitly) defined by the MR image acquisition method. The in-plane sampling functions w_1 and w_2 are defined by the rectangular part of k-space that is sampled and thus are Dirichlet, or periodic sinc functions. For multi-slice MR images, w_3 depends on the slice selection excitation, which usually is either a (windowed) sinc or a Gaussian shaped RF pulse. In our experiments, with windowed sinc slice excitation, w_3 was modeled with a smoothed box function :

$$w_3(y) = \begin{cases} 1 & |y| \leq \frac{1}{3} \\ \frac{1}{2} - \frac{1}{2} \sin\left(3\pi\left(|y| - \frac{1}{2}\right)\right) & \frac{1}{3} < |y| < \frac{2}{3} \\ 0 & \frac{2}{3} \leq |y| \end{cases} \tag{4}$$

Due to the aliasing caused by the slice selection function, it is easier to describe the acquisition in image space (Eq. (II)), rather than in the k-space.

Note that, in this model, the transformation $(T_j(\mathbf{x}))_2$ is not necessarily just a rotation or even only an affine transformation, but it might also contain other deformations. For example, when the images are recorded with Echo Planar Imaging (EPI), it might also describe displacements in the phase encoding direction due to inhomogeneities of the main magnetic field.

2.3 Reconstruction of the Object

When the MR images \mathbf{S} are acquired, the reconstruction of the object intensities \mathbf{o} can be stated as a regularized least squares problem:

$$\hat{\mathbf{o}} = \arg \min_{\mathbf{o}} (\mathbf{X}\mathbf{o} - \mathbf{S})^T (\mathbf{X}\mathbf{o} - \mathbf{S}) + \lambda \mathbf{o}^T \mathbf{K} \mathbf{o}, \tag{5}$$

where \mathbf{K} specifies the regularization term, which will be explained below, and λ is a scalar constant that scales the regularization. In general, the solution of this regularized least squares problem is given by:

$$\hat{\mathbf{o}} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{K})^{-1} \mathbf{X}^T \mathbf{S} \tag{6}$$

2.4 Regularization

Regularization is a standard technique to solve under-determined or badly conditioned problems [7]. The current problem is badly conditioned due to the high resolution of the grid on which the object intensities are reconstructed. This grid should contain all spatial frequencies present in the MR images. Thus, it will most likely also contain spatial frequencies that are (almost) not sampled by any of the MR images. As there is no information about these (high) spatial frequencies, it is best to force the amplitude of them towards zero by adding the power in the (high) frequencies to the minimization criterium. In this work, this is achieved with the square of the discrete second derivative of the reconstructed \mathbf{o} :

$$\mathbf{o}^T \mathbf{K} \mathbf{o} = \left(\frac{\partial^2 \mathbf{o}}{\partial x_1^2} \right)^2 + \left(\frac{\partial^2 \mathbf{o}}{\partial x_2^2} \right)^2 + \left(\frac{\partial^2 \mathbf{o}}{\partial x_3^2} \right)^2. \quad (7)$$

The strength of the regularization is controlled by the variable λ in Eq. (6). Increasing λ increases the bias in the solution, while reducing the variance. A good value for λ is a value that (approximately) minimizes the Root Mean Square Error (RMSE).

For realistic image dimensions, the matrices in the general solution given in Eq. (6) are too large to actually store, even as sparse matrices. Furthermore, the solution of the linear system with QR or LU decomposition would consume way to much computation time for any realistic size of images. Therefore, the conjugated gradients method [8,9], which is an efficient iterative method to solve linear systems, was used to obtain the solution.

2.5 Affine Transformation

It is possible to evaluate the matrix vector multiplications that are required for the conjugated gradient method explicitly. However, this will require a large amount of computation time. When the transform T_j is an affine transform, or a subset of an affine transform, such as a combination of translation, scaling, and rotation, the acquisition of the MR images can be reformulated as an affine transform of the object o , followed by subsequent filter operations with the three orthogonal sampling functions w_i .

An affine transform of the initial coordinate \mathbf{x} to the final coordinate \mathbf{y} is given by

$$\mathbf{y} = \mathbf{T}\mathbf{x} + \mathbf{C}, \quad (8)$$

where \mathbf{T} ($n \times n$) specifies the affine transformation matrix in n dimensions, and \mathbf{C} ($n \times 1$) specifies the translation part of the transformation. Then, the affinely transformed image \tilde{o} of a continuous image o is given by

$$\tilde{o}(\mathbf{x}) = o(\mathbf{T}\mathbf{x} + \mathbf{C}). \quad (9)$$

This transform of a continuous image specifies the ideal transform that we try to approximate with the discretely sampled images. Different affine transformation methods exist. The most common methods interpolate the source image,

with linear interpolation or (approximate) sinc interpolation. However, for multi-dimensional images, a better transformation method is given below.

In order to properly evaluate the different transformation methods, both the spatial and the frequency domain should be studied. This is easy since an affine transform of an image also causes an affine transformation and a frequency dependent phase shift in the frequency domain:

$$\tilde{\mathcal{O}}(\mathbf{f}) = \mathcal{O}(\mathbf{f}\mathbf{T}^{-1})e^{i\mathbf{f}\mathbf{C}}, \tag{10}$$

where $\mathcal{O} = \mathcal{F}(o)$, $\tilde{\mathcal{O}} = \mathcal{F}(\tilde{o})$ are the Fourier transformed o and \tilde{o} , and \mathbf{f} ($1 \times n$) is a position in the frequency domain. Obviously, the spatial transformations induced by a transformation method should be as specified by the affine transform. However, this is the case for all methods that are discussed and displayed in Fig. 1, so it cannot be used for discrimination. Also, the magnitudes of the spatial frequencies should not be modified by the transformation method. Linear interpolation fails in this respect. An stronger discriminating feature is aliasing. The energy in spatial frequencies of the source image that are aliased to different (thus: wrong) spatial frequencies should be minimal. The aliasing in multi dimensional ($n > 1$) images might arise from three subtly different sources. First, spatial frequencies which are below the Nyquist frequency in the continuous source image o , might be transformed to above the Nyquist frequency in the destination image \tilde{o} . Secondly, spatial frequencies which are above the Nyquist frequency in the source image o , might be transformed to below the Nyquist frequency in the destination image \tilde{o} . Thirdly, spatial frequencies that are above the Nyquist frequency in both the continuous source and destination images might alias to below the Nyquist frequency of the destination image due to the sampling of this image.

The new transformation method proposed in this paper applies the affine transform T_f efficiently by splitting it into a **S**et of **S**Hear transformations (SSH). The shear transformations $\tilde{\mathbf{T}}_j$, $j \in \{1, \dots, 2n\}$ satisfy $\mathbf{T}_f = \prod_{j=1}^{2n} \tilde{\mathbf{T}}_j$. Each $\tilde{\mathbf{T}}_j$ differs from the identity matrix only in row d_j , so the image is only deformed along the d_j^{th} main axis, allowing efficient interpolation with a 1D low pass filter. An optimal set of these $\tilde{\mathbf{T}}_j$ can be constructed with:

- Initialize all $\tilde{\mathbf{T}}_j$ to the $(n + 1) \times (n + 1)$ identity matrix.
- By computing the total cost of the transformations defined in the following steps, search for the optimal permutation p of the numbers $\{1, \dots, n\}$.
- Initialize the remaining transform to the inverse transformation: $\mathbf{R}_0 = \mathbf{T}_f^{-1}$
- for $j = 1 \dots n$:

$$(\tilde{\mathbf{T}}_j)_{p_j, p_{1 \dots j}} = \left(((\mathbf{R}_{j-1})_{p_{1 \dots j}, p_{1 \dots j}})^{-1} \right)_{j, 1 \dots j} u_j \tag{11}$$

$$\mathbf{R}_j = \tilde{\mathbf{T}}_j \mathbf{R}_{j-1}, \tag{12}$$

- for $j = n + 1 \dots 2n$

$$(\tilde{\mathbf{T}}_j)_{p_{j-n}, 1 \dots n+1} = (\mathbf{R}_{j-1}^{-1})_{p_{j-n}, 1 \dots n+1} \tag{13}$$

$$\mathbf{R}_j = \tilde{\mathbf{T}}_j \mathbf{R}_{j-1}. \tag{14}$$

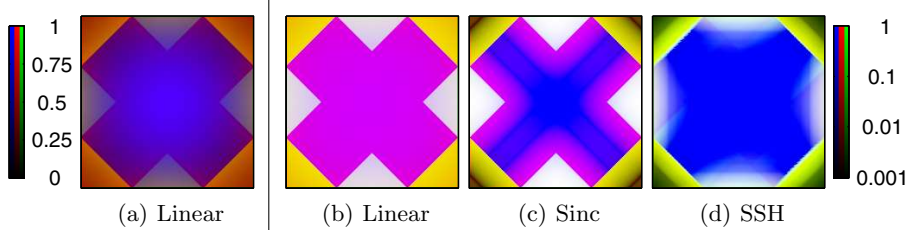


Fig. 1. Elementary frequency cell (zero spatial frequency at center, Nyquist frequency at edges) after rotating over 45 degrees with several methods. Colors: non constant **Blue** indicates spectral distortions. white: First type of aliasing, **Yellow**: Second type of aliasing, **Purple**: Third type of aliasing. Ideal transformation is **Blue** octagon and black triangles in the corners. Subfigure (a) and (b) display the result of the linear interpolator, scaled differently. Of the methods considered, linear interpolation is the only that has significant spectral distortions. (c) Sinc interpolation with a 11×11 neighborhood. (d) Rotation when the transform is split in a set of shear transforms. The length of the filters of (d), which is 15, was chosen such that the number of computations was (approximately) equal to (c).

- Find the lowest up-sampling factors u_j such that no distortion or aliasing is present in the final image and derive the cutoff frequency of the low-pass filter from the maximum frequency that needs to be transferred to the output image.

See Fig. 1 for a comparison of this SSH method to linear interpolation and a windowed sinc interpolation, when rotating a 2D image over 45 degrees. Note that, for 2D images, with the same computational complexity, the quality of the SSH method is substantially better than windowed sinc interpolation. When transforming 3D instead of 2D images, this method requires only approximately 50% more computations per voxel, which compares favorably to sinc interpolation where the number of computations is multiplied by the number of samples in the windowed sinc function.

3 Experiments

Several datasets of a bird were recorded with a Bruker small animal scanner. The resulting MR images were $192 \times 192 \times 32$ with voxel dimensions $0.125mm \times 0.125mm \times 0.75mm$. MR images in $N = 36$ different orientations were recorded and the reconstructed volume spanned the whole head in a volume of $21mm \times 21mm \times 22mm$ with isotropic voxel dimensions $0.1mm \times 0.1mm \times 0.1mm$.

As preprocessing step, the MR images were aligned by computing the required translations from the projections of the MR images to the object space,

$$\mathbf{b}_j = \mathbf{X}_j^T \mathbf{S}_j. \quad (15)$$

Since the images are assumed to have the same contrast, the optimal translation between two images was computed with a mean square difference measure,

$$\hat{\Delta}_{j,k} = \arg \min_{\Delta_{j,k} \in \mathcal{R}^3} \sum_{\mathbf{x}} (b_j(\mathbf{x}) - b_k(\mathbf{x} + \Delta_{j,k}))^2. \quad (16)$$

The position of each image j was adjusted by $1/N \sum_{k=1}^N \hat{\Delta}_{j,k}$.

After the alignment of the images, the conjugated gradient method with affine transformation with the SSH method was used to approximately solve Eq. (6). In our implementation each iteration of the conjugated gradient method took approximately 5m:35s of CPU time on one core of a Intel Core 2 Quad CPU @ 3.0 GHz, with 8GB of RAM.

4 Results

One of the original MR images and the high resolution reconstructed image are shown in Fig. 2. The 3D MR image is displayed in the read-phase, read-slice, and phase-slice directions, which, for this image, coincide with the coronal, sagittal and transversal directions in which the high resolution reconstructed image is shown. Note the substantially lower resolution of the MR image in the

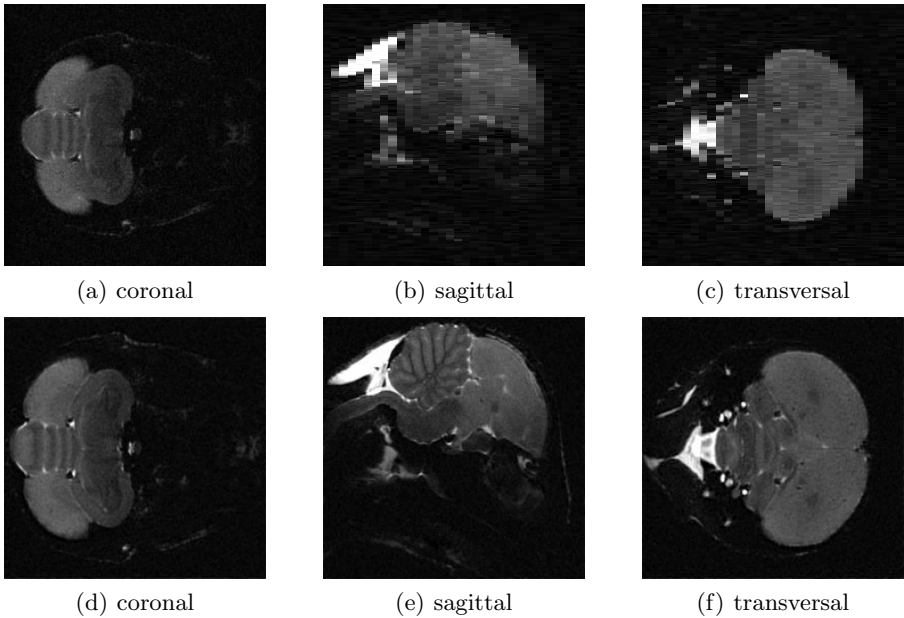


Fig. 2. (a), (b), and (c) show one of the 36 original MR images. (d), (e), and (f) show the high resolution reconstructed image.

slice direction, especially when compared to the high resolution reconstructed image. Also, observe the factor 2 reduction of the noise standard deviation. Furthermore, note the improved detail in the coronal views, which is due to the reduced blurring in the slice direction.

The conjugated gradient method converged quickly. After 15 iterations the update to \mathbf{o} was approximately 0.001 of the magnitude of \mathbf{o} , which is below the noise magnitude, indicating that 15 iterations are sufficient. Note that this number of iterations needed for convergence is substantially lower than the approximately 1000 iterations required by the methods in [6]. This is due to the high rate of convergence of the Conjugated Gradient method.

5 Conclusion

In this paper, a method was developed by which a high resolution isotropic image can be reconstructed from a set of anisotropic multi-slice MR images, recorded with different slice orientations. In contrast to previous reconstruction methods, this new method does not constrain the slice orientations. The reconstruction method uses an improved model of the MR acquisition, but does not require any prior knowledge about the imaged object. The high resolution image of the object is accurately reconstructed by the conjugated gradient method in a small number of iterations, substantially less than previous methods. The experiments show that the quality of the reconstructed isotropic image is substantially better, both in resolution and SNR, than any of the original MR images.

References

1. Zimmerman, R.A., Gibby, W.A., Carmody, R.F.: Neuroimaging: clinical and physical principles. Springer, New York (2000)
2. Peled, S., Yeshurun, Y.: Superresolution in MRI; application to human white fibre vizualization by diffusion tensor imaging. *Magn. Reson Med.* 45, 29–35 (2001)
3. Carmi, E., Liu, S., Alon, N., Fiat, A., Fiat, D.: Resolution enhancement in MRI. *Magn. Reson Imaging* 24, 133–154 (2006)
4. Scheffler, K.: Superresolution in MRI? *Magn. Reson Med.* 48, 408 (2002)
5. Greenspan, H., Oz, G., Kiryati, N., Peled, S.: MRI inter-slice reconstruction using super-resolution. *Magn. Reson Imaging* 20, 437–446 (2002)
6. Shilling, R.Z., Robbie, T.Q., Bailloeul, T., Mewes, K., Mersereau, R.M., Brummer, M.E.: A super-resolution framework for 3-D high-resolution and high-contrast imaging using 2-D multislice MRI. *IEEE T. Med. Imaging* 28(5), 633–644 (2009)
7. Engl, H.W., Hanke, M., Neubauer, A.: Regularization of inverse problems. Kluwer Academic Publishers, Dordrecht (2000)
8. Hestenes, M.R., Stiefel, E.: Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Stand* 49(6), 409–436 (1952)
9. Shewchuk, J.R.: An introduction to the conjugate gradient method without the agonizing pain. Technical report, Carnegie Mellon University, Pittsburgh, PA, USA (1994)

In-Vivo Estimates of Axonal Characteristics Using Optimized Diffusion MRI Protocols for Single Fibre Orientation

Torben Schneider¹, Claudia A.M. Wheeler-Kingshott¹,
and Daniel C. Alexander²

¹ NMR Unit, Department of Neuroinflammation, UCL Institute of Neurology,
University College London, UK

T.Schneider@ion.ucl.ac.uk

² Centre for Medical Image Computing, Department of Computer Science, University
College London, UK

Abstract. This work presents diffusion MR protocols that allow estimation of axonal parameters like diameter and density in the live human brain. Previous approaches demand very high field experimental systems or suffer from long acquisition times and are therefore impractical for use in clinical studies. We propose a method that significantly reduces scan time by making use of the a-priori known fibre orientation in structures with well defined single fibre (SF) organisation like the corpus callosum (CC) and produces protocols that can be performed in under 25 minutes on a standard clinical system. Results from a computer simulation experiment show that our SF protocols can generate parameter estimates with similar precision to previously proposed orientation invariant (OI) protocols. Furthermore, we acquire the 20 minute long SF protocol and the 1 hour long OI protocol in a scan/rescan study on two healthy subjects and compare the axonal parameter maps from both protocols.

1 Introduction

Diffusion weighted imaging (DWI) is an MRI technique that is sensitive to the random motion of water molecules and thus provides an insight into the microstructural properties of biological tissue. In recent years, numerous studies have demonstrated the great potential of DWI measures for the investigation of pathological and developmental changes of the human central nervous system. One of the most commonly used DWI method is diffusion tensor imaging (DTI) [1]. DTI allows the computation of useful indices, e.g., mean diffusivity or anisotropy of diffusion in each voxel but those measures are unspecific to individual microstructural characteristics.

Recently, Assaf et al [2] developed the AxCalibre model of cylindrical axons with gamma distributed radii to estimate axon diameter distributions in white matter tissue. The method is validated in in-vitro optic and sciatic nerve samples and estimated parameters show good correlation with corresponding histology. Barazany et al [3] applies the AxCalibre approach to image axon size distributions

in the corpus callosum of live rat brain. However, scan times are long and the high 7T magnetic field and maximum gradient strength (400 mT/m) are impossible to achieve on a live human scanners, that typically operate at 1.5-3T and can achieve maximum gradient strength between 30-60 mT/m. Alexander et al [4] demonstrates measurements of axon diameter and density in the live human brain on a standard clinical scanner using a simplified AxCalibre model and multi shell high angular resolution diffusion imaging (HARDI). The method relies on an experiment design optimization [5] to achieve sensitivity without the need of high gradient strengths and long acquisition times. The use of HARDI permits orientation invariant (OI) estimation of microstructure parameters over large areas of the brain with a range of different fibre orientations. However, with scan time close to an hour this approach is still impractical for clinical studies. Furthermore, with increasing scan time subjects are more likely to move during the acquisition, which results in a loss of signal quality and decreases the accuracy of parameter estimation.

Our aim in this study is to reduce total scan time to a clinical feasible maximum of 20-25 minutes while maintaining accuracy of the parameter estimation. To attain the reduction we discard the requirement for orientational invariance and focus on structures with known fibre orientation such as the corpus callosum similar to earlier studies [2,3]. We modify the existing protocol optimization framework [5] to incorporate this a-priori information about fibre organisation and define single fibre (SF) protocols that measure only the diffusion signal perpendicular and parallel to the known fibre orientation. We use computer simulations to compare the SF and the OI approach and perform a scan/rescan experiment on two healthy volunteers to investigate feasibility of estimating microstructural parameters in-vivo.

2 Methods

2.1 Tissue Model

As in Alexander et al [4], we model diffusion in a white matter voxel as a combination of water particles trapped inside three different compartments:

1. Intra-axonal water experiencing diffusion restricted inside cylindrical axons with equal radius R as in [8]
2. Extra-axonal water that is hindered due to the presence of adjacent axons. Diffusion is approximated by a diffusion tensor, with parallel diffusion coefficient d_{\parallel} in the direction of the cylinders and symmetric diffusion d_{\perp} in the perpendicular directions.
3. Water that experiences unhindered diffusion, e.g., in the cerebro-spinal fluid (CSF), modeled by an isotropic Gaussian distribution of displacements with diffusion coefficient d_I .

To reduce the number of free parameters in the model, we express d_{\perp} by using the tortuosity model of Szafer et al [9].

The diffusion weighted MR signal is measured by the pulsed gradient spin echo (PGSE) method, which has the following free parameters: diffusion sensitising gradient vector \mathbf{G} of strength $|\mathbf{G}|$ and duration δ , and diffusion time Δ . We combine the analytic expressions for the PGSE signal of each compartment $S_i(\mathbf{G}, \delta, \Delta)$ with i in $1, \dots, m$ to approximate the total signal S in each voxel by:

$$S(\mathbf{G}, \delta, \Delta) = S_0 \sum_{i=1}^3 f_i S_i(\mathbf{G}, \delta, \Delta), \tag{1}$$

where S_0 is the MR signal with no diffusion weighting and f_i is the volume fraction of each tissue compartment with $0 \leq f_i \leq 1$ and $\sum_{i=1}^3 f_i = 1$.

2.2 Protocol Optimisation

The aim of this work is to use a model-based imaging approach in each voxel to fit a set of MR signals acquired with different combinations of PGSE parameters to the tissue model parameters. A minimum of 6 different diffusion MR signals are required to estimate all model parameters but usually we acquire more signals to overdetermine the solution and add noise control. We define a set of PGSE parameters that acquires N different MR signals as protocol

$$\mathcal{P} = \{\mathbf{G}_1, \delta_1, \Delta_1 \dots, \mathbf{G}_N, \delta_N, \Delta_N\}. \tag{2}$$

As mentioned above, the aim of the optimisation algorithm [5] is to find the protocol \mathcal{P} , that allows the most accurate estimation of the tissue model parameters under given hardware and time constraints. The Fisher information matrix (FIM) provides a lower bound on the inverse covariance matrix of parameter estimates, i.e., the \mathcal{P} that maximizes the FIM will maximize the precision of those estimates. We use the d-optimality criterion [6], which is defined as the determinant of the inverse FIM of protocol \mathcal{P} and tissue model parameters θ :

$$D(\theta, \mathcal{P}) = \det[(\mathbf{J}^T \boldsymbol{\Omega} \mathbf{J})^{-1}], \tag{3}$$

where \mathbf{J} is the $N \times 6$ Jacobian matrix with the $ijst$ element $\partial S(\mathbf{G}_i, \delta_i, \Delta_i) / \partial \theta_j$. In the original approach $\boldsymbol{\Omega} = \text{diag}\{1, \dots, 1\}$. Following [5], we use a stochastic optimization algorithm [7] that returns \mathcal{P}' with minimal D among all possible \mathcal{P} with respect to the given scanner hardware limits.

In the original approach, N is chosen to account for the acquisition time limit. The acquisition is divided in M sets of different PGSE settings with gradient directions in each set being fixed. However, in this approach a decrease of the total number of acquisitions N must reduce the angular resolution of gradient direction sampling, which will increase the uncertainty in fibre direction estimates and thus tissue parameter measures. The OI approach also requires that N and M are known so that for every combination of N and M the full optimization algorithm has to be performed. Furthermore this method does not reward protocols that sample more important measurements more heavily.

In this work we introduce the asymptotic SF protocol optimisation. We focus on specific structures with known fibre orientation like the CC. This avoids the need for high angular resolution, which potentially dramatically reduces the number of required measurements. As in [23], we constrain most measurements in the protocol to have gradient direction perpendicular to the fibre bundles, but we include one measurement in the parallel direction for the estimation of diffusivity along the axons. We extend the algorithm to optimize M different PGSE settings and include $\Omega = \text{diag}\{w_1, \dots, w_M\}$ with $\sum_{m=1}^M w_m = 1$ in the optimisation of Eq. 3. The weighting factors w_m reflect how important each measurement is, i.e. how often it should be sampled relative to the other measurements. For any given N we can calculate the number of measurements for each element of \mathcal{P} by $N_m = w_m N$, hence the resulting protocols are independent of N .

2.3 Model Fitting

We use the three stage fitting algorithm as described by Alexander et al [4], to fit the tissue model to the acquired MR signal in each voxel. We increase stability by fixing d_{\parallel} to $1.7 \cdot 10^{-9} m^2 s^{-1}$ and d_{\perp} is fixed to $3.0 \cdot 10^{-9} m^2 s^{-1}$ [23,4]. The objective function is defined as the maximum likelihood of model parameters given the observed MR signals under Rician noise ($\sigma = 0.05$). An initial estimation is found using a coarse grid search algorithm over a set of physiologically possible parameters. Then a gradient descent algorithm further refines the parameter estimates. Finally a Markov Chain Monte Carlo (MCMC) algorithm with a burn-in of 2000, 50 samples at an interval of 200 provides posterior distributions of the parameters f_1 , f_2 and the axon radius r . An average over the MCMC samples provides the final parameter estimates. We report the axon diameter index $a = 2r$ and the axon density index $\rho = 4f_1\pi^{-1}r^{-2}$.

3 Experiments and Results

We motivated our protocol optimization by the aim to produce protocols that can be performed on a typical human scanner in a clinically feasible time. To validate our approach, we generate optimized protocols for a clinical 3T Philips Achieva scanner with a maximum gradient strength of $|\mathbf{G}_{max}| = 60 mT/m$. We use the asymptotic optimization to generate SF protocols that can be performed in 20 minutes with a total number of acquisitions $N = 90$ (SF₉₀). For comparison with previous studies, we also generate an OI protocol using $N = 360$ (OI₃₆₀) and an SF protocol with the same number of acquisitions (SF₃₆₀). The resulting protocols are presented in table 1. For the SF₉₀ and SF₃₆₀ protocols we set $M = 8$ but only sequences with $w > 0$ are reported. The OI₃₆₀ protocol optimisation uses $M = 4$ and report the three unique PGSE parameter settings.

3.1 Simulations

We use the free diffusion simulation of Hall and Alexander [10], which performs a Monte Carlo (MC) simulation of water particles in packed cylinders. We use

Table 1. PGSE settings of SF₉₀, SF₃₆₀ and OI₃₆₀ protocols. ⊥ and ∥ mark acquisitions perpendicular and parallel to the fibre bundles.

(a) SF ₃₆₀ and SF ₉₀ protocols						(b) OI ₃₆₀ protocol					
	Δ	δ	G	b		Δ	δ	G	b		
N_m	[ms]	[ms]	[mT/m]	[s/mm ²]		N_m	[ms]	[ms]	[mT/m]	[s/mm ²]	
70	18	0	0	0		71	0	0	0	0	
72	17	33.0	14.5	36.8	550	∥	101	19.2	11.7	60.0	540
38	10	22.4	15.9	60.0	1114	⊥	107	38.2	12.5	47.8	870
45	11	29.3	22.8	60.0	2908	⊥	81	29.1	21.6	60.0	2634
68	17	48.0	26.6	43.7	3666	⊥	360				
67	17	40.5	34.0	60.0	8692	⊥					
360 90											

the 44 synthetic white matter substrates from Alexander et al [4] with diameter distributions and packing densities similar to previously reported histology studies. We perform the MC simulation with 50000 walkers and 20000 time steps for each protocol. For each substrate we generate 10 sets of noise-free MR signals and add Rician noise of $\sigma = 0.05$, resulting in total of 440 sets of noisy MR signals. For each protocol we apply the model fitting procedure to the 440 sets of MR signals and retrieve the tissue model parameters.

To compare the axon distributions with the estimated axon diameter index a we have to take into consideration that the contribution of each axon to the MR signal depends its volume and is proportional to the square of its diameter. As in [4] we correlate the estimated axon diameter index a with the weighted axon diameter average $\hat{a} = \hat{f} / \int p(\alpha)\alpha^3 d\alpha$, where p is the true distribution of axon diameter α and \hat{f} is the intracellular volume fraction $\hat{f} = \int p(\alpha)\alpha^2 d\alpha$.

3.2 MRI Experiment

The SF₉₀ and OI₃₆₀ protocols (see table I) are implemented on a 3T Philips Achieva scanner to test the clinical viability of the 20 minute SF₉₀ protocol and compare it to the three times longer OI₃₆₀ protocol. Diffusion weighted MR images of two healthy volunteers (male 32yo, female 25yo) are acquired using a cardiac-gated EPI sequence with the following imaging parameters: 10 slices, slice thickness=5mm, in-plane resolution=128x128 (FOV=35x35mm²), TR=7RR, TE=125ms/TE=100ms for SF₉₀ and OI₃₆₀ respectively. We position the centre slice so that it is aligned with the mid-sagittal body of the CC to be able to acquire DWI measurements perpendicular and parallel to the fibres of the CC. SF₉₀ acquisition is repeated twice on two separate days for each subject to investigate the reproducibility of the estimated parameter maps.

3.3 Results

Figure I presents the results from fitting the model to the synthetic MC data sets as described above. For all three protocols we plot the fitted axon diame-

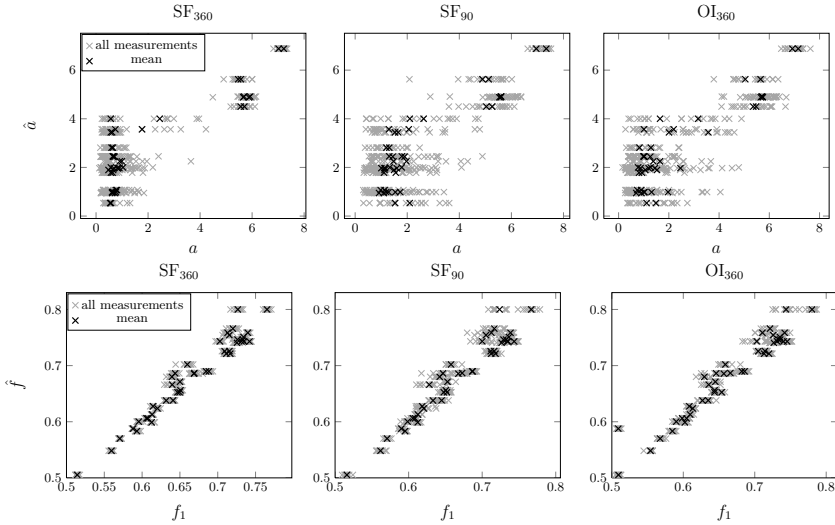


Fig. 1. Scatter plots of estimated tissue model parameters a and f_1 (grey) and mean a and f_1 over 10 replications (black) against true \hat{a} and \hat{f}_1 of the MC substrates

ter index a against \hat{a} and the intra-cellular volume fraction f_1 against the true intra-cellular volume fraction \hat{f}_1 for all 440 noisy sets of MR signals. We also compute the mean over the 10 replications for each of the 44 unique substrates and display them in the same plot. The bottom row of Fig. 1 shows that all protocols estimated the volume fraction accurately with little variance. Further, all protocols estimate larger a that agree with \hat{a} . The estimated a varies arbitrarily between $0 - 2\mu\text{m}$ for $\hat{a} \lesssim 3\mu\text{m}$. Thus smaller \hat{a} can be distinguished from larger ones but not accurately measured. This is because the limited maximal gradient strength that does not attenuate the signal from water inside axons of diameter $< 2\mu\text{m}$. Despite the limitation, the trends of a agree with the true values for \hat{a} and suggest that the index a is a useful discriminator of axon diameter distributions. SF₃₆₀ estimates both indices more accurately than OI₃₆₀ and variations among the 10 estimates in each substrate are smaller. SF₉₀ and OI₃₆₀ appear to have similar accuracy and precision in estimating \hat{a} and \hat{f}_1 . This suggests that we can reduce by a third by exploiting a-priori known fibre orientation while maintaining similar quality of parameter estimates.

Figure 2 shows maps of a and ρ in the centre slice of the CC for all acquisitions in two volunteers. From previous histological studies [11] we expected low axon diameter and high density in the splenium and genu and higher axon diameters with lower density in the body of the CC. As predicted by the MC simulations (see also [4]), all protocols overestimated a because of the lack of sensitivity to lower diameters. The high-low-high trend in a and low-high-low trend of ρ can be observed in both subjects in OI₃₆₀ results but are less apparent in SF₉₀ scans. The worst case is SF₉₀ of subject 1, which presents very noisy parameter maps. This is likely to be caused by a misalignment with the true

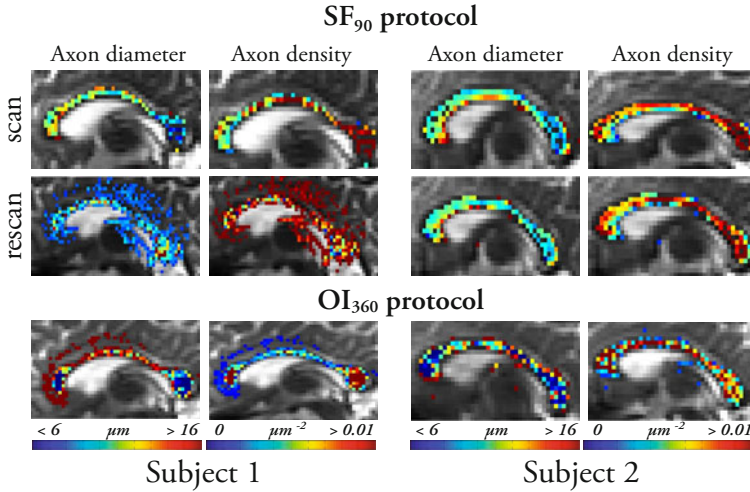


Fig. 2. Color coded parameter maps of a and ρ in the centre slice of the CC in two subjects. Scan and rescan results for the SF₉₀ are shown together with results from the OI₃₆₀ acquisition.

fibre direction of the CC and the gradient directions, which demonstrates the sensitivity of the SF protocol to accurate positioning. Furthermore, all SF scans consistently produce larger estimates of a than OI₃₆₀. Variation in true fibre orientation is again the likely explanation. Unlike the SF protocols, the OI protocol can better compensate for this variation because of the high angular gradient sampling. However, despite the limitations, the results of subject 2 demonstrate reproducible estimates of a and ρ . This suggests that with accurate positioning, the 20 minute SF₉₀ protocol is able to produce comparable parameter maps to OI₃₆₀, which requires more than three times the scan time.

4 Conclusion

In this work we propose optimized diffusion MRI protocols that use the known fibre orientation in specific structures like the CC and allow us to estimate indices of axon diameter and density in the live human brain. We develop a new optimization algorithm that overcomes several limitations of previous approaches and produces DWI protocols that can be acquired in under 20 minutes. While previous protocols were too time consuming for clinical practise, the short acquisition time of our protocols opens the possibility to be included in a variety of studies. Experiments on synthetic data show that our protocols can provide axon diameter and density indices with similar variance to those from longer orientational invariant protocols. In-vivo scans on two healthy volunteers show the potential of our method to produce parameter maps of axon diameter and density that agree with the general histologic trend but also reveal the limitations caused by misalignment and variation in fibre orientation compared to the

longer OI protocol. If such protocols are to be used, great care must be taken to align gradient directions with the fibre orientation. Future work aims to account for uncertain or erroneous fibre orientation by incorporating some tolerance for fibre orientation variation in the optimisation.

Acknowledgements. Thanks to Dr Hubbard for support in implementing the protocols. TS is sponsored by the ISRT. The NMR Unit at the UCL Institute of Neurology is supported by the MS society of Great Britain and Northern Ireland. DCA is funded by EPSRC EP/E007748 and the EU CONNECT consortium www.brain-connect.eu.

References

1. Basser, P., Mattiello, J., LeBihan, D.: MR diffusion tensor spectroscopy and imaging. *Biophys. J.* 66 (1994)
2. Assaf, Y., Blumenfeld-Katzir, T., Yovel, Y., Basser, P.: AxCaliber: A Method for Measuring Axon Diameter Distribution from Diffusion MRI. *MRM* 59 (2008)
3. Barazany, D., Assaf, Y., Basser, P.: In vivo measurement of axon diameter distribution in the corpus callosum of rat brain. *Brain* 132 (2009)
4. Alexander, D.C., Hubbard, P.L., Hall, M.G., Moore, E.A., Ptito, M., Parker, G.J.M., Dyrby, T.B.: Orientationally invariant indices of axon diameter and density from diffusion MRI. *NeuroImage* (in press, 2010)
5. Alexander, D.C.: A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features. *MRM* 60 (2008)
6. O'Brien, T.E., Funk, G.M.: A gentle introduction to optimal design for regression models. *Am. Stat.* 57 (2003)
7. Zelinka, I., Lampinen, J.: SOMA - self-organizing migrating algorithm. In: 6th International Conference on Soft Computing (2002)
8. Gelderen, P.V., Despres, D., Zijl, P.C.M.V., Moonen, C.T.W.: Evaluation of restricted diffusion in cylinders phosphocreatine in rabbit leg muscle. *JMR* 103 (1994)
9. Szafer, A., Zhong, J., Gore, J.: Theoretical model for water diffusion in tissues. *MRM* 33 (1995)
10. Hall, M.G., Alexander, D.C.: Convergence and parameter choice for Monte-Carlo simulations of diffusion MRI. *IEEE TMI* 28 (2009)
11. Aboitiz, F., Scheibel, A., Fisher, R., Zaidel, E.: Fiber composition of the human corpus callosum. *Brain Research* 598 (1992)

Diffusion-Based Population Statistics Using Tract Probability Maps

Demian Wassermann¹, Efstathios Kanterakis², Ruben C. Gur³,
Rachid Deriche¹, and Ragini Verma²

¹ Athena project-team, INRIA Sophia Antipolis-Méditerranée, 2004 rt des Lucioles, 06902, FR

² Section of Biomedical Image Analysis, Radiology, UPENN, PA 19104, USA

³ Department of Psychiatry, University of Pennsylvania Medical Center, PA 19104, USA

Abstract. We present a novel technique for the tract-based statistical analysis of diffusion imaging data. In our technique, we represent each white matter (WM) tract as a tract probability map (TPM): a function mapping a point to its probability of belonging to the tract. We start by automatically clustering the tracts identified in the brain via tractography into TPMs using a novel Gaussian process framework. Then, each tract is modeled by the skeleton of its TPM, a medial representation with a tubular or sheet-like geometry. The appropriate geometry for each tract is implicitly inferred from the data instead of being selected *a priori*, as is done by current tract-specific approaches. The TPM representation makes it possible to average diffusion imaging based features along directions locally perpendicular to the skeleton of each WM tract, increasing the sensitivity and specificity of statistical analyses on the WM. Our framework therefore facilitates the automated analysis of WM tract bundles, and enables the quantification and visualization of tract-based statistical differences between groups. We have demonstrated the applicability of our framework by studying WM differences between 34 schizophrenia patients and 24 healthy controls.

1 Introduction

Diffusion tensor imaging (DTI) has become of modality of choice for studying the white matter (WM) pathology due to its unique *in vivo* characterization of WM microstructure [1]. A wide range of paradigms have been developed for studying group differences between healthy subjects and a population with pathology using DTI. Of these, the tract-based spatial statistics (TBSS) [2] made a break from voxel-wise techniques by performing statistics on the skeletonization of the full WM [3]. This technique has rapidly gained popularity for finding group differences in WM. However, TBSS relies on a whole-brain approach disregarding anatomical specificity. Recently, techniques tried to alleviate this problem by analysing WM tracts individually by representing each tract as a lower-dimensional structure, using tubular geometries [3,4,5] or sheet-like parametrizations [6,7]. However, there is an issue in these tract-specific techniques: the need to *a priori* choose between a tubular or a sheet-like representation of the WM tracts which introduces an artificial bias as some fasciculi as inferior-fronto

¹ In this work we follow the same definition of skeleton as in Smith et al. [2].

occipital have a spreading structure which is not accurately represented by a tube and the cingulum are tube-like and not well represented by a sheet.

In this paper, we propose a technique to perform tract-specific statistical analyses overcoming the previously mentioned limitations. For this, we model each WM tract as a function mapping any point in space to the probability of its belonging to such a tract. This model is called a tract probability map (TPM) and can be robustly calculated using a recently proposed mathematical framework for WM bundles based on Gaussian processes (GPs) [8]. Within this framework, we develop a skeleton-based statistical analysis [2,6,7] technique for finding differences among populations. In developing this technique, we present two contributions: first, a general approach for tract-based statistical analysis which is not bound to a sheet-or-tube representation; and second, the applicability of the mathematical framework for the segmentation of WM structures presented by [8] for statistical analysis.

2 Computing Tract Probability Maps

The first step in developing our statistical analysis technique is the calculation of tract probability maps for a set of WM tracts obtained from whole brain DTI-tractography [9,8]. DTI-tractography is a technique to trace pathways followed by axons in the WM, representing them as smooth three dimensional trajectories at super-voxel resolution [1]. These trajectories provide the means to work on elements which are more anatomically-oriented than isolated voxels and can be grouped into WM tracts like the cingulum or the fornix. We show examples of TPMs for WM tracts in fig. 1.

In order to calculate TPMs for WM tracts from these trajectories while preserving super-voxel resolution, we use a GP framework [8]. The advantages of this representation are two-fold: first, the TPMs calculated using GPs are continuous functions which can be sampled at any desired resolution; second, being a parametrical representation the GP framework allows us to work robustly on its parameter space instead of performing operations in image space. In the remainder of this section we detail the calculation of TPM for a WM tract: we start by describing the representation of a trajectory within

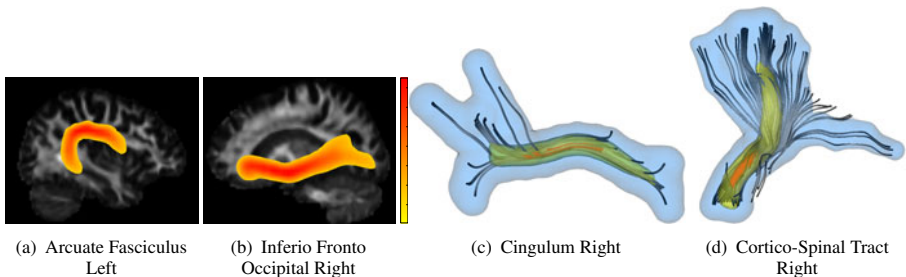


Fig. 1. Population-averaged tract probability maps for four WM tracts. These maps are shown over FA (a-b) and as iso-probability surfaces (c-d). The probability at each voxel is calculated using eq. (2). Color code for fig. (c-d) is as follows, Blue: $\text{TPM}(\mathbf{p}) = .01$, Yellow: $\text{TPM}(\mathbf{p}) = .2$, Red: $\text{TPM}(\mathbf{p}) = .6$.

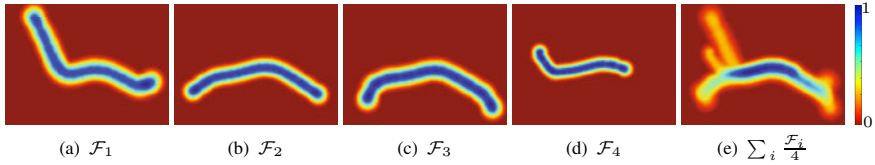


Fig. 2. Mean indicator function for four fiber tracts (a-d) and mean indicator function for the bundle formed by averaging them according to our framework (e). Color code ranges from blue when it is likely that a voxel belongs to the bundle of fibres to red when it does not belong.

the GP framework; next, we show how to go from the GP representation of a single trajectory to the representation of a bundle of trajectories; and finally, using the GP framework, we describe the calculation of the TPM for a bundle of trajectories.

We model each trajectory \mathcal{F} as a blurred indicator function $y(\cdot)$, such that if the three dimensional point \mathbf{p} belongs to \mathcal{F} , the value of $y(\mathbf{p})$ is 1 and it decays to 0 as \mathbf{p} is further away from \mathcal{F} . Several examples of blurred indicator for a trajectory are shown in fig. 2 (a-d). Formulating the hypothesis that the least curved a blurred indicator function is the better it represents a smooth trajectory, it is possible to model the indicator function $y(\cdot)$ for a trajectory \mathcal{F} as a GP [8]. Then, using the properties of the GPs, the value of $y(\cdot)$ for a trajectory can be characterized at each point in space \mathbf{p} as an univariate Gaussian: $y(\mathbf{p}) \sim \mathcal{G}(y^*(\mathbf{p}), \sigma^2(\mathbf{p}))$. Particularly, the mean and the variance of this univariate Gaussian are inferred from a sampling of \mathcal{F} , $\mathbf{f} = \{\mathbf{f}_1 \dots \mathbf{f}_N\} \subset \mathcal{F}$:

$$y^*(\mathbf{p}) = [C_{\mathbf{f}}(\mathbf{p})]^T C_{\mathbf{ff}}^{-1} \mathbf{1} \quad \text{and} \quad \sigma^2(\mathbf{p}) = c(\mathbf{p}, \mathbf{p}) - C_{\mathbf{f}}(\mathbf{p})^T C_{\mathbf{ff}}^{-1} C_{\mathbf{f}}(\mathbf{p}) \quad (1)$$

where $[C_{\mathbf{f}}(\mathbf{p})]_i = [c(\mathbf{f}_i, \mathbf{p})]_i$, $[C_{\mathbf{ff}}]_{ij} = [c(\mathbf{f}_i, \mathbf{f}_j)]_{ij}$, $\mathbf{1} = [1 \dots 1]^T$ and

$$c(\mathbf{p}, \mathbf{p}') := \psi(\|\mathbf{p} - \mathbf{p}'\|), \quad \psi(r) = \begin{cases} 2|r|^3 - 3Rr^2 + R^3 & r \leq R \\ 0 & r > R \end{cases}$$

Once we have represented single trajectories as GPs, obtaining the representation for a bundle of trajectories $\mathcal{B} = \{\mathcal{F}_1, \dots, \mathcal{F}_N\}$ is straightforward: the fact that at each point the value of the indicator function for each trajectory \mathcal{F} , $y_{\mathcal{F}}(\mathbf{p})$, has a univariate Gaussian distribution, enables us to calculate the mean and covariance of the Gaussian distribution characterizing the indicator function for \mathcal{B} at a point \mathbf{p} as [8]

$$y_{\mathcal{B}}^*(\mathbf{p}) = \frac{1}{N} \sum_{i=1}^N y_{\mathcal{F}_i}^*(\mathbf{p}) \quad \text{and} \quad \sigma_{\mathcal{B}}(\mathbf{p}) = \frac{1}{N^2} \sum_{\mathcal{F} \in \mathcal{B}} \sigma_{\mathcal{F}}^2(\mathbf{p}).$$

The result of the combination of the indicator functions representing various trajectories into a bundle is illustrated in fig. 2(e).

Finally, according to [8], the TPM for a WM tract formed by a bundle of trajectories \mathcal{B} is calculated using the GP framework as

$$\text{TPM}(\mathbf{p}) := \text{prob}\{\mathbf{p} \in \mathcal{B}\} \propto \mathbb{E}[y_{\mathcal{B}}(\mathbf{p}) = 1] = \frac{1}{2\sqrt{\pi(1 + \sigma_{\mathcal{B}}^2(\mathbf{p}))}}, \quad (2)$$

where $\mathbb{E}[\cdot]$ stands for the expected value of a random variable. Figure 1 illustrates the *tract probability map* calculated from a GP for a bundle using color-coded surfaces and probability maps overlaid on the FA image of the spatially normalized subject.

In the above, we have introduced our GP-based framework for representing white matter fiber bundles and its two operations: combination of fibers into a bundle and calculation of the TPM. These tools are fundamental to performing a tract-based statistical analysis of the WM, which we develop in the next section.

3 Skeleton-Based WM Tract Analysis

Our statistical analysis framework is designed to help identify precise differences in tracts. The algorithm to analyse a particular fibre bundle starts by generating its TPM as in section 2. By construction, the $\text{TPM}(\cdot)$ function has a ridge (or valley) of high probability at the centre of the tract that decays going outwards. Consequently, the main idea is to find a representation of the tract as a sheet or a line and to project diffusivity information on that representation. For this, we adapt the thinning algorithm of [2] which is used extensively in recent literature. However, our work can be easily restated using a different thinning algorithm. Then, we project a scalar diffusivity measure (e.g. the FA) on the skeleton and perform group analysis on it. This process is shown in fig. 4.

Skeleton Calculation. The first step in the algorithm is to obtain the skeleton. For this, we start by identifying the voxels contained in such a skeleton. We compute the centre of gravity (CofG) of a sphere of radius r around the voxel \mathbf{p} , $S(\mathbf{p}, r) \subset \mathbb{R}^3$,

$$\text{CofG}(\mathbf{p}; r) = \frac{\int_{\mathbf{q} \in S(\mathbf{p}, r)} \mathbf{q} \text{TPM}(\mathbf{q}) d\mathbf{q}}{\int_{\mathbf{q} \in S(\mathbf{p}, r)} \text{TPM}(\mathbf{q}) d\mathbf{q}}. \quad (3)$$

If the centre of gravity is sufficiently close to \mathbf{p} , it is reasonable to assume that \mathbf{p} is very close to the ridge or valley which we will call skeleton from now on. This is illustrated in fig. 3(a) where \mathbf{p} is a cross, $S(\mathbf{p}, r)$ a circle and the $\text{CofG}(\mathbf{p}, r)$ a square. Then, to select a voxel in the skeleton, we start from each voxel that is sufficiently close to its centre of gravity and move in a direction perpendicular to the skeleton until we reach a local maximum. The voxels containing such a maximum are considered to be in the skeleton.

As we have seen previously, in order to extract the skeleton, we need to calculate the direction that is perpendicular to it. If the voxels are sufficiently far away from it, this direction is provided by the gradient of the TPM. This is due to the fact that the probability of a voxel belonging to the underlying tract of the TPM decreases when we move away from the skeleton. However, if we are on the skeleton, meaning on a ridge or a valley, the gradient will have a 0 norm and no defined orientation. In these cases it is reasonable to assume that the gradient changes more rapidly in the direction perpendicular to the skeleton [10]. Thus, we take the eigenvector corresponding to the largest eigenvalue of the Hessian of the TPM, as the direction perpendicular to the skeleton. This is illustrated in fig. 3. Then, the direction perpendicular to the skeleton at \mathbf{p} is given by

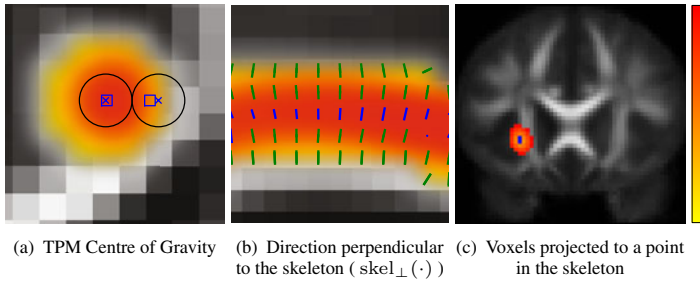


Fig. 3. Finding the direction perpendicular to the tract probability map skeleton. Fig. (a) the deviation of the TPM’s centre of gravity (square) from the centre of a neighbourhood (cross) indicates the distance from the skeleton. Fig. (b) when the voxel is far enough from the skeleton, the direction of the gradient (green) is perpendicular; when it is close enough, the gradient norm is close to zero, in these cases the principal direction of the Hessian (blue) indicates the perpendicular direction. Fig. (c) skeleton of a WM tract (blue) and the voxels which are closest to it.

$$\text{skel}_{\perp}(\mathbf{p}) := \begin{cases} \nabla \text{TPM}(\mathbf{p}) / \|\nabla \text{TPM}(\mathbf{p})\| & \text{if } \|\text{CofG}(\mathbf{p}; r) - \mathbf{p}\| > \epsilon \\ \mathbf{e}_1(\nabla^2 \text{TPM}(\mathbf{p})) & \text{if } \|\text{CofG}(\mathbf{p}; r) - \mathbf{p}\| \leq \epsilon \end{cases} \quad (4)$$

In this equation ϵ is taken to be a tenth of the voxel width and $\mathbf{e}_1(M)$ is the eigenvector corresponding to the largest eigenvalue of M , and ∇ and ∇^2 are the gradient and the Hessian operators. Moreover, due to the properties of our GP framework, the gradient and Hessian of the TPM (eq. (2)) can be calculated analytically.

Finally, it is important to emphasize that this is just one technique to calculate the skeleton. The statistical analysis algorithm is not dependent on it, but is general and can be applied to any skeleton calculated from a TPM. Most of the ridge and valley calculation algorithms are based on the gradient and Hessian of the analysed function (10), hence it is possible to reformulate them in terms of the GP framework for white matter bundles.

Voxel Projections and Statistical Analysis. Having created the skeleton for a tract, we project all the voxels within the thresholded TPM to their closest point on the skeleton. This is achieved by starting at every voxel and following the direction perpendicular to the skeleton at that voxel, calculated in eq. (4), until the skeleton is reached. Then, we calculate the expected scalar diffusivity measure at every point of the skeleton. In fig. 3(c), we show a point on the inferio-fronto occipital tract (IFO) skeleton in green and the voxels in the image which are closest to it. The values of these voxels are projected to the skeleton and averaged according to their probability of being in the bundle. This procedure for the IFO and the fractional anisotropy (FA), is as follows: for each point in the skeleton of IFO, \mathbf{s} , we calculated the set of voxels $\mathcal{N}(\mathbf{s})$ within the TPM having \mathbf{s} as the closest on the skeleton. This is done by following the direction perpendicular to the skeleton calculated in the previous section. Then, the expected value of FA on each skeleton voxel \mathbf{s} , is the expectation of the value of the FA on $\mathcal{N}(\mathbf{s})$ projected to \mathbf{s} :

$$E_{\text{FA}}^{\text{IFO}}(\mathbf{s}) = \frac{\int_{\mathcal{N}(\mathbf{s})} \text{TPM}^{\text{IFO}}(\mathbf{p}) \text{FA}(\mathbf{p}) d\mathbf{p}}{\int_{\mathcal{N}(\mathbf{s})} \text{TPM}^{\text{IFO}}(\mathbf{p}) d\mathbf{p}} \quad (5)$$

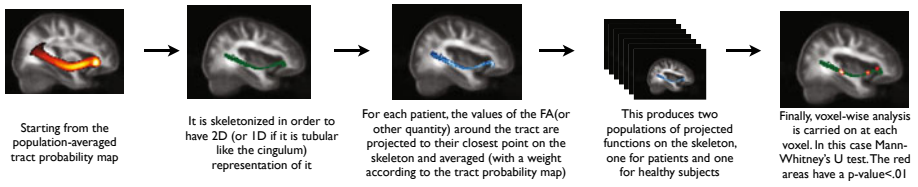


Fig. 4. Skeleton-based tract-specific statistical analysis on a WM tract

Finally, for each voxel in the skeleton, we perform voxel-based analysis. In this case we use the Mann-Whitney U test to find voxels where the distribution of the fractional anisotropy or the axial or radial diffusivities among patients and controls is different, and we report trends indicated by the analysed t-score.

4 Data

Whole-brain DWI datasets were acquired on 58 volunteers (35.60 ± 10.48 years, 28 Male, 24 healthy, 34 diagnosed with schizophrenia) on a Siemens Trio 3T scanner with $1.71 \times 1.71 \text{mm}^2$ in-plane resolution, 2mm slice thickness. Six unweighed B0 images and 64 diffusion weighted images ($b=1000 \text{s/mm}^2$) were acquired with non-collinear diffusion sensitizing gradients.

Data Preparation. DTI images for each subject were computed and deformably registered, using DTI-DROID [11], to a DTI atlas [9]. Full brain tractography was performed by seeding in sub-voxel resolution by taking every voxel with linear anisotropy higher than .3, dividing it into $0.25 \times 0.25 \times 0.25 \text{mm}^3$ sub-voxels and seeding from each sub-voxel. On average, 10,000 fibre tracts were obtained for each subject.

Identification of WM Tracts. In order to identify white matter tracts, the clustering algorithm presented by [8] was applied to every subject individually. *Population-averaged tract probability maps* for major tracts were obtained for each extracted white matter tract: on both hemispheres the arcuate, inferio-fronto occipital, uncinate fasciculi and cingulum bundle along with the fornix; also, the frontal and posterior forceps were extracted. In order to perform the statistical analysis, we extracted the skeleton of each *population-averaged tract*.

Statistical Analysis of Tracts. To characterize diffusion properties of the white matter structures, four scalar quantities derived from the diffusion tensors were used: Fractional Anisotropy (FA) [1], axial diffusivity ($\lambda_{\parallel} = \lambda_1$) and radial diffusivity ($\lambda_{\perp} = \frac{1}{2}(\lambda_2 + \lambda_3)$). The FA measure was used due to its biological interpretation as a measure of tract integrity and its popularity in the literature. We chose the other two measures because there are several reports relating differences on axial and radial diffusivity with differences in myelination and axonal calibre [12].

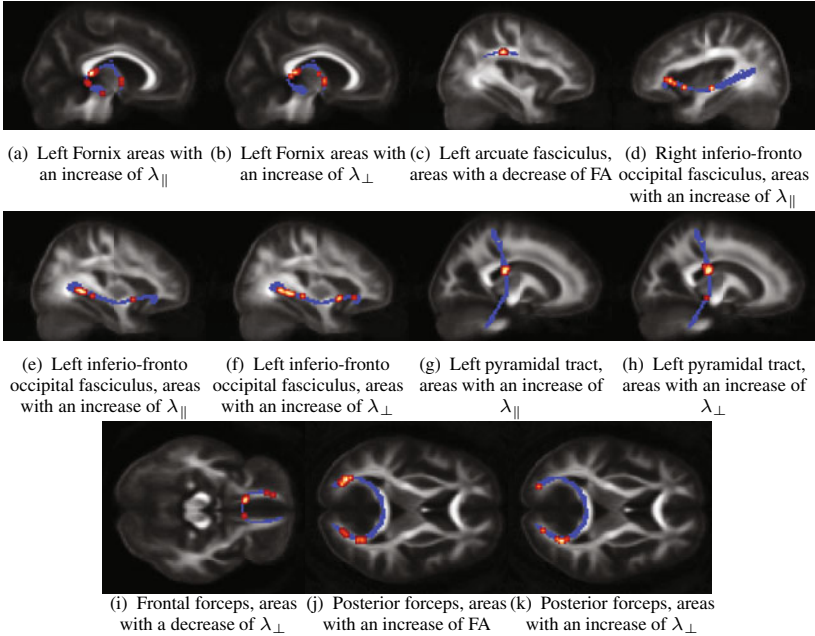


Fig. 5. Differences between healthy and schizophrenic populations in tracts obtained using the statistical analysis described in section 3. We show tracts where more than 10% of the number of voxels in the whole skeleton, shown in blue, have a p -value < 0.001 . For the regions defined by these voxels, shown in red, we report trend of the difference between patients and controls indicated by the t -statistic.

5 Results

Results of our skeleton-based analysis are shown in fig. 5. In this figure we show tracts where more than 10% of the number of voxels in the whole skeleton have a p -value < 0.001 . For the regions of the tracts defined by these voxels, we report the trend of the difference between patients and controls indicated by the t -statistic in order to characterize the detected change in diffusivity properties. In the left fornix, there are regions with an increase of λ_{\parallel} and λ_{\perp} ; the right arcuate fasciculus shows regions with a decrease of FA; the right inferio-fronto occipital fasciculus shows areas of decreasing λ_{\parallel} ; in the left inferio-fronto occipital fasciculus there are areas of increased axial and λ_{\perp} ; the frontal forceps shows areas of decreased λ_{\perp} ; the posterior forceps has areas of increased FA and λ_{\perp} ; and the left pyramidal tract exhibits an increase of λ_{\parallel} and λ_{\perp} .

6 Discussion

Our results are in agreement with current literature. The predominance of differences in tracts within the left frontal and temporal lobes has been reported by a recent meta-analysis of DTI studies in schizophrenia [13] and a detailed survey [14]. Particularly, the

decrease of the FA on the left arcuate fasciculus, shown by our analysis was reported by [15]. Regarding our studies analysing FA changes in tract skeletons, we found an increased FA in left IFO which was previously shown by [16] and decreased FA in the fornix is consistent with [17]. This agreement with literature demonstrates that our statistical analysis is suited for white matter studies, therefore encouraging us to carry on with a more comprehensive study on larger databases.

Regarding the analyses of parallel (λ_{\parallel}) and perpendicular (λ_{\perp}) diffusivities, even if recent work shows that these measures are easier to correlate with biological changes [12], there are no studies in schizophrenia using DTI that consider them [14,13]. However, our results agree with [13] in that the differences found are predominantly in the temporal and frontal lobes of the left hemisphere. In future work, we expect that an in-depth study of the differences in these quantities will shed some light into the biological characteristics of white matter changes in schizophrenia.

Finally, the method presented is not bound to a single scalar diffusivity measure, it is even generalizable to full-tensor studies as done, for instance, by [6]. Moreover, it is possible to correlate these changes with age and cognitive benchmarks in order to better analyse the relation between white matter changes and the different degrees of disease progression.

Acknowledgements. This work was supported by NIH grants #R01MH079938 and #R01MH060722, the French ANR NucleiPark and the France-Parkinson Association.

References

1. Pierpaoli, C., Barnett, A., Pajevic, S., Chen, R., Penix, L., Virta, A., Basser, P.: Water diffusion changes in Wallerian degeneration and their dependence on white matter architecture. *Neuroimage* 13(6), 1174–1185 (2001)
2. Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., Behrens, T.E.J.: Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31(4), 1487–1505 (2006)
3. O'Donnell, L.J., Westin, C.F., Golby, A.J.: Tract-based morphometry for white matter group analysis. *NeuroImage* 45(3), 832–844 (2009)
4. Maddah, M., Grimson, W.E.L., Warfield, S.K., Wells, W.M.: A unified framework for clustering and quantitative analysis of white matter fiber tracts. In: *MIA*, pp. 191–202 (2008)
5. Corouge, I., Fletcher, P.T., Joshi, S., Gouttard, S., Gerig, G.: Fiber tract-oriented statistics for quantitative diffusion tensor mri analysis. *MIA* 10(5), 786–798 (2006)
6. Yushkevich, P.A., Zhang, H., Simon, T.J., Gee, J.C.: Structure-specific statistical mapping of white matter tracts. *NeuroImage* 41(2), 448–461 (2008)
7. Zhang, H., Awate, S., Das, S., Woo, J., Melhem, E., Gee, J., Yushkevich, P.: A Tract-Specific Framework for White Matter Morphometry Combining Macroscopic and Microscopic Tract Features. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI* 2009. LNCS, vol. 5762, pp. 141–149. Springer, Heidelberg (2009)
8. Wassermann, D., Bloy, L., Kanterakis, E., Verma, R., Deriche, R.: Unsupervised white matter fiber clustering and tract probability map generation: Applications of a gaussian process framework for white matter fibers. *NeuroImage* 51(1), 228–241 (2010)

9. Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D.S., Calabresi, P.A., Pekar, J.J., van Zijl, P.C.M., Mori, S.: Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification. *NeuroImage* 39(1), 336–347 (2008)
10. Kindlmann, G., Estepar, R., Smith, S., Westin, C.: Sampling and Visualizing Creases with Scale-Space Particles. *IEEE Trans. on Viz. and Comp. Graph.* 15(6), 1415–1424 (2009)
11. Yang, J., Shen, D., Davatzikos, C., Verma, R.: Diffusion Tensor Image Registration Using Tensor Geometry and Orientation Features. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part II. LNCS*, vol. 5242, pp. 905–913. Springer, Heidelberg (2008)
12. Beaulieu, C.: *The Biological Basis of Diffusion Anisotropy*. In: *Diffusion MRI*. Academic Press, London (2009)
13. Ellison-Wright, I., Bullmore, E.: Meta-analysis of diffusion tensor imaging studies in schizophrenia. *Schizophrenia Research* 108(1-3), 3–10 (2009)
14. Kubicki, M., Shenton, M.E.: *Diffusion Tensor Imaging and Its Application to Schizophrenia and Related Disorders*. In: *Diffusion MRI*. Academic Press, London (2009)
15. Burns, J., Job, D., Bastin, M.E., Whalley, H., Macgillivray, T., Johnstone, E.C., Lawrie, S.M.: Structural disconnectivity in schizophrenia: a diffusion tensor magnetic resonance imaging study. *The British journal of psychiatry: the journal of mental science* 182, 439 (2003)
16. Ashtari, M., Cottone, J., Ardekani, B., Cervellione, K., Szeszko, P., Wu, J., Chen, S., Kumra, S.: Disruption of white matter integrity in the inferior longitudinal fasciculus in adolescents with schizophrenia as revealed by fiber tractography. *Arch. of Gen. Psych.* 64(11), 1270 (2007)
17. Fitzsimmons, J., Kubicki, M., Smith, K., Bushell, G., Estepar, R., Westin, C., Nestor, P., Niznikiewicz, M., Kikinis, R., McCarley, R., et al.: Diffusion tractography of the fornix in schizophrenia. *Schizophrenia Research* 107(1), 39–46 (2009)

Axon Diameter Mapping in the Presence of Orientation Dispersion with Diffusion MRI

Hui Zhang and Daniel C. Alexander

Microstructure Imaging Group, Department of Computer Science, University College London, London WC1E 6BT, United Kingdom

Abstract. Direct measurement of tissue microstructure with diffusion MRI offers a new class of biomarkers, such as axon diameters, that give more specific information about tissue than measures derived from diffusion tensors. The existing techniques of this kind assume a single axon orientation in the tissue model, which may be a reasonable approximation only for the most coherent brain white matter, such as the corpus callosum. For most other areas, orientation dispersion is not negligible and, if unaccounted for, leads to overestimation of the axon diameters, prohibiting their accurate mapping over the whole brain. Here we propose a new model that captures the effect of orientation dispersion explicitly. An efficient numerical scheme is developed to enable the axon diameter estimation by fitting the proposed model. Synthetic data experiments demonstrate that the new model provides an axon diameter index that is robust to the presence of orientation dispersion. Results on *in vivo* human data show reduced axon diameter index and better agreement with histology compared to previous methods suggesting improvements in the axon diameter estimate.

1 Introduction

Diffusion MRI reveals tissue microstructure by measuring patterns of water diffusion as influenced by the microscopic environment of the tissue. The most popular technique of this kind is the diffusion-tensor imaging (DTI) [1]. It provides simple biomarkers, such as mean diffusivity (MD) and fractional anisotropy (FA), that are effective at indicating major microstructural changes in tissue during normal development or due to disease. However, a limitation of these biomarkers is that they are inherently non-specific [2], e.g. changes in FA may be attributed to changes in a combination of the underlying microstructure features, such as axon density and diameter, and cannot isolate their individual contributions. Direct measurement of these features can shed new light into development and disease mechanisms by offering a new class of highly-specific biomarkers.

Previous work, e.g. [3,4,5,6,7], shows the feasibility of using diffusion MRI to measure microstructure features directly in excised biological tissues as well as *in vivo*. A particularly successful approach, exemplified by [3,4], is the model-based strategy in which a geometric model of the microstructure of interest predicts the MR signal from water diffusing within. In estimating axon diameters using this

approach, the earlier attempts, e.g. [5,7], assume a single and known orientation of axon in the tissue model. Alexander [6] demonstrates in simulation that it is possible to estimate axon diameters of unknown orientation using a multi-shell high angular resolution diffusion-weighted imaging (HARDI) protocol optimized with realistic in vivo diffusivities and constraints typical of a clinical scanner. Later work [8] gives promising results in both fixed monkey brains and live human brains that show trends in the midsagittal corpus callosum (CC) and cortico-spinal tract (CST) that we expect from histology.

The model in [8] simplifies real white matter (WM) in two important ways: it assumes a single axon diameter rather than a distribution; it assumes all the axons within single voxels have the same single orientation. The first simplification means that the fitted model provides a single index (summary statistic) of the axon diameter distribution rather than the full distribution. The index correlates with the mean diameter weighted by volume so is still useful. The second simplification holds for all previous methods of axon diameter estimation [3,4,5,7] and limits estimates only to the most coherent structures, e.g. CC; even there the assumption of coherent orientation is questionable. Most brain WM has significant orientation dispersion. This leads to the overestimation of the axon diameter index or distribution because axons oblique to the assumed single orientation appear to have larger cross section, as shown in Fig. 2(a). The limitation prevents the accurate mapping of their microstructure features over the whole brain and casts doubt on estimates even from CC and CST.

Here we aim to ameliorate this limitation by introducing a new tissue model that explicitly represents the dispersion in axon orientation. An efficient numerical scheme makes the proposed model feasible for axon diameter estimation. Experimental results demonstrate that the proposed model estimates axon diameter indices from synthetic data and human brain data, and other microstructure features, more accurately than previous ones. The rest of the paper is organized as follows: Section 2 describes the proposed tissue model and its numerical implementation; Section 3 gives the experimental design and results; Section 4 summarizes the contribution and discusses future work.

2 Tissue and Signal Model

The proposed model is based on the simplified version [6] of Assaf et al.'s composite hindered and restricted model of diffusion [4]. It models WM as a population of cylindrical axons with a single radius and impermeable cell walls embedded in a homogeneous medium as in [3,4,9,5,6,7]. To account for the CSF contribution through the partial volume effect, an isotropic compartment is included as in [7]. The normalized MR signal $A(\mathbf{G}, \Delta, \delta)$ from the standard pulsed-gradient spin-echo (PGSE) sequence, with gradient pulses of length δ , strength and direction \mathbf{G} , and separation Δ , can be written as

$$(1 - \nu_{iso})(\nu_{ic}A_{ic}(\mathbf{G}, \Delta, \delta) + (1 - \nu_{ic})A_{ec}(\mathbf{G}, \Delta, \delta)) + \nu_{iso}A_{iso}(\mathbf{G}, \Delta, \delta), \quad (1)$$

where $\nu_{iso} \in [0, 1]$ is the volume fraction of the isotropic compartment, $\nu_{ic} \in [0, 1]$ the relative volume fraction of the intra-cellular compartment among WM, and

A_{ic} , A_{ec} , and A_{iso} the normalized signals from the intra-cellular, extra-cellular and isotropic compartments respectively.

Intra-Cellular Model. Let the distribution of axon orientation be $\rho : \mathbb{S}^2 \mapsto \mathbb{R}^+$, where $\rho(\mathbf{n}) = \rho(-\mathbf{n})$ and $\int \rho(\mathbf{n}) d\mathbf{n} = 1$. The intra-cellular signal is then

$$A_{ic}(\mathbf{G}, \Delta, \delta|\rho, R) = \int \rho(\mathbf{n}) A_{cyl}(\mathbf{G}, \Delta, \delta|\mathbf{n}, R) d\mathbf{n}, \tag{2}$$

where A_{cyl} is the signal from water restricted by a cylinder of radius R and orientation \mathbf{n} for which we use the Gaussian phase distribution approximation [10], as in [6].

Here we consider two models for ρ :

- The *Delta model* where $\rho(\mathbf{n})$ is a δ -function, which reduces (2) to the form of previous models in [4][6].
- The *Watson model* where we model ρ with a Watson distribution [11] so that

$$\rho(\mathbf{n}) = f(\mathbf{n}|\boldsymbol{\mu}, \kappa) = M\left(\frac{1}{2}, \frac{3}{2}, \kappa\right)^{-1} e^{\kappa(\boldsymbol{\mu} \cdot \mathbf{n})^2}, \tag{3}$$

where M is a confluent hypergeometric function, $\boldsymbol{\mu}$ is the unit vector about which the distribution is cylindrically symmetric, and κ , the concentration parameter, controls the extent of orientation dispersion.

The Watson distribution is the spherical analog of the Gaussian distribution with cylindrical symmetry. It allows us to capture the essential characteristics of nonparallel axons with κ , the single extra parameter compared to the Delta model. Although the general admissible values for κ range from $-\infty$ to $+\infty$, here we constrain it to be strictly positive such that $\boldsymbol{\mu}$ represents the mean orientation of the distribution with larger κ corresponding to lower dispersion about $\boldsymbol{\mu}$.

Extra-Cellular Model. This compartment exhibits hindered diffusion and is modeled with simple (Gaussian) anisotropic diffusion, such that $A_{ec}(\mathbf{G}, \Delta, \delta) = e^{-(\Delta-\delta/3)\gamma^2\delta^2\mathbf{G}^T\mathbb{D}_{ec}\mathbf{G}}$, where \mathbb{D}_{ec} is the (apparent) diffusion tensor of the compartment. We define \mathbb{D}_{ec} in terms of the distribution of axon orientation ρ as

$$\mathbb{D}_{ec} = \int \rho(\mathbf{n}) \mathbb{D}_h(\mathbf{n}) d\mathbf{n}, \tag{4}$$

where $\mathbb{D}_h(\mathbf{n})$ is the diffusion tensor for axons with a single orientation \mathbf{n} for which we use the definition $\mathbb{D}_h(\mathbf{n}) = (d - d_{\perp})\mathbf{n}\mathbf{n}^T + d_{\perp}\mathbb{I}$, with d_{\perp} being the apparent diffusion coefficient perpendicular to axons and \mathbb{I} being the identity tensor, as in [4]. For the Delta model, $\mathbb{D}_{ec} = \mathbb{D}_h$. For the Watson model, $\mathbb{D}_{ec} = (d' - d'_{\perp})\boldsymbol{\mu}\boldsymbol{\mu}^T + d'_{\perp}\mathbb{I}$, where

$$d' = d + (d - d_{\perp}) \left(\frac{1}{2\sqrt{\kappa}DawsonF(\sqrt{\kappa})} - \frac{1}{2\kappa} - 1 \right) \tag{5}$$

$$d'_{\perp} = d_{\perp} + (d - d_{\perp}) \left(\frac{1}{4\kappa} - \frac{1}{4\sqrt{\kappa}DawsonF(\sqrt{\kappa})} + \frac{1}{2} \right) \tag{6}$$

and $DawsonF$ is the Dawson’s integral [12]. We use Szafer et al.’s tortuosity model [13] and set $d_{\perp} = d(1 - \nu_{ic})$.

Isotropic Model. The isotropic compartment does not depend on the distribution of axon orientation and is modeled as a simple isotropic diffusion process such that $A_{iso} = e^{-(\Delta-\delta/3)\gamma^2\delta^2\|G\|^2D}$ with D being the intrinsic diffusion coefficient of the compartment.

Numerical Implementation. The integral in (2) does not have a closed-form analytic solution while brute-force numerical integration is prohibitively expensive, up to 1000 times longer than computing the Delta model. To make parameter estimation using the Watson model feasible, we express the Watson distribution function by its spherical harmonic (SH) expansion which makes the integral in (2) analytic leading to the following summation:

$$A_{ic} = \sum_{l=0}^{\infty} f_{l0}(\kappa) \sqrt{\frac{2l+1}{4\pi}} P_l(\mathbf{q} \cdot \boldsymbol{\mu}) G_l(d, \nu_{ic}, R), \quad (7)$$

where $f_{l0}(\kappa)$ is the spherical harmonic (SH) coefficient (order l and 0) of the *canonical* Watson distribution $f(\mathbf{n}|\boldsymbol{\mu} = \mathbf{z}, \kappa)$ and has a closed-form analytic expression for all l , P_l is the Legendre polynomial of order l , G_l is a real function, nonzero only for even l , and can be calculated analytically. In the current implementation, the summation is truncated at $l = 12$ to give a good approximation for κ up to 128. The implementation is orders of magnitude more efficient than numerical integration, taking approximately only 50% longer than computing the Delta model.

3 Experiments and Results

This section describes the synthetic and in vivo human data experiments and results that compare the proposed Watson model to the Delta model in terms of their ability to estimate axon diameters and other microstructure parameters. The synthetic experiments compare the accuracy with which known tissue parameters can be recovered by each model while the in vivo experiments compare each model's performance in real human data.

Data Acquisition. In vivo imaging uses a clinical 3T Philips scanner to acquire 360 diffusion-weighted images of a healthy volunteer in about an hour with the protocol in [8]. The protocol, determined via the optimization procedure in [6], divides the measurements into 4 HARDI shells each with 90 gradient directions and includes 4 $b = 0$ images. It uses sagittal echo-planar imaging with in-plane resolution 128×128 with $1.8 \times 1.8\text{mm}^2$ voxels and thickness 3.9mm, no gap, resulting in white matter signal-to-noise (SNR) at $b = 0$ about 20.

Model Fitting. For both experiments, we fit the models to data with the Rician Markov Chain Monte Carlo (MCMC) procedure in [6], after an initial grid search and gradient descent to determine the maximum likelihood (ML) estimates of the parameters. The MCMC procedure refines the ML estimates for R and ν_{ic} , for both models, and κ , for the Watson model, by computing their

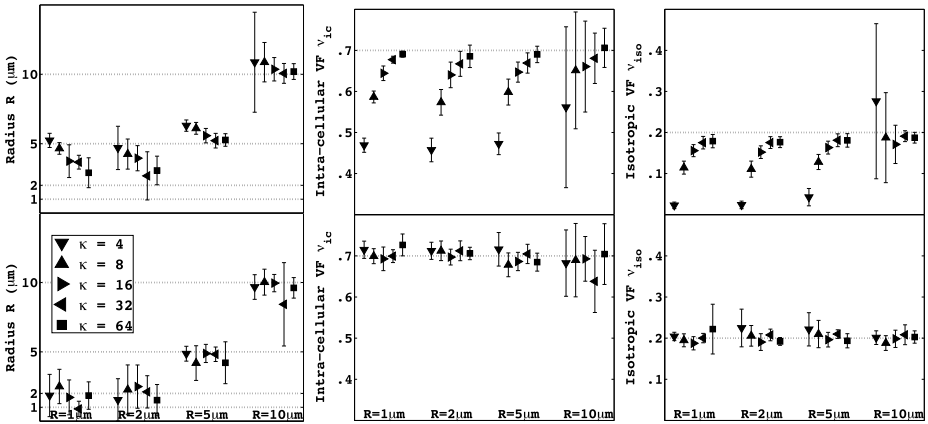


Fig. 1. The parameter estimates from the synthetic experiments using the Delta (top) and Watson (bottom) models. The dotted lines indicate the true parameter values.

posterior distributions, the means of which provide the final estimates for these parameters. Throughout, we fix d and D to their expected values in human in vivo data, which are 1.7 and $3.0 \times 10^{-9} \text{ m}^2\text{s}^{-1}$, respectively. The MCMC setting is 40 samples at intervals of 200 iterations after a burn in of 2000 iterations.

Synthetic Experiments. We synthesize MR data from the proposed Watson model using the imaging protocol of the human data described above and adds synthetic Rician noise with $\sigma = 0.05$ to match the SNR of the human data. The true model parameters are $\nu_{ic} = 0.7$, $\nu_{iso} = 0.2$, $R \in \{1, 2, 5, 10\} \mu\text{m}$, and $\kappa \in \{4, 8, 16, 32, 64\}$. We test 10 different orientations, uniformly distributed over the sphere. For each parameter, we report the mean and standard deviation of its 10 estimates over the different orientations.

Fig. 1 shows the recovered model parameters for both the Delta and Watson models. The Delta model consistently overestimates the axon radius, as expected, with the effect more pronounced for small radii and large orientation dispersion (i.e., lower κ). It also consistently underestimates ν_{ic} with increasing bias for decreasing κ . The estimate for ν_{iso} is similarly biased downward except for the largest value of R which is biased upward. For the very high κ values, 32 and 64, the bias in the estimation becomes negligible for all the parameters.

In contrast, the Watson model provides consistently more accurate estimates for these parameters with just a few exceptions. It overestimates the smallest value of R , making it indistinguishable from $R = 2 \mu\text{m}$, but the overestimation is slight, which does not prevent it from being distinguished from $R = 5 \mu\text{m}$. The model slightly underestimates the radius for some κ and R values but there is no systematic pattern.

The Watson model also supports the direct estimation of orientation dispersion via κ , the results of which are shown in Fig. 2(b). The estimates are accurate

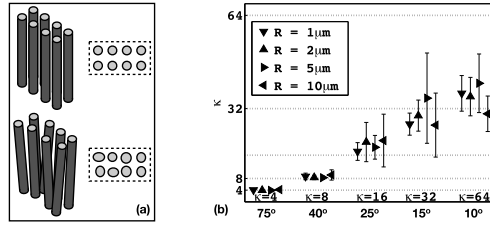


Fig. 2. (a) The schematic illustration of the overestimation of axon diameters due to orientation dispersion. (b) The estimated κ using the Watson model from the synthetic experiments. The dotted lines indicate the true κ values. The numerics underneath the x-axis indicate the polar angle value containing 95% of orientation variation in the corresponding Watson distribution.

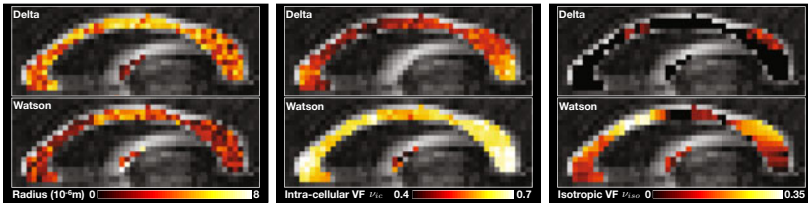


Fig. 3. The parameter estimates for the midsagittal slice of CC using the Delta and Watson models

for smaller κ values but become increasingly biased downward for higher κ . The highest values, 32 and 64, are more difficult to distinguish.

In Vivo Human Brain Data Experiments. Here we fit each model to real brain data. To avoid regions with fiber crossing, we compute the linearity and planarity [14] maps of the best-fit diffusion tensor and select only the voxels with high linearity, greater than 0.6, and low planarity, less than 0.2.

Fig. 3 shows the estimated parameters from both models in the midsagittal slice of CC. The corresponding κ estimates are shown in Fig. 4 (bottom row & middle column). The results agree strongly with the trends observed in the synthetic experiments, i.e., for κ lower than 32, the Delta model overestimates R and underestimates ν_{ic} and ν_{iso} . They are also much more consistent with histology [15] than results from the Delta model, i.e., axons have smaller radii, are more densely packed (higher ν_{ic}) and coherent (higher κ) in the genu and splenium of CC than its midbody. Fig. 4 shows additional results of the estimates of R from both models and the estimates of κ from the Watson model for two lateral slices in addition to the full midsagittal slice. Besides confirming the trend observed above, it shows that the estimates of κ in the genu and the splenium of CC, as expected, are the highest for the whole brain.

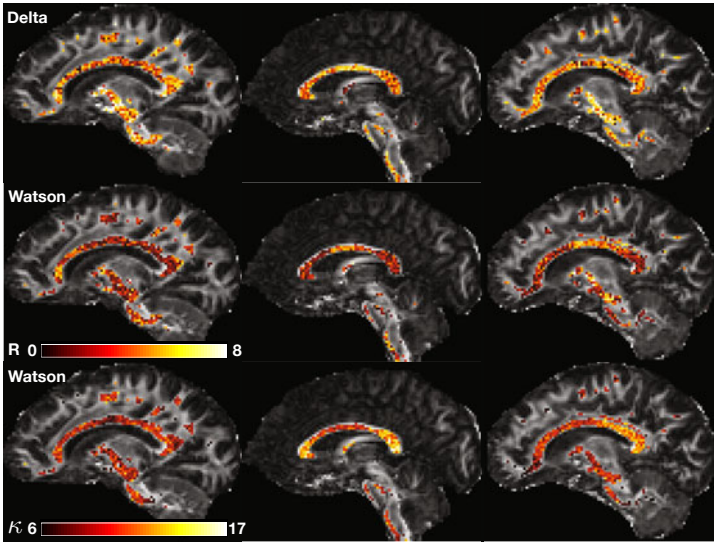


Fig. 4. The estimates of R for 3 sagittal slices using the Delta (top) and Watson (middle) models, with the corresponding κ estimates (bottom) from the Watson model

4 Discussion

In this paper, we proposed a new white matter tissue model for the direct estimation of axon diameters and other microstructural features using diffusion MRI. The proposed model advances the state-of-the-art by explicitly accounting for the presence of orientation dispersion in axon bundles to mitigate the overestimation bias in estimating axon diameters. Synthetic experiments demonstrate that, in the presence of orientation dispersion, models assuming a single orientation not only overestimates axon diameters but also underestimates intra-cellular volume fractions and biases isotropic volume fractions. In contrast, the proposed model can both estimate these key tissue parameters more accurately without systematic bias and provide estimates to the extent of orientation dispersion. Results from in vivo human data experiments agree strongly with the findings from simulation, which suggests the orientation dispersion must be taken into account even for the most coherent structures like CC and CST.

One possible limitation of the present approach is its use of the Watson distribution to model the underlying orientation dispersion in brain data. For voxels with significant spanning or bending axon bundles, the cylindrical symmetry of the Watson distribution may prove overly simplistic and introduce potential bias in estimating microstructural parameters. Future work will explore the use of more complex orientation distributions, such as the Bingham distribution [11], to model such effects.

Acknowledgment. We thank Penny Hubbard and Geoffrey Parker from University of Manchester for providing the in-vivo human dataset. The future and emerging technologies (FET) program of the EU FP7 framework fund the CONNECT consortium (www.brain-connect.eu), which supports this work. EPSRC fund DCA under grant EP/E007748.

References

1. Basser, P.J., Mattiello, J., Bihan, D.L.: MR diffusion tensor spectroscopy and imaging. *Biophys. J.* 66 (1994)
2. Pierpaoli, C., Jezzard, P., Basser, P.J., Barnett, A., Chiro, G.D.: Diffusion tensor MR imaging of the human brain. *Radiology* 201 (1996)
3. Stanisz, G.J., Szafer, A., Wright, G.A., Henkelman, M.: An analytical model of restricted diffusion in bovine optic nerve. *MRM* 37, 103–111 (1997)
4. Assaf, Y., Freidlin, R.Z., Rhode, G.K., Basser, P.J.: New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *MRM* 52, 965–978 (2004)
5. Assaf, Y., Blumenfeld-Katzir, T., Yovel, Y., Basser, P.J.: AxCaliber: a method for measuring axon diameter distribution from diffusion MRI. *MRM* 59, 1347–1354 (2008)
6. Alexander, D.C.: A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features. *MRM* 60, 439–448 (2008)
7. Barazany, D., Basser, P.J., Assaf, Y.: In-vivo measurement of the axon diameter distribution in the corpus callosum of a rat brain. *Brain* 132, 1210–1220 (2009)
8. Alexander, D.C., Hubbard, P.L., Hall, M.G., Moore, E.A., Ptito, M., Parker, G.J.M., Dyrby, T.B.: Orientationally invariant indices of axon diameter and density from diffusion MRI. *NeuroImage* (2010), doi:10.1016/j.neuroimage.2010.05.043
9. Jespersen, S.N., Kroenke, C.D., Ostergaard, L., Ackerman, J.J.H., Yablonskiy, D.A.: Modeling dendrite density from magnetic resonance diffusion measurements. *NeuroImage* 34, 1473–1486 (2007)
10. Murday, J.S., Cotts, R.M.: Self-diffusion coefficient of liquid lithium. *J. Chem. Phys* 48, 4938–4945 (1968)
11. Mardia, K.V., Jupp, P.E.: Directional statistics. Wiley series in probability and statistics. John Wiley & Sons, Ltd., Chichester (1990)
12. Abramowitz, M., Stegun, I.A.: Handbook of mathematical functions with formulas, graphs, and mathematical tables, 9th edn. Dover, New York (1972)
13. Szafer, A., Zhong, J.H., Gore, J.C.: Theoretical model for water diffusion in tissues. *MRM* 33, 697–712 (1995)
14. Westin, C.F., Maier, S.E., Mamata, H., Nabavi, A., Jolesz, F.A., Kikinis, R.: Processing and visualization for diffusion tensor MRI. *MIA* 6, 93–108 (2002)
15. Aboitiz, F., Scheibel, A.B., Fisher, R.S., Zaidel, E.: Fiber composition of the human corpus callosum. *Brain research* 598, 143–153 (1992)

Model-Free, Regularized, Fast, and Robust Analytical Orientation Distribution Function Estimation

Jian Cheng^{1,2}, Aurobrata Ghosh², Rachid Deriche², and Tianzi Jiang¹

¹ Center for Computational Medicine, LIAMA, Institute of Automation, Chinese Academy of Sciences, China

² Athena Project Team, INRIA Sophia Antipolis – Méditerranée, France
j.iancheng@nlpr.ia.ac.cn

Abstract. High Angular Resolution Imaging (HARDI) can better explore the complex micro-structure of white matter compared to Diffusion Tensor Imaging (DTI). Orientation Distribution Function (ODF) in HARDI is used to describe the probability of the fiber direction. There are two type definitions of the ODF, which were respectively proposed in Q-Ball Imaging (QBI) and Diffusion Spectrum Imaging (DSI). Some analytical reconstructions methods have been proposed to estimate these two type of ODFs from single shell HARDI data. However they all have some assumptions and intrinsic modeling errors. In this article, we propose, almost without any assumption, a uniform analytical method to estimate these two ODFs from DWI signals in q space, which is based on Spherical Polar Fourier Expression (SPFE) of signals. The solution is analytical and is a linear transformation from the \mathbf{q} -space signal to the ODF represented by Spherical Harmonics (SH). It can naturally combines the DWI signals in different Q-shells. Moreover It can avoid the intrinsic Funk-Radon Transform (FRT) blurring error in QBI and it does not need any assumption of the signals, such as the multiple tensor model and mono/multi-exponential decay. We validate our method using synthetic data, phantom data and real data. Our method works well in all experiments, especially for the data with low SNR, low anisotropy and non-exponential decay.

1 Introduction

High Angular Resolution Diffusion Imaging (HARDI) is used to probe non-Gaussian diffusion which represents more intricate micro-structure in the tissue. Orientation Distribution Function (ODF) [12] was proposed to describe the fiber directions. There are two type of ODFs. One is denoted as Φ_t , proposed using radial projection by Tuch in QBI [1]. Another one is denoted as Φ_w , proposed as the marginal probability of the Ensemble Average Propagator (EAP) $P(\mathbf{Rr})$ by Wedeen in DSI [2]. Φ_t need to be normalized and Z is the normalization factor. While Φ_w is naturally normalized.

$$\Phi_t(\mathbf{r}) = \frac{1}{Z} \int_0^\infty P(\mathbf{Rr})dR \quad \Phi_w(\mathbf{r}) = \int_0^\infty P(\mathbf{Rr})R^2dR = \frac{1}{2} \int_{-\infty}^\infty P(\mathbf{Rr})R^2dR \quad (1)$$

where $\mathbf{R} = \mathbf{Rr}$ is the displacement in 3D space. It has been shown that Φ_w is more sharper than Φ_t [234], which means Φ_w is more discriminative for fiber detection.

Historically, Funk-Radon Transform (FRT) was used in QBI to estimate Φ_t numerically [1] or analytically [5]. However, the intrinsic blurring effect of FRT can bring some errors [1]. Φ_w was firstly proposed in DSI and was estimated from numerical radial integral after the numerical Fourier Transform of the signals [2]. Most recently, several similar analytical reconstruction methods were proposed separately to estimate Φ_w from single shell HARDI data [6,3,4]. Elegant analytical solutions were found [6,3,4] based on the mono-exponential decay assumption [7] which gives the full information about $E(\mathbf{q})$ in the whole 3D \mathbf{q} -space from the $E(\mathbf{q}_0)$ only in a single shell. The approximated EAP $\tilde{P}(R\mathbf{r})$ actually is the true EAP $P(R\mathbf{r})$ convolved by the Fourier transform of the function $E(q, \mathbf{u})^{q^2/q_0^2} E(q, \mathbf{u}^{-1})$ [7], where $\mathbf{q} = q\mathbf{u}$, $q = \|\mathbf{q}\|$. It was shown surprisingly that the estimated $\tilde{P}(R\mathbf{r})$ and $\tilde{\Phi}_w$ are sharper than the real $P(R\mathbf{r})$ and Φ_w [6,3,4] in the synthetic data generated from mixture tensor model. However, since this surprising results come from the intrinsic modeling error from the unrealistic kernel smooth, it is still not clear if the methods based on that assumption can work well in the complex real data with non-exponential decay, low anisotropy and low SNR. Similarly with [7], the authors in [3] extended mono-exponential model to multi-exponential model so that it can reduce the modeling error and work for the data in multiple shells. However, it is impractical because a nonlinear fitting is needed for every direction [7], suffering from limited samples, local minima, computational complexity, and an analytic solution exists only when three b values satisfy an arithmetic process.

$$E(\mathbf{q}) = \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l a_{n,l,m} R_n(\|\mathbf{q}\|) Y_l^m(\mathbf{u}) \quad B_{n,l,m}(\mathbf{q}) = R_n(\|\mathbf{q}\|) Y_l^m(\mathbf{u}) \quad (2)$$

$$R_n(\|\mathbf{q}\|) = \kappa_n(\zeta) \exp\left(-\frac{\|\mathbf{q}\|^2}{2\zeta}\right) L_n^{1/2}\left(\frac{\|\mathbf{q}\|^2}{\zeta}\right) \quad \kappa_n(\zeta) = \left[\frac{2}{\zeta^{3/2}} \frac{n!}{\Gamma(n+3/2)}\right]^{1/2} \quad (3)$$

In [8], the Spherical Polar Fourier Expression (SPFE) was proposed to sparsely represent $E(\mathbf{q})$. See formulae (2), (3), where $Y_l^m(\mathbf{u})$ is the l order m degree Spherical Harmonic (SH) basis and $R_n(q)$ is the Gaussian-Laguerre polynomial basis. Since $B_{n,l,m}(\mathbf{q})$ is the orthonormal basis in R^3 , any type of $E(\mathbf{q})$ could be represented by a linear combination of $\{B_{n,l,m}\}$. After the coefficients $\{a_{n,l,m}\}$ of the signal are estimated from a least square fit or a nonlinear robust estimation [8], Φ_t could be calculated through an inner product of the coefficients $a_{n,l,m}$ and a kernel $b_{n,l,m}$. The problem in [8] is that $b_{n,l,m}$ needs to be calculated numerically from FFT for every direction or calculated for one direction then rotated by Wigner rotation matrix for other directions. That is inefficient and can bring some numerical error, especially for these kernels which have some delta functions inside, e.g. the kernels for Φ_t and Φ_w . And it can not provide an elegant analytical parametrized result like analytical QBI [5].

In this paper, instead of adding strong assumptions for single shell data in [6,3,4] and numerical solution using FFT and Wigner matrix in [8], we propose a uniform analytical estimation method for Φ_t and Φ_w based on SPFE, which includes two linear transformations from the coefficients $\{a_{n,l,m}\}$ of $E(\mathbf{q})$ to the coefficients $\{c_{l,m}^t\}$ of Φ_t and $\{c_{l,m}^w\}$ of Φ_w represented by SHs. First we deduce the transformations for Φ_t and Φ_w . Next, we perform the method in some non-exponential synthetic data and a challenging phantom data. At last we test our methods in a real monkey data with several b values.

2 Analytical ODF Estimation Based On SPF

It has been shown that the line integral of $P(\mathbf{R}\mathbf{r})$ in \mathbf{R} -space in (1) is equivalent to the integral in the plane in \mathbf{q} -space which is orthogonal to the line in \mathbf{R} -space [3,4,6]. See formula (4), where Δ_b is the Laplace-Beltrami operator. Our contribution is to deduce the elegant analytical solution for data in multiple shells based on these previous studies in [3,4,6,8]. Our analytical estimation methods almost do not need any assumption about the signal. The only assumption we need is that the signal $E(\mathbf{q})$ can be sparsely represented by SPF in formula (2), which has been validated in [8].

$$\Phi_t(\mathbf{r}) = \frac{1}{Z} \iint_{\Pi_r} E(\mathbf{q})q\delta(\mathbf{r}^T\mathbf{u})dq\mathbf{u} \quad \Phi_w(\mathbf{r}) = \frac{1}{4\pi} - \frac{1}{8\pi^2} \iint_{\Pi_r} \frac{\Delta_b E(\mathbf{q})}{q} \delta(\mathbf{r}^T\mathbf{u})dq\mathbf{u} \quad (4)$$

2.1 Estimation of Φ_t

Put the formula (2) into (4), we can easily get the solution.

$$\begin{aligned} \Phi_t(\mathbf{r}) &= \frac{1}{Z} \iint_{\Pi_r} \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l a_{n,l,m} R_n(q) Y_l^m(\mathbf{u}) q \delta(\mathbf{r}^T\mathbf{u}) dq \mathbf{u} \\ &= \frac{1}{Z} \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l a_{n,l,m} \left(\int_0^{2\pi} Y_l^m(\mathbf{u}) \delta(\mathbf{r}^T\mathbf{u}) d\mathbf{u} \right) \left(\int_0^\infty R_n(q) q dq \right) \\ &= \frac{1}{Z} \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l a_{n,l,m} (2\pi P_l(0) Y_l^m(\mathbf{r})) \left(\frac{\kappa_n(\zeta)\zeta}{2} \int_0^\infty \exp\left(-\frac{x}{2}\right) L_n^{1/2}(x) dx \right) \quad (5) \\ &= \boxed{\frac{2\pi\zeta}{Z} \sum_{l=0}^L \sum_{m=-l}^l \left(\sum_{n=0}^N \sum_{i=0}^n \kappa_n(\zeta) \binom{i-0.5}{i} (-1)^{n-i} P_l(0) a_{n,l,m} \right) Y_l^m(\mathbf{r})} \quad (6) \end{aligned}$$

where $P_l(0)$ is the Legendre polynomial of order l at 0. We get (5) because SH is the eigenfunction of the FRT [5]. From (5) to (6), we use the property of Laguerre polynomial [9]. Thus here we have a linear transformation from the coefficients $a_{n,l,m}$ of $E(\mathbf{q})$ to the coefficients $c_{l,m}^i = \sum_{n=0}^N \sum_{i=0}^n \kappa_n(\zeta) \binom{i-0.5}{i} (-1)^{n-i} P_l(0) a_{n,l,m}$. Please note that the author in [10] gave a solution for Φ_t in page 122. Unfortunately, the integrand there was wrong because of wrong volume element. Here we give the right analytical formulae.

We also give the result of the integral in a given disk $\Pi_{(r,C)}$ whose radius is C . In the formula (4), the integral $\Phi_t(\mathbf{r})$ gives the same weight for $E(\mathbf{q})$ with large q and for $E(\mathbf{q})$ with small q . However, if we just have several b values, the error of estimated signal $\hat{E}(\mathbf{q})$ will be small if q is between these b values and will be large if q is large than all b values. Thus if an approximate C is given, the disk integral $\Phi_t(\mathbf{r}, C)$ may have better angular resolution than $\Phi_t(\mathbf{r})$ [4]. Considering $L_n^{1/2}(x) = \sum_{i=0}^n l_n^i x^i$, $l_n^i = (-1)^i \binom{n+0.5}{n-i} \frac{1}{i!}$, and the lower incomplete gamma function $\gamma(i, x) = \int_0^x t^{i-1} \exp(-t) dt$, we have

$$\begin{aligned} \Phi_t(\mathbf{r}, C) &= \frac{2\pi\zeta}{Z} \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l a_{n,l,m} (P_l(0) Y_l^m(\mathbf{r})) \left(\frac{\kappa_n(\zeta)}{2} \int_0^{C^2/\zeta} \exp\left(-\frac{x}{2}\right) L_n^{1/2}(x) dx \right) \quad (7) \\ &= \boxed{\frac{2\pi\zeta}{Z} \sum_{l=0}^L \sum_{m=-l}^l \left(\sum_{n=0}^N \sum_{i=0}^n \kappa_n(\zeta) \binom{n+0.5}{n-i} \frac{(-2)^i}{i!} \gamma(i+1, 0.5C^2/\zeta) P_l(0) a_{n,l,m} \right) Y_l^m(\mathbf{r})} \end{aligned}$$

2.2 Estimation of Φ_w

Similarly, put the formula (2) into (4) we can get the analytical expression for Φ_w .

$$\Phi_w(\mathbf{r}) = \frac{1}{4\pi} - \frac{1}{8\pi^2} \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l a_{n,l,m} \left(\int_0^{2\pi} \Delta_b Y_l^m(\mathbf{u}) \delta(\mathbf{r}^T \mathbf{u}) d\mathbf{u} \right) \left(\int_0^\infty \frac{R_n(q)}{q} dq \right) \quad (8)$$

However, we can not solve it just like what we did for Φ_l , because the division by q introduces a pole. It is a little hard to find the analytical solution for Φ_w . And the author in [10,8] did not give any solution for that. We solve this problem by considering $E(0) = 1$, which is a true fact for any DWI data. That means, for our basis, the identity $E(0) = \sum_{n,l,m} a_{n,l,m} R_n(0) Y_l^m(\mathbf{u}) = \sum_{n,l,m} a_{n,l,m} \kappa_n(\zeta) Y_l^m(\mathbf{u}) = 1$ holds for any $\mathbf{u} \in S^2$. Also keep in mind that a constant addition inside Δ_b does not change the final result. First we consider the integral inside a given disk $\Pi_{(\mathbf{r},C)}$, then we have

$$\Phi_w(\mathbf{r}, C) = \frac{1}{4\pi} - \frac{1}{8\pi^2} \iint_{\Pi_{(\mathbf{r},C)}} \frac{\Delta_b(E(\mathbf{q}) - E(0))}{q} \delta(\mathbf{r}^T \mathbf{u}) dq d\mathbf{u} \quad (9)$$

$$\begin{aligned} &= \frac{1}{4\pi} - \frac{1}{8\pi^2} \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l a_{n,l,m} \left(\int_0^{2\pi} \Delta_b \left(\int_0^C \frac{R_n(q) - R_n(0)}{q} dq \right) Y_l^m(\mathbf{u}) \delta(\mathbf{r}^T \mathbf{u}) d\mathbf{u} \right) \\ &= \frac{1}{4\pi} - \frac{1}{8\pi^2} \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l a_{n,l,m} \kappa_n(\zeta) \left(\int_0^{2\pi} \Delta_b I_n(C) Y_l^m(\mathbf{r}) \delta(\mathbf{r}^T \mathbf{u}) d\mathbf{u} \right) \end{aligned} \quad (10)$$

Now there is no pole! For $I_n(C)$ we have

$$I_n(C) = \int_0^C \frac{R_n(q) - R_n(0)}{\kappa_n(\zeta) q} dq \quad (11)$$

$$= \frac{1}{2} \int_0^{C^2/\zeta} \left(\frac{\exp(-x/2) - 1}{x} + \sum_{i=1}^n l_n^i x^{i-1} \exp(-\frac{x}{2}) \right) dx \quad (12)$$

$$= \underbrace{0.5(-\gamma - E_1(0.5C^2/\zeta) - \log(0.5C^2/\zeta))}_{I_n^1(C)} + 0.5 \underbrace{\sum_{i=1}^n l_n^i 2^i \gamma(i, 0.5C^2/\zeta)}_{I_n^2(C)} \quad (13)$$

where $\gamma \simeq 0.5772$ is the Euler–Mascheroni constant, $E_1(x) = \int_x^\infty \frac{\exp(-t)}{t} dt$ is the exponential integral. Although there are two parts in $I_n(C)$ and $I_n^1(C)$ tends to infinity, it actually has no contribution for $\Phi_w(\mathbf{r}, C)$, because $\sum_{n,l,m} a_{n,l,m} \kappa_n(\zeta) Y_l^m(\mathbf{u}) I_n^1(C) = I_n^1(C)$ is a constant inside Δ_b . Then considering $Y_l^m(\mathbf{u})$ is the eigenfunction of FRT and Δ_b , we have the analytical result for $\Phi_w(\mathbf{r}, C)$ and $\Phi_w(\mathbf{r}) = \lim_{C \rightarrow \infty} \Phi_w(\mathbf{r}, C)$, $c_{l,m}^w = \lim_{C \rightarrow \infty} c_{l,m}^w(C)$

$$\Phi_w(\mathbf{r}, C) = \frac{1}{4\pi} - \frac{1}{8\pi} \sum_{n,l,m} a_{n,l,m} \kappa_n(\zeta) \sum_{i=1}^n l_n^i 2^i \gamma(i, 0.5C^2/\zeta) P_l(0) (-l)(l+1) Y_l^m(\mathbf{r}) \quad (14)$$

$$c_{l,m}^w(C) = \frac{1}{\sqrt{4\pi}} \delta(l) \delta(m) - \frac{1}{8\pi} \sum_{n=1}^N \sum_{i=1}^n (-1)^j \kappa_n(\zeta) \binom{n+0.5}{n-i} \frac{2^i}{i!} \gamma(i, 0.5C^2/\zeta) P_l(0) (-l)(l+1) a_{n,l,m} \quad (15)$$

$$c_{l,m}^w = \frac{1}{\sqrt{4\pi}}\delta(l)\delta(m) - \frac{1}{8\pi} \sum_{n=1}^N \sum_{i=1}^n (-1)^i \kappa_n(\zeta) \binom{n+0.5}{n-i} \frac{2^i}{i} P_l(0) (-l)(l+1) a_{n,l,m} \tag{16}$$

Now we have two estimations for the true Φ_w . One is the integral in the whole plane, which is similar with [3], and another one is the integral in a given disk, which is similar with [4]. However, the mono-exponential decay model was assumed during the disk in [4], and in the whole plane in [3], so that the integral for the radial integral could be approximated just using the data in the \mathbf{q} -circle. While our method does not need any assumption on the data and it can handle the data in different q -shells. Also please note three important points in the formulae above. *First*, we get the the $E_1(x)$ in the derivation process, but it is negligible. While in [4], it is indispensable. *Second*, the formulae (15), (16) tell us that at least order 1 of R_n is needed to represent an anisotropic ODF. That is true because if we just use the radial basis of order zero, it is easily seen that the estimated signal is just an isotropic one, which means the estimated ODF is isotropic. Thus we need to use at least two shells to get a reasonable results, although our methods can be performed in single shell data. *Third*, since in (4) $1/q$ gives small weight for $E(\mathbf{q})$ with large q and large weight for $E(\mathbf{q})$ with small q , that means the error in large q may be negligible! Thus Φ_w is more robust to estimation error of $E(\mathbf{q})$. How to choose an approximate C for $\Phi_w(\mathbf{r}, C)$ and $\Phi_t(\mathbf{r}, C)$ is still an open question out of the scope of this paper. And since the improvement of $\Phi_w(C)$ in [4] over Φ_w in [3] is very subtle [4], here we just consider the Φ_w and Φ_t , not $\Phi_w(C)$, $\Phi_t(C)$.

2.3 Implementation of Methods

The Implementation includes two steps. The first step is to estimate coefficients $\{a_{n,l,m}\}$ from the observed signals $\{E(\mathbf{q}_i)\}$. The second step is the uniform linear analytical solution proposed above from $\{a_{n,l,m}\}$ to $\{c_{l,m}^t\}$ and $\{c_{l,m}^w\}$, which is actually independent of the first step. The whole estimation error is just from the first step, since the second step is analytical and compact. The authors in [8] suggested two methods to estimate $\{a_{n,l,m}\}$, a linear least square (LS) fitting with regularization in the radial and spherical parts, and a non-linear PDE based optimization process, which considers the Rician noise. Here we choose the LS method in the first step since it is more faster. We suggest that the Rician correction could be performed directly on the DWI data as a pre-processing step [11][12], although in our experiments to perform an appropriate comparison of methods we did not do any Rician correction. For LS estimation, let's denote the signal vector by $E = [E(\mathbf{q}_i)]_{S \times 1}$, the coefficient vector by $A = [a_{n,l,m}]_{0.5(L+1)(L+2)(N+1) \times 1}$, the basis matrix by $M = [R_n(\mathbf{q})Y_l^m(\mathbf{u})]_{S \times 0.5(L+1)(L+2)(N+1)}$, and the spherical and radial regularization diagonal matrices respectively by $L = [l(l+1)]$ and $N = [n(n+1)]$, where λ_l and λ_n are the regularization terms for spherical and radial parts. Then $A = (M^T M + \lambda_l L^T L + \lambda_n N^T N) M^T E$. For the second step, the linear transformations in (6), (7), (15), (16) could be also implemented as a matrix multiplication. Thus the whole process is just a linear matrix multiplication on the data vector E . Similarly with analytical Q-ball in [5], the matrix is independent of E and needs to be calculated only once for the whole data set. It makes our method extremely fast.

There are two important points to consider in the implementation. The first one is about $E(0)$. If we have a data set with several b values, b_1, b_2, \dots, b_N , we actually use

$N + 1$ b values, considering $E(0) = 1$ for any $\mathbf{u} \in S^2$, which makes our estimation more reasonable and accurate. Otherwise, there is no warranty for the estimated signal $\tilde{E}(0) = 1$. For the single shell HARDI data, considering $b = 0$ can let us have 2 shells, which will improve the results. The second one is how to determine the parameter ζ in (3). The authors in [8] proposed a strategy for it, which is dependent on the radial truncation order N . However, we think the parameter should be just dependent on the signal, not on the basis order. Considering $E(\mathbf{q}) = \exp(-4\pi^2\tau q^2 D)$, $b = 4\pi^2\tau q^2$, and a typical diffusion coefficient of $D = 0.7 \times 10^{-3} \text{mm}^2/\text{s}$, a typical b-value $b = 3000 \text{s}/\text{mm}^2$, we set $\zeta = \frac{1}{8\pi^2\tau \times 0.7 \times 10^{-3}}$. If $4\pi^2\tau = 1$, then ζ is about 700. In our experiments we always set $\zeta = 700$.

3 Results on Synthetic, Phantom and Real Data

Synthetic data. Gaussian mixture model $S(\mathbf{q}_s) = \sum_{i=1}^M p_i G_i(\mathbf{q}_s)$, $G_i(\mathbf{q}_s) = \exp(-q_s^2 \mathbf{u}_s^T \mathbf{D}_i \mathbf{u}_s)$ has been used widely to generate synthetic data [11, 5, 6, 3, 4]. However, that could bias the results in favor of those methods assuming a model based on Gaussian mixture or mono/multi-exponential decay. Here we choose both Gaussian mixture and non-Gaussian mixture to validate our methods. We set $S(\mathbf{q}) = \sum_{i=1}^M p_i f_i(\mathbf{q})$, $f(\mathbf{q}) = G(\mathbf{q})$ for a Gaussian mixture model and $f(\mathbf{q}) = 0.5G(\mathbf{q}) + 0.5T(\mathbf{q})$, $T(\mathbf{q}) = \exp(-\sqrt{2}q^2 \mathbf{u}^T \mathbf{D} \mathbf{u})$ for a non-Gaussian mixture model. It could be proved that the Φ_t and Φ_w of $T(\mathbf{q})$ are the same as the ODFs of $G(\mathbf{q})$, although they have different EAPs [9]. Thus we have the analytical ground truth of ODFs. We set the eigenvalues of \mathbf{D} as $[0.3, 0.3, 1.7] \times 10^{-3} \text{mm}^2/\text{s}$. and use the same way in [5] to add Rician noise with $SNR = 1/\sigma$, which is defined as the ratio of maximal signal intensity of $S(0) = 1$ to the standard deviation σ of complex Gaussian noise. We test the methods with $SNR = 10$ in four configurations of ODF: one fiber or two orthogonal fibers with Gaussian model or non-Gaussian model. For each configuration, data in 4 shells ($b=500, 1000, 2000, 3000 \text{s}/\text{mm}^2$) were generated. For the single shell methods, e.g. Φ_t estimation in [5] and Φ_w estimation in [3], 4 order SH with $\lambda = 0.006$ was chosen. For the proposed methods, we use all data from four shells and chose $N = 2$, $L = 4$, $\lambda_n = 5e - 8$, $\lambda_l = 1e - 7$, $\zeta = 700$. We recorded the percentage of correct number of detected ODF maximum and the mean of angular error over 1000 trails [5]. See table 1. The experiments showed: 1) Normally, the Φ_t is more robust to Rician noise than Φ_w , although Φ_w is more sharper and has better angular resolution. It is the similar conclusion in [4]. 2) our methods for Φ_t and Φ_l both got better performance than the methods for single shell data in [5, 3]. 3) the method in [3] will get worse results for the data with non-exponential decay or much noise.

Phantom data. We performed our methods in a public phantom data with 3 shells, where b-value is 650, 1500, 2000 s/mm^2 respectively. This data has been used in the fiber cup contest in MICCAI 2009 to evaluate tracking methods [13]. The anisotropy of this data is very low, which makes it hard to detect the fibers. We believe that it is complex enough to evaluate different reconstruction methods and tracking methods. We compare our reconstruction method using 3 shells with the method in [3] using one shell ($b=2000$), since the result of $b=2000$ is better than the results of $b = 650$ and 1500 for the method in [3]. For our method, we choose $L = 4$, $\lambda_l = 5e - 8$ in the

Table 1. Each column shows the percentage of correct number of detected ODF maximum and the mean of angular errors under an given ODF configuration. The left part and the right part in each column are respectively for the estimation methods for Φ_t and Φ_w . The first four rows recorded the performance of previous works on single shell data in [5] and in [3] with 81 gradient directions on the hemisphere. The last row shows the results of our methods using 4 shells.

b value	1 fiber (Gaussian)	1 fiber (non-Gaussian)	2 fiber (Gaussian)	2 fiber (non-Gaussian)
500	(100%,4.8°)(91.8%,7.0°)	(100%,5.2°)(79.6%,7.6°)	(71.0%,17.5°)(57.4%,19.5°)	(74.6%,16.0°)(43.6%,18.7°)
1000	(100%,3.1°)(99.9%,4.6°)	(100%,4.0°)(95.6%,6.2°)	(88.6%,9.5°)(80.9%,10.5°)	(87.7%,11.4°)(63.0%,12.1°)
2000	(100%,2.6°)(99.9%,3.5°)	(100%,3.7°)(90.1%,4.8°)	(98.1%,6.9°)(87.5%,7.0°)	(93.1%,9.5°)(60.7%,9.9°)
3000	(100%,3.1°)(87.6%,4.0°)	(99.4%,4.6°)(52.0%,5.2°)	(95.6%,7.7°)(56.0%,7.9°)	(85.3%,11.2°)(40.4%,12.1°)
4 shells	(100%,2.0°)(100%,2.5°)	(100%,3.1°)(98.5%,3.4°)	(99.8%,4.5°)(94.7%,5.7°)	(98.5%,5.5°)(78.5%,7.5°)

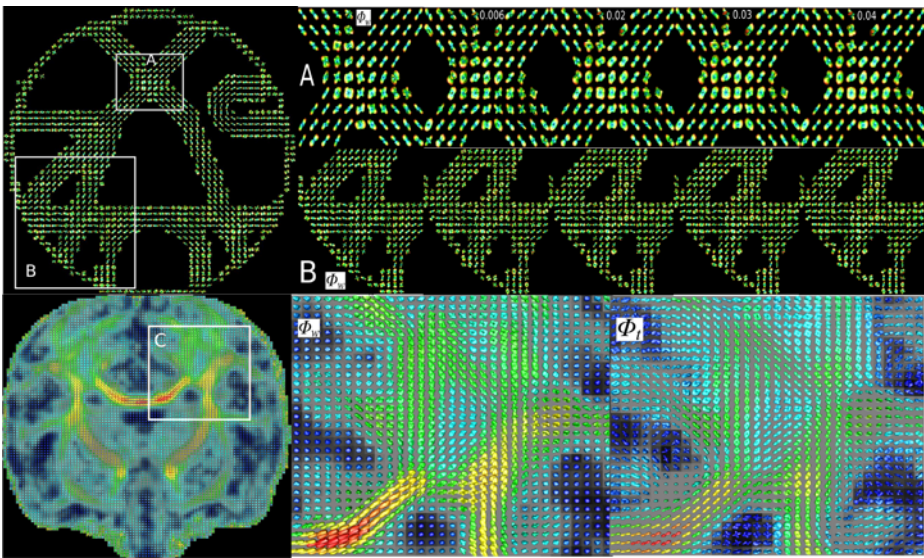


Fig. 1. First row: phantom data, from left to right: whole view of Φ_w from our method, Φ_w in region A and B from our method, method in [3] with $\lambda = 0.006, 0.02, 0.03, 0.04$; Second row: real data result from our method, from left to right: whole view of Φ_w , Φ_w and Φ_t in region C

spherical part and $N = 1, \lambda_n = 1e-9$ in the radial part [8]. To perform a fair comparison, we choose $L = 4$ and tune the Laplacian regularization term λ from 0.006 (suggested in [5][3]) to 0.02, 0.03 and 0.04 for the method in [3]. Two crossing areas were chosen for visualization using min-max normalization [1]. The results were shown in Fig. 1. It shows that the method using 3 shells is better. The bad performance of the method in [3] probably comes from the error of the mono-exponential assumption.

Real data. We perform our method in a real monkey data with 3 shells ($b = 500, 1500, 3000 \text{ s/mm}^2$), 30 directions at each b value, TE/TR/matrix=120ms/6000ms/128 × 128. We set $L = 4, N = 2, \lambda_l = 5e-8, \lambda_n = 1e-9$ and show, in Fig. 1, the results of Φ_t

with min-max normalization and Φ_w without normalization, since Φ_w is sharper than Φ_r . The glyphs were colored by GFA calculated from ODF [11].

4 Conclusion

We proposed a uniform model-free fast robust analytical ODF reconstruction method based on Spherical Polar Fourier (SPF) expression of the signal in \mathbf{q} -space. The coefficients of the two kinds of ODF under SH could be linearly and analytically calculated from the coefficients of the signal under SPF. It is a linear transformation that is independent of the data. This transformation matrix is just calculated only once for a whole data set, which makes the method very fast. Our method can avoid the error from unrealistic assumptions and can naturally combine data from different Q-shells. The results in synthetic data phantom data and real data show that our method can get better results compared with previous single shell HARDI methods in [5] and in [3], especially for the data with low anisotropy, low SNR and non-exponential decay.

Acknowledgment. This work was partly supported by the Natural Science Foundation of China (30730035), the National Key Research and Development Program of China (2007CB512305), the National High Technology Research and Development Program of China (2009AA02Z302), the External Cooperation Program of the Chinese Academy of Sciences (GJHZ200826), the French ANR “Neurological and Psychiatric diseases” NucleiPark and the France-Parkinson Association.

References

1. Tuch, D.S.: Q-ball imaging. *Magnetic Resonance in Medicine* 52, 1358–1372 (2004)
2. Wedeen, V.J., Hagmann, P., Tseng, W.Y.I., Reese, T.G., Weisskoff, R.M.: Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance In Medicine* 54, 1377–1386 (2005)
3. Aganj, I., Lenglet, C., Sapiro, G., Yacoub, E., Ugurbil, K., Harel, N.: Reconstruction of the orientation distribution function in single and multiple shell q-ball imaging within constant solid angle. *Magnetic Resonance in Medicine* (2009)
4. Tristán-Vega, A., Westin, C.F., Aja-Fernández, S.: A new methodology for the estimation of fiber populations in the white matter of the brain with the Funk-Radon transform. *NeuroImage* 49, 1301–1315 (2010)
5. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Regularized, fast and robust analytical q-ball imaging. *Magnetic Resonance in Medicine* 58, 497–510 (2007)
6. Canales-Rodriguez, E.J., Melie-Garcia, L., Iturria-Medina, Y.: Mathematical description of q-space in spherical coordinates: Exact q-ball imaging. *Magnetic Resonance In Medicine* 61, 1350–1367 (2009)
7. 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). *NeuroImage* 31, 1086–1103 (2006)
8. Assemlal, H.E., Tschumperlé, D., Brun, L.: Efficient and robust computation of PDF features from diffusion MR signal. *Medical Image Analysis* 13, 715–729 (2009)
9. Gradshteyn, I., Ryzhik, I.: *Table of Integrals, Series, and Products*. Elsevier, Amsterdam (2007)

10. Assemblal, H.E.: Diffusion MR image analysis for the estimation of tissues local architecture. PhD thesis, the University of Caen (2010)
11. Tristán-Vega, A., Aja-Fernández, S.: DWI filtering using joint information for DTI and HARDI. *Medical Image Analysis* 14, 205–218 (2010)
12. Descoteaux, M., Wiest-Daessle, N., Prima, S., Barillot, C., Deriche, R.: Impact of rician adapted non-local means filtering on HARDI. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part II. LNCS*, vol. 5242, pp. 122–130. Springer, Heidelberg (2008)
13. Poupon, C., Rieul, B., Kezele, I., Perrin, M., Poupon, F., Mangin, J.: New diffusion phantoms dedicated to the study and validation of high-angular-resolution diffusion imaging (HARDI) models. *Magnetic Resonance In Medicine* 60(6), 1276–1283 (2008)

Biomarkers for Identifying First-Episode Schizophrenia Patients Using Diffusion Weighted Imaging

Yogesh Rathi^{1,2,4}, James Malcolm², Oleg Michailovich³, Jill Goldstein¹,
Larry Seidman¹, Robert W. McCarley⁴, Carl-Fredrik Westin¹,
and Martha E. Shenton^{1,4}

¹ Harvard Medical School, Boston

² Georgia Institute of Technology, Atlanta

³ University of Waterloo, Canada

⁴ VA Clinical Neuroscience Division, Brockton

Abstract. Recent advances in diffusion weighted MR imaging (dMRI) has made it a tool of choice for investigating white matter abnormalities of the brain and central nervous system. In this work, we design a system that detects abnormal features (biomarkers) of first-episode schizophrenia patients and then classifies them using these features. We use two different models of the dMRI data, namely, spherical harmonics and the two-tensor model. The algorithm works by first computing several diffusion measures from each model. An affine-invariant representation of each subject is then computed, thus avoiding the need for registration. This representation is used within a kernel based feature selection algorithm to determine the biomarkers that are statistically different between the two populations. Confirmation of how well these biomarkers identify each population is obtained by using several classifiers such as, k-nearest neighbors, Parzen window classifier, and support vector machines to separate 21 first-episode patients from 20 age-matched normal controls. Classification results using leave-many-out cross-validation scheme are given for each representation. This algorithm is a first step towards early detection of schizophrenia.

1 Introduction

A recent World Health Organization (WHO) report estimates that nearly 1% of the population in the US is affected by schizophrenia. A growing body of evidence suggests that early detection and treatment of schizophrenia (and many other brain disorders) is critical in forming and predicting the course and outcome of the disorder [1]. The tools proposed in this work can serve as a first step towards early detection of schizophrenia, which may result in better prognosis and functional outcome. However, very little work has been done on developing a biomarker that characterizes first-episode schizophrenia or other subtle psychiatric disorders such as mild to moderate traumatic brain injury.

There has been some work done on classifying patients with **chronic** schizophrenia using structural MRI [2,3]. The authors in [4,5] use dimensionality reduction followed by linear discriminant analysis for classification of patients with schizophrenia (chronic). They, however, only use the fractional anisotropy (FA) images derived from single tensor estimation as a discriminant feature. Another related work is by [6], where

the authors use kernel methods for discriminating schizophrenia patients. Recent work has also focussed on using other imaging modalities, such as, functional MRI for detection of schizophrenia in prodromals [7]. The work presented in this paper, can provide complementary anatomical input to such fMRI based techniques for early detection of schizophrenia.

2 Our Contribution

In this work, we propose to design an algorithm that locates abnormal features of first-episode (FE) schizophrenia patients. These features are then used within a classification system to determine their potential use as biomarkers. While several studies have reported statistical differences in diffusion measures for FE patients [8,9], to the best of our knowledge, this is the first study that uses them to perform classification.

Existing work on distinguishing **chronic** schizophrenia used the single tensor model, which is known to be inadequate in regions of crossing and branching - a common configuration occurring throughout the brain [10]. In this work, we use a nonparametric spherical harmonics model [11], as well as a parametric two-tensor model [12] to detect biomarkers and perform classification of FE patients. These models can better capture multi-fiber configurations and hence the abnormality in the underlying anatomy.

Another novel aspect in this work is the use of an affine invariant probabilistic representation of each subject, which avoids the computational cost of registration along with errors due to mis-registration. Finally, we use a kernel based method [13] to locate statistically different diffusion measures (biomarkers) followed by classification of FE patients using several classifiers, i.e., Parzen window classifier, k-nearest neighbor and support vector machines.

3 Preliminaries

Schizophrenia is characterized by several symptoms such as, hallucinations, delusions, suspiciousness, etc. These occur in varying degrees in people affected by this disorder. **Chronic** schizophrenia typically implies that the patient has had psychotic symptoms for atleast five years, while **first-episode** schizophrenics are patients who have recently had their first psychotic episode (thus they are in the early stage of the development of the disease). Thus, knowing anatomical biomarkers that can reliably distinguish FE patients can be quite useful in determining the risk for prodromal subjects (subjects with high risk of schizophrenia). Recent studies have shown that around 30-40% of prodromals convert into schizophrenics. Thus, a tool that can provide the probability of a prodromal subject being anatomically close to FE patients can be immensely useful for early detection of schizophrenia. This is the main motivation behind our work.

In diffusion weighted imaging, image contrast is related to the strength of water diffusion along fiber orientation. At each image voxel, diffusion is measured along a set of distinct gradients, $\mathbf{u}_1, \dots, \mathbf{u}_n \in \mathbb{S}^2$ (on the unit sphere), producing the corresponding signal, $\mathbf{s} = [s_1, \dots, s_n]^T \in \mathbb{R}^n$. One of the simplest model that explains \mathbf{s} is the diffusion tensor model, which provides a Gaussian estimate of the fiber orientation [14]. However, this model is highly inadequate in regions of crossings and

branching fibers [15]. To overcome this limitation, several other models have been proposed [15][16][17][18][19][20][12]. Of these, we use the nonparametric spherical harmonics (SH) model of computing the orientation distribution function (ODF) [11] and the unscented Kalman filter (UKF) based two-tensor model proposed in [12]. These choices are guided by the following considerations: 1). The SH model can represent an arbitrary number of fiber configurations at each voxel, 2). The UKF based two-tensor model estimation incorporates the correlation in diffusion of water along the fiber direction, and thus is a robust estimator of the diffusion profile for one and two fiber configurations.

Spherical harmonics (SH): This nonparametric model is one of the popular techniques for estimating the diffusion ODF [11]. The method works by first computing the coefficients of the spherical harmonic (SH) basis of order L that best fits the measured signal and subsequent analytical computation to obtain the ODF. Given any bandlimited signal s defined on the sphere, one can write it as an expansion in terms of the SH basis as: $s_i = \sum_{l=0}^L \sum_{m=-l}^l c_{l,m} Y_{l,m}$, where $Y_{l,m}$ are the spherical harmonic basis functions. The above equations can be written as a linear system of equations and $c_{l,m}$ can be computed using least squares.

Filtered Two-tensor (F2T): In this case, the signal is assumed to be generated by a mixture of two Gaussians. Thus, $s_i = \sum_{j=1}^2 \exp(-b\mathbf{u}_i^T D_j \mathbf{u}_i)$, where $\{D_1, D_2\}$ are the two diffusion tensors estimated recursively within an unscented Kalman filtering framework [12]. The method works by starting tractography from the seed region and diffusion tensors are estimated as a fiber is traced from seed to termination. In this work, we perform whole brain tractography, by seeding the entire brain (except CSF). Thus, the F2T model is estimated at every voxel of the brain (in terms of the fibers that pass through each voxel), apart from CSF.

4 Methods

Validation studies have indicated the correlation between de-myelination, cellular packing, and axonal damage to diffusion measures such as fractional anisotropy (FA), trace (TR), norm (N), etc., [21][22]. Thus, these measures are potential candidates for being used as biomarkers. In the case of the SH model, generalized fractional anisotropy (GFA) and generalized norm (GN) are the diffusion measures of interest [23]. These measures can be readily computed in the SH basis as follows [1]:

$$\text{GFA} = \sqrt{\frac{n \sum_{i=1}^n (S_i - \bar{S})^2}{(n-1) \sum_{i=1}^n S_i^2}}, \quad \text{GN} = \|\mathbf{c}\|_2, \quad \text{with } \mathbf{S} = [S_1 S_2 \dots S_n],$$

where \mathbf{S} is the estimated signal using the SH model, $\mathbf{c} = [c_1 c_2 \dots]$ are the coefficients in the SH basis, and n is the number of samples. GFA captures the anisotropy of the signal, while GN measures the “size” or amount of diffusion. Note that, these measures are typically used for ODF’s, but given the linear relation between the signal and the ODF in the SH basis [11], the measures computed above are equivalent (upto a linear transformation) to those computed for ODF’s. Computing these measures directly for the signal avoids computation of the ODF.

¹ The acquired scanner signal is denoted by s , while the estimated signal is denoted by \mathbf{S} .

The F2T model allows for computation of a different set of orthogonal diffusion measures such as the FA, Mode (MD) and norm (N) [24]. These measures capture different (orthogonal) aspects of the shape of the tensor. Thus, FA measures the anisotropy while norm captures the amount of diffusion. Mode distinguishes between planar, ellipsoidal and spherical shapes. Given, a diffusion tensor D , these measures can be computed as follows:

$$N = \|D\|, \quad FA = \frac{\sqrt{3}\tilde{D}}{\sqrt{2}\|D\|}, \quad MD = 3\sqrt{6} \left| \left(\frac{\tilde{D}}{\|\tilde{D}\|} \right) \right|, \quad \tilde{D} = D - \frac{1}{3}tr(D)I$$

where, $|\cdot|$ denotes the determinant, $tr(\cdot)$ is the trace, I is the identity matrix and $\|\cdot\|$ denotes the frobenius norm of a matrix. These measures can be computed for each tensor, and thus six features are obtained.

4.1 Affine Invariant Representation

The next step in our algorithm is computing an affine invariant representation, i.e., a representation that does not change even if an affine transformation is applied to the underlying data. We achieve this by computing a probability density function (PDF) of each diffusion measure defined above. A nonparametric estimate of the PDF can be computed using the following expression [25]: $p(z) = \frac{1}{Mh} \sum_{i=1}^M G\left(\frac{z-I(x)}{h}\right)$, $z \in \{\text{Range of } I\}$, where $I(x)$ is a scalar value at spatial location x , M is the number of data points, G is a Gaussian kernel and h denotes the bandwidth of the kernel. Notice that, the spatial position x is arbitrary and applying an affine transformation to it does not change the PDF.

For each of the diffusion measures discussed earlier, we compute a PDF from values estimated throughout the brain. Thus, for the SH model, an affine invariant representation of a subject is given by: $\mathbf{P}_{sh} = [p_{gfa} \ p_{gn}]$, where p_{gfa} and p_{gn} are the PDF's of the GFA and GN respectively. The PDF's are computed at n_b bins, and thus \mathbf{P}_{sh} is an $n_b \times 2$ matrix. Similarly, for the F2T model, we compute PDF's for all the 6 diffusion measures to obtain a probabilistic representation $\mathbf{P}_{f2t} = [p_{fa1} \ p_{md1} \ p_{n1} \ p_{fa2} \ p_{md2} \ p_{n2}]$ of each subject ($n_b \times 6$ matrix).

4.2 Biomarker Detection and Classification

Once a PDF of each diffusion measure is computed for each subject, the goal is to determine, which of these measures characterize FE patients. Since each PDF is a high (n_b) dimensional vector, appropriate methods have to be used. As such, we will use the kernel based method of [13] for statistical hypothesis testing. This method has several advantages, chief among them are : **a)** It can be used with high dimensional data without sacrificing robustness and accuracy. **b)** The data need not necessarily lie in a Euclidean vector space, i.e., any type of data with an appropriate kernel can be used. **c)** This method computes statistical differences without any assumption on the distribution from which the samples are drawn (the popular t-test assumes a Gaussian distribution for the samples in each population). Thus, subtle differences can be captured using the kernel based method.

This method tests the hypothesis of two distributions being equal $p = q$. The test statistic used is the maximum mean discrepancy (MMD) between the two samples. Let

\mathcal{F} be a class of functions $f : \mathcal{X} \rightarrow \mathbb{R}$ and p, q be probability distributions (with domain \mathcal{X}), then MMD is defined as: $MMD[\mathcal{F}, p, q] = \sup_{f \in \mathcal{F}} (\mathbb{E}_{x \sim p}[f(x)] - \mathbb{E}_{y \sim q}[f(y)])$, where \mathbb{E} represents the expected value. Computing MMD involves, mapping the data to a reproducing kernel Hilbert space (RKHS) and computing the inner product (between high dimensional features) in this space using an appropriate kernel (Gaussian in our case). If MMD is greater than a certain threshold, the null hypothesis ($p = q$) is rejected. The hypothesis threshold is selected based on significance level α , typically set to 0.05.

We will use this kernel method for feature selection (biomarker detection). Thus, in the case of F2T model, each of the PDF's $p_i \in \mathbf{P}_{f2t}$ is tested, and the ones which are statistically different are used for classification. Classification involves learning a function that minimizes a particular metric so as to best separate the groups in the training data set. Several classifiers have been proposed in the literature. We will use three popular ones: Parzen window classifier [26], k-nearest neighbor [27] and support vector machines (SVM) [28]. A typical way to ensure robustness to overfitting for any classifier is to perform a leave-many-out cross validation. In this technique, a certain percentage of the available data are randomly selected (without replacement) for training the classifier. Testing is then performed on the remaining data and classification error computed. This process is repeated several thousand times and the overall performance of the classifier is computed in terms of its sensitivity and specificity.

Briefly, the entire algorithm can be summarized as follows:

Algorithm 1. Biomarker detection and classification

- 1: Compute diffusion measures and their PDF's for each subject to obtain $\mathbf{P}_{sh}, \mathbf{P}_{f2t}$.
 - 2: Randomly select $x\%$ of the subjects for training the classifier C . This process is repeated to obtain M training data sets $\{T_i\}_{i=1}^M$.
 - 3: For each training data set T_i , use the kernel based hypothesis testing method [13] to find the PDF's $p_b \in \mathbf{P}_{sh}$ or $p_b \in \mathbf{P}_{f2t}$ which are statistically different. This is the biomarker detection part of the algorithm.
 - 4: Train the classifier C on the training data set T_i using the selected features p_b (PDF's).
 - 5: Test the classifier on the remaining data (corresponding to T_i) and compute correct detection rate and false positives.
 - 6: Perform steps (3)-(5) for all training data sets $\{T_1, \dots, T_M\}$ and compute the overall sensitivity and specificity of the classifier.
-

In step 3 above, a counter (corresponding to each diffusion measure) is incremented each time a diffusion measure is selected as statistically different. Thus the measure that best characterizes a given population will be chosen frequently, thereby indicating its potential use as a biomarker.

5 Results

We applied the above algorithm for detecting biomarkers of FE schizophrenia patients. The data set consisted of 21 FE patients (17 males, 4 females, average age: 21.21 ± 4.56 years) and 20 normal controls (15 males, 5 females, average age: 22.47 ± 3.48 years)

Table 1. Classification accuracy for SH and F2T models

Classifier	SH model								F2T model							
	40%		60%		80%		98%		40%		60%		80%		98%	
	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se
knn	0.71	0.72	0.76	0.76	0.78	0.77	0.80	0.78	0.83	0.82	0.85	0.85	0.85	0.85	0.85	0.86
SVM	0.70	0.69	0.75	0.71	0.77	0.73	0.77	0.77	0.79	0.78	0.83	0.80	0.84	0.81	0.85	0.81
PWC	0.70	0.72	0.70	0.76	0.68	0.79	0.70	0.84	0.79	0.77	0.80	0.85	0.82	0.87	0.80	0.85

with the p-value for age being 0.34. dMRI data was acquired for all the subjects on a 3T scanner with 51 gradient encoding directions and 8 baseline images. The spatial resolution was $1.66 \times 1.66 \times 1.7mm^3$ with a b-value of 900.

The SH model was estimated throughout the brain, while F2T model estimation was done in terms of whole brain tractography [12]. Thus, F2T was not estimated in CSF areas, where no fiber tracts exist. Estimation of both these models was done for all 41 subjects and the proposed algorithm was applied as described before (Algorithm 1).

For the SH model, the generalized norm (GN) was consistently chosen as the biomarker by the kernel hypothesis testing method [13] for all the training data sets. The reason is evident from the plots of PDF's of all subjects (see Figure 1). In the case of F2T model, norm of the two tensors (N1, N2) were the distinguishing features that separated the two groups. No differences in FA or mode were observed. Thus, for both the models, a measure of the “amount of diffusion” was chosen as the biomarker that characterized FE patients, i.e., overall diffusivity in FE patients is higher than NC. Biologically, this implies that, in the FE patients there is demyelination of the axons or lower density of

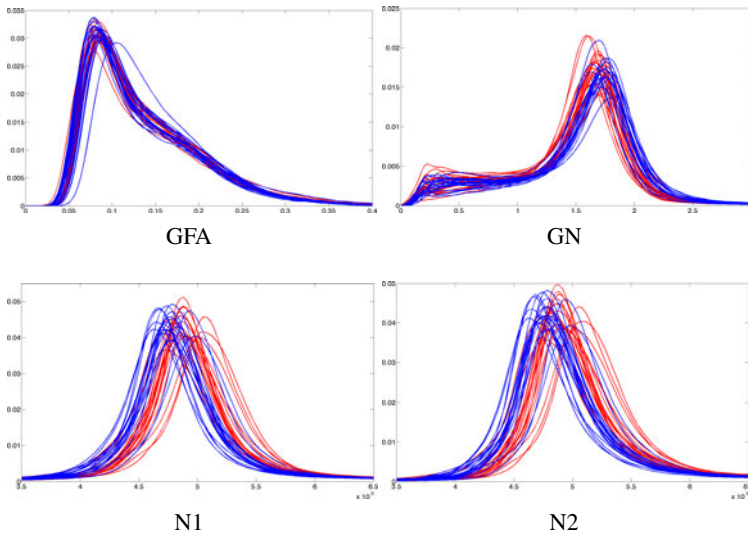


Fig. 1. PDF's of diffusion measures for 21 FE patients (red) and 20 NC (blue). Top row: GFA and GN are shown, but only GN was chosen as biomarker for the SH model. Bottom: Norms of both the tensors in F2T model show significant difference between the groups.

cellular packing, allowing for more diffusion of water. Another point to note, as evident in Figure 1, is that the tensor norms (N1, N2) measure the diffusivity, which is higher for FE patients, while the signal norm (GN) is lower for FE patients. This is because, higher diffusivity D implies lower signal S : $S(\mathbf{u}_i) = S_0 \exp(-bD(\mathbf{u}_i))$.

Table 1 gives the sensitivity (Se) and specificity (Sp) of each of the classifier's (for both models) with different number of samples ($x = \{40\%, 60\%, 80\%, 98\%\}$) in the training data set. For each x , 10000 training data sets were randomly generated from the original data and testing was done on the remaining samples. This method of cross-validation is a very good estimator of the generalization property of any classifier [29].

For the k-nearest neighbor (knn) classifier, we used 6 nearest neighbors, with cosine of the angle between the PDF's as a measure of similarity (the PDF's can be thought of as n_b dimensional vectors). In the case of SVM and Parzen window classifier (PWC), the kernel width was chosen so as to minimize the error during training.

In general, the k-nearest neighbor (knn) classifier gave the best performance for this data set. Also, the F2T model performed much better in terms of the classification accuracy than the SH model. Notice that, the performance of knn is close to optimal even when only 60% of the data is used in the training data set. Combining the features from both the models did not improve the performance of any classifier.

6 Discussion and Application

In this work, we proposed a system for detecting biomarkers of FE schizophrenia patients using two representations; spherical harmonics and two-tensor. The effectiveness of using these biomarkers to characterize FE patients was obtained by testing their classification accuracy (85% specificity and 86% sensitivity). *Future application involves using these biomarkers to determine the probability of a prodromal subject being at risk of developing schizophrenia.* This can be easily done by computing $\hat{P} = [p_{n_1}, p_{n_2}]$ for a prodromal subject and then using a nonparametric density estimator to compute the probability of \hat{P} being a FE patient. This will be the focus of our future work. Thus, the proposed method can be of great clinical significance for early detection of schizophrenia.

Acknowledgements

This work was supported in part by a Department of Veteran Affairs Merit Award (Dr. M Shenton, Dr. R McCarley), the VA Schizophrenia Center Grant (RM, MS) and NIH grants:1P50MH080272-01 (MS,RM), P41 RR13218 (MS), R01 MH 52807 (RM), R01 MH 50740 (MS), R01MH074794 (Dr. Westin) and NA-MIC (NIH) grant U54 GM072977-01 (Dr. Ron Kikinis).

References

1. McGlashan, T.: Early detection and intervention in schizophrenia: editors introduction. *Schizophr Bull* 22(2), 197–199 (1996)
2. Davatzikos, C., Shen, D., Gur, R., Wu, X., Liu, D., Fan, Y., Hughett, P., Turetsky, B., Gur, R.: Whole-brain morphometric study of schizophrenia revealing a spatially complex set of focal abnormalities. *Archives of general psychiatry* 62(11), 1218–1227 (2005)

3. Pohl, K.M., Sabuncu, M.R.: A unified framework for mr based disease classification. In: Prince, J.L., Pham, D.L., Myers, K.J. (eds.) IPMI 2009. LNCS, vol. 5636, pp. 300–313. Springer, Heidelberg (2009)
4. Caan, M., Vermeer, K., van Vliet, L., Majoie, C., Peters, B., den Heeten, G., Vos, F.: Shaving diffusion tensor images in discriminant analysis: A study into schizophrenia. *Medical Image Analysis* 10(6), 841–849 (2006)
5. Caprihan, A., Pearlson, G., Calhoun, V.: Application of principal component analysis to distinguish patients with schizophrenia from healthy controls based on fractional anisotropy measurements. *Neuroimage* 42(2), 675–682 (2008)
6. Khurd, P., Verma, R., Davatzikos, C.: Kernel-based manifold learning for statistical analysis of diffusion tensor images. In: Karssemeijer, N., Lelieveldt, B. (eds.) IPMI 2007. LNCS, vol. 4584, p. 581. Springer, Heidelberg (2007)
7. Schobel, S., Lewandowski, N., Corcoran, C., Moore, H., Brown, T., Malaspina, D., Small, S.: Differential targeting of the CA1 subfield of the hippocampal formation by schizophrenia and related psychotic disorders. *Archives of General Psychiatry* 66(9), 938 (2009)
8. Friedman, J., Tang, C., Carpenter, D., Buchsbaum, M., Schmeidler, J., Flanagan, L., Golembos, S., Kanellopoulou, I., Ng, J., Hof, P., et al.: Diffusion tensor imaging findings in first-episode and chronic schizophrenia patients. *American Journal of Psychiatry* 165(8), 1024 (2008)
9. Kubicki, M., McCarley, R., Westin, C.F., Park, H.J., Maier, S., Kikinis, R., Jolesz, F., Shenton, M.: A review of diffusion tensor imaging studies in schizophrenia. *J. of Psychiatric Research* 41, 15–30 (2007)
10. Behrens, T., Johansen-Berg, H., Jbabdi, S., Rushworth, M., Woolrich, M.: Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage* 34, 144–155 (2007)
11. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Regularized, fast, and robust analytical Q-ball imaging. *Magnetic Resonance in Medicine* 58, 497–510 (2007)
12. Malcolm, J., Shenton, M., Rathi, Y.: Neural tractography using an unscented Kalman filter. In: Prince, J.L., Pham, D.L., Myers, K.J. (eds.) *Information Processing in Medical Imaging*. LNCS, vol. 5636, pp. 126–138. Springer, Heidelberg (2009)
13. Gretton, A., Borgwardt, K., Rasch, M., Scholkopf, B., Smola, A.: A kernel method for the two-sample-problem. *Journal of Machine Learning Research* 1, 1–10 (2008)
14. Basser, P., Mattiello, J., LeBihan, D.: MR diffusion tensor spectroscopy and imaging. *Biophysical Journal* 66(1), 259–267 (1994)
15. Tuch, D.: Q-ball imaging. *Magnetic Resonance in Medicine* 52, 1358–1372 (2004)
16. Anderson, A.: Measurement of fiber orientation distributions using high angular resolution diffusion imaging. *Magnetic Resonance in Medicine* 54(5), 1194–1206 (2005)
17. Jian, B., Vemuri, B.: A unified computational framework for deconvolution to reconstruct multiple fibers from diffusion weighted MRI. *TMI* 26(11), 1464–1471 (2007)
18. Jansons, K., Alexander, D.: Persistent angular structure: New insights from diffusion MRI data. *Inverse Problems* 19, 1031–1046 (2003)
19. Tournier, J.D., Calamante, F., Gadian, D., Connelly, A.: Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage* 23, 1176–1185 (2004)
20. Barmpoutis, A., Hwang, M.S., Howland, D., Forder, J., Vemuri, B.: Regularized positive-definite fourth order tensor field estimation from DW-MRI. *Neuroimage* (2009)
21. Budde, M., Kim, J., Liang, H., Schmidt, R., Russell, J., Cross, A., Song, S.: Toward accurate diagnosis of white matter pathology using diffusion tensor imaging. *Magnetic resonance in medicine* 57(4), 688 (2007)
22. Chenevert, T., Brunberg, J., Pipe, J.: Anisotropic diffusion in human white matter: demonstration with MR techniques in vivo. *Radiology* 177(2), 401–405 (1990)

23. Ozarslan, E., Vemuri, B., Mareci, T.: Generalized scalar measures for diffusion MRI using trace, variance, and entropy. *Magnetic Resonance in Medicine* 53(4), 866–876 (2005)
24. Kindlmann, G., Ennis, D., Whitaker, R., Westin, C.: Diffusion tensor analysis with invariant gradients and rotation tangents. *TMI* 26(11), 1483–1499 (2007)
25. Parzen, E.: On estimation of a probability density function and mode. *The annals of mathematical statistics*, 1065–1076 (1962)
26. Jain, A., Ramaswami, M.: Classifier design with Parzen windows. *Pattern Recognition and artificial intelligence*, 211–228 (1988)
27. Cover, T., Hart, P.: Nearest neighbor pattern classification. *IEEE Transactions on Information Theory* 13(1), 21–27 (1967)
28. Schölkopf, B., Burges, C., Smola, A.: *Advances in kernel methods: support vector learning*. MIT Press, Cambridge (1999)
29. Picard, R., Cook, R.: Cross-validation of regression models. *Journal of the American Statistical Association* 79(387), 575–583 (1984)

A Novel White Matter Fibre Tracking Algorithm Using Probabilistic Tractography and Average Curves

Nagulan Ratnarajah¹, Andrew Simmons^{2,3}, Oleg Davydov⁴, and Ali Hojjatoleslami¹

¹ Medical Image Computing, School of BioSciences, University of Kent, U.K.

² Neuroimaging Department, Institute of Psychiatry, Kings College London, U.K.

³ NIHR Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust and Institute of Psychiatry, King's College London, U.K.

⁴ Department of Mathematics and Statistics, University of Strathclyde, Glasgow, U.K.

Abstract. This paper presents a novel white matter fibre tractography approach using average curves of probabilistic fibre tracking measures. We compute “representative” curves from the original probabilistic curve-set using two different averaging methods. These typical curves overcome a number of the limitations of deterministic and probabilistic approaches. They produce strong connections to every anatomically distinct fibre tract from a seed point and also convey important information about the underlying probability distribution. A new clustering algorithm is employed to separate fibres into branches before applying averaging methods. The performance of the technique is verified on a wide range of seed points using a phantom dataset and an *in vivo* dataset.

1 Introduction

Fibre tractography using diffusion tensor magnetic resonance imaging (DT-MRI) is a promising method for reconstructing the pathways of white matter fasciculi in the human brain noninvasively. A variety of algorithms have been proposed aiming to generate fibre-tract trajectories [1-5]. Generally these algorithms can be categorised into two main types, deterministic and probabilistic. Deterministic approaches are capable of creating anatomically reliable reconstructions of major white matter tracts. However, they do not correctly deal with branching of white matter tracts as such techniques produce only one path per seed point and there is no measure describing the confidence or uncertainty of the reconstructed trajectories. Probabilistic tractography algorithms have been developed to overcome these shortcomings.

The aim of probabilistic tracking methods is to provide a natural approach for modelling uncertainty and generate multiple curves originating from a seed point. Probabilistic methods have also been developed to resolve fibre crossings at the intravoxel level [6] and these methods allow branching of white matter tracts. However, the deterministic tractography approaches have several advantages over probabilistic tractography for some applications such as neurosurgery. Firstly, visualization of the deterministic streamline trajectories is similar to the expected *in vivo* white matter fibre tracts, whereas the output of probabilistic methods is a connectivity map, which is not a single well-defined trajectory, but rather a spatial distribution. These connectivity maps contain dense 3D volumes of potential connectivities, which cannot be easily inspected. The determination of a connectivity map is also a time-consuming process and requires

large amounts of memory. Connectivity maps derived using frequency of connection methods demonstrate high frequency connections close to the seed point and low frequency connections at distance from the seed point. This can lead to difficulty in interpreting tracking results, because the derived connection probabilities are not comparable at different distances from the seed point [7]. The connectivity maps from probabilistic tractography are no more than an indication of the number of times that a range of trajectories pass through a voxel from the seed point. The aim of any tractography algorithm is to reconstruct tracts that accurately correlate with the underlying white matter pathways. Secondly, output tracts from probabilistic methods can leak into unexpected regions producing incorrect white matter connections [8].

In this study, we present a novel fibre tractography algorithm using average curves of the output of a probabilistic method, which overcomes a number of the limitations of deterministic and probabilistic tractography techniques described above. The performance of the algorithm has been evaluated using images of a phantom and an *in vivo* data.

2 Methodology

We generate a number of tracts from a seed point based on a probabilistic algorithm; we then divide the curves into two sections, forward and backward from the seed point, then each of the two sets of curves are separated further using a clustering algorithm to find all branch sets. Curves that are very short/long compared to the average arc-length of curves of each branch and curve branches with small number of tracts are deleted from the set. Finally the average curves methods were applied for each set of branched curves from the seed point and the resultant curves concatenated. The tractography algorithm connecting the above steps is described in Algorithm 1. Figure 1 illustrates the concept of our method with the example of a seed point from the phantom data using random walk curves and mean averaging.

We investigated the use of three probabilistic algorithms (a wild bootstrapping [5], a Bayesian [4] and a Random walk [3]) in this work. Two pair-wise distances between curves γ_i and γ_j were implemented for clustering and median averaging algorithms:

1. The Hausdorff distance: $H(\gamma_i, \gamma_j) = \max [H'(\gamma_i, \gamma_j), H'(\gamma_j, \gamma_i)],$ (1)

$$\text{where } H'(\gamma_i, \gamma_j) = \max_{x \in \gamma_i} \min_{y \in \gamma_j} \|x - y\|, \quad (2)$$

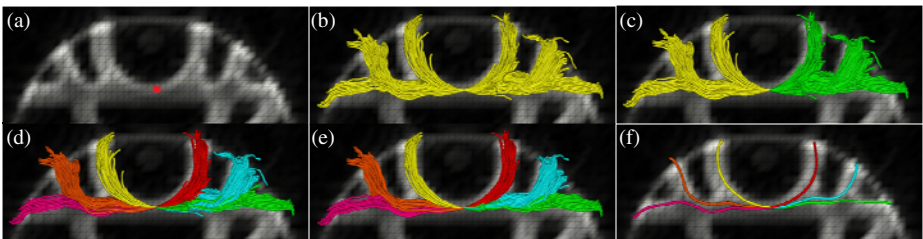


Fig. 1. Overview of the framework of our tractography method (a) Seed point, (b) probabilistic tracking from the seed point, (c) forward and backward tracks, (d) clustered branches, (e) branches after pre-processing and (f) average curves

2. The average closest distance: $G(\gamma_i, \gamma_j) = \text{mean} \left(G'(\gamma_i, \gamma_j), G'(\gamma_j, \gamma_i) \right)$, (3)
 where $G'(\gamma_i, \gamma_j) = \text{mean} \min_{x \in \gamma_i, y \in \gamma_j} \|x - y\|$ (4)

Here, we define the mathematical framework for representing and averaging probabilistic fibre tracking curves in \mathbb{R}^3 . Let Γ be a set of N probabilistic fibre tracking curves from a seed point, where each curve γ_i is defined by a set of k_i ordered points in $m = 3$ dimensions, represented by a $k_i \times m$ matrix.

$\Gamma = \{\gamma_i, 1 \leq i \leq N, \gamma_i \in \mathcal{M}_{k_i, m}\}$, Here $\mathcal{M}_{n, m}$ denotes the set of all $n \times m$ matrices.

2.1 Clustering

We developed a distance based divisive hierarchical clustering scheme that uses various curve distance metrics to find all branches. The fibre tracking algorithm and the

Algorithm 1. The tractography algorithm.

Given a DTI volume, a probabilistic method and predefined parameters: l for branching and r, t, c for pre-processing.

1. Using a given probabilistic tracking method, generate N tracts as a set of curves Γ from a seed point.
2. Divide the curves Γ based on their direction forward and backward from the seed point.

$$\Gamma_1 = \{\gamma_i^1, 1 \leq i \leq N, \gamma_i^1 \in \mathcal{M}_{k_i^1, m}\} \text{ and } \Gamma_2 = \{\gamma_i^2, 1 \leq i \leq N, \gamma_i^2 \in \mathcal{M}_{k_i^2, m}\},$$

where $k_i = k_i^1 + k_i^2 - 1$

3. Define the branch sets $\Gamma_{1,p}, p = 1 \dots P$ and $\Gamma_{2,q}, q = 1 \dots Q$ for Γ_1 and Γ_2 , using the clustering algorithm (threshold l), where P and Q are the number of branches.

$$\Gamma_{1,p} = \{\gamma_i^{1,p}, 1 \leq i \leq N_{1,p}, \gamma_i^{1,p} \in \mathcal{M}_{k_i^{1,p}, m}\}, \sum N_{1,p} = N \text{ and}$$

$$\Gamma_{2,q} = \{\gamma_i^{2,q}, 1 \leq i \leq N_{2,q}, \gamma_i^{2,q} \in \mathcal{M}_{k_i^{2,q}, m}\}, \sum N_{2,q} = N$$

4. Pre-processing
 - a. Delete branch set if number of curves in branch $< r\%$ of N
 - b. Delete curves in every branch set if the arc-length of the curve $< t\%$ or $> c\%$ of the average arc-length of the curves in the branch
 - c. Output: $\Gamma_b = \{\gamma_i^b, 1 \leq i \leq N_b, \gamma_i^b \in \mathcal{M}_{k_i^b, m}\}$,
 $b = 1, \dots, B$, where $\sum N_b \leq N$ and $B \leq P + Q$
5. Estimate average curve γ_{avg}^b for each branch set Γ_b and concatenate each resulting average curve γ_{avg}^b giving output γ_s from seed point S .
6. Repeat from step 1 for a new seed point.

separation at the seed point provides a set $\Gamma_k, k = 1$ or 2 of 3D curves γ_i . Given a pair-wise distance d and a fibre γ_i , d is computed between γ_i and γ_j for all γ_j in $\Gamma_k, j \neq i$. A similarity matrix D is defined which organizes the pair-wise distances d between each pair of curves. Matrix D is then used in a classical divisive hierarchical clustering algorithm, described by the following steps.

- 1) Clustering algorithms begins with all the curves as a single cluster.
- 2) Select the maximum value in the D and find the corresponding curves γ_{c1} and γ_{c2} .
- 3) Divide the curves $\gamma_i, i \neq c1, c2$, into two clusters, one related to γ_{c1} and the other related to γ_{c2} using a measure $Min(d(\gamma_i, \gamma_{c1}), d(\gamma_i, \gamma_{c2}))$.

Repeat steps 2 and 3 for every cluster and successively divide the curves into a hierarchy of smaller and smaller clusters until the maximum value $< l$, where l is a threshold to be chosen.

2.2 Average Curves

We consider a representative curve from a given collection of curves in space as the average curve of the collection. Clearly the representative curve needs to be as close as possible to all curves in the collection. This can be achieved by ensuring that the average curve is that which minimises the difference from all the other curves. Two types of average measures are implemented for this work.

Mean Curve: The mean curve is calculated using an arc-length re-parameterisation method, which re-parameterises the curves by placing a high number of points at equal arc-length steps on each curve. In our implementation, we used a parameter for deciding the constant arc-length step based on the average arc-length of the set of curves in the branch to reparameterize the curves. We consider the collection of parametric curves as being an independent realization of a probabilistic process $\Gamma(\mathbf{t})$ that has mean $E\{\Gamma(\mathbf{t})\} = \mu(\mathbf{t})$. A random curve from the population may then be expressed as

$$\Gamma(\mathbf{t}) = \mu(\mathbf{t}) + \varepsilon(\mathbf{t}) \tag{5}$$

where $\varepsilon(\mathbf{t})$ are independent and $E(\varepsilon(\mathbf{t})) = 0$. For non parametrical estimation of the overall mean function $\mu(\mathbf{t})$, we use the least squares estimate of μ , which is obtained by averaging the data values separately at each parameter value \mathbf{t} .

$$\hat{\mu}(\mathbf{t}) = \frac{1}{p(\mathbf{t})} \sum_{i=1}^{p(\mathbf{t})} x_i(\mathbf{t}), \quad i = 1, 2, \dots, \text{Max} \tag{6}$$

where $\mu \in \mathcal{M}_{Max,m}$, $p(\mathbf{t})$ is the number of curves involved in the calculation and Max is the maximum number of points placed in the reparameterization process.

Median Curve: The median curve is selected from the collection as the curve which differs least from all other curves. This is computationally easier than constructing a new curve. We use the matrix of pair-wise distances D , described in section 2.1, to identify the best curve(s) from the set of likely curves in the branch-set. We use a hierarchical algorithm to find the median curve with the following steps.

- 1) Select the maximum value in the D and find the correspondent curves γ_{C1} and γ_{C2} .
- 2) Remove the two curves γ_{C1} and γ_{C2} from the set.
- 3) Repeat steps 1 and 2 for every remaining curve-set and successively remove the curves into a hierarchy of smaller and smaller numbers of curves until one or two curves remain in the set.

If the number of resultant curves is two, the mean curve approach described above is applied to the resultant curves to produce a single curve.

2.3 Data Acquisition

Diffusion-weighted data were acquired from a physical phantom [9] on a 3T MRI system with $3 \times 3 \times 3 \text{ mm}^3$ voxel resolution, b value=1500 s/mm² and 64 diffusion directions. We applied the average curves algorithm to the phantom data, using three probabilistic tracking methods, described above, from 16 pre-defined seed positions with 1000 iterations. The same fibre structures were then extracted using the FACT [2] deterministic algorithm using the same parameters.

Quantitative Error Analysis: The ground truth curves γ_T is the ideal trajectory and γ_S the resultant curves using the average curves method of most probable branches and FACT method from the 16 seed points of the phantom data. The two most probable branches from the forward and backward curve sets are selected from the set of different branches from the seed point as that containing the highest number of curves. Performance measures (ξ) were calculated using the average closest distance G' from γ_S to γ_T , and the Hausdorff distance H' from γ_S to γ_T .

Diffusion weighted images were acquired from a healthy human on a 1.5 Tesla scanner with 64 diffusion encoding gradients and a b-value of 1000 s/mm² ($2 \times 2 \times 2 \text{ mm}^3$ voxels). Random trajectories were initiated from a seed point in the corpus callosum and two points in the right/left internal capsule using Bayesian and wild bootstrapping probabilistic methods respectively.

3 Results

Figure 2 shows the typical results of average curves of the most probable branches for each seed point of the phantom data. The average curve results are consistent with the ground truth except where some paths meet crossing regions. A comparison of our method with an implementation of the FACT deterministic method using the same seed points showed that our algorithm is more robust in the presence of complex pathways. The FACT trajectories show unusual tract behaviour, sharp bends and loops, while the average curve trajectories generally did not. The lengths of the resulting FACT tracking curves were also smaller than the average curve lengths.

Table 1 presents mean and standard deviation values for the performance measures (ξ) for the 16 pre-defined seed points of the phantom data. The average curve methods are more accurate than the streamline tracking methods with significantly lower errors. The mean curve results are considerably lower than the results of the two distance median curves. Here we used asymmetric measures to estimate the error, in order to avoid accumulation of errors due to early stopping curves or mis-directed paths.

Table 1. The performance measures (ξ) in *mm* (F -FACT method, D-Distance, A-Average closest distance, H-Hausdorff distance, W-Wild bootstrap, B-Bayesian, R-Random Walk).

D	F	Mean curve			Median Curve(A)			Median Curve(H)		
		W	B	R	W	B	R	W	B	R
A	6.25 ± 7.6	2.82 ± 1.3	2.95 ± 1.6	3.18 ± 1.9	3.08 ± 0.8	3.31 ± 1.6	3.67 ± 2.1	3.51 ± 2.1	3.96 ± 1.6	4.08 ± 2.9
H	17.6 ± 27.8	8.75 ± 8.3	9.74 ± 8.2	11.07 ± 10.1	8.73 ± 7.6	11.04 ± 10.2	13.04 ± 11.4	11.07 ± 11.1	11.08 ± 11.1	14.63 ± 14.1

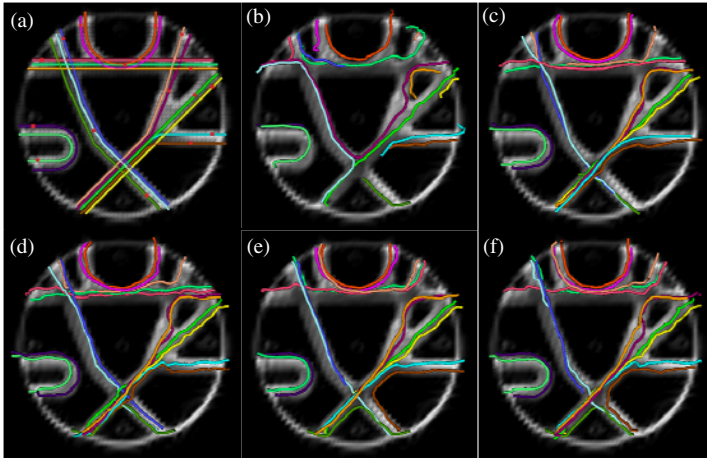


Fig. 2. (a) Ground truth and seed points and results of (b) FACT (c) mean and (d) median curves (A) of wild bootstrapping and (e) mean and (f) median curves (A) of Bayesian tracking

The desired results observed from performing clustering of the corpus callosum (Figure 3) are its division into anatomic regions according to the fibre projections. Mean and median curves of the clusters show representative connections from the seed point to different regions. Using more than one seed point will increase the number of branches and average curves. However, we have illustrated the results using a single seed point in order to provide a clear view of the average curves and to simplify the evaluation.

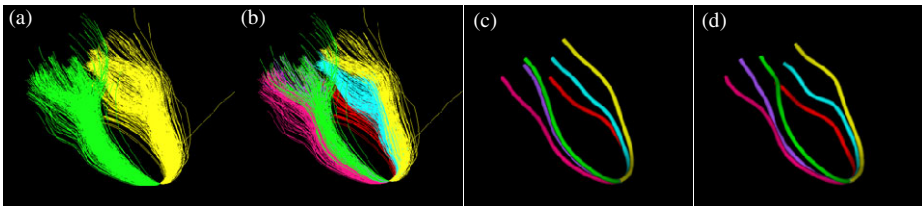


Fig. 3. Tractography results using Bayesian probabilistic tracking from a seed point in the corpus callosum (a) Forward and backward tracts (b) clustered tracts (c) mean curves (d) median curves (A)

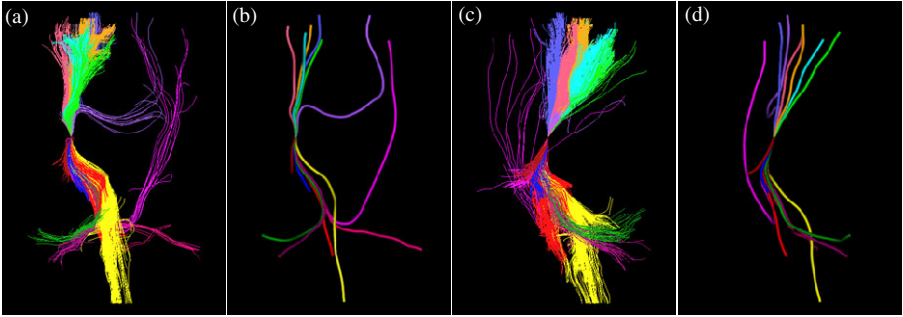


Fig. 4. Tractography results using wild bootstrap tracking obtained from a seed point placed in the cortico-spinal tract. (a), (c) Clustered trajectories and (b), (d) mean curves. (a), (b) front view and (c), (d) lateral view.

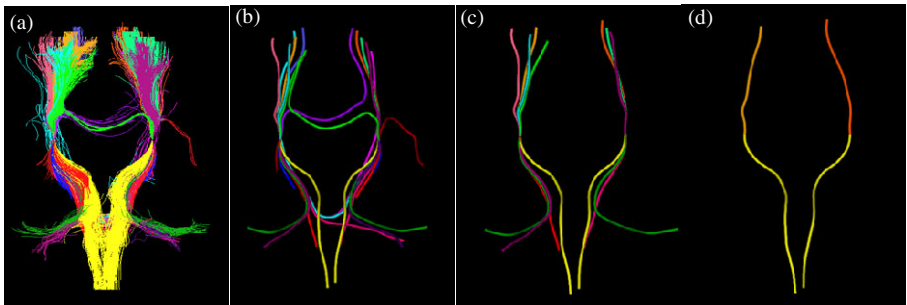


Fig. 5. Tractography results using wild bootstrap tracking obtained from a seed point placed in the left/right cortico-spinal tract from a front view. (a) Clustered trajectories, (b) mean curves, (c) mean curves of branches which contain more than 20% of the total number of generated curves and (d) mean curve of most probable branches.

The estimated corticospinal tracts (Figure 4 and 5) propagate inferiorly through the internal capsule and the results show that our average curves properly reconstruct the fibres to the different motor areas. The clustering result shows that the tracts were grouped into different plausible bundles. However, the outputs of the wild bootstrapping method showed that a number of deterministic curves erroneously cross the pons and project into the contralateral hemisphere. Figures 5 (c) and (d) show that anatomically implausible pathways are mostly represented by branches with low number of curves. When thresholded at $> 20\%$ number of curves, the results demonstrate plausible routes.

4 Discussion

In this paper, we have introduced a novel fibre tractography method based on average curves of probabilistic methods, which results in a single well-defined trajectory for every strong anatomically distinct connection from a seed point by combining the

advantages of both deterministic and probabilistic algorithms. We evaluated the average curves algorithm on both phantom and *in vivo* data. The overall shape of the fibre tract trajectories has been shown to correspond with known anatomy, providing quantitatively useful data. However, there are conceptual and practical issues that must be understood when choosing this approach. The results of our average curves approach depend on the probabilistic methods which we applied. While the probabilistic methods still suffer from the general problems of tractography, relying only on the overall shape of trajectories gives an inherent degree of protection against the effects of noise and partial volume.

The accuracy of fibre tractography is influenced by the diffusion tensor measurements' sensitivity to image noise and various other factors. Several studies have investigated the accuracy of deterministic tractography algorithms. We have used average curves as a tool to analyse errors in these probabilistic methods.

The results of our tractography algorithm show that the method handles branching correctly and addresses many of the difficulties faced by traditional probabilistic methods. The average curves were also shown to be good representations of optimal fibre paths of strong connections using both phantom and *in vivo* data. We have also presented new techniques for clustering probabilistic curves in 3D, to find anatomically distinct branches and remove outlier curves. The average curve methods we describe here are fast and relatively easy to implement. One limitation of our study is that we use some semi-automated parameters to identify the short curves and separate the curves from generated curves.

References

1. Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A.: In vivo fiber tractography using DT-MRI data. *Magn. Reson. Med.* 44, 625–632 (2000)
2. Mori, S., Crain, J., Chacko, V.P., Van Zijl, P.C.M.: Three dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *An. Ne.* 45(2), 265–269 (1999)
3. Hagmann, P., Thiran, J.P., Jonasson, L., Vandergheynst, P., Clarke, S., Maeder, P., Meuli, R.: DTI mapping of human brain connectivity: statistical fibre tracking and virtual dissection. *NeuroImage* 19(3), 545–554 (2003)
4. Friman, O., Farnback, G., Westin, C.F.: A Bayesian approach for stochastic white matter tractography. *IEEE Trans. Med. Imaging* 25(8), 965–978 (2006)
5. Jones, D.K.: Tractography gone wild: probabilistic fibre tracking using the wild bootstrap with diffusion tensor MRI. *IEEE Trans. Med Imaging* 27, 1268–1274 (2008)
6. Behrens, T.E., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., Clare, S., Matthews, P.M., Brady, J.M., Smith, S.M.: Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn. Reson Med.* 50(5), 1077–1088 (2003)
7. Morris, D., Embleton, K., Parker, G.: Probabilistic fibre tracking: differentiation of connections from chance events. *NeuroImage* 42, 1329–1339 (2008)
8. Descoteaux, M., Deriche, R., Lenglet, C.: Diffusion Tensor Sharpening Improves White Matter Tractography. *SPIE Image Processing: Medical Imaging, USA* (February 2007)
9. Poupon, C., Rieul, B., Kezele, I., Perrin, M., Poupon, F., Mangin, J.F.: New diffusion phantoms dedicated to the study and validation of high-angular-resolution diffusion imaging (HARDI) models. *Magnetic Resonance in Medicine* 60(6), 1276–1283 (2008)

Multi-Diffusion-Tensor Fitting via Spherical Deconvolution: A Unifying Framework

Thomas Schultz¹, Carl-Fredrik Westin², and Gordon Kindlmann¹

¹ Computer Science Department and Computation Institute,
University of Chicago, Chicago IL, USA

² Laboratory of Mathematics in Imaging, Brigham and Women’s Hospital,
Harvard Medical School, Boston MA, USA*

Abstract. In analyzing diffusion magnetic resonance imaging, multi-tensor models address the limitations of the single diffusion tensor in situations of partial voluming and fiber crossings. However, selection of a suitable number of fibers and numerical difficulties in model fitting have limited their practical use. This paper addresses both problems by making spherical deconvolution part of the fitting process: We demonstrate that with an appropriate kernel, the deconvolution provides a reliable approximative fit that is efficiently refined by a subsequent descent-type optimization. Moreover, deciding on the number of fibers based on the orientation distribution function produces favorable results when compared to the traditional F-Test. Our work demonstrates the benefits of unifying previously divergent lines of work in diffusion image analysis.

1 Introduction

The diffusion tensor model [1] is widely used for analyzing data from diffusion weighted magnetic resonance imaging (DW-MRI), but is inadequate in situations of partial voluming and fiber crossings. Multi-compartment models provide a natural extension by combining multiple diffusion tensors. They have been used to study the effects of partial voluming [2], and to analyze the diffusion weighted signal in voxels with multiple fiber contributions [3].

Fitting multi-tensor models requires nonlinear optimization, for which previous work has used descent-type algorithms [3,4,5,6]. However, these methods only find the global optimum when provided with an initial guess that is sufficiently close to the final solution. At the state of the art, numerous randomized restarts are tried to reach a suitable optimum with high probability [3,7,6], which incurs an excessive computational cost. Alternatively, regularization over spatial neighborhoods [8] requires numerical solution of a partial differential equation.

We present a novel way of applying spherical deconvolution [9] to “kick-start” model fitting, as summarized in Figure 1. We derive a deconvolution kernel to

* The authors would like to thank Alfred Anwander (MPI CBS, Leipzig, Germany) for providing data used in this work. T.S. is supported by a fellowship within the Postdoc Program of the German Academic Exchange Service (DAAD). C.-F.W. is supported by NIH R01MH074794, NIH P41RR013218.

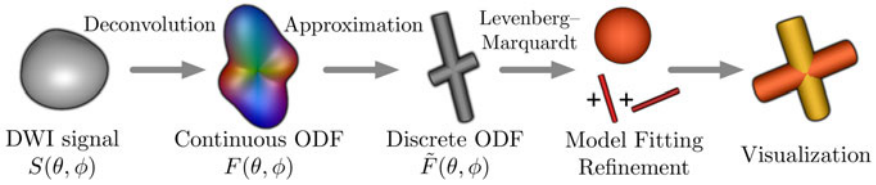


Fig. 1. In our framework, deconvolution and ODF approximation make Levenberg-Marquardt fitting of ball-and-stick models faster and more reliable

approximate the ball-and-stick model [10,11], a common variant of the multi-tensor model. The deconvolution result is analyzed by a recent method for fitting discrete models to continuous orientation distribution functions [12]. This decides the number of fibers and provides a starting point to make subsequent optimization [13,14,15,16] more reliable and efficient.

Diffusion image analysis approaches based on spherical deconvolution have generally been distinct from those involving explicit fitting of multi-fiber models. Our main contribution is showing how these two approaches can be adapted and combined to create a unified algorithmic solution that offers advantages of both.

2 Related Work

2.1 Multi-tensor Models

Multi-tensor models assume k diffusion compartments with little or no exchange during measurement time [2], each parametrized by a symmetric 3×3 diffusion tensor \mathbf{D}_i per compartment, with eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$. The signal fractions $f_i \in [0, 1]$ sum to unity. With this, the signal $S(\mathbf{g})$ is predicted as

$$S(\mathbf{g}) = S_0 \sum_{i=1}^k f_i e^{-b\mathbf{g}^T \mathbf{D}_i \mathbf{g}}, \tag{1}$$

where k is the number of compartments, S_0 is the non diffusion-weighted signal, b is the diffusion weighting, and \mathbf{g} is the diffusion-sensitizing gradient.

A full k -tensor model has $7k - 1$ degrees of freedom, but additional constraints are imposed in practice. When a single non-zero b -value is used, a mathematical indeterminacy prevents simultaneous estimation of the isotropic part of \mathbf{D}_i and its volume fraction f_i [7]. Most authors constrain the isotropic part, for example by assuming equal eigenvalues on all \mathbf{D}_i [3,14,5], but some prefer to fix the f_i instead [7,15]. A very common assumption is axial symmetry ($\lambda_2 = \lambda_3$) [3,7,8,10,14,15,16]. In many cases, k is limited to $k \leq 2$ [3,7,16,5].

In this work, we focus on the ball-and-stick model [10], which assumes that all \mathbf{D}_i have equal λ_1 ; a single “ball” compartment is completely isotropic ($\lambda_1 = \lambda_2 = \lambda_3$), the remaining “stick” compartments are perfectly linear ($\lambda_2 = \lambda_3 = 0$). For n fiber terms, this leads to $k = n + 1$ compartments and $3n + 1$ degrees of freedom. We consider the model up to $n = 3$.

2.2 Spherical Deconvolution

Rather than assuming a fixed number of compartments, spherical deconvolution reconstructs an orientation distribution function (ODF) $F(\theta, \phi)$, which specifies a continuous density of volume fractions on the unit sphere. The predicted signal $S(\theta, \phi)$ is then defined as the convolution of the ODF $F(\theta, \phi)$ with an axially symmetric single-fiber response function $R(\gamma)$,

$$S(\theta, \phi) = \int_0^{2\pi} \int_0^\pi F(\theta', \phi') R(\gamma') \sin(\theta') d\theta' d\phi' \quad (2)$$

where γ' is the angle between directions given by (θ, ϕ) and (θ', ϕ') . Typically, $S(\theta, \phi)$ and $R(\gamma)$ are estimated from the data and modeled in spherical harmonics and rotational harmonics, respectively. This reduces spherical deconvolution to simple scalar division, and yields $F(\theta, \phi)$ [9].

It has been pointed out [10] that Equation (2) describes a continuous version of the ball-and-stick model when substituting

$$R(\gamma) = S_0 e^{-bd \cos^2 \gamma} \quad (3)$$

and adding $S_0 f_{\text{iso}} \exp(-bd)$ outside the integral, where f_{iso} is the volume fraction of the ball compartment. We use this relation to compute an ODF $F(\theta, \phi)$ that corresponds to the continuous ball-and-stick model. FORECAST [14] has followed a similar goal, though corresponding to a multi-tensor model with non-zero perpendicular diffusivity ($\lambda_2 = \lambda_3 > 0$) in the individual compartments.

2.3 Discrete Approximations of Continuous ODFs

From a continuous ball-and-stick model, the discrete case is recovered by replacing the continuous function $F(\theta, \phi)$ with a discrete ODF $\tilde{F}(\theta, \phi)$ that is a finite sum of weighted delta peaks. Even though it is common to recover discrete directions by locating maxima in $F(\theta, \phi)$ [9,17], this is not accurate. It ignores the fact that delta peaks represented by finite-order spherical harmonics no longer have negligible width, so the individual peaks interfere with each other.

Therefore, we employ nonlinear optimization based on higher-order tensor representations to find a discrete approximation $\tilde{F}(\theta, \phi)$, as described in [12]. Selecting the maximum spherical harmonics order involves a tradeoff between increasing peak sharpness and reducing the influence of noise. Unlike maximum extraction, the optimization in [12] explicitly accounts for the blurring of ODF peaks at low orders. We found that because of this, a good tradeoff is already achieved at maximum order four.

3 Using Spherical Deconvolution for Model Fitting

3.1 Fitting the Ball-and-Stick Model

The previous section described how fitting the ball-and-stick model can theoretically be formulated as a deconvolution problem with a discrete ODF $\tilde{F}(\theta, \phi)$. We now describe an efficient and reliable algorithm based on this insight.

Initial per-voxel estimates of the diffusivity d and the isotropic volume fraction f_{iso} of the ball-and-stick model are obtained from the maximum apparent diffusion coefficient d_{max} and the average diffusion-weighted intensity \bar{S}

$$d_{\text{max}} = -\frac{1}{b} \min_i \log \frac{S_i}{S_0} \quad \bar{S} = \frac{1}{n} \sum_{i=1}^n S_i, \tag{4}$$

where S_i are the diffusion-weighted values. Integrating Equation (1) over the unit sphere shows that \bar{S} varies linearly with f_{iso} between $\bar{S} = S_0 \sqrt{\pi/(4bd)} \text{erf}(\sqrt{bd})$ (at $f_{\text{iso}} = 0$) and $\bar{S} = S_0 \exp(-bd)$ (at $f_{\text{iso}} = 1$), where erf is the Gauss error function. Assuming that $d \approx d_{\text{max}}$, this allows us to compute f_{iso} from \bar{S} . These initial estimates are refined further as part of the final descent-based fitting.

Now, the predicted isotropic part $S_0 f_{\text{iso}} \exp(-bd)$ is subtracted from S_i , and a spherical harmonics representation of $S(\theta, \phi)$ is fit to the remainder. Like [12], we deconvolve to a non-ringing cosine power lobe ($\cos^4 \gamma$) instead of a truncated delta peak. Using computer algebra software, the order-four coefficients of the deconvolution kernel that corresponds to $R(\gamma)$ in Equation (3) are found as:

$$R_0 = \frac{5\sqrt{\pi}S_0}{2\sqrt{bd}} \text{erf}(\sqrt{bd}); \quad R_2 = \frac{-35S_0}{32\sqrt{(bd)^3}} \left(6\sqrt{bd}e^{-bd} + \sqrt{\pi}(2bd - 3)\text{erf}(\sqrt{bd}) \right) \tag{5}$$

$$R_4 = \frac{105S_0}{512\sqrt{(bd)^5}} \left(-30\sqrt{bd}(2bd + 21)e^{-bd} + 9\sqrt{\pi}(4bd(bd - 5) + 35)\text{erf}(\sqrt{bd}) \right)$$

Dividing the order- n spherical harmonics coefficients of $S(\theta, \phi)$ by R_n gives $F(\theta, \phi)$ [9], which is then approximated by a discrete ODF $\tilde{F}(\theta, \phi)$ with the algorithm in [12]. This involves nonlinear optimization, but it is much simpler than fitting Equation (1) directly: Instead of fitting to 50–100 DWI values, it considers only 15 coefficients of $F(\theta, \phi)$ (at maximum order four). Moreover, both the objective function and its derivatives involve only additions and multiplications.

The peaks in $\tilde{F}(\theta, \phi)$ approximate fiber directions, their weights provide relative volume fractions. In a final step, these estimates are refined by fitting Equation (1) to the original DWI data, using Levenberg-Marquardt (LM) [13].

3.2 Selecting the Number of Sticks

For best quality, the correct number of compartments needs to be selected in the multi-tensor model: If it is chosen too low, the model might indicate fiber directions that do not align with any true tract. On the other hand, overestimating the true number reduces the accuracy of the result due to overfitting.

In spherical deconvolution, it is common to estimate the number of fiber compartments from the maxima in the ODF $F(\theta, \phi)$ [17]. Since our framework uses deconvolution, we explore a similar strategy based on the discrete ODF $\tilde{F}(\theta, \phi)$. We use a three-stage test with thresholds $t_{\{0,1,2\}}$: If $f_{\text{iso}} > t_0$, no fiber

¹ <http://www.ics.forth.gr/~lourakis/levmar/> [Accessed on 22 Feb 2010].

is detected. Otherwise, $F(\theta, \phi)$ is normalized to integrate to unity and a discrete ODF with two delta peaks is extracted, whose weights are $w_1 > w_2$. If $w_2 < t_1$, a single fiber is used. Otherwise, a three-peak approximation is found with $w_1 > w_2 > w_3$. If $w_3 < t_2$, we assume two fibers, else three compartments are used.

4 Results

4.1 Synthetic Data

Synthetic diffusion-weighted data with 60 directions, $b = 3000 \text{ s/mm}^2$, and Rician noise at $\text{SNR}_0 = 30$ was created from a multi-cylinder model with non-zero $\lambda_2 = \lambda_3$. Eigenvalues were sampled from Gaussian distributions with parameters according to estimates from real DW-MRI data. Principal eigenvectors and volume fractions were sampled uniformly at random. Cases in which any volume fraction was below 0.2 or any pair of vectors was closer than 30° were rejected.

We created 5000 one-, two-, and three-fiber configurations each and used Levenberg-Marquardt with 100 random restarts, sampled in analogy to data generation, to find the global optimum for the predetermined, correct number of compartments. We then counted the number of times a single randomized run and a run that was kick-started by spherical deconvolution found the correct optimum. Table 1 shows that in the two- and three-fiber cases, the prediction raised the chances of finding the optimum by around 10%. On average, the cost of deconvolution, finding the discrete ODF, and final refinement by LM, is about half the computational cost of finding the optimum by LM alone.

Table 2 lists the mean and median angular deviation of the individual stick compartments, sorted by volume fraction, from the ground truth. It confirms that spherical deconvolution comes close enough to the final result to be a useful seed, but subsequent optimization still improves upon its accuracy, especially in the three-fiber case. To validate our choice of [12] over ODF peak finding, we also

Table 1. Percentage of cases in which the global optimum was found by a single randomized run of LM, versus our kick-start method, with “speed” in voxels per second

	1 Fiber		2 Fibers		3 Fibers	
	% Found	Speed	% Found	Speed	% Found	Speed
Randomized Run	99.9	3430	90.3	1496	82.6	805
Kick-Started Run	100	5635	98.6	3264	93.1	1527

Table 2. Mean and median (in parentheses) of angular deviation from ground truth confirms that spherical deconvolution seeds the optimization close to the final result

	1 Fiber	2 Fibers		3 Fibers		
	Angle	Angle 1	Angle 2	Angle 1	Angle 2	Angle 3
Kick-Started Run	1.2 (1.0)	2.3 (1.9)	5.9 (3.9)	3.9 (3.0)	6.8 (4.9)	8.8 (6.3)
Prediction alone	1.3 (1.1)	3.0 (2.3)	7.5 (5.0)	5.8 (4.4)	9.5 (7.3)	12.0 (9.6)
ODF Peaks	1.5 (1.3)	3.5 (2.6)	12.4 (6.3)	5.7 (4.2)	12.4 (8.1)	16.2 (11.1)

Table 3. In synthetic data, our deconvolution-based test for the number of fiber compartments produces better results than standard statistical tests

Truth	Prediction											
	BIC				F-Test				Our Test			
	0	1	2	3	0	1	2	3	0	1	2	3
No Fiber	98%	2%	0%	0%	98%	2%	0%	0%	100%	0%	0%	0%
1 Fiber	0%	29%	43%	28%	0%	82%	6%	12%	0%	94%	5%	1%
2 Fibers	0%	5%	74%	21%	0%	15%	53%	32%	0%	5%	83%	12%
3 Fibers	0%	5%	23%	73%	2%	31%	13%	54%	0%	0%	16%	84%

list the most accurate result we were able to achieve using that more traditional technique (order 6, truncated delta peak, Laplace-Beltrami regularization [17]).

Finally, we used the synthetic data to compare our criterion for selecting the fiber number to two statistical tests previously applied to diffusion model selection, the Bayesian Information Criterion (BIC) [18], and the F-Test [19,6,18]. Automatic relevance determination [11] is not included in our comparison, since it aims at Bayesian model averaging rather than at making a hard decision. As shown in Table 3, the parameter-free BIC exhibited a strong bias towards selecting two or three fibers. The thresholds $t_{\{0,1,2\}}$ of the other two tests were set to balance sensitivity and specificity at each stage. The deconvolution-based test achieved best results for all ground truth configurations.

We also repeated all experiments with $b = 1000 \text{ s/mm}^2$. Without providing the quantitative results, we qualitatively state that the probability of finding the correct optimum increased, but average accuracy with respect to ground truth decreased. Our test for model selection became less reliable, while results of the F-Test improved. Apparently, the F-Test benefits from a low effective noise level, while the deconvolution-based test requires the better separation of the individual compartments afforded by higher b values.

4.2 Real Data

In order to identify the voxels in a real dataset (60 directions, $b = 1000 \text{ s/mm}^2$, isotropic voxel size 1.72 mm) in which the individual tests are most likely to use two- and three-fiber models, we calibrated the F-Test and our deconvolution-based test to generate 25% no-fiber, 40% one-fiber, 25% two-fiber, and 10% three-fiber voxels within a brain mask.

A detail of the result is shown in Figure 2, the intersection of corpus callosum and corticospinal tract on a coronal slice. For visualization, we map the stick compartments to tensors whose perpendicular diffusivity is scaled by the volume fraction of the ball compartment ($\lambda_2 = \lambda_3 = f_{\text{iso}}d$). The fiber fractions are renormalized to unity and color-coded. In (a), standard single diffusion tensors are shown for reference. We found that the F-Test tends to fit multiple stick compartments to voxels that likely just contain a single bending or spreading bundle, e.g., in the body of the corpus callosum (CC). The deconvolution-based test is more robust to such cases, as long as they lead to a clear single ODF peak.

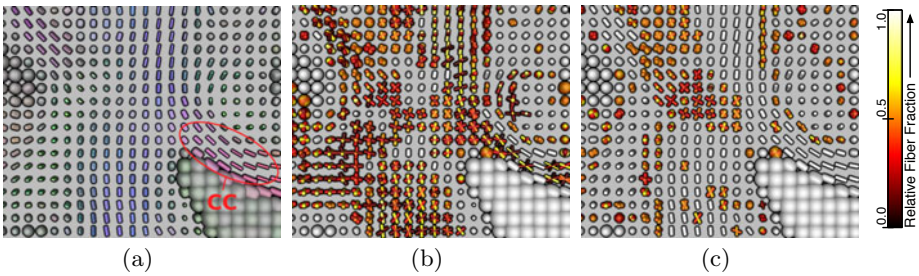


Fig. 2. In real data (a), our deconvolution-based test (c) produced more plausible results than the F-Test (b), which frequently fits multiple sticks in regions that are generally thought to contain a single bending or spreading compartment.

We expect that this will allow us to track through crossings like the one presented in Figure 2, while avoiding spurious tracts. However, design of a tractography method that makes use of our novel framework is left as future work.

With the deconvolution-based test, Levenberg-Marquardt is only run once per voxel, which provides an additional speedup. Our complete pipeline, including testing and fitting the final result, processed 1050 voxels per second with the F-Test, 2950 voxels per second with the deconvolution-based test, on a single CPU core of a 2.7 GHz workstation. Finally, 20 randomized restarts of Levenberg-Marquardt improved upon the optimum found by our method in less than 0.2% of all voxels. Therefore, we conclude that our pipeline offers a reliable and efficient solution for fitting ball-and-stick models.

5 Conclusion

Traditionally, multi-fiber models and spherical deconvolution are used as competing methods, each with its own set of advantages and disadvantages: Linear spherical deconvolution is extremely fast and does not require pre-specification of an expected number of fibers. On the other hand, multi-tensor models offer higher accuracy for applications like multi-fiber streamline tractography [4,7,17,15], where it is the primary goal to estimate the most likely fiber directions.

We have presented a framework that combines the best of both worlds: Based on spherical deconvolution, a plausible number of fiber compartments is found automatically. Initializing the fitting with the deconvolution result doubles the speed of the computation, while at the same time increasing the probability of finding the global optimum to more than 95% in the two-fiber case and to more than 90% for three fibers. As a result, we achieve a fully integrated, reliable and efficient algorithmic solution for multi-tensor fitting.

References

1. Basser, P., Mattiello, J., Bihan, D.L.: Estimation of the effective self-diffusion tensor from the NMR spin echo. *J. Magn. Reson.* B(103), 247–254 (1994)
2. Alexander, A., Hasan, K., Lazar, M., Tsuruda, J., Parker, D.: Analysis of partial volume effects in diffusion-tensor MRI. *Magn. Reson. Med.* 45, 770–780 (2001)

3. Tuch, D., Reese, T., Wiegell, M., Makris, N., Belliveau, J., Wedeen, V.: High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magn. Reson. Med.* 48, 577–582 (2002)
4. Parker, G., Alexander, D.: Probabilistic Monte Carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. In: Taylor, C.J., Noble, J.A. (eds.) *IPMI 2003. LNCS*, vol. 2732, pp. 684–695. Springer, Heidelberg (2003)
5. Peled, S., Friman, O., Jolesz, F., Westin, C.F.: Geometrically constrained two-tensor model for crossing tracts in DWI. *Magn. Reson. Imaging* 24(9), 1263–1270 (2006)
6. Nedjati-Gilani, S., Parker, G., Alexander, D.C.: Mapping the number of fibre orientations per voxel in diffusion MRI. In: *Proc. Intl. Soc. Mag. Reson. Med.*, p. 3169 (2006)
7. Kreher, B., Schneider, J., Mader, I., Martin, E., Hennig, J., Il'yasov, K.: Multitensor approach for analysis and tracking of complex fiber configurations. *Magn. Reson. Med.* 54, 1216–1225 (2005)
8. Pasternak, O., Assaf, Y., Intrator, N., Sochen, N.: Variational multiple-tensor fitting of fiber-ambiguous diffusion-weighted magnetic resonance imaging voxels. *Magn. Reson. Imaging* 26, 1133–1144 (2008)
9. Tournier, J.D., Calamante, F., Gadian, D., Connelly, A.: Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage* 23, 1176–1185 (2004)
10. 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. *Magn. Reson. Med.* 50, 1077–1088 (2003)
11. Behrens, T., Johansen-Berg, H., Jbabdi, S., Rushworth, M., Woolrich, M.: Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage* 34, 144–155 (2007)
12. Schultz, T., Seidel, H.P.: Estimating crossing fibers: A tensor decomposition approach. *IEEE Trans. Vis. Comp. Graphics* 14(6), 1635–1642 (2008)
13. Marquardt, D.: An algorithm for least-squares estimation of nonlinear parameters. *SIAM J. Applied Math.* 11(2), 431–441 (1963)
14. Anderson, A.: Measurement of fiber orientation distributions using high angular resolution diffusion imaging. *Magn. Reson. Med.* 54(5), 1194–1206 (2005)
15. Malcol, J., Michailovich, O., Bouix, S., Westin, C.F., Shenton, M., Rathi, Y.: A filtered approach to neural tractography using the Watson directional function. *Medical Image Analysis* 14, 58–69 (2010)
16. Hosey, T., Williams, G., Ansorge, R.: Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magn. Reson. Med.* 54, 1480–1489 (2005)
17. Descoteaux, M., Deriche, R., Knösche, T., Anwander, A.: Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE Trans. Med. Imaging* 28(2), 269–286 (2009)
18. Freidlin, R., Özarslan, E., Komlosh, M., Chang, L.C., Koay, C., Jones, D., Basser, P.: Parsimonious model selection for tissue segmentation and classification applications: A study using simulated and experimental DTI data. *IEEE Trans. Med. Imaging* 26(11), 1576–1584 (2007)
19. Alexander, D., Barker, G., Arridge, S.: Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data. *Magn. Reson. Med.* 48, 331–340 (2002)

Statistical Analysis of Tensor Fields

Yuchen Xie, Baba C. Vemuri, and Jeffrey Ho

Department of Computer and Information Sciences and Engineering
University of Florida

Abstract. In this paper, we propose a Riemannian framework for statistical analysis of tensor fields. Existing approaches to this problem have been mainly voxel-based that overlook the correlation between tensors at different voxels. In our approach, the tensor fields are considered as points in a high-dimensional Riemannian product space and accordingly, we extend Principal Geodesic Analysis (PGA) to the product space. This provides us with a principled method for linearizing the problem, and coupled with the usual log-exp maps that relate points on manifold to tangent vectors, the global correlation of the tensor field can be captured using Principal Component Analysis in a tangent space. Using the proposed method, the modes of variation of tensor fields can be efficiently determined, and dimension reduction of the data is also easily implemented. Experimental results on characterizing the variation of a large set of tensor fields are presented in the paper, and results on classifying tensor fields using the proposed method are also reported. These preliminary experimental results demonstrate the advantages of our method over the voxel-based approach.

1 Introduction

Tensor field data sets are quite commonly encountered in diffusion tensor imaging (DTI) [1] and Tensor-Based Morphometry (TBM) [2]. Most tensor fields that have been reported in recent medical image analysis literature are fields of symmetric positive-definite matrices (SPDs), and this paper proposes a framework for statistical analysis on the space of these tensor fields.

The Riemannian geometry of the SPD matrices and its applications to medical image analysis problems that require statistical analysis of ensembles of SPD matrices has been the focus of intensive study in the past several years, e.g. [3][5][9]. In [5], Pennec et al. developed a Riemannian framework for computing statistics on SPD tensors. In [9], Schwartzman discussed the geometry of positive-definite matrices and studied probability distributions defined on SPD matrices. Principal Geodesic Analysis (PGA), as a generalization of Principal Component Analysis (PCA) for data on a Riemannian manifold, was introduced in [4]. In [3], Fletcher and Joshi described PGA-based methods for statistical analysis of diffusion tensors, in the context of computing (Karcher) mean of a collection of SPD matrices and characterizing their variance.

We remark that all these earlier works have focused on the statistical analysis of symmetric positive-definite matrices, and to the best of our knowledge there is no literature on statistical analysis of tensor fields where each field is treated in its entirety as a single entity. Although the methods for SPD matrices can be applied to study the

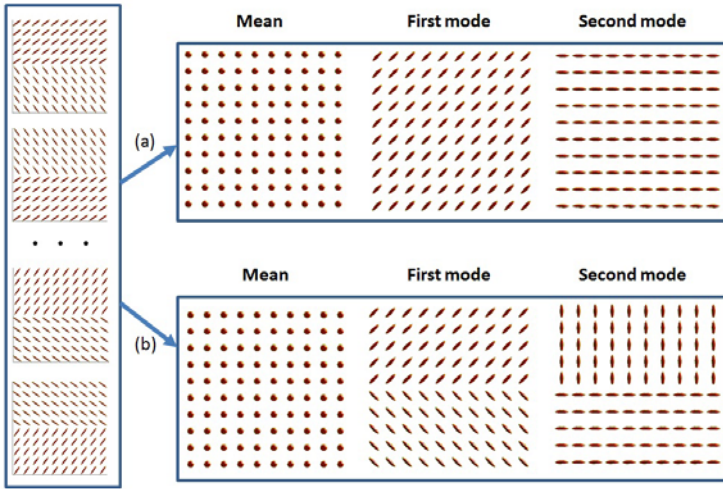


Fig. 1. Left column: Eight input tensor fields. **(a)** Mean, first mode and second mode of tensor fields computed using a pixel-based approach: means and modes are determined at each pixel independently using eight tensors. Note that the two modes are constant fields. **(b)** Mean, first mode and second mode of tensor fields computed using the proposed method.

statistics of tensor fields using a voxel-based approach where tensor fields are aligned and the statistics are gathered independently for every voxel, it is clearly insufficient and inadequate as it fails to capture global interaction patterns in the tensor fields. This point can be best illustrated using a simple example shown in Figure 1. Here, we generate four pairs of 10×10 tensor fields. Each tensor field contains two constant subfields occupying the top and bottom regions of the domain. Tensor fields in each pair differ by a reflection that swaps the top and bottom regions. For this collection of eight tensor fields, pixel-wise statistics fails to capture the global patterns in the tensor fields as the first and second modes are all constant fields. This result is not surprising as a pixel-based approach only considers tensors at each location independently, and it completely ignores the possible correlation between tensors at different locations.

In this paper, we propose a Riemannian framework for statistical analysis of a set of tensor fields that is capable of capturing the global correlation within the fields. Specifically, each tensor field can be represented by a point in the Riemannian product space. We extend Principal Geodesics Analysis (PGA) to Riemannian symmetric products, and this provides a principled method for linearizing the problem by mapping data (tensor fields) to a Euclidean space that is the tangent space at one specific point. The global correlations of the tensor field are then captured using Principle Component Analysis (PCA) in the tangent space, and the modes of variation for the tensor fields can be determined first in the tangent space followed by the exponential map. In addition, dimensionality reduction of data can also be efficiently implemented using PCA in the Euclidean space. This is particularly important as one major difficulty of working with the space of tensor fields is its dimension. For example, the space of $100 \times 100 \times 100$

tensor-valued images has 6×10^6 dimension, and dimensionality reduction is necessary for most applications and analysis. The proposed method is evaluated using OASIS dataset. We characterized the variation within a set of deformation tensor fields and applied our method to the tensor field classification problem. Preliminary experimental results have demonstrated the superiority of our method compared to the voxel-based approach.

2 Statistical Analysis in the Space of Tensor Fields

In this section, we present the details of statistical analysis in the space of tensor fields, and we will consider only fields of symmetric positive-definite matrices. Let $P(n)$ denote the space of $n \times n$ symmetric positive-definite matrices and $Sym(n)$ denote the space of $n \times n$ symmetric matrices. A tensor field defined on a domain Ω in \mathbb{R}^K is treated as a function $f : \Omega \rightarrow P(n)$. Since both diffusion tensor fields and deformation tensor fields used in medical imaging are almost always defined over a grid (pixels or voxels), Ω will be a collection of m points in \mathbb{R}^K , and we identify the space of tensor fields on Ω with the product $P(n)^m = \underbrace{P(n) \times P(n) \times \dots \times P(n)}_m$. Thus a tensor

field \mathbf{X} in $P(n)^m$ is represented as an m -tuple (X_1, X_2, \dots, X_m) , where each X_i is a symmetric positive-definite matrix, the value of \mathbf{X} at one point in Ω .

2.1 Geometry of Tensor Fields

The space $P(n)$ is a symmetric Riemannian manifold [3] with $GL(n)$ as the symmetry group. This can be generalized directly to product spaces $P(n)^m$ using product Riemannian structure, and in particular, the Riemannian geodesic distances, log and exponential maps have closed-form expressions. Specifically, the group $GL(n)^m$ acts transitively on $P(n)^m$ with the action given by $\phi_G(\mathbf{X}) = (G_1 X_1 G_1^T, \dots, G_m X_m G_m^T)$, where each $G_i \in GL(n)$ is a $n \times n$ invertible matrix and X_i is a $n \times n$ positive-definite matrix. The tangent space of $P(n)^m$ at any point can be identified with $Sym(n)^m$ since the tangent space of a product manifold is the product of tangent spaces. Let $\mathbf{Y}, \mathbf{Z} \in T_{\mathbf{M}}P(n)^m$ be two tangent vectors at $\mathbf{M} \in P(n)^m$. The product Riemannian metric gives the inner product between the two vectors as

$$\langle \mathbf{Y}, \mathbf{Z} \rangle_{\mathbf{M}} = \sum_{i=1}^m \text{tr}(Y_i M_i^{-1} Z_i M_i^{-1}). \tag{1}$$

Using this metric, the Riemannian exponential map at \mathbf{M} maps the tangent vector \mathbf{Y} to a point in $P(n)^m$

$$\text{Exp}_{\mathbf{M}}(\mathbf{Y}) = (G_1 \exp(G_1^{-1} Y_1 G_1^{-T}) G_1^T, \dots, G_m \exp(G_m^{-1} Y_m G_m^{-T}) G_m^T) \tag{2}$$

where $G_i \in GL(n)$ such that $\mathbf{M} = (G_1 G_1^T, \dots, G_m G_m^T)$.

Given $\mathbf{X} \in P(n)^m$, the log map at \mathbf{M} is

$$\text{Log}_{\mathbf{M}}(\mathbf{X}) = (G_1 \log(G_1^{-1} X_1 G_1^{-T}) G_1^T, \dots, G_m \log(G_m^{-1} X_m G_m^{-T}) G_m^T). \tag{3}$$

Using this definition of log map in $P(n)^m$, the geodesic distance between two tensor fields \mathbf{M} and \mathbf{X} is computed by

$$d(\mathbf{M}, \mathbf{X}) = \|\text{Log}_{\mathbf{M}}(\mathbf{X})\| = \sqrt{\sum_{i=1}^m \text{tr}(\log^2(G_i^{-1} X_i G_i^{-T}))}. \tag{4}$$

2.2 Statistics on the Space of Tensor Fields

Using the formula above for the geodesic distance, we define the (intrinsic) mean of N tensor fields as the tensor field that minimizes the sum of squared geodesic distances:

$$\mathbf{M} = \arg \min_{M \in P(n)^m} \frac{1}{N} \sum_{i=1}^N d(M, \mathbf{X}_i)^2. \tag{5}$$

The sum of squares on the RHS above can be re-written as a sum over all points in Ω . This implies that the value of $\mathbf{M}(p)$ of \mathbf{M} at one point $p \in \Omega$ is the usual Karcher mean in $P(n)$ of $\mathbf{X}_1(p), \dots, \mathbf{X}_N(p)$. In particular, since the Karcher mean is unique on $P(n)$ [3], this shows that \mathbf{M} will be unique as well, and it can be computed using an iterative algorithm similar to the one in [3].

After obtaining the intrinsic mean \mathbf{M} of the input tensor fields $\mathbf{X}_1, \dots, \mathbf{X}_N$, we will determine the modes of variation using PGA. Specifically, we use the log map to map all the tensor fields to the tangent space at \mathbf{M} , $x_i = \text{log}_{\mathbf{M}}(X_i)$. This is a Euclidean space in which we can analyze the data points x_1, \dots, x_N using principal component analysis. We define the principal vectors $\mathbf{V}_1, \dots, \mathbf{V}_k$ in $T_{\mathbf{M}}P(n)^m$ according to the following equations:

$$\mathbf{V}_1 = \arg \max_{\|\mathbf{V}\|=1} \sum_{i=1}^N \langle \mathbf{V}, \text{Log}_{\mathbf{M}}(\mathbf{X}_i) \rangle_{\mathbf{M}}^2,$$

$$\mathbf{V}_k = \arg \max_{\|\mathbf{V}\|=1} \sum_{i=1}^N \sum_{j=1}^{k-1} \langle \mathbf{V}_j, \text{Log}_{\mathbf{M}}(\mathbf{X}_i) \rangle_{\mathbf{M}}^2 + \langle \mathbf{V}, \text{Log}_{\mathbf{M}}(\mathbf{X}_i) \rangle_{\mathbf{M}}^2.$$

The orthonormal vectors \mathbf{V}_i spanned a K -dimensional subspace S_K that best approximates x_1, \dots, x_N in the least-squares sense, and they can be computed using PCA. By exponentiating vectors in S_K , $\text{Exp}_{\mathbf{M}}\left(\sum_{k=1}^d \alpha_k \mathbf{V}_k\right)$, where α_k tells the variation of k th mode, we obtain the geodesic submanifold $S_K \subset P(n)^m$ that can serve as a good approximation of the input tensor fields.

There are two important details that differ from the usual application of PGA [3]. First, except at the identity, the inner product defined in Equation 1 does not correspond to the standard Euclidean inner product, which is required for the familiar PCA algorithm. Therefore, we first transform the data to the tangent space at the identity, which is accomplished via the following transform, $\mathbf{X} \in T_{\mathbf{M}}P(n)^m$

$$\phi_{\mathbf{G}^{-1}}(\mathbf{X}) : (X_1, \dots, X_m) \longrightarrow (G_1^{-1} X_1 G_1^{-T}, \dots, G_m^{-1} X_m G_m^{-T}), \tag{6}$$

where $\mathbf{G} = (G_1, \dots, G_m)$ is such that $\mathbf{M} = (G_1 G_1^T, \dots, G_m G_m^T)$. Once the data have been mapped to $T_{\mathbf{I}}P(n)^m$, we can apply the usual PCA algorithm to obtain principal

vectors $\mathbf{U}_i, i = 1, \dots, K$. They are then transformed back to $T_{\mathbf{M}}P(n)^m$ using $\mathbf{V}_i = \phi_{\mathbf{G}}(\mathbf{U}_i)$. Due to the high-dimensionality of $P(n)^m$, we use the Gram matrix instead of the usual covariance matrix when computing the principal vectors in the tangent space, which is the approach used in many computer vision applications such as the Eigenfaces [8]. The complete algorithm is summarized in Algorithm One.

Algorithm 1. PGA for Tensor Fields

- 1: **Input** N tensor fields $\mathbf{X}_1, \dots, \mathbf{X}_N \in P(n)^m$.
 - 2: Compute intrinsic mean \mathbf{M} of input tensor fields.
 - 3: Compute $\mathbf{Y}_i = \text{Log}_{\mathbf{M}}(\mathbf{X}_i)$ for $i = 1, \dots, N$.
 - 4: Translate \mathbf{Y}_i to the tangent space of identity \mathbf{I} .
 - 5: Perform PCA in $T_{\mathbf{I}}P(n)^m$ and get eigenvectors \mathbf{U}_i .
 - 6: Translate \mathbf{U}_i back to get \mathbf{V}_i in the tangent space of \mathbf{M} .
-

2.3 Tensor Fields Classification

We can formulate a tensor field classification algorithm using the principal directions and geodesic submanifolds. One common method for solving classification problems on Riemannian manifold is to map input data to the tangent space and do the classification in the tangent space [6]. However, *this approach does not respect the geometry of the manifold as the geodesic distance between two points on the manifold are usually not the same or even commensurate with the distance between their images in the tangent space.* A more principled approach is to use the distances to geodesic submanifolds as the feature for classification.

Assume a binary classification problem, and the training tensor fields are labelled as one of the two classes. For each label $k(k = 1, 2)$, we compute a low-dimensional geodesic submanifold \mathbf{S}_k using training tensor fields with label k . For a test tensor field \mathbf{X} , we can determine its class by comparing the geodesic distances $d_k = \min_{\mathbf{Y} \in \mathbf{S}_k} d(\mathbf{X}, \mathbf{Y})$. A tensor field is classified as belonging to class k if d_k is smaller than the other geodesic distance. The key step in this algorithm is to find the minimizer in \mathbf{S}_k that gives the geodesic distance d_k . Since any point in the geodesic submanifold \mathbf{S}_k can be written as $\text{Exp}_{\mathbf{M}}\left(\sum_{i=1}^d \alpha_i \mathbf{V}_i\right)$, where \mathbf{M} is the mean and $\alpha_1, \dots, \alpha_d$ are real coefficients, d_k can be solved via the following optimization problem in \mathbb{R}^d ,

$$\min_{\alpha_1, \dots, \alpha_d} d\left(\mathbf{X}, \text{Exp}_{\mathbf{M}}\left(\sum_{i=1}^d \alpha_i \mathbf{V}_i\right)\right)^2. \tag{7}$$

Unfortunately, minimizing Equation 7 can be time-consuming for large tensor fields. Therefore, we approximate d_k by the geodesic distance $d(\mathbf{X}, \mathbf{Z})$ between \mathbf{X} and \mathbf{Z} defined by

$$\mathbf{Z} = \text{Exp}_{\mathbf{M}}\left(\sum_{i=1}^d \mathbf{V}_i \langle \mathbf{V}_i, \text{Log}_{\mathbf{M}}(\mathbf{X}) \rangle_{\mathbf{M}}\right). \tag{8}$$

That is, we obtain \mathbf{Z} by first map \mathbf{X} to the tangent space at \mathbf{M} using Log map and project it onto the principal subspace S_d . The projection is then mapped down to the

manifold using the exponential map to get Z . In our experiments discussed in the next section, the differences between the approximated distance and the one computed by the optimization are less than 1% and have no major influence on the classification results. We summarize the tensor fields classification algorithm in Algorithm Two.

Algorithm 2. Tensor Fields Classification

- 1: **Training** compute geodesic submanifolds S_1 and S_2 by using PGA on training tensor fields of different classes separately.
 - 2: **Testing** for each test tensor field, compute their geodesic distances d_1 and d_2 to submanifolds S_1 and S_2 respectively. If $d_1 < d_2$, classify the test data to class1, otherwise, set it to class2.
-

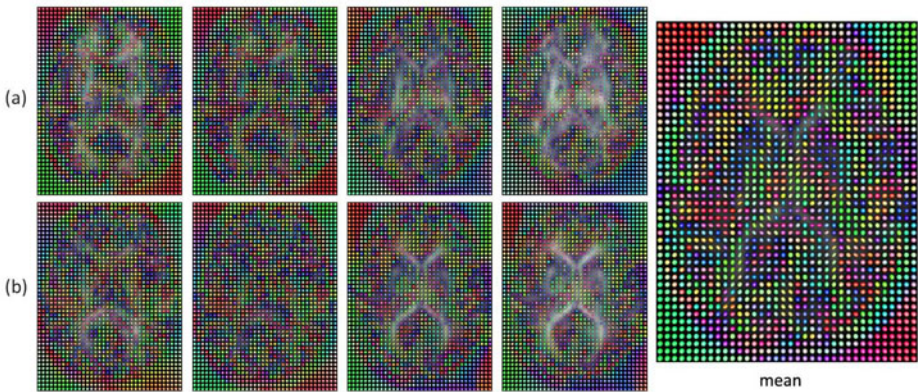


Fig. 2. Statistical Analysis for deformation tensor fields from old age group. For better visualization, we downsample the images, only axial view is shown and set FA as the background in the display. **(a)** Tensor field variation along the first principal direction. From left to right, the coefficient α_1 is $-2\sigma_1, -\sigma_1, \sigma_1, 2\sigma_1$. **(b)** Tensor field variation along the second principal direction. From left to right, the coefficient α_2 is $-2\sigma_2, -\sigma_2, \sigma_2, 2\sigma_2$. **Right column:** Mean tensor field.

3 Experimental Results

The data used in our experiments are from the freely available Open Access Series of Imaging Studies (OASIS) MRI data set, which contains a cross-sectional collection of 416 subjects aged between 18-96 [7]. Each brain image has a resolution of $176 \times 208 \times 176$ voxels. We divided 416 subjects three groups: young subjects (40 or younger), middle-aged subjects (between 40 and 60 years of age) and old subjects (60 or older). There is a subset of old subjects that were diagnosed with probable Alzheimer’s Disease (AD). We compute the atlas for all the MR images in the OASIS data set using a group-wise nonrigid registration [10], and this also gives the the deformation field from each image to the atlas. For each voxel, we compute the Jacobian matrix J of the deformation field and build the deformation strain tensor $S = (J^T J)^{1/2}$. This gives a strain tensor field for every subject.

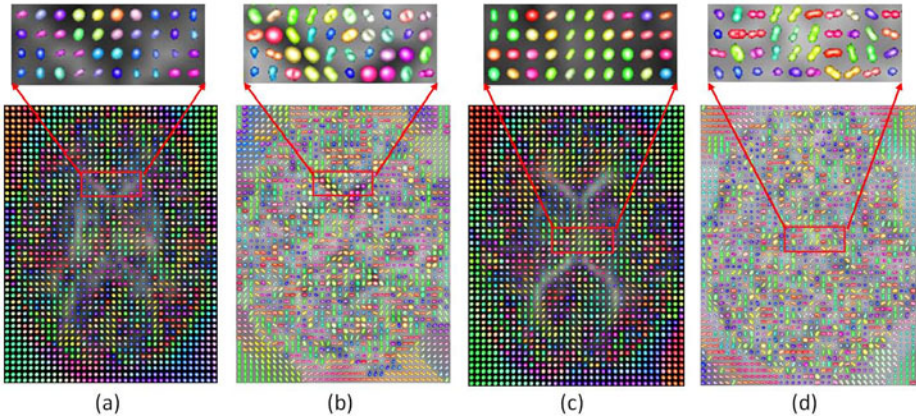


Fig. 3. Comparisons with voxel-based method. (a) and (c) are the first mode (σ_1) and second mode (σ_2) computed using Algorithm One. (b) and (d) are the first mode and the second mode computed using voxel-wise PGA. FA is used as the background in the display.

Table 1. Tensor Fields Classification on OASIS

	Old vs. Young	Old vs. Middle	Middle vs. Young	AD vs. Control
Nearest Neighbor	92.43%	87.74%	78.42%	84.29%
Submanifold Projection	96.43%	90.23%	84.32%	88.57%

In the first experiment, we characterize the variation in the tensor fields within an age group by computing the modes of variation using Algorithm-1. The dimension of the geodesic submanifold is set at 20 after examining the eigenvalue distribution. Figure 2 displays the mean tensor fields and the variations along the first two principal directions for the old group. The comparison with the mean and modes computed using the voxel-based approach is shown in Figure 3. We can clearly see that the modes computed using voxel-based method are fragmentary and they do not reflect the global structures of the tensor fields. This is not surprising because voxel-based method does not consider correlations between different voxels.

For the second experiment, we test our tensor-field classification algorithm. We randomly divide the brain images for each age group into four subsets. Images from one of the subsets are the test images, while other three subsets give the training images. The training images are used to compute the geodesic submanifolds S_d , and we classify the test images for every pair of age groups using Algorithm 2. We use a four-fold cross-validation in the experiments to fully evaluate the algorithm on OASIS data set. We compared the performance of our algorithm with the nearest neighbor method that maps each tensor field to the tangent space of the mean. Low-dimensional feature vectors are generated using PCA projection, and the classification is done using nearest neighbor of these feature vectors. We have also tested the proposed algorithm on classification for healthy and diseased (Alzheimer) brain images. All results are tabulated in Table 1. The experimental results indicate that the deformation strain tensor fields

do capture the subtle structural changes in the brain images across different age groups (and also for diseased samples), and the good classification results show the importance and value of doing statistical analysis on the space of tensor fields.

4 Conclusions

In this paper, we have presented the geometry of the space of tensor fields and proposed a framework for statistical analysis of a set of tensor fields. We have extended the PGA framework to the space of tensor fields considered as a Riemannian product space, and the modes of variation computed by the proposed algorithm capture the correlation between tensors at different locations. We have also proposed a novel tensor field classification algorithm using distances to the geodesic submanifolds as the main features for classification. Experimental results have shown that our approach provides a better characterization of the variation within a collection of tensor fields when compared to the voxel-based approach. In addition, good classification results on a large population of brain images have further validated the proposed framework.

Acknowledgement

This research is in part supported by the NIH grant EB007082 to BCV and the NSF grant IIS 0916001 to JH and BCV.

References

1. Basser, P.J., Mattiello, J., Le Bihan, D.: MR diffusion tensor spectroscopy and imaging. *Biophysics Journal* 66, 295–267 (1994)
2. Lepore, N., Brun, C., Chou, Y., Chiang, M., Dutton, R.A., Hayashi, K.M., Luders, E., Lopez, O.L., Aizenstein, H.J., Toga, A.W., Becker, J.T., Thompson, P.M.: Generalized Tensor-Based Morphometry of HIV/AIDS Using Multivariate Statistics on Deformation Tensors. *IEEE Transactions on Medical Imaging* (2008)
3. Fletcher, P.T., Joshi, S.: *Riemannian Geometry for the Statistical Analysis of Diffusion Tensor Data*. Signal Processing (2007)
4. Fletcher, P.T., Lu, C., Pizer, S.M., Joshi, S.: Principal Geodesic Analysis for the Study of Nonlinear Statistics of Shape. *IEEE Transactions on Medical Imaging* (2004)
5. Pennec, X., Fillard, P., Ayache, N.: A Riemannian Framework for Tensor Computing. *International Journal of Computer Vision* (2006)
6. Wu, J., Smith, W.A.P., Hancock, E.R.: Gender Classification using Shape from Shading. *British Machine Vision Conference* (2007)
7. Marcus, D.S., Wang, T.H., Parker, J., Csernansky, J.G., Moris, J.C., Buckner, R.L.: Open Access Series of Imaging Studies (OASIS): Cross-Sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults. *Journal of Cognitive Neuroscience* (2007)
8. Turk, M.A., Pentland, A.P.: Face Recognition Using Eigenfaces. In: *IEEE Conference on Computer Vision and Pattern Recognition* (1991)
9. Schwartzman, A.: *Random Ellipsoids and False Discovery Rates: Statistics for Diffusion Tensor Imaging Data*. Ph.D. dissertation, Dept. of Statistics, Stanford University (2006)
10. Joshi, S., Davis, B., Jomier, M., Gerig, G.: Unbiased Diffeomorphic Atlas Construction for Computational Anatomy. *NeuroImage* (2004)

Multivariate Varying Coefficient Models for DTI Tract Statistics^{*}

Hongtu Zhu, Martin Styner, Yimei Li, Linglong Kong, Yundi Shi, Weili Lin, Christopher Coe, and John H. Gilmore

Department of Biostatistics, Radiology, Psychiatry and Computer Science, and Biomedical Research Imaging Center, University of North Carolina at Chapel Hill

Abstract. Diffusion tensor imaging (DTI) is important for characterizing the structure of white matter fiber bundles as well as detailed tissue properties along these fiber bundles *in vivo*. There has been extensive interest in the analysis of diffusion properties measured along fiber tracts as a function of age, diagnostic status, and gender, while controlling for other clinical variables. However, the existing methods have several limitations including the independent analysis of diffusion properties, a lack of method for accounting for multiple covariates, and a lack of formal statistical inference, such as estimation theory and hypothesis testing. This paper presents a statistical framework, called VCMTS, to specifically address these limitations. The VCMTS framework consists of four integrated components: a varying coefficient model for characterizing the association between fiber bundle diffusion properties and a set of covariates, the local polynomial kernel method for estimating smoothed multiple diffusion properties along individual fiber bundles, global and local test statistics for testing hypotheses of interest along fiber tracts, and a resampling method for approximating the p -value of the global test statistic. The proposed methodology is applied to characterizing the development of four diffusion properties along the splenium and genu of the corpus callosum tract in a study of neurodevelopment in healthy rhesus monkeys. Significant time effects on the four diffusion properties were found.

1 Introduction

In the existing literature, there are three major approaches to the group analysis of diffusion imaging data including region-of-interest (ROI) analysis, voxel-wise analysis, and fiber tract based analysis [1], [2], [3]. The region-of-interest (ROI) method primarily computes averages diffusion properties in some manually drawn ROIs, generates various summary statistics per ROI, and then carries

^{*} This work was supported in part by NSF grant BCS-08-26844 and NIH grants UL1-RR025747-01, MH086633, P01CA142538-01 and AG033387 to Dr. Zhu, NIH grants MH064065, HD053000, and MH070890 to Dr. Gilmore, NIH grants R01NS055754 and R01EB5-34816 to Dr. Lin, Lilly Research Laboratories, the UNC NDRC HD 03110, Eli Lilly grant F1D-MC-X252, and NIH Roadmap Grant U54 EB005149-01, NAMIC to Dr. Styner.

out statistical analysis on these summary statistics [3]. ROI analysis suffers from identifying meaningful ROIs, particularly the long curved structures common in fiber tracts, the instability of statistical results obtained from ROI analysis, and the partial volume effect in relative large ROIs.

Compared with the other two methods, voxel-wise analysis has been widely used in neuroimaging studies [4]. The voxel-wise analysis involves two sequential steps. The first step is to fit a statistical model to diffusion properties at each voxel and generating a parametric map of test statistics (or p -values). The second step includes a correction to the multiple comparisons across the many voxels of imaging volume [5]. However, the voxel-wise analysis suffers from the issues of misalignment and arbitrary smoothing extent [6, 7]. As pointed out in [7], the final statistical results of voxel-wise analysis can strongly depend on the amount of smoothing applied to the diffusion tensor imaging data.

The third method is to develop fiber tract based analysis of diffusion properties, such as eigenvalues and fractional anisotropy (FA) values [1, 2, 8, 9]. In [1], a *tract-based spatial statistics* framework is proposed to carry out a pointwise analysis along the white matter skeleton. In [8], a model-based framework is developed for the analysis of diffusion properties on the medial manifolds of fiber tracts followed by testing pointwise hypotheses on the medial manifolds. In [9], a functional principal component analysis (PCA) is used to compare a univariate diffusion property, such as fractional anisotropy, across two (or more) populations for a single hypothesis test per tract [9]. Furthermore, in [10], a constrained PCA method is proposed to fit age-related changes white matter diffusion of fiber tracts. The functional and constrained PCA methods suffer from the issues of the independent analysis of diffusion properties, a lack of method for accounting for multiple covariates, and a lack of formal statistical inference, such as estimation theory and hypothesis testing. In [11], a functional regression framework, called FRATS, is proposed for the analysis of multiple diffusion properties along fiber bundle as functions and their association with a set of covariates of interest in real applications.

The goal of this paper is to develop a multivariate varying coefficient model framework, called VCMTS, to completely address the issues of the functional and constrained PCA. Compared with the existing literature including [10], [11], and [9], we have made several novel contributions. We develop a multivariate varying coefficient model to statistically characterize the association between multiple fiber bundle diffusion properties and a set of covariates of interest. We use the local polynomial kernel method to regularize multiple diffusion properties along individual fiber bundles. We propose both local and global test statistics for testing hypothesis of interest along and on fiber tracts. We approximate the p -value of the global test statistic using a resampling method.

2 Method

A schematic overview of VCMTS is given in Figure 1. We have proven that each component of VCMTS is statistically sound under some mild conditions and the

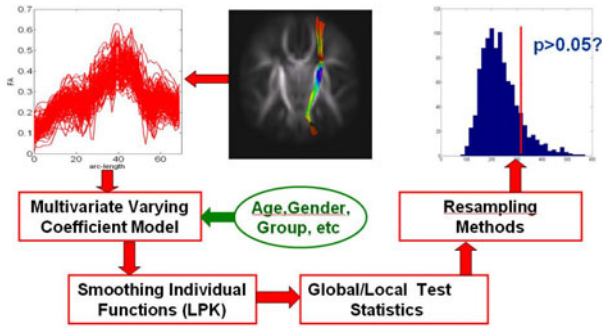


Fig. 1. A schematic overview of VCMTS, in which the right internal capsule tract is used as an illustration

detailed proof can be found in [12]. The associated software for implementing VCMTS will be available in <https://bios.unc.edu/~hzhu/> and disseminated to imaging researchers through <http://www.nitrc.org/>. We describe each of these components in detail below.

2.1 Multivariate Varying Coefficient Model

We develop a *multivariate varying coefficient model* to characterize the relationship between multiple diffusion properties along fiber tracts and a set of covariates of interest, such as age, group status, and gender. For the i -th subject, we consider an $m \times 1$ vector of diffusion properties, denoted by $\mathbf{y}_i(s_j) = (y_{i,1}(s_j), \dots, y_{i,m}(s_j))^T$, and its associated arc length s_j for the j -th location grid point on the fiber bundle for $j = 1, \dots, n_G$ and $i = 1, \dots, n$, where n_G and n denote the numbers of grid points and subjects, respectively. We assume that

$$y_{i,k}(s) = \mathbf{x}_i^T B_k(s) + \eta_{i,k}(s) + \epsilon_{i,k}(s), \tag{1}$$

where $B_k(s) = (b_{k1}(s), \dots, b_{kp}(s))^T$ is a $p \times 1$ vector of functions of s , \mathbf{x}_i is a $p \times 1$ vector of covariates of interest, $\epsilon_{i,k}(s)$ are measurement errors and $\eta_{i,k}(s)$ characterize individual curve variations from $\mathbf{x}_i^T B_k(s)$. The varying coefficient matrix $B(s) = [B_1(s), \dots, B_m(s)]$ characterizes the association between fiber bundle diffusion properties and the covariates of interest \mathbf{x}_i . Model (1) is a multivariate varying coefficient model [13]. Let $SP(\mu, \Sigma)$ denote a stochastic process vector with mean function $\mu(t)$ and covariance function $\Sigma(s, t)$. We assume that $\boldsymbol{\epsilon}_i(s) = (\epsilon_{i,1}(s), \dots, \epsilon_{i,m}(s))^T$ and $\boldsymbol{\eta}_i(s) = (\eta_{i,1}(s), \dots, \eta_{i,m}(s))^T$ are independent, and $\boldsymbol{\eta}_i(s)$ and $\boldsymbol{\epsilon}_i(s)$ are independent and identical copies of $SP(\mathbf{0}, \Sigma_\eta)$ and $SP(\mathbf{0}, \Sigma_\epsilon)$, respectively. Moreover, $\epsilon_i(s)$ and $\epsilon_i(s')$ are assumed to be independent and thus $\Sigma_\epsilon(s, t)$ takes the form of $\Sigma_\epsilon(s, s)\mathbf{1}(s = t)$, where $\mathbf{1}(\cdot)$ is an indicator function. Finally, the covariance structure of $\mathbf{y}_i(s)$, denoted by $\Sigma_y(s, t)$, takes the form of $\Sigma_y(s, t) = \text{Cov}(\mathbf{y}_i(s), \mathbf{y}_i(t)) = \Sigma_\eta(s, t) + \Sigma_\epsilon(s, s)\mathbf{1}(s = t)$.

As an illustration, in our clinical study on early rhesus monkey brain development, we are interested in studying the evolution of the three eigenvalues λ_i

of diffusion tensor ($\lambda_1 \geq \lambda_2 \geq \lambda_3$) along two selected fiber tracts in 24 healthy rhesus monkeys (Fig. 2 (a)-(c)). See clinical data for details. We consider a model of λ_1 and $\lambda_{(2,3)} = (\lambda_2 + \lambda_3)/2$ along a specific tract as follows:

$$\begin{aligned} \lambda_{i,1}(s) &= \beta_{11}(s) + \beta_{12}(s) \times g_i + \beta_{13}(s) \times \text{age}_i + \beta_{14}(s) \times \text{age}_i^2 + \eta_{i1}(s), \\ \lambda_{i,(2,3)}(s) &= \beta_{21}(s) + \beta_{22}(s) \times g_i + \beta_{23}(s) \times \text{age}_i + \beta_{24}(s) \times \text{age}_i^2 + \eta_{i2}(s), \end{aligned} \tag{2}$$

where $\lambda_{i,k}$ are the three eigenvalues the i -th subject for $k = 1, 2, 3$, and g_i and age_i denote gender and age, respectively. In this case, $m = 2$, $B(s) = (\beta_{jk}(s))$ is a 2×4 matrix, and $\mathbf{x}_i = (1, g_i, \text{age}_i, \text{age}_i^2)^T$. It is trivial to extend model (2) to other nonlinear and nonparametric functions of age [13].

To estimate the coefficient functions in $B(s)$, we develop an adaptive *local polynomial kernel smoothing* technique [14], whereas it is possible to use spline type of methods including B-spline and smoothing spline. Specifically, using Taylor’s expansion, we can expand $B_k(s_j)$ at s to obtain $B_k(s_j) = B_k(s) + \dot{B}_k(s)(s_j - s) = A_k(s)\mathbf{z}_{h_{n_G,k}}(s_j - s)$, where $\mathbf{z}_{h_{n_G,k}}(s_j - s) = (1, (s_j - s)/h_{n_G,k})^T$ and $A_k(s) = [B_k(s), h_{n_G,k}\dot{B}_k(s)]$ is a $p \times 2$ matrix, in which $\dot{B}_k(s) = (\dot{b}_{k1}(s), \dots, \dot{b}_{kp}(s))$ is a $p \times 1$ vector and $\dot{b}_{kl}(s) = db_{kl}(s)/ds$ for $l = 1, \dots, p$. We calculate a weighted least squares estimate of $A_k(s)$ as follows. Let $K(\cdot)$ be a kernel function, such as the Gaussian and uniform kernels [14]. For a fixed bandwidth h and each k , we estimate $A_k(s)$ by minimizing an objective function given by

$$\sum_{i=1}^n \sum_{j=1}^{n_G} [y_{i,k}(s_j) - \mathbf{x}_i^T A_k(s)\mathbf{z}_{h_{n_G,k}}(s_j - s)]^2 K_{h_{n_G,k}}(s_j - s), \tag{3}$$

where $K_{h_{n_G,k}}(\cdot) = K(\cdot/h_{n_G,k})/h_{n_G,k}$ is a rescaled kernel function. For each k , we pool the data from all n subjects and select an optimal bandwidth $h_{n_G,k}$, denoted by $\hat{h}_{k,o}^{(1)}$, by minimizing the cross-validation score. Based on $\hat{h}_{k,o}^{(1)}$, we can obtain an estimate of $B_k(s)$, denoted by $\hat{B}_{k,o}(s)$.

2.2 Smoothing Individual Functions and Covariance Estimating

To simultaneously construct all individual functions $\eta_{i,k}(s)$, we also employ the local polynomial kernel smoothing technique [14]. Specifically, using Taylor’s expansion, we can expand $\eta_{i,k}(s_j)$ at s to obtain $\eta_{i,k}(s_j) = \mathbf{d}_{i,k}(s)^T \mathbf{z}_{h_{n_G,k}^{(2)}}(s_j - s)$, where $\mathbf{d}_{i,k}(s) = (\eta_{i,k}(s), h_k^{(2)}\dot{\eta}_{i,k}(s))^T$ is a 2×1 vector. We develop an algorithm to estimate $\mathbf{d}_{i,k}(s)$ as follows. For each k and a fixed bandwidth $h_{n_G,k}^{(2)}$, we estimate $\mathbf{d}_{i,k}(s)$ by minimizing an objective function given by

$$\sum_{j=1}^{n_G} [y_{i,k}(s_j) - \mathbf{x}_i^T \hat{B}_{k,o}(s_j) - \mathbf{d}_{i,k}(s)^T \mathbf{z}_{h_{n_G,k}^{(2)}}(s_j - s)]^2 K_{h_{n_G,k}^{(2)}}(s_j - s). \tag{4}$$

For each k , we pool the data from all n subjects and select the optimal bandwidth $h_{n_G,k}$, denoted by $\hat{h}_{k,o}^{(2)}$, by minimizing the generalized cross-validation score.

Based on $\hat{h}_{k,o}^{(2)}$, we can estimate $\eta_{i,k}(s)$ and $\boldsymbol{\eta}_i(s)$, denoted by $\hat{\eta}_{i,ko}(s)$ and $\hat{\boldsymbol{\eta}}_{i,o}(s)$, respectively, for all i and k .

After obtaining $\hat{\boldsymbol{\eta}}_{i,o}(s)$, we can estimate the mean function $\boldsymbol{\eta}(s)$ and the covariance function $\Sigma_\eta(s, t)$ of $\boldsymbol{\eta}_i(s)$. Specifically, we estimate $\boldsymbol{\eta}(s)$ and $\Sigma_\eta(s, t)$ by using their empirical counterparts of the estimated $\hat{\boldsymbol{\eta}}_{i,o}(s)$ as follows: $\hat{\boldsymbol{\eta}}_o(s) = n^{-1} \sum_{i=1}^n \hat{\boldsymbol{\eta}}_{i,o}(s)$ and $\hat{\Sigma}_\eta(s, t) = (n - m)^{-1} \sum_{i=1}^n \hat{\boldsymbol{\eta}}_{i,o}(s) \hat{\boldsymbol{\eta}}_{i,o}(t)^T$.

We construct a nonparametric estimator of the covariance matrix $\Sigma_\epsilon(s, s)$ as follows. Let $\hat{B}_o(s) = [\hat{B}_{1,o}(s), \dots, \hat{B}_{m,o}(s)]$ and $\hat{\boldsymbol{\epsilon}}_i(s_j) = \mathbf{y}_i(s_j) - \hat{B}_o(s_j)^T \mathbf{x}_i - \hat{\boldsymbol{\eta}}_{i,o}(s_j)$ be estimated residuals for $i = 1, \dots, n$ and $j = 1, \dots, n_G$. We consider an estimate of $\Sigma_\epsilon(s, s)$ given by $\hat{\Sigma}_\epsilon(s, s) = (n - m)^{-1} \sum_{i=1}^n \sum_{j=1}^{n_G} \tilde{K}_{h^{(3)}}(s_j - s) [\hat{\boldsymbol{\epsilon}}_i(s_j)]^{\otimes 2}$, where $\tilde{K}_{h^{(3)}}(s_j - s) = K_{h^{(3)}}(s_j - s) / \sum_{j=1}^{n_G} K_{h^{(3)}}(s_j - s)$. To select the optimal bandwidth $h^{(3)}$, denoted by $\hat{h}_o^{(3)}$, we minimize the cross-validation score. Based on $\hat{h}_o^{(3)}$, we can estimate $\Sigma_\epsilon(s, s)$, denoted by $\hat{\Sigma}_{\epsilon o}(s, s)$.

2.3 Test Statistics and Resampling Method

In neuroimaging studies, most scientific questions require the comparison of fiber bundle diffusion properties along fiber bundles across two (or more) diagnostic groups and the assessment of the development of fiber bundle diffusion properties along time. Such questions can often be formulated as linear hypotheses of $B(s)$ as follows: $H_0 : \mathbf{Cvec}(B(s)) = \mathbf{b}_0(s)$ for all s vs. $H_1 : \mathbf{Cvec}(B(s)) \neq \mathbf{b}_0(s)$, where \mathbf{C} is a $r \times mp$ matrix of full row rank and $\mathbf{b}_0(s)$ is a given $r \times 1$ vector of functions. We propose both *local and global* test statistics. The local test statistic can identify the exact location of significant grid point on a specific tract. At a given grid point s_j on a specific tract, we test the local null hypothesis $H_0(s_j) : \mathbf{Cvec}(B(s_j)) = \mathbf{b}_0(s_j)$ using a local test statistic $S_n(s_j) = \mathbf{nd}(s_j)^T [\mathbf{C}(\hat{\Sigma}_\eta(s_j, s_j) \otimes \hat{\Omega}_X^{-1})\mathbf{C}^T]^{-1} \mathbf{d}(s_j)$, where $\hat{\Omega}_X = n^{-1} \sum_{i=1}^n \mathbf{x}_i^{\otimes 2}$ and $\mathbf{d}(s) = \mathbf{Cvec}(\hat{B}_o(s) - \mathbf{b}_0(s))$. We test the null hypothesis $H_0 : \mathbf{Cvec}(B(s)) = \mathbf{b}_0(s)$ for all s using a global test statistic $S_n = n \int_0^{L_0} \mathbf{d}(s)^T [\mathbf{C}(\hat{\Sigma}_\eta(s, s) \otimes \hat{\Omega}_X^{-1})\mathbf{C}^T]^{-1} \mathbf{d}(s) ds$, where L_0 is the whole arc length of a specific fiber bundle. In order to use S_n as a test statistic, we can show that S_n has appropriate asymptotic distribution as $n \rightarrow \infty$. We develop a *resampling method* (or wild bootstrap method) to approximate the p -value of S_n . The key ideas are to fit model (II) under the null hypothesis H_0 , which yields $\hat{B}_o^*(s_j)$, $\hat{\boldsymbol{\eta}}_{i,o}^*(s_j)$ and $\hat{\boldsymbol{\epsilon}}_{i,o}^*(s_j)$ for $i = 1, \dots, n$ and $j = 1, \dots, n_G$, and then to generate random samples from the fitted model in order to approximate the null distribution of S_n .

3 Results

Clinical Data. Twenty four healthy rhesus monkeys (male and female included) at the Harlow Primate Laboratory with age between 10 to 72 months were scanned on a 3 Tesla GE scanner (SIGNA Excite) with a high-resolution 3DSPGR sequence ($0.2344 \times 0.2344 \times 0.4980mm^3$), a T2-weighted spin-echo sequence ($0.2344 \times 0.2344 \times 1.5mm^3$) and a 12-direction diffusion-weighted EPI sequence ($0.5469 \times 0.5469 \times 2.5mm^3$). After DTI estimation, a nonlinear fluid deformation based high-dimensional, unbiased atlas computation method was used

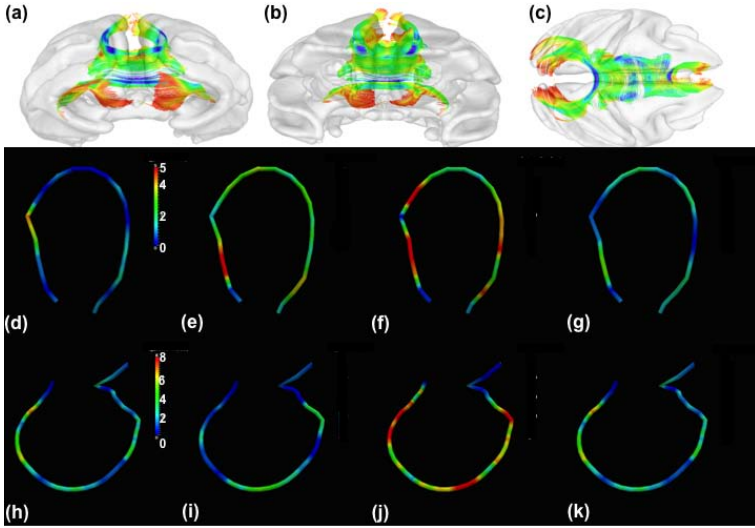


Fig. 2. Results from a study of neurodevelopment in healthy rhesus monkeys: panels (a)-(c): (a) anterior, (b) posterior, and (c) superior views of corpus callosum tracts; panels (d)-(g): $-\log_{10}(p)$ -values of $S_n(s_j)$ for testing time effect in the genu tract: (d) λ_1 , (e) $\lambda_{(2,3)}$, (f) FA, (g) MD; panels (h)-(k): $-\log_{10}(p)$ -values of $S_n(s_j)$ for testing time effect in the splenium tract: (d) λ_1 , (e) $\lambda_{(2,3)}$, (f) FA, (g) MD

to carry out a large deformation non-linear registration [15]. Detailed information regarding the DTI atlas building procedure has been described in [9]. Major fiber bundles are tracked in the atlas space within 3D Slicer (www.slicer.org). With the fiber bundles in atlas space, each subject's DTI data is transformed into the atlas space. For each subject at a given time point, the data within the fiber bundle is parameterized as a sampled function of equidistance steps along the fiber. The result of the procedure is thus a set of corresponding sampled functions, including FA, MD, etc. parameterized by arc length from the atlas fiber tract for each individual subject using invert of the atlas-building transformation. These sampled functions at each point along the fiber tract were then used to study the effect of age, gender and other covariates on neural development.

For the sake of space, we chose two tracts of interest including the splenium and genu of the corpus callosum tract and then computed fractional anisotropy (FA), mean diffusivity (MD), and λ_1 and $\lambda_{(2,3)}$ of diffusion tensors at each grid point on both tracts for each of the 24 monkeys. FA denotes the inhomogeneous extent of local barriers to water diffusion, while MD measures the averaged magnitude of local water diffusion. The three eigenvalues of diffusion tensor may, respectively, reflect the magnitude of water diffusivity along and perpendicular to the long axis of white matter fibers [16].

We applied VCMTS to the joint analysis of λ_1 , $\lambda_{(2,3)}$, FA, and MD values along the splenium tract as follows. We fitted the functional linear model (2) to these

four diffusion properties from all 24 subjects, in which $\mathbf{x}_i = (1, g_i, \text{age}_i, \text{age}_i^2)^T$ and $m = 2$, and then we estimated the function of regression coefficient vector $\hat{B}(s)$. Secondly, we constructed the global test statistic S_n to test the effects of all the age effect for each of the four diffusion properties, and performed hypothesis testing on the whole splenium and genu tracts. The p -value of S_n was approximated using the resampling method with $G = 10,000$.

We considered the genu tract and performed hypothesis testing on time effect for the whole tract. The p -values of S_n corresponding to λ_1 , $\lambda_{(2,3)}$, FA, and MD equal 0.31, 0.19, 0.007, and 0.29, respectively. This indicates a significant change of the degrees of anisotropy, not the degree of diffusivity, along the genu tract. We further performed hypothesis testing at each grid point along the splenium tract (Figs. 2). For λ_1 , $\lambda_{(2,3)}$, and MD, no significant effect of time effect was found, even though the $-\log_{10}(p)$ values of $S_n(s)$ for age at several single grid points were slightly greater than 2 (Fig. 2 (d), (e), (g)). For FA alone, the effects of time were significant in the middle and tail of the genu tract (Fig. 2 (f)).

We considered the splenium tract and performed hypothesis testing on time effect for the whole tract. The p -values of S_n corresponding to λ_1 , $\lambda_{(2,3)}$, FA, and MD equal 0.001, 0.002, 0.000, and 0.004, respectively. This indicates a significant change of the degrees of diffusivity and anisotropy, along the splenium tract. We further performed hypothesis testing at each grid point along the splenium tract (Fig. 2 (h)-(k)). For all diffusion properties, the effects of time were significant in most of grid points along the splenium tract (Fig. 2 (h)-(k)).

4 Discussion

We have developed VCMTS for statistically analyzing multiple diffusion properties along fiber bundle and assessing their association with a set of covariates in real applications. The proposed methodology is demonstrated in a study of neurodevelopment in rhesus monkey. Significant time effect on multiple diffusion properties were examined and localized in two representative tracts. VCMTS is able to delineate the complex inhomogeneous spatial-temporal maturation patterns as the apparent changes in FA, MD, and the eigenvalues of diffusion tensors. Specifically, our results suggest that white matter maturation patterns are different in different white matter regions. We expect that this novel statistical tool will lead to new findings in our clinical applications.

References

1. Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M., Matthews, P., Behrens, T.E.: Tractbased spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage* 31, 1487–1505 (2006)
2. O'Donell, L.J., Westin, C.F., Golby, A.J.: Tract-based morphometry for white matter group analysis. *NeuroImage* 45, 832–844 (2009)
3. Snook, L., Plewes, C., Beaulieu, C.: Voxel based versus region of interest analysis in diffusion tensor imaging of neurodevelopment. *NeuroImage* 34, 243–252 (2007)

4. Snook, L., Paulson, L.A., Roy, D., Phillips, L., Beaulieu, C.: Diffusion tensor imaging of neurodevelopment in children and young adults. *NeuroImage* 26, 1164–1173 (2005)
5. Worsley, K.J., Taylor, J.E., Tomaiuolo, F., Lerch, J.: Unified univariate and multivariate random field theory. *Neuroimage* 23, 189–195 (2004)
6. Hecke, W.V., Sijbers, J., Backer, S.D., Poot, D., Parizel, P.M., Leemans, A.: On the construction of a ground truth framework for evaluating voxel-based diffusion tensor mri analysis methods. *NeuroImage* 46, 692–707 (2009)
7. Jones, D.K., Symms, M.R., Cercignani, M., Howard, R.J.: The effect of filter size on vbm analyses of dt-mri data. *NeuroImage* 26, 546–554 (2005)
8. Yushkevich, P.A., Zhang, H., Simon, T., Gee, J.C.: Structure-specific statistical mapping of white matter tracts. *Neuroimage* 41, 448–461 (2008)
9. Goodlett, C.B., Fletcher, P.T., Gilmore, J.H., Gerig, G.: Group analysis of dti fiber tract statistics with application to neurodevelopment. *NeuroImage* 45, S133–S142 (2009)
10. Gouttard, S., Prastawa, M., Bullitt, E., Lin, W.L., Goodlett, C., Gerig, G.: Constrained data decomposition and regression for analyzing healthy aging from fiber tract diffusion properties. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5761, pp. 321–328. Springer, Heidelberg (2009)
11. Zhu, H.T., Styner, M., Tang, N.S., Liu, Z.X., Lin, W.L., Gilmore, J.: Frats: Functional regression analysis of dti tract statistics. *IEEE Transactions on Medical Imaging* 29, 1039–1049 (2010)
12. Zhu, H.T., Li, R.Z., Kong, L.N., Styner, M., Gilmore, J.: Multivariate varying coefficient models with applications for dti tract statistics. Technical report, University of the North Carolina (2010)
13. Fan, J., Zhang, W.: Statistical methods with varying coefficient models. *Statistics and Its Interface* 2 (2008)
14. Fan, J., Gijbels, I.: *Local Polynomial Modelling and Its Applications*. Chapman and Hall, London (1996)
15. Joshi, S., Davis, B., Jomier, M., Gerig, G.: Unbiased diffeomorphic atlas construction for computational anatomy. *Neuroimage* 23, S151–S160 (2004)
16. Song, S.K., Sun, S.W., Ju, W.K., Lin, S.J., Cross, A.H., Neufeld, A.H.: Diffusion tensor imaging detects and differentiates axon and myelin degeneration in mouse optic nerve after retinal ischemia. *Neuroimage* 20, 1714–1722 (2003)

Author Index

- Abolmaesumi, Purang II-68, III-311
Abràmoff, Michael D. III-33, III-603
Abu El-Ghar, Mohammed I-10
Abugharbieh, Rafeef II-331
Achanta, Radhakrishna II-463
Adebar, Troy K. II-626
Adeshina, Steve A. II-635
Aganj, Iman II-84
Ahmidi, Narges III-295
Aja-Fernandez, S. I-518
Aksoy, Murat I-259
Al-Sanawi, Hisham III-359
Alexander, Daniel C. I-183, I-534,
I-623, I-640, II-404
Alhonnoro, Tuomas I-45
Aljabar, P. III-1
Allain, Baptiste II-514
Altinay, Murat III-105
An, Hongyu II-274
Anderson, Jeffrey S. II-363
André, Barbara II-480
Andrews, Shawn III-9
Angelini, Elsa D. II-34
Arbel, Tal I-127, II-290, II-643, III-41
Armspach, Jean-Paul II-117
Arnold, Douglas L. II-290, III-41
Asano, Takaharu II-50
Ashraf, Haseem I-37
Atasoy, Selen II-437
Atkins, M. Stella I-582
Audière, Stéphane II-34
Ausubel, F.M. III-634
Avants, Brian I-324, III-105
Ayache, Nicholas I-111, II-151,
II-420, II-480, II-652
Ayles, Helen III-650
Azernikov, Sergei III-555

Bab-Hadiashar, Alireza II-193
Baka, Nora I-452
Bakker, N.H. I-526
Balachandran, Ramya III-587
Balda, Michael III-547
Balicki, Marcin III-303

Balocco, Simone II-59
Baloch, Sajjad III-113, III-555
Bammer, Roland I-259
Barber, David II-380
Barbu, Adrian I-28
Barkovich, A. James II-339
Barmpoutis, Angelos I-582
Barnes, Josephine II-125
Barnes, Nick II-266
Baronnet, Flore I-200
Barron, John III-205
Batchelor, Philip G. I-418
Baust, Maximilian II-586
Bax, Jeff II-17
Bellani, M. II-177
Ben Ayed, Ismail I-409
Ben-Bashat, Dafna I-103
Ben-Sira, Liat I-103
Benali, Habib I-316
Berezney, Ronald II-530
Bernardis, Elena I-119
Beuthien, Björn II-546
Biesdorf, Andreas I-444
Biller, Lisa I-151
Blum, Christian I-291
Blum, Tobias III-400
Bó, Antônio P.L. I-267
Boardman, J.P. III-1
Bocan, Thomas I-308, III-57
Bock, Jelena III-416
Boctor, Emad M. II-9
Boese, Jan I-375, I-476
Boisvert, Jonathan II-68
Boomen, R.v.d. I-526
Boone, Darren III-497
Borschneck, Dan II-68
Bourgeat, Pierrick II-125, II-185
Bove, Susan III-57
Brambilla, P. II-177
Breteler, Monique M.B. II-101
Brockmann, Gernot I-476
Brown, Tracy I-308
Brugada, Josep II-1
Brunenberg, Ellen I-175

- Buchner, Anna M. II-480
 Buhmann, Joachim M. II-209
 Buonaccorsi, G.A. III-121
 Burdette, Clif III-311
- Caan, Matthan W.A. I-167, II-249
 Cabeen, Ryan P. I-357
 Cagniard, Cedric III-237
 Camara, Oscar II-1
 Carpenter, A.E. III-634
 Carvajal-Gonzalez, Santos III-57
 Castellani, U. II-177
 Cavallaro, Alexander I-95
 Chakrapani, Shruthi I-357
 Chan, Kap Luk II-522
 Chan, Tony II-323
 Chaney, Ed III-335
 Chang, Yu-Bing III-278
 Changizi, Neda III-17, III-563
 Charbit, Maurice II-34
 Chatelin, Simon I-235
 Chelikani, Sudhakar I-53
 Chen, Elvis III-205
 Chen, Hanbo II-412
 Chen, Hua-mei I-340
 Chen, Mei I-1, I-209
 Chen, Sean Jy-Shyang II-92
 Chen, Taoyi III-473
 Chen, Terrence III-269
 Chen, Ting III-65
 Chen, Yasheng II-274
 Chen, Zhe I-53
 Cheng, Jack C.Y. III-538
 Cheng, Jian I-590, I-648
 Chertok, Michael III-642
 Cheung, Carling L. III-408
 Chiao, Ping-Chun III-57
 Chinchapatnam, P. II-420
 Chitphakdithai, Nicha I-367
 Cho, Daniel S. III-205
 Chowdhury, Ananda S. III-89
 Chu, Winnie C.W. III-538
 Chu, Xinqi II-522
 Chung, Adrian I-69
 Chung, Moo K. III-505
 Ciompi, Francesco II-59
 Clarkson, Matthew J. I-534, II-125
 Clatz, Olivier I-111
 Coe, Christopher I-690
 Cointepas, Yann I-550
- Collins, D. Louis II-92, II-290, II-643,
 III-41, III-129, III-181
 Colliot, Olivier I-316
 Comaniciu, Dorin I-28, I-95, I-218,
 I-383, I-460, I-476, III-269
 Comas, Olivier II-371
 Combès, Benoît II-594
 Commowick, Olivier III-25, III-155
 Conery, A.L. III-634
 Constantini, Shlomi I-103
 Cook, Philip A. I-324
 Cook, Richard J. II-514
 Cootes, Timothy F. II-635
 Corani, G. II-488
 Cotin, Stéphane II-371
 Coulthard, A. II-185
 Counsell, S. III-1
 Coupé, Pierrick III-129
 Cowan, Brett I-383
 Craige, Caryne III-105
 Criminisi, Antonio I-111
 Crozier, Stuart II-185, II-388
 Cuingnet, Rémi I-316
- D'hooge, Jan II-1
 Dabbah, M.A. I-300
 Dabool, Elad III-457
 Daghli, M. II-185
 Dalal, Pahal I-349
 Danilchenko, Andrei III-587
 Darzi, Ara W. III-245, III-319
 Das, Sandhitsu III-105
 Dassopoulos, T. II-454
 Davatzikos, Christos II-160,
 II-257, II-676
 Davydov, Oleg I-666
 Dawson, Rodney III-650
 de Boer, Renske II-101
 de Bruijne, Marleen I-37, I-452,
 II-193, III-595
 De Camilli, Pietro II-315
 De Craene, Mathieu II-1
 de Groot, Marius II-101
 de Jong, Pim A. II-396, III-650
 de La Gorce, Martin II-668
 De Nigris, Dante II-643
 De Silva, Tharindu III-213
 Dehghan, Ehsan I-283
 Del Maestro, Rolando II-643
 Delingette, Hervé I-235, II-420, II-652

- Delmaire, Christine II-217
 Delmas, Patrice III-481
 Demiralp, Çağatay I-542
 den Dekker, Arjan I-167
 Deriche, Rachid I-590, I-631, I-648
 Descoteaux, Maxime I-550
 Díaz, Alejandro III-163
 Dietemann, Jean-Louis I-574, II-355
 Dinov, Ivo I-357, III-49, III-81
 Dirksen, Asger I-37, II-193
 Doltra, Adelina II-1
 Dong, Bin I-143
 Doria, V. III-1
 Dormont, Didier I-316
 dos Santos, Thiago R. I-251, II-660
 Douglas, Tania S. III-619
 Dowson, N. II-185
 Duchateau, Nicolas II-1
 Duda, Jeffrey T. II-282
 Duin, Robert P.W. I-37
 Duits, Remco I-175
 Duncan, James S. I-53, I-367,
 I-502, II-538, II-315
 Duriez, Christian II-371
 Durr, Alexandra II-217
- Ecabert, O. I-400
 Edwards, A.D. III-1
 Eisenschenk, Stephan J. III-65
 El-Baz, Ayman I-10
 Elhabian, Shireen III-626
 Elliott, Colm II-290
 Ellis, Randy E. III-359
 Elson, Daniel III-245
 Engin, Kayihan III-137
 Eskildsen, Simon F. III-181
 Essafi, Salma III-189
 Esses, S.J. III-73
 Estroff, Judy A. II-109
- Falk, Robert III-626
 Fallavollita, Pascal III-311
 Fang, Tong III-113, III-555
 Farag, Aly III-626
 Farag, Amal III-626
 Fay, M. II-185
 Fenster, Aaron II-17, III-213
 Feußner, Hubertus III-400
 Feulner, Johannes I-95
- Fichtinger, Gabor I-283, II-68,
 III-295, III-311, III-383, III-424
 Fillard, Pierre I-200, I-550
 Finkelstein, Yehuda III-642
 Fischer, Bernd II-546
 Fitzpatrick, J. Michael I-251, III-587
 Flanagan, Ronan I-45
 Fleming, Ioana N. II-9
 Fletcher, P. Thomas II-363, III-529
 Flöry, Simon III-555
 Folkerts, Michael III-449
 Fonov, Vladimir III-129
 Forman, Christoph I-259
 Foroughi, Pezhman II-9
 Foskey, Mark III-335
 Fox, Nick C. I-534, II-125
 Francis, Simon J. II-290, III-41
 Frangi, Alejandro F. I-468, I-518, II-1
 Franz, A.M. I-251
 Freiman, M. III-73
 Frigstad, Sigmund I-510
 Friman, Ola III-416
 Fripp, Jurgen II-125
 Fritz, Andrew II-530
 Fua, Pascal I-291, II-463
 Fuchs, Siegfried III-113
 Fuchs, Thomas J. II-209
- Gallia, Gary L. III-295
 Gamage, Pavan III-481
 Gambardella, L. II-488
 Gammon, Braden III-359
 Gangeh, Mehrdad J. III-595
 Ganor, Roy III-642
 Gao, Dashan II-446
 Gao, Fei III-571
 Gao, Hang II-1
 Gao, Wei II-298
 Gardi, Lori II-17
 Garvin, Gregory J. I-340
 Garvin, Mona K. III-33
 Gaser, Christian II-169
 Gateno, Jaime III-278
 Gatta, Carlo II-59
 Gauvrit, Jean-Yves II-92
 Gee, James C. I-324, II-282
 Gehlbach, Peter III-303, III-465
 Geng, Xiujuan I-598
 Georgescu, Bogdan I-218, I-383,
 I-460, I-476

- Geremia, Ezequiel I-111
 Gerig, Guido II-602
 Gholipour, Ali II-109
 Ghosh, Aurobrata I-590, I-648
 Ghotbi, Reza III-237
 Gianaroli, L. II-488
 Gill, Jaswinder I-391
 Gill, Sean II-68
 Gilmore, John H. I-690, II-602
 Gimel'farb, Georgy I-10
 Giovanello, Kelly II-298
 Giusti, A. II-488
 Glenn, Orit A. II-339
 Glocker, Ben II-676
 Goela, Aashish I-340
 Gogin, Nicolas I-391
 Golby, Alexandra J. II-225
 Goldstein, Jill I-657
 Golland, Polina I-191, II-151, III-634
 Gong, Ren Hui II-68
 González, Germán I-291
 Gorbunova, Vladlena II-193
 Graham, J. I-300,
 Graham, James III-626
 Grbić, Saša I-218
 Grigis, Antoine II-117
 Grimbergen, Kees I-167
 Grossman, Murray I-324, II-282
 Grzeda, Victor III-424
 Gu, Hong I-598
 Gu, Xianfeng II-323, III-538
 Gu, Xuejun III-449
 Guehring, Jens I-383
 Guevara, Pamela I-550, II-217, II-347
 Guion, Peter III-383
 Guo, Lei II-143, II-412
 Gur, Ruben C. I-558, I-631
 Gutman, Boris I-357
 Gyacskov, Igor II-17

 Ha, Linh II-602, III-529
 Habas, Piotr A. II-339
 Hager, Gregory D. II-9, II-454,
 III-295, III-465
 Hajnal, J.V. III-1
 Hall, Matt G. II-404
 Halligan, Steve III-497
 Hamamci, Andac III-137
 Hamarneh, Ghassan II-331, III-9,
 III-17, III-563

 Hameeteman, Reinhard III-97
 Handa, James III-303, III-465
 Harloff, Andreas III-416
 Hassan Zahraee, Ali III-432
 Hawkes, David J. I-435, II-514,
 III-497
 Hawkins, Maria A. I-69
 He, Huiguang III-489
 He, Tiancheng III-392
 Hege, H.-C. I-227
 Heibel, Hauke III-343
 Heimann, Tobias I-235
 Heismann, Björn III-547
 Heitz, Fabrice II-117
 Hellier, Pierre II-92
 Hemetsberger, Rayyan II-59
 Hennemuth, Anja III-416
 Hicks, Carol I-308
 Ho, Jeffrey I-682
 Ho, Khek Yu II-522
 Hoffman, Eric A. II-578
 Hogeweg, Laurens II-396, III-650
 Højgaard, Liselotte III-253
 Hojjatoleslami, Ali I-666
 Holm, Darryl D. II-610
 Honnorat, Nicolas III-440
 Hornegger, Joachim I-95, I-151, I-259,
 I-460, III-547
 Hose, Rod II-380
 Htwe, That Mon II-522
 Hu, Jiayi I-332
 Hu, Mingxing II-514, III-497
 Hu, Xintao II-143, II-412
 Hu, Zhihong III-33
 Hua, Jing I-332, III-489, III-538
 Huang, Junzhou I-135
 Huang, Xiaolei I-86
 Huber, Martin I-95, I-460

 Iglesias, Juan Eugenio III-81, III-147
 Ikram, M. Arfan II-101
 Ingallhalikar, Madhura I-558
 Ionasec, Razvan Ioan I-218, I-460
 Iordachita, Iulian III-303
 Ishaq, Omer III-17
 Ishii, Lisa III-295
 Ishii, Masaru III-295
 Ishikawa, Hiroshi I-1
 Islam, Ali I-409

- Jackson, A. III-121
 Jacobs, Sander S.A.M. II-193
 Jacques, Robert III-521
 Jähne, B. I-251
 James, David R.C. III-319
 Janoos, F. II-201
 Jayson, G.C. III-121
 Ji, Songbai II-274
 Jia, Hongjun II-570, II-618, II-684
 Jia, Xun I-143, III-449
 Jiang, Steve B. I-143, III-449
 Jiang, Tianzi I-590, I-648
 John, Matthias I-375, I-476
 Jolly, Marie-Pierre I-383
 Joshi, Anand A. I-357
 Joshi, Sarang II-602, III-529
 Joshi, Shantanu H. I-357
 Joskowicz, Leo I-103, III-73, III-457
 Jürgens, Philipp I-61
- Kadoury, Samuel III-579
 Kadowaki, Takashi II-50
 Kainmueller, D. I-227
 Kainz, Bernhard I-45
 Kamel, Mohamed S. III-595
 Kamen, Ali II-546, II-586
 Kanade, Takeo I-209
 Kanterakis, Efstathios I-558, I-631
 Karamalis, Athanasios I-243
 Karimaghloo, Zahra III-41
 Karnik, Vaishali V. II-17
 Karpikov, Alexander II-538
 Kaushal, Aradhana III-383
 Kaynig, Verena II-209
 Kazhdan, Michael III-521
 Kempfert, Jörg I-375, I-476
 Khalifa, Fahmi I-10
 Khallaghi, Siavash II-68
 Khedoe, Ganesh I-167
 Kim, Hyungmin I-61
 Kim, Jinman II-562
 Kim, Kio II-339, II-355
 Kim, Minjeong II-306
 Kim, Sungeun III-611
 Kindlmann, Gordon I-674, III-163
 King, Alex III-521
 King, Andy P. I-391
 King, Richard III-529
 Kirisli, Hortense I-452
- Kirschstein, Uwe I-375, I-476
 Klein, Stefan I-452
 Kleinschmidt, Andreas I-200
 Klinder, Tobias III-327
 Kneser, R. I-400, I-526
 Knopp, M.V. II-201
 Koizumi, Norihiro II-50
 Kong, Linglong I-690
 Konukoglu, Ender I-111
 Koob, Meriam I-574, II-355
 Korosoglou, G. I-526
 Kozerke, Sebastian I-418
 Krieger, Axel III-383
 Krishnan, Arun I-19
 Kronman, A. III-73
 Kroon, Dirk-Jan III-221
 Kruecker, Jochen II-42
 Kubicki, Marek I-191
 Kubota, Naoto II-50
 Kucuk, Nadir III-137
 Kumar, R. II-454
 Kurhanewicz, John III-666
 Kwok, Ka-Wai III-229, III-319
 Kyujin Paik, Jamie III-432
- Ladikos, Alexander III-237
 Lai, Rongjie III-49
 Lai, Zhaoqiang I-332
 Laidlaw, David H. I-542
 Lallemand, Joe II-437
 Lamata, Pablo II-380
 Lamecker, H. I-227
 Laporte, Catherine I-127
 Larsen, Rasmus III-253
 Lashkari, Danial II-151
 Lasso, Andras III-383
 Lauritsch, Günter I-151
 Lee, Huai-Ping III-335
 Lee, Jack II-380
 Lee, Junghoon I-283
 Lee, Kyungmoo III-33
 Lee, Su-Lin I-69
 Lee, Tzumin II-472
 Leff, Daniel R. III-319
 Lehéricy, Stéphane I-316, II-217
 Lelieveldt, Boudewijn I-452
 Lenglet, Christophe II-84
 Lepetit, Vincent II-463
 Lerotic, Mirna I-69
 Lessick, J. I-400

- Leung, Kelvin K. I-435, I-534, II-125
 Levy, Josh III-335
 Lewis, John H. III-449
 Li, Hongsheng I-86
 Li, Kaiming II-143
 Li, Kang I-209
 Li, Liyuan II-522
 Li, Ruijiang III-449
 Li, Shuo I-340, I-409
 Li, Yang II-133
 Li, Yimei I-690
 Liachenko, Serguei III-57
 Liang, Liang II-315
 Liao, Hongen II-50
 Liao, Rui I-375, I-476
 Liao, Wei-shing III-269
 Lilja, Mikko I-45
 Lim, Joo Hwee II-522
 Lin, Ching-Long II-578
 Lin, Weili I-690, II-133, II-274, II-298
 Linguraru, Marius George III-89
 Linte, Cristian A. III-205
 Litjens, G.J.S. II-396
 Littmann, Arne I-383
 Liu, Chang I-332
 Liu, Cheng-Yi III-147
 Liu, David I-28
 Liu, Huafeng III-571
 Liu, Jianfei II-505
 Liu, Jiang II-522
 Liu, Qingshan I-484
 Liu, Tianming II-143, II-412
 Liu, Wei II-363
 Liu, Yinxiao III-172
 Liu, Yunlong III-172
 Liu, Yu-Ying I-1
 Ljosa, V. III-634
 Lo, Pechin I-37, II-193
 Loog, Marco I-37, III-595
 Lorenz, Cristian III-327
 Lou, Yifei I-143
 Lovat, Laurence B. II-514
 Lu, Chao I-53
 Lu, Le I-19
 Lu, Xiaoguang I-383
 Lucchi, Aurélien II-463
 Lui, Lok Ming II-323, III-538
 Lv, Jinglei II-143
 Lythgoe, Mark F. II-404
 Ma, Jun I-19
 Ma, YingLiang I-391
 Maal, Thomas J.J. III-221
 MacFarlane, D. II-185
 Machiraju, R. II-201
 Madabhushi, Anant III-197,
 III-658, III-666
 Magli, C. II-488
 Magnenat-Thalmann, Nadia II-562
 Mahapatra, Dwarikanath I-493
 Mahdavi, S. Sara II-76
 Maier-Hein, Lena I-251, II-660
 Maitrejean, Serge II-668
 Majumdar, Angshul III-513
 Malandain, Grégoire III-155
 Malcolm, James G. I-657, II-233
 Malik, R.A. I-300
 Malone, Ian B. I-534
 Mangin, Jean-François I-550,
 II-217, II-347
 Manjón, José V. III-129
 Mansi, T. II-652
 Marchal, Maud II-92
 Marchesseau, Stéphanie I-235
 Markl, Michael III-416
 Marrakchi-Kacem, Linda I-550, II-217
 Marron, J. Stephen III-529
 Martin-Fernandez, M. I-518
 Marvasty, Idean I-484
 Mateus, Diana II-437
 Mayrhauser, Ursula I-45
 McCarley, Robert W. I-657
 McClelland, Jamie III-497
 McKay, Richard II-446
 McMillan, Corey I-324, II-282
 McNutt, Todd III-521
 Meining, Alexander II-437
 Meinzer, Hans-Peter I-251, II-660
 Melkisetoglu, Rupen III-555
 Men, Chunhua III-449
 Menze, Bjoern H. I-111, II-151
 Mercier, Laurence II-643
 Meriaudeau, Fabrice II-125
 Merrifield, Robert III-261
 Metaxas, Dimitris N. I-135, I-484
 Metz, Coert I-452
 Michailovich, Oleg I-607, I-657
 Michel, Fabrice III-189
 Mochizuki, Takashi II-50
 Modat, Marc I-534, III-497

- Mol, Christian III-650
 Mollet, Nico I-452
 MomayyezSiahkal, Parya I-566
 Monaco, James P. III-197
 Montreuil, Jacques II-17
 Moore, John III-205, III-408
 Moradi, Mehdi I-283, II-76
 Morandi, Xavier II-92
 Morel, Guillaume III-432
 Mórocz, I.Á. II-201
 Morris, William J. II-76
 Morton, Daniel I-308
 Mountney, Peter II-496
 Mousavi, Parvin II-68
 Mouton, André III-619
 Muehl, Judith I-45
 Mueller, Susanne III-105
 Mullin, G. II-454
 Murgasova, M. III-1
 Murino, V. II-177
 Myers, Eugene II-472
 Mylonas, George P. III-319
- Nachmani, Ariela III-642
 Narasimhamurthy, Anand I-510
 Navab, Nassir I-218, I-243, II-26, II-437,
 II-586, III-237, III-343, III-400
 Neeffjes, Lisan I-452
 Nenadic, Igor II-169
 Newton, Richard III-245
 Ng, Bernard II-331
 Nguan, Christopher Y. II-626
 Nho, Kwangsik III-611
 Nie, Jingxin II-412
 Niederer, Steven II-380
 Nielsen, Mads II-193
 Niemeijer, Meindert III-33, III-603
 Niessen, Wiro J. I-452, II-101, III-97
 Noblet, Vincent II-117
 Nolte, Lutz-Peter I-61
 Noonan, David P. III-245
 Norsletten, David II-380
 Norton, Isaiah II-225
 Nöttling, Alois I-375
- O'Connor, J.P.B. III-121
 O'Donnell, Lauren J. II-225
 Oentoro, Anton III-359
 Olabarriaga, Silvia I-167
 Olesen, Oline Vinter III-253
- Ong, Eng Hui II-522
 Orihuela-Espina, Felipe III-319
 Ostermann, Jörn III-327
 Ou, Yangming II-676
 Oubel, Estanislao I-574
 Oudry, Jennifer II-34
 Ourselin, Sébastien I-435, I-534, II-125,
 II-388, II-514, III-497
- Padfield, Dirk I-510, II-446
 Pai, Darshan I-332
 Palencia, C. I-518
 Panagiotaki, Eleftheria II-404
 Papademetris, Xenophon I-53
 Paragios, Nikos II-668, II-676,
 III-189, III-440, III-579
 Parker, G.J.M. III-121
 Patel, Rajni III-205
 Paulsen, Rasmus R. III-253
 Pauly, Olivier III-343
 Pautler, Stephen E. III-408
 Pavlidis, I. III-351
 Payne, Christopher J. III-245
 Pearlman, Paul C. I-502
 Peitgen, Heinz-Otto III-416
 Peng, Hanchuan II-472
 Pennec, X. II-652
 Perina, A. II-177
 Peters, J. I-400, I-526
 Peters, Terry M. III-205, III-408
 Petropoulos, I. I-300
 Pichora, David R. II-68, III-359
 Piella, Gemma II-1
 Pieper, Steve II-554
 Pike, G. Bruce III-181
 Pitcher, Richard D. III-619
 Platel, Bram I-175
 Pluta, John III-105
 Poh, Chee Khun II-522
 Poignet, Philippe I-267
 Poline, Jean-Baptiste II-241
 Pollari, Mika I-45
 Poot, Dirk I-167, I-615
 Portugaller, Horst I-45
 Poupon, Cyril I-550, II-217, II-347
 Poupon, Fabrice II-217
 Prabhu, Sanjay P. II-109
 Prastawa, Marcel II-602
 Pratt, Philip I-77, I-275
 Precel, Ronit I-103

- Preston, J. Samuel III-529
 Prima, Sylvain II-594
 Propper, Ruth II-225
 Pruessner, Jens III-129
 Punithakumar, Kumaradevan I-409
 Pura, John A. III-89

 Qi, Yuan III-611
 Qian, Zhen I-484
 Qu, Zhenshen III-473

 Radeva, Petia II-59
 Rajagopalan, Vidya II-339
 Rambaldelli, G. II-177
 Ramus, Liliane III-155
 Rangarajan, Anand III-65
 Rathi, Yogesh I-191, I-607, I-657, II-233
 Ratnarajah, Nagulan I-666
 Razavi, Reza I-391, I-435, II-420
 Rehg, James M. I-1
 Reiber, Johan H.C. III-367
 Reiser, Maximilian III-237
 Relan, J. II-420
 Renard, Félix II-117
 Resnick, Susan M. II-160
 Reyes, Mauricio I-61
 Rhode, Kawal S. I-391, I-435, II-420
 Richa, Rogério I-267
 Richardson, John B. III-181
 Ridgway, Gerard R. II-125
 Rigolo, Laura II-225
 Riklin Raviv, T. III-634
 Rinaldi, C. Aldo I-391
 Rinehart, Sarah I-484
 Risacher, Shannon L. III-611
 Risholm, Petter II-554
 Risser, Laurent II-610
 Rittscher, Jens II-446
 Rivaz, Hassan II-9
 Rivière, Denis I-550, II-347
 Roberts, C. III-121
 Roberts, Timothy P.L. I-558
 Robles, Montserrat III-129
 Roca, Pauline II-217, II-347
 Rodriguez Leor, Oriol II-59
 Roed, Bjarne III-253
 Rohkohl, Christopher I-151
 Rohling, Robert N. II-626
 Rohr, Karl I-444
 Romagnoli, Cesare II-17

 Rose, C.J. III-121
 Rose, S. II-185
 Rosen, Mark III-666
 Ross, Ian G. I-409
 Ross, James C. III-163
 Ross, Thomas J. I-598
 Rosso, Charlotte I-316
 Roth, Holger III-497
 Rother, Diego III-465
 Rousseau, Francois I-574, II-339, II-355
 Rowe, Matthew C. I-183
 Rueckert, Daniel II-610, III-1
 Rumbach, Lucien II-117
 Rutherford, M.A. III-1
 Ryan, Natalie I-534

 Saad, Ahmed III-9
 Saboo, Rohit III-335
 Saha, Punam K. III-172
 Sahin, Mustafa II-109
 Sakuma, Ichiro II-50
 Salcudean, Septimiu E. I-283, II-76,
 II-429, II-626
 Salganicoff, Marcos I-19
 Salvado, Olivier II-125, II-185
 Samarabandu, Jagath III-213
 Sammet, S. II-201
 Samset, Eigil II-554
 Samson, Yves I-316
 Sánchez, C.I. III-603
 San José Estépar, Raúl III-163
 Sandrin, Laurent II-34
 Sapiro, Guillermo II-84
 Sauer, Frank III-287
 Sauvage, Vincent III-245
 Savadjiev, Peter II-233
 Savinaud, Mickael II-668
 Saykin, Andrew J. III-611
 Schaap, Michiel I-452, II-101, III-97
 Schievano, Silvia I-460
 Schilham, A.M.R. II-396
 Schmid, Jérôme II-562
 Schmidt, M. I-251
 Schneider, Torben I-623
 Schultz, Thomas I-674
 Schuman, Joel S. I-1
 Scott, Julia II-339
 Seidman, Larry I-657
 Seim, H. I-227
 Seitel, Alexander I-251, II-660

- Seo, Seongho III-505
 Sermesant, Maxime I-418, II-420, II-652
 Seshamani, S. II-454
 Shah, Mohak III-41
 Shahid, Muhammad Waseem II-480
 Shaker, Saher B. III-595
 Shamir, Reuben R. III-457
 Shamoin, Denis III-367
 Shang, Jianzhong III-245
 Shen, Chunhua II-266
 Shen, Dinggang I-349, II-133, II-274,
 II-306, II-570, II-618, II-684
 Shen, Hongying II-315
 Shen, Kai-Kai II-125
 Shen, Li III-611
 Shen, Tian I-86
 Shen, Weijia II-522
 Shenton, Martha E. I-657, II-233
 Sherbondy, Anthony J. I-183
 Shi, Feng I-349, II-133
 Shi, Lin III-538
 Shi, Pengcheng I-159, III-571
 Shi, Yonggang III-49
 Shi, Yundi I-690
 Shin, Wanyong I-598
 Shoshan, Yigal III-457
 Siddiqi, Kaleem I-566
 Sijbers, J. I-615
 Silva, Cláudio T. II-602
 Silva, Etel II-1
 Silverman, Edwin K. III-163
 Simari, Patricio III-521
 Simmons, Andrew I-666
 Singh, Jaskaran III-81
 Singh, Nikhil III-529
 Sinusas, Albert J. I-502
 Siow, Bernard II-404
 Sitges, Marta II-1
 Slabaugh, Greg III-497, III-555
 Slump, Cornelis H. III-221
 Smith, Erin Janine III-359
 Smith, J. II-185
 Smith, Kevin II-463
 Smith, Nic II-380
 Song, Danny III-311
 Song, Qi III-172
 Sonka, Milan III-172
 Sørensen, Lauge I-37, III-595
 Sosna, J. III-73
 Sotiras, Aristeidis II-676
 Sparks, Rachel III-658
 Srinivasan, L. III-1
 St. John, Paul III-359
 Staib, Lawrence H. I-53
 Staring, Marius III-367
 Stephenson, Diane I-308
 Stiegler, Philipp I-45
 Stoeck, Christian T. I-418
 Stoel, Berend C. III-367
 Stojkovic, Branislav II-530
 Stolk, Jan III-367
 Stoyanov, Danail I-77, I-275, III-261
 Studholme, Colin I-574, II-339, II-355
 Styner, Martin I-690
 Subramanian, Kalpathi R. II-505
 Subramanian, Navneeth I-510
 Suehling, Michael I-28
 Sukno, Federico M. I-468
 Summers, Ronald M. III-89
 Sun, Bo I-357, III-49
 Sun, Hui I-468
 Sun, Ying I-493
 Sundar, Hari III-287
 Szewczyk, Jérôme III-432
 Sznitman, Raphael III-465
 Tagare, Hemant D. I-502
 Taimouri, Vahid I-332, III-489
 Tait, Diana I-69
 Tam, Roger III-17
 Tamir, Idit III-457
 Tang, Hui III-97
 Tansella, M. II-177
 Tardif, Christine L. III-181
 Tavakoli, M. I-300
 Taylor, Andrew I-460
 Taylor, Charles A. I-426, III-375
 Taylor, Russell III-303, III-465, III-521
 Taylor, Zeike A. II-388
 Teboul, Olivier III-189
 ter Haar Romeny, Bart I-175
 Thirion, Bertrand I-200, I-550, II-241
 Thiruvenkadam, Sheshadri I-510
 Thomas, P. II-185
 Thompson, Paul M. II-169,
 II-323, III-147
 Tian, Zhen I-143
 Tiwari, Pallavi III-666
 Tobon-Gomez, Catalina I-468
 Toga, Arthur W. I-357, III-49

- Tong, Gregory III-81
 Toomre, Derek II-538
 Tori, Fufa. L. III-97
 Toussaint, Nicolas I-418
 Tscheliessnigg, Karlheinz I-45
 Tsekos, N. III-351
 Tsiamyrtzis, P. III-351
 Tsin, Yanghai III-113, III-287
 Tsymbal, Alexey I-460
 Tu, Zhuowen III-81, III-147
 Tucholka, Alan II-217, II-347
 Türetken, Engin I-291
 Tustison, Nicholas J. I-324

 Unal, Gozde III-137
 Uneri, Ali III-303

 Vaillant, Régis III-440
 van den Bouwhuijsen, Quirijn J.A.
 III-97
 van der Graaff, Maaïke II-249
 van der Lijn, Fedde II-101
 van der Lugt, Aad II-101, III-97
 van Ginneken, Bram II-396,
 III-603, III-650
 Van Leemput, Koen II-151
 Van Meir, V. I-615
 van Noorden, Sander II-249
 van Onkelen, Robbert S. III-97
 van Velsen, Evert F.S. II-101
 van Vliet, Lucas I-167, II-249, III-97
 van Walsum, Theo I-452, III-97
 Varoquaux, Gaël I-200, II-241
 Vavylonis, Dimitrios I-86
 Vecerova, Jaromira II-530
 Vegas-Sanchez-Ferrero, G. I-518
 Vembar, M. I-400
 Vemuri, Baba C. I-682, III-65
 Venkataraman, Archana I-191
 Vercauteren, Tom II-480, II-514
 Verma, Ragini I-558, I-631
 Vernooij, Meike W. II-101
 Vialard, François-Xavier II-610
 Viergever, M.A. II-396
 Vikal, Siddharth III-383
 Visentini-Scarzanella, Marco I-77,
 I-275, III-261
 Vitanovski, Dime I-218, I-460
 Vitiello, Valentina III-229
 Voigt, Ingmar I-218

 von Tengg-Kobligh, Hendrik I-444
 Voros, Szilard I-484
 Vorperian, Hourii K. III-505
 Vos, Frans I-167, II-249
 Vrooman, Henri A. II-101

 Wachinger, Christian II-26
 Wächter, I. I-400, I-526
 Wählby, C. III-634
 Wallace, Michael B. II-480
 Walther, Thomas I-375, I-476
 Wan, Jing III-611
 Wang, Changhong III-473
 Wang, Chaohui III-189
 Wang, Defeng III-538
 Wang, Hongzhi I-468, III-105
 Wang, Lei II-266
 Wang, Linwei I-159
 Wang, Peng III-269
 Wang, Qian II-570, II-618, II-684
 Wang, Song I-349
 Wang, Yang I-218
 Wang, Yaping II-133
 Wang, Ying II-160
 Ward, Aaron D. II-17, III-17, III-213
 Ward, Rabab K. III-513
 Warfield, Simon K. II-109, III-25
 Washko, George R. III-163
 Wassermann, Demian I-631
 Watson, Y. III-121
 Weber, Marc-André II-151
 Wedlake, Chris III-205, III-408
 Weese, J. I-400, I-526
 Wein, Wolfgang I-243
 Weiner, Michael III-105
 Weizman, Lior I-103
 Weldelesassie, Yonas T. I-582
 Wells III, William M. II-554
 Wen, Xu I-283
 Westin, Carl-Fredrik I-191, I-657, I-674,
 II-225, II-233, III-163
 Whalen, Stephen II-225
 Wheeler-Kingshott, Claudia A.M. I-623
 Whitcomb, Louis L. III-383
 Wiener, Michael III-529
 Wiles, Andrew D. III-587
 Willinger, Rémy II-235
 Winter, C. II-185
 Witteman, Jacqueline C.M. III-97
 Wolf, Lior III-642

- Wollstein, Gadi I-1
 Wolz, Robin II-610, III-1
 Wong, Kelvin K. III-473
 Wong, Ken C.L. I-159
 Wong, Stephen T.C. III-392, III-473
 Wong, Tsz Wai II-323
 Woods, Roger P. I-357
 Wörz, Stefan I-444
 Wu, Binbin III-521
 Wu, Guorong II-306, II-570,
 II-618, II-684
 Wu, Xiaodong III-172
 Wu, Xunlei II-274

 Xia, James J. III-278
 Xiao, Changyan III-367
 Xie, Jun II-472
 Xie, Sheng Quan III-481
 Xie, Weixin III-392
 Xie, Yuchen I-682
 Xie, Zhiyong I-308, III-57
 Xiong, Guanglei I-426, III-375
 Xiong, Zixiang III-278
 Xu, Helen III-383
 Xu, Jianrong I-332
 Xu, Jinhui II-530
 Xu, Jun III-197
 Xu, Wei Liang III-481
 Xu, Xun I-28
 Xue, Zhong II-133, III-392, III-473

 Yan, Pingkun II-42
 Yan, Shuicheng II-522
 Yang, Dewen I-308
 Yang, Guang-Zhong I-69, I-77, I-275,
 II-437, II-496, III-229, III-245, III-261,
 III-319
 Yang, Jianfei II-143
 Yang, Qian II-538
 Yang, Yihong I-598
 Yap, Pew-Thian II-306
 Yau, Shing-Tung II-323, III-538
 Yedidya, Tamir III-642
 Yelnik, Jérôme II-217

 Yeniaras, E. III-351
 Yigitsoy, Mehmet II-26
 Yin, Youbing II-578
 Yin, Zhaozheng I-209
 Yip, Michael C. II-626
 Yoo, Terry S. II-505
 Yotter, Rachel Aine II-169
 Young, Alistair I-383
 Yu, Stella X. I-119
 Yuhashi, Kazuhito II-50
 Yurgelun-Todd, Deborah II-363
 Yushkevich, Paul A. I-468, III-105

 Zachow, S. I-227
 Zamanyan, Alen I-357
 Zeitz, Michael J. II-530
 Zeng, Wei III-538
 Zhan, Yiqiang I-19
 Zhang, Degang II-143
 Zhang, Heye I-159
 Zhang, Hui I-640, II-404
 Zhang, Pei II-635
 Zhang, Shaoting I-135
 Zhang, Tianhao II-257
 Zhang, Tuo II-143, II-412
 Zhao, Ting II-472
 Zheng, Yefeng I-375, I-476
 Zheng, Yuanjie I-324
 Zhou, Luping II-266
 Zhou, S. Kevin I-28, I-95, I-460,
 I-476, III-269
 Zhou, Xiang I-19
 Zhou, Xiaobo III-278
 Zhou, Y. III-351
 Zhou, You II-50
 Zhu, Hongtu I-690, II-274, II-298
 Zhu, Jiong I-332
 Zhu, Mengchen II-429
 Zhu, Peihong II-363
 Zhu, Ying III-287
 Zhu, Yongding II-530
 Zhuang, X. I-435
 Zikic, Darko II-586
 Zouhar, Alexander III-113, III-555